COMPUTATIONAL INVESTIGATIONS INTO INTERACTIONS BETWEEN METAL IONS AND NUCLEIC ACIDS

BRIANA BOYCHUK
Bachelor of Science, Brandon University, 2017

A thesis submitted
in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

THEORETICAL AND COMPUTATIONAL SCIENCE

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

COMPUTATIONAL INVESTIGATIONS INTO INTERACTIONS BETWEEN METAL IONS AND NUCLEIC ACIDS

BRIANA BOYCHUK

Date of Defense: May 13, 2025

Dr. Stacey Wetmore Supervisor	Professor	Ph.D.
Dr. Marc Roussel Thesis Examination Committee Member	Professor	Ph.D.
Dr. Paul Hayes Thesis Examination Committee Member	Professor	Ph.D.
Dr. Georg Schreckenbach External Examiner University of Manitoba Winnipeg, Manitoba	Professor	Ph.D.
Dr. Michael Gerken Chair, Thesis Examination committee	Professor	Ph.D.

DEDICATION

This thesis is dedicated to my late Uncle David, who ignited my lifelong interest in science and sparked my love for *Star Trek*. His love and support over the years have been a constant source of encouragement in my ongoing pursuit of scientific discovery.

ABSTRACT

Metal ion contamination is an ongoing global crisis that negatively affects human and environmental health. As a result, metal ion detection and remediation techniques are required. Nucleic acids are promising tools as biosensors for metal detection. Although nucleic acid sensors have been designed for some metal ions (e.g., Cs^{+}, Pb^{2+}), there are metal ion contaminants that require sensing and remediation for which no sensors currently exist. Specifically, with increasing mining and usage of lithium in lithium-ion batteries, rising levels of lithium ions are present in the environment. To design nucleic acid sensing and remediation platforms, it is important to understand how metal ions interact with nucleic acids. Computer modeling can provide valuable insights into structural and binding information of metal–nucleic-acid interactions, thereby aiding the rational design of nucleic-acid-based sensing and remediation solutions for metal ions. This thesis uses different computational techniques, including ab initio (CCSD(T)/CBS), density functional theory (DFT), molecular dynamics (MD), and ab initio molecular dynamics (AIMD) to identify reliable methods to describe metal–nucleic-acid interactions, understand how metal ions interact with nucleic acids, and explore modifications that can capture lithium. The information gained from this thesis lays the groundwork for designing a lithium-specific nucleic acid platform for lithium extraction, which can also support future development of a nucleic acid sensor for lithium. Additionally, insights gained can be applied to designing sensing and extraction platforms for other metal ions, and aid development of therapeutics and materials.

CONTRIBUTIONS OF AUTHORS AND PREFACE

Chapters 2, 3, and 5 are published, and permissions to include these works in this thesis have been obtained from those publishers.

Chapter 2 – **Boychuk, B. T. A.**; Jeong, R. Y. E; Wetmore, S. D.* Assessment of the Accuracy of DFT-Predicted Li^+–Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408.

Chapter 3 – **Boychuk, B. T. A.**; Meyer, S. P.; Wetmore, S. D.* Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* **2023**, *11*.

Chapter 4 – **Boychuk, B. T. A.**; Meyer, S. P.; Wetmore, S. D.* Building a Data Set of Group II Metal–Nucleic Acid Interactions and Evaluating the Performance of Density Functional Theory Methods. Unpublished.

Chapter 5 – **Boychuk, B. T. A.**; Wetmore, S. D.* Assessment of Density Functional Theory Methods for the Structural Prediction of Transition and Post-Transition Metal–Nucleic Acid Complexes. *J. Chem. Theory Comput.* **2023**, *19* (15), 5273-5288.

Chapter 6 – **Boychuk, B. T. A.**; Wagner, L. W.; Lea, M. J.; Wetmore, S. D.* A DFT Investigation into the Structure and Strength of Hydrated Li^+–Nucleic Acid Interactions. Unpublished.

Chapter 7 – **Boychuk, B. T. A.**; Deak, T. K.; Wetmore, S. D.* Manipulating Metal Binding to Nucleic Acids Through Chemical Modification: A Molecular Dynamics Simulation Study of PNA. Unpublished.

Chapter 8 – **Boychuk, B. T. A.**; Wetmore, S. D.* Investigating a Crown Ether–Nucleic Acid Adduct for Targeting Lithium: A Density Functional Theory and Ab Initio Molecular Dynamics Approach. Unpublished.

For all projects, I was involved in project conceptualization, methodology, data curation, formal analysis, investigation, validation, visualization, and writing (original draft, and review and editing). S. D. Wetmore was involved in project conceptualization, methodology, project administration, formal analysis, funding acquisition, resources, supervision, and writing (reviewing and editing). The details outlining the contributions of calculations performed in each chapter are provided below. Chapters 5 and 8 are not discussed below since I conducted all calculations for those chapters. All published chapters have been reformatted according to thesis requirements.

For Chapter 2, I performed all optimization, frequency, CCSD(T)/CBS, and XDM calculations. R. Y. E Jeong contributed to project investigation and data curation by carrying out most of the DFT single-point calculations, in addition to reviewing and editing the manuscript. For Chapter 3, I conducted all calculations, and S. P. Meyer carried out all data extraction and organization, as well as reviewing the manuscript. In Chapter 4, I performed all calculations, except for DFT calculations on complexes involving Be^{2+}. S. P. Meyer performed single-point calculations using DFT methods on Be^{2+} complexes, and carried out data extraction and organization for all complexes. For Chapter 6, L. W. Wagner performed DFT calculations on the implicit and explicit data of Li^+–G/P with 0 to 5 Li^+-coordinated water molecules. I conducted implicit and explicit calculations of hydrated Li^+ at all nucleobase sites. M. J. Lea carried out calculations on the hydrated Li^+–WCF base pair complexes. For Chapter 7, T. K. Deak

performed simulations on Li^+–DNA systems for the GC-rich sequence in each salt concentration (100, 250, 500, and 750 mM). Generative AI was only used to suggest synonyms during the writing process.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Stacey Wetmore for supporting my research interests and providing valuable insights. I have learned a great deal about scientific research from her and grown as a scientist under her supervision. Thank you to my committee members, Dr. Marc Roussel and Dr. Paul Hayes for your guidance, support, and valuable feedback over the years. I would also like to thank Dr. Michael Gerken for chairing my defense. I would like to express my gratitude to Dr. Georg Schreckenbach for serving as my external examiner and offering his valuable time to participate in my thesis defense. I greatly appreciate the funding from NSERC and the UofL that I have received over the course of my graduate journey, which allowed me to pursue this degree.

Overall, my graduate journey has been shaped by the remarkable people I have had the privilege to work with. First, I want to thank the past students I collaborated with, including Rebecca Jeong, Luke Wagner, Mark Lea, Jackson Laurie, Brooklyn Palmer, Sarah Meyer, Megan Ryan, and Trinity Deak. Next, I would like to express how incredibly grateful I am to have worked with many amazing members of the lab, both past and present. Specifically, I would like to thank current and past post-doctoral fellows (Dr. Mohamed Aboelnga, Dr. Preethi Seelam, Dr. James McFarlane, Dr. Indu Negi, Dr. Austin Pounder, and Dr. Basel Mansour), graduate students (Priya Bhutani, Dr. Stefan Lenz, Dr. Katie Wilson, Nathania Takyi, Prakash Chukka, Cynthia Fonderson, Dr. Ryan Kung, Dr. Rajwinder Kaur, Rebecca Jeong, Dylan Nikkel, Makay Murray, Mark Lea, Angela Frederickson, Ridwan Tajudeen, and Umer Yaqoob). It has

been an incredible experience working with all of you; thank you for the valuable insights over the years, the laughs, the fun, and the memories.

Outside of lab, I was fortunate to have met some amazing people. Davinder Dhalla, Elizabeth Trofimenkoff, and Elaura Lilienthal, thank you all for your friendship and the fun times. I would also like to thank JT and Jin from Coffee Company, your coffee always tastes great and kept me going when I needed it most. I would also like to thank my Zumba (Gabrielle) and Interval Training (Elizabeth) instructors, you both kept my physical and mental health from collapsing.

There are a few people who I would like to thank specifically. First, I would like to thank Dr. Preethi Seelam and Dr. Indu Negi for being strong lighthouses at shore for me, I am always appreciative of your insight and support. Rebecca Jeong and Rajwinder Kaur, words cannot even begin to describe how much our friendship over the years has meant to me, thank you for your kindness and support.

To my family, my mother, thank you for your unwavering love and support of me over the years. Finally, to my partner, Edgar Schuldeis, thank you for joining me on this journey, always supporting my dreams, and your unconditional love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D-IR	two-dimensional infrared spectroscopy
A	adenine
AIMD	ab initio molecular dynamics
ALS	amyotrophic lateral sclerosis
AMPE	average mean percent error
AMUE	average mean unsigned error
ARMSD	average root-mean-squared deviation
ARMSE	average root-mean-squared error
B	bridging
BE	binding energy
BJ	Becke–Johnson
BSSE	basis set superposition error
C	cytosine
CBS	complete basis set limit
CCSD(T)	coupled cluster with singles, doubles, and perturbative triples
CE	crown ether
CEC	crown ether cytosine
CECBP	crown ether cytosine base pair model
CECNBP	crown ether cytosine nucleoside base pair model
CECPBP	crown ether cytosine peptide base pair model
CI	Configuration Interaction
CPCM	conductor-like polarizable continuum model
CSD	Cambridge structural database
DFT	density functional theory
DNA	deoxyribonucleic acid
DZ	double-zeta

ECP	effective core potential
FOD	fractional occupation number weighted density
G	guanine
GGA	general gradient approximation
GH	global hybrid
HF	Hartree–Fock
IEF-PCM	integral equation formulism-polarizable continuum model
IRMPD	infrared multiple photon dissociation
LDA	local density approximation
Lithium	lithium ion
LNA	locked nucleic acid
Max	maximum
Metal	metal ion
MD	molecular dynamics
miRNA	microRNA
MM	molecular mechanics
MPE	mean percent error
MUE	mean unsigned error
NB	non-bridging
NFOD	number of fractional occupation differences
NL	nonlocal
NMR	nuclear magnetic resonance
P	phosphate moiety
PE	percent error
PDB	protein data bank
PNA	peptide nucleic acid
PNA-FISH	PNA probe coupled with fluorescence in situ hybridization

QM	quantum mechanics
QZ	quadruple-zeta
RDF	radial distribution function
RMSD	root-mean-square deviation
RMSE	root-mean-square error
RNA	ribonucleic acid
RSH	range-separated hybrid
SELEX	systematic evolution of ligands by exponential enrichment
SDD	Stuttgart Dresden
T	thymine
TCID	threshold collision-induced dissociation
TMP	trimethyl phosphate
TEP	triethyl phosphate
TZ	triple-zeta
U	uracil
UE	unsigned error
VV10	Vydrov and Van Voorhis
WCF	Watson-Crick-Franklin
XDM	exchange-hole dipole moment
ZORA	zeroth-order regular approximation

Chapter 1: Introduction

1.1. General Overview

Metal ion (simply referred to as 'metal' throughout this thesis) contamination is an ongoing global challenge that negatively affects our health and environment. Therefore, it is important to develop metal detection and extraction techniques to monitor and remediate metal contamination. Among metals, lithium is an emerging contaminant in the environment, largely due to the production and disposal of lithium-ion batteries.[1] Nucleic acid sensing and extraction platforms are an attractive green onsite measurement technique for detecting and remediating metals.[2] Many nucleic acid sensing platforms have been designed to date.[2-4] However, there remains limitations with some sensors, primarily due to metal competition effects,[2, 5] structural instability in different environments,[6] and inability to pass through cellular membranes to sense metal contaminants in the cell.[2, 5, 6] Furthermore, there are challenges with the design process of nucleic acid sensors such as tuning the binding affinity of sensors for high sensitivity and selectivity,[2, 7] which is largely due to limitations in our understanding of how metals interact with nucleic acids. While there have been an abundance of experimental studies exploring metal–nucleic acid interactions,[8-12] gaps remain in the literature on the structural and binding interactions between metals and specific nucleic acid binding sites, with these limitations also extending to lithium. To address these gaps in the literature, computational modeling can provide atomic-level insights into structural and energetic information about metal–nucleic acid interactions, which have been otherwise challenging to obtain experimentally. This thesis uses computer modeling to understand how a broad range of metals interact with nucleic acids, an provide direction on the rational design of a nucleic acid extraction platform to

target lithium, with the potential to extend this platform into developing a future lithium sensor.

1.1. Metal Contamination

Metals are of indispensable value to our lives due to their utilization in biology, materials, and biotechnologies. In biology, metals are required by many proteins as co-factors to aid protein folding and maintain proper function in biological processes such as catalyzing chemical reactions within the body.[13] For example, Zn^{2+} is required to fold zinc finger proteins, which play important roles in gene regulation,[14] while Mg^{2+} or Mn^{2+} is required by some endonucleases to catalyze phosphodiester bond cleavage to repair damaged DNA.[15] In materials, nucleic acid–metal composites have been constructed as green catalysts.[16] In biotechnologies, a nucleic acid nanoarchitecture that enhances antibiotic efficiency for treating tuberculosis is formed through Fe^{2+}-assisted coordination-driven assembly.[17]

Although metals are important in these different areas, excess levels of metals contaminate the environment through industrial processes, such as mining, battery manufacturing, and agricultural activities.[18, 19] As a result, metals can bioaccumulate in ecosystems, our food sources, and the human body, leading to adverse effects on the environment and human health. Excess levels of metals in the environment can disrupt ecosystems by altering redox activity,[20] reducing microbial and plant growth,[20] as well as causing deformities and organ damage in organisms.[21] Plants are particularly susceptible to lithium toxicity, which reduces water retention and cellular membrane integrity.[22] In humans, an overabundance of metals can cause neurological disorders,[19, 23, 24] renal

dysfunction,[23, 24] cardiovascular disease,[23, 24] and various cancers.[19, 23, 24] As a particular example, high concentrations of lithium in the body can result in renal and cerebellar dysfunction.[25] Since metal contamination negatively affects human and ecosystem health, metal detection and extraction platforms are required for metal contaminants and serve as an extraction method for environmental remediation. Nucleic acids are a viable solution for developing metal sensing and extraction technologies, which are discussed in more detail in the next section.

1.3. Nucleic Acids as Metal Sensors

Metal sensing platforms are useful analytical tools that provide a cost-effective complement to current instrumental techniques (e.g., atomic absorption spectroscopy and inductively coupled plasma mass spectrometry) for detecting metal contaminants in the human body, our food sources, and drinking water. There have been various types of metal sensing platforms developed using metal-organic frameworks, proteins, or antibodies.[26-28] However, most of these approaches experience limitations, such as complicated syntheses procedures, low sensitivity or specificity, and a narrow range of environmental conditions where the material remains functional. Alternatively, nucleic acids offer several advantages as metal sensing and extraction platforms, which include high specificity for metals, multiple binding sites, high structural stability, and ease of chemical modification.[2, 29] Additionally, nucleic acids consist of chemically diverse structural components, which enable the formation of unique 3D structures that can create metal-specific binding pockets.[2]

Several experimental studies have worked towards designing nucleic acid sensors to detect metals, primarily using the systematic evolution of ligands by exponential enrichment (SELEX).[30, 31] While multiple sensors have been designed to detect metals (e.g., Cs^{+}, Pb^{2+}, Hg^{2+}),[30] there are drawbacks with existing sensors, including limited specificity for a given metal and difficult to isolate high-quality nucleic acid probes.[2, 7] For example, a nucleic acid sensor was developed to selectively bind Mg^{2+} over a strong competitor (Pb^{2+}), but this sensor experiences interference in the presence of Ca^{2+}.[2] Another nucleic acid sensor initially identified to be specific for Pb^{2+} also shows activity in the presence of other metals such as Zn^{2+} and Mg^{2+}.[2, 32] Although many studies have exploited nucleic acids for metal sensing,[2] these platforms can be tuned and utilized for metal capture, aiding environmental remediation efforts. Understanding how the structures of nucleic acids are impacted by the presence of metals can provide insights that can guide the rational design of nucleic acid sensing and extraction platforms for specific metal capture. The subsequent section discusses nucleic acid structures that enable diverse functionality for metal–nucleic acid sensors.

1.4. Nucleic Acid Structure

Nucleic acids are biomolecules that consist of long chains of repeating subunits called nucleotides. Nucleotides are composed of three basic components: a negatively charged phosphate backbone, a sugar, and a nitrogenous base (nucleobase). The sugar is deoxyribose in DNA or ribose in RNA. There are four canonical nucleobases, with three found in both DNA and RNA, namely adenine (A), guanine (G), and cytosine (C), while the fourth nucleobase is thymine (T) in DNA and uracil (U) in RNA (Figure 1.1). These

nucleobases can form complementary pairs through Watson-Crick-Franklin hydrogen bonding (G:C and A:T in DNA or A:U in RNA) to create different types 3D architectures. In nature, canonical DNA forms a double-stranded duplex, while RNA typically exists in the single-stranded form, with complementary base pairing resulting in folding of RNA into structures such as hairpins and loops. There are two grooves in DNA duplexes, known as the major and minor groove (Figure 1.2), with the major groove width being wider (~12 Å), and the minor groove width being narrower (~6 Å). In contrast, the major groove in RNA duplexes is narrower (~4 Å), while the minor groove is wider (~11 Å).[33-35]

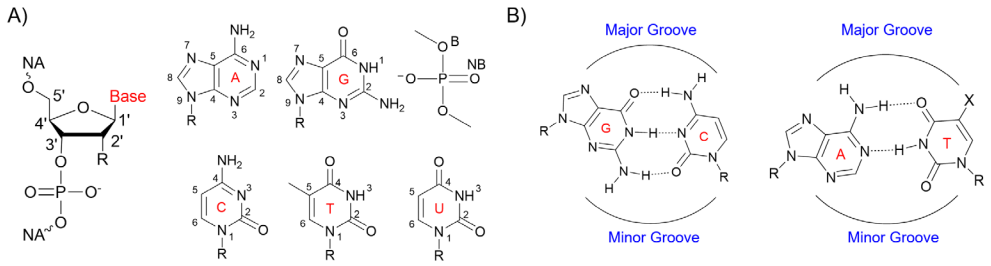


Figure 1.1. A) The nucleotide component in nucleic acids, consisting of the phosphate moiety (B = bridging oxygens and NB = nonbridging oxygens), nucleobase, and sugar (R = H/OH in DNA/RNA). B) Complementary Watson-Crick-Franklin base pairs, where R = A:T (X = CH_3) in DNA and A:U (X = H) in RNA.

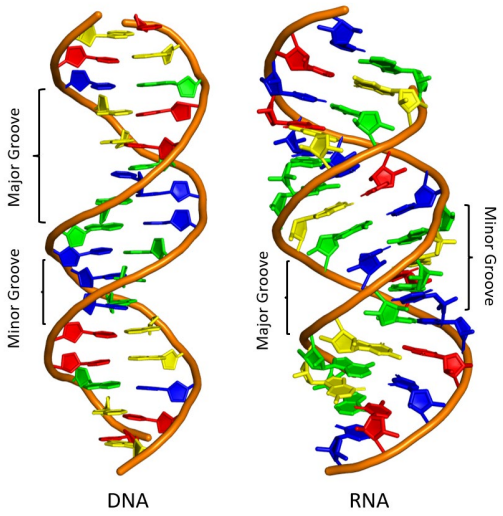


Figure 1.2. Major and minor groove widths of DNA and RNA.

The structural components in nucleic acids permit the formation of unique structural motifs with specialized functions. For example, RNA contains a 2′-OH group on the sugar subunit, which permits RNA to fold into diverse structures (e.g., pseudoknots, loops, or hairpins) that are utilized in gene expression, pharmaceuticals, and biosensing.[36-38] Although DNA typically adopts a duplex formation, double-stranded DNA can form four-way junctions that have been employed as logic gates for cancer detection and as therapeutics.[39] DNA can also form hybrid duplexes with RNA to yield functional nucleic acids like DNAzymes[40] and aptamers,[41] which can be used as biosensors or therapeutics to treat diseases.[4, 40, 41] For example, the L-shaped fold of the 8–17 DNAzyme (Figure 1.3) permits Pb^{2+} binding with greater specificity than other divalent metals.[42] Another example is the Cd^{2+}-specific aptamer (Figure 1.3), which contains a binding loop that adopts a compact and double-twisted arrangement.[43] Nucleic acids can also assemble into G-quadruplexes (Figure 1.3) that are stabilized by K^{+}, with G-quadruplexes adapted as sensors

for targeting K^+.[44] Overall, the folding patterns and single-stranded or double-stranded nature of nucleic acids, which are influenced by variations in the backbone, nucleobases, and sugar, result in unique functions, highlighting the potential nucleic acids have towards developing biosensing and extraction platforms with enhanced metal specificity.

Figure 1.3. Examples of nucleic acid sensors designed to target specific metals, namely the Cd^{2+}-specific (Cd-2-1) aptamer (left), Pb^{2+}-sensing (8–17) DNAzyme (middle), and K^+-specific thrombin binding aptamer (right), highlighting the structural diversity of nucleic acids and associated unique functions.

1.5. Chemical Modifications

The functionality of metal–nucleic acid sensors can be further enhanced by introducing chemical modifications to the backbone, nucleobase, or sugar (Figure 1.4). Chemical alterations of nucleic acids can change the metal binding environment. For example, addition of bulky hydrophobic groups can influence stacking interactions and hydrogen bonding with surrounding nucleotides to create metal-specific pockets.[45] There are various modifications that can be introduced that range from small (e.g., 2′-fluoro, 2′–OMe, phosphorothioate) to large (e.g., O6-pyridyloxobutyl guanine, 3-fluoro-6-

methylaniline) alterations (Figure 1.4), each yielding unique functional outcomes, such as catalysis or specific binding depending on the modification type and attachment site.[46-48] For example, a phosphorothioate modification to the backbone in the Ce13d DNAzyme results in site-specific backbone cleavage in the presence of thiophilic metals.[49] Alteration of G nucleobases within a G-quadruplex by incorporating imidazole has altered the metal-specificity of the nucleic acid from K^+ towards transition metals (i.e., Cu^{2+}, Zn^{2+}, Co^{2+}, Ni^{2+}).[50] While modifications to the backbone and nucleobases are more effective for targeting metals than modifications to the sugar, alterations to the sugar component can enhance nucleic acid stability and further increase the stability in the presence of specific metals.[51] For example, locked nucleic acids (LNA), which feature a bridge connecting C4′ to O2′ and effectively locking the sugar conformation to increase the structural stability of nucleic acids,[52] are utilized to selectively target siRNA for cancer detection.[52] Another type of LNA modification featuring *N*,*N*-bis(2-pyridylmethyl)-β-alanyl-2′-amino-LNA has been shown to further enhance nucleic acid stability in the presence of Cu^{2+}, highlighting the importance of chemical modifications to tune structural and functional properties of nucleic acids,[51] which could be used in pharmaceutical and biosensor development. While there have been many studies that have explored modified nucleic acids binding to metals, there are other modifications that have yet to be investigated that show potential for tuning metal binding affinity. For example, crown ethers are widely used to selectively capture alkali metals, with the 18-crown-6 attachment to nucleic acids binding Na^+ that produces a current signal, enabling the detection of abasic sites in DNA.[53] Peptide nucleic acids (PNA) contain a neutral, peptide-based backbone is resistant to enzyme degradation and experiences less electrostatic repulsion than in DNA or RNA.[54] Although PNA has not been explored for binding to metals, its structural and chemical properties demonstrate potential for

developing nucleic acid sensing and extraction platforms that function in different environments. Overall, chemically altering nucleic acids enables the development of improved nucleic acid sensors for metal detection. To design nucleic acid sensing and extraction platforms with precise metal recognition capabilities, it is important to understand how metals interact with canonical and modified nucleic acids. The subsequent section discusses previous literature that has uncovered key information about the properties of metal–nucleic acid interactions.

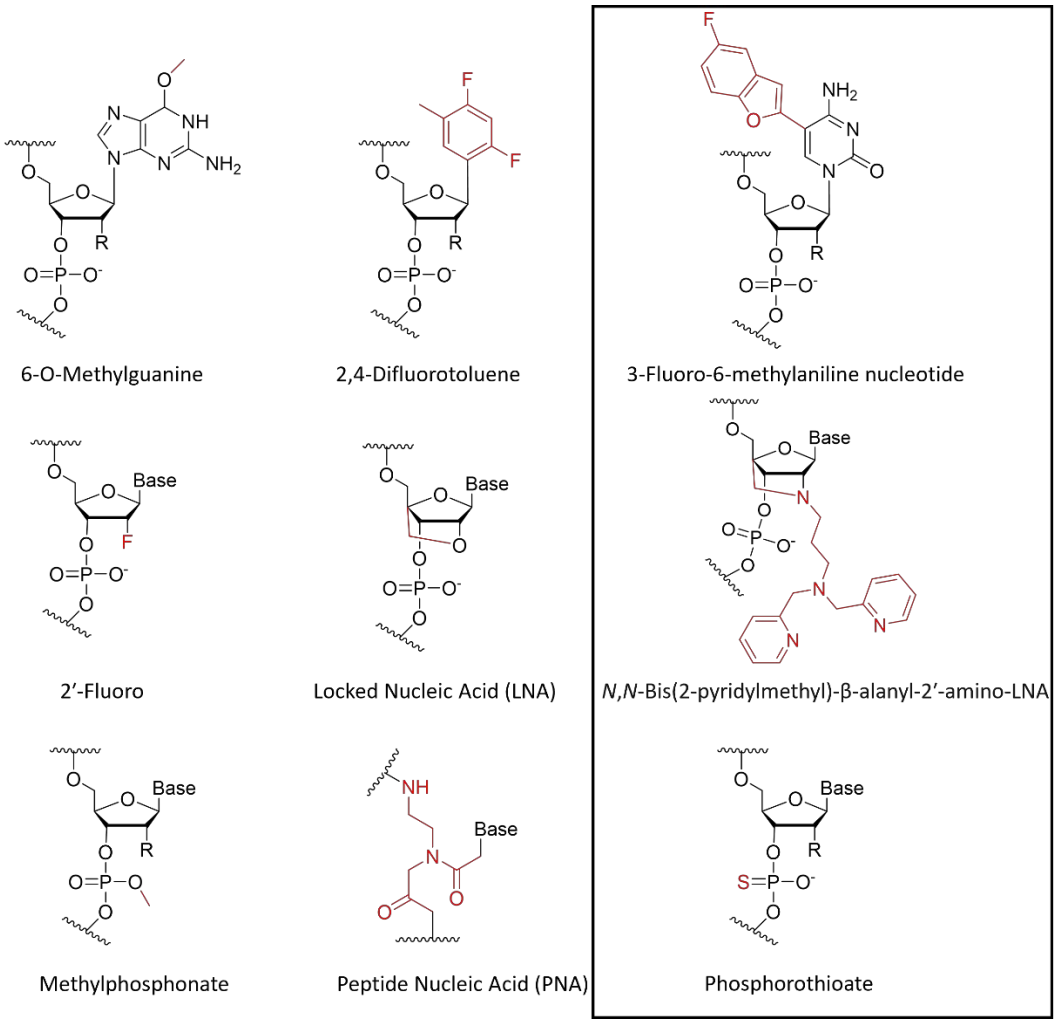


Figure 1.4. Examples of chemical modifications that can be made to the nucleobases (top), sugar (middle), and backbone (bottom) that can impart new function with metals. The modifications outlined in the box have been studied with metals.[49, 51, 55]

1.6. Experimental Studies Investigating Metal–Nucleic Acid Interactions

Multiple experimental studies have used mass spectrometry,[12] X-ray crystallography,[56, 57] and threshold collision-induced dissociation (TCID)[8] to uncover the structural and thermodynamic properties of metals interacting with canonical nucleic

acids.[2, 58] From these studies, structural details, bond dissociation energies, kinetic stability, spectroscopic properties, and stoichiometries of select metal–nucleic acid complexes have been identified.

There have been several experimental studies using TCID and the kinetic method approach to determine the binding strengths of nucleic acid subcomponents (i.e., A, C, G, T, U, or the phosphate moiety) to lithium.[8, 59-63] Despite these studies, there is missing data for G and the phosphate moiety.[8, 59-63] Additionally, no structural data was obtained as TCID experiments do not provide structural information. Similar gaps exist in the literature on nucleic acids interacting with other metals (e.g., alkali and alkaline earth metals, transition metals).[58, 64]

Another important factor to consider in nucleic acids binding to metals is the role water plays in these interactions. Metal-coordinated waters can stabilize nucleic acids by forming hydrogen bonds with nucleic acids,[64-66] highlighting the importance of understanding how water influences metal–nucleic acid interactions. Two-dimensional infrared (2D-IR) spectroscopic[67, 68] and infrared multiphoton dissociation (IRMPD) spectroscopic[69-73] studies have been utilized to investigate the dynamics of hydrated metal–nucleic acid interactions by providing vibrational and fragmentation data. In the case of Li^+, multiple IRMPD experiments have been performed to uncover the vibrational spectra of hydrated Li^+–nucleobase interactions.[69-71, 73] However, these experiments do not provide 3D structural information, such as the metal-coordination geometry or the nucleobase binding site. As a result, key details into how water affects metal–nucleic acid interactions remain unclear.

Aside from canonical nucleic acids, several experimental studies have synthesized chemical modifications to nucleic acids to tune nucleic acid functionality for binding metals in applications such as materials, therapeutics, and biosensing.[2, 74-76] As a specific example, modifying G:C base pairs in DNA to P:C (P = 1H-imidazo[4,5-f][1,10]phenanthroline) base pairs enables selective binding of Ag^+ to form metal-mediated base pairs (P-Ag-C).[76, 77] These interactions were studied using different techniques such as circular dichroism and NMR spectroscopy,[76, 77] with this information used for designing new materials and biosensors to detect Ag^+.[76, 77] Nevertheless, only select nucleic acid modifications with specific metals have been investigated, with no modifications explored with the purpose to capture lithium. There are clear gaps in the literature regarding how certain metals interact with nucleic acids.

Computational modeling is a powerful tool that can provide atomic-level insights into how metals interact with canonical and modified nucleic acids, in addition to how water affects these interactions. The next section discusses in more detail computational investigations that have been done on metal–nucleic acid interactions.

1.7. Computational Studies Investigating Metal–Nucleic Acid Interactions

1.7.1. Ab Initio and Density Functional Theory Studies

To complement experimental studies and fill gaps in the literature regarding how metals interact with nucleic acids, computational chemistry has proven useful. Prior to this thesis, several computational studies had uncovered fundamental structural and energetic information on metal–nucleic acid interactions.[59, 60, 62, 78-91] The most accurate

computational approach for investigating metal–nucleic acid interactions is coupled-cluster with single, double, and perturbative triples excitation at the complete basis set limit (CCSD(T)/CBS). CCSD(T)/CBS offers highly accurate binding strengths of molecules; however, this method is computationally expensive, limiting its application to small systems (~30–50 atoms). Therefore, CCSD(T)/CBS has been confined to studying bare metals interacting with nucleic acid subcomponents in the gas phase. To overcome these limitations, computationally cheaper methods such as density functional theory (DFT) have been used to provide detailed information on the bond lengths, coordination numbers, geometries, as well as binding strengths, of metal complexes.

Several DFT studies have investigated the structural and binding information of metal–nucleic acid interactions,[62, 69, 82, 85, 87-90] For example, B3LYP was used to explore the binding of metals (Li^+, Na^+, K^+, Be^{2+}, Mg^{2+}, Ca^{2+}) to all possible nucleobase sites.[82] Other studies have modeled the phosphate backbone using trimethyl phosphate (TMP) and triethyl phosphate (TEP) to gain insights into the structures and binding strengths of group I metals interacting with the phosphate moiety.[62, 91] However, these models do not account for the naturally occurring charge of nucleic acids. Separate studies have investigated transition metal (Sc^+–Cu^+, Cd^{2+}, Cu^{2+}) binding to G or A using B3LYP.[78, 81, 90] Despite numerous DFT studies on metal–nucleic acid interactions, not all nucleobase binding sites have been explored for each metal, and the phosphate backbone model remains incomplete. Due to the approximations used in DFT, it is important to validate its reliability against experimental data or highly accurate computational methods like CCSD(T).[86] The accuracy of certain functionals used to investigate metal–nucleic acid interactions is uncertain, which

raises concerns about the reliability of the reported results. This underscores the need for assessing the accuracy of DFT results.

1.7.2. Molecular Dynamics Simulations

To investigate the impact of solvent and the surrounding nucleic acid environment on metal–nucleic acid complexes, molecular dynamics (MD) simulations can be used. Due to the limited number of atoms that can be explored using quantum chemical techniques, MD simulations that utilize classical physics capture the time-dependent behavior of metal–nucleic acid interactions in solvated environments, revealing the role of hydration and ion mobility. These approaches have proven effective for understanding metal binding preferences with nucleic acids and the structural impact metals have on nucleic acids.[92-99] For example, an MD study on group II metals (Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}) interacting with DNA uncovered that Mg^{2+} retains its hydration shell and primarily interacts with DNA through Mg^{2+}-coordinated waters (i.e., an outer-sphere complex), while Ca^{2+}, Sr^{2+}, or Ba^{2+} tend to lose a water molecule and bind directly to DNA (i.e., an inner-sphere complex).[92] Another MD study explored metal competition among Li^{+}, Na^{+}, K^{+}, and Mg^{2+} to bind to DNA, identifying that these metals form outer-shell contacts through metal-coordinated waters to multiple nucleic acid binding sites, and larger populations of metals surround G nucleobases.[93] However, this study only looked at total interactions (within 6 Å) and did not dissect the data into direct or indirect metal coordination to nucleic acids.[93] MD simulations exploring Li^{+}, Na^{+}, K^{+}, Cs^{+}, Ca^{2+}, Sr^{2+}, and Ba^{2+} binding patterns to RNA or DNA identified that metal interactions with RNA differ from DNA.[94] Specifically, DNA generally displaces a single metal-coordinated water, while more metal-coordinated waters

are displaced by RNA, permitting metals to form more direct contacts with RNA and create folded enclosures.[94] Overall, while MD simulations have been used to investigate metal–nucleic acid interactions in water,[92-99] these studies focused on either the structural impact of metal binding on the nucleic acid or the general distribution in metal binding location (e.g., major versus minor groove, nucleobases versus phosphate backbone). Thus, the impact of water on the structure and binding in metal–nucleic acid complexes remain only partially explored. Additionally, the application of classical force fields in MD simulations can oversimplify the treatment of polarization and electronic effects of hydrated metal–nucleic acid interactions,[100] with higher-level computational techniques (e.g., DFT) important for uncovering structure and thermodynamics of these interactions.

Overall, despite the valuable information gained from these studies, uncertainties remain, such as how metals bind to specific nucleic acid sites (i.e., phosphate backbone versus nucleobases), as well as the role of the metal identity or oxidation on metal–nucleic acid interactions. Furthermore, the impact of water and solvation effects on the structural and binding information of metal–nucleic acid interactions remain unclear. Additionally, the impact of chemical modifications on metal–nucleic acid interactions is not understood. The focus of this thesis is to use a combination of CCSD(T)/CBS, DFT, and MD computational techniques to gain a fundamental understanding of how metals interact with canonical and modified nucleic acids.

1.8. Scope of thesis

This thesis uses multiple computational techniques to gain fundamental structural and binding information of metal–nucleic acid interactions to work towards developing

nucleic acid extraction platforms for capturing metals, where these platforms can be used for future research into designing nucleic acid metal sensors. While there have been computational studies on Li^+–nucleic acid interactions,[59, 60, 62] the accuracy of methods used to study these complexes is in question and not all nucleic acid components were explored. Therefore, Chapter 2 uses small models featuring Li^+ interacting with each binding site of each nucleobase (A, C, G, T, or U) and the phosphate moiety to uncover structural and binding information using the gold-standard computational technique (CCSD(T)/CBS). Subsequently, the performance of more cost-effective approaches (DFT) is assessed for the ability to reproduce the CCSD(T)/CBS values and thereby identify reliable methods that can be used in later chapters. This chapter uncovers fundamental structural and binding information of metal–nucleic acid interactions and underscores the importance of testing the reliability of computational methods for providing an accurate description of metal–nucleic acid properties. The DFT method (MN15) identified in Chapter 2 is used for other chapters in this thesis that involve Li^+–nucleic acid interactions.

Chapter 3 investigates the impact of the metal identity on metal–nucleic acid interactions. Like Chapter 2, small models featuring group I metals (Li^+, Na^+, K^+, Rb^+, and Cs^+) interacting with each nucleic acid component (A, C, G, T, U or the phosphate moiety) are used to uncover the structural and binding information of these interactions. Although accurate DFT methods for Li^+ were identified in Chapter 2, reliable DFT methods for describing group I metal–nucleic acid interactions remain uncertain. Using the same approach as Chapter 2, the most robust DFT methods for consistently and accurately describing group I metal–nucleic acid interactions are identified against CCSD(T)/CBS results. Furthermore, Chapter 3 compares different metals interacting with nucleic acid

binding sites, which provide important insights for designing nucleic acid extraction platforms with enhanced metal specificity.

Chapter 4 further explores reliable methodologies for investigating how group II metals interact with nucleic acids. Using the methodology from Chapter 3, the structural and binding information of group II metal–nucleic acid interactions is elucidated. Subsequently, the same DFT methods explored in Chapter 3 are tested against the CCSD(T)/CBS results on group II metal–nucleic acid interactions. Overall, this chapter provides fundamental structural and binding information using a high level of theory for group II metal–nucleic acid interactions in addition to outlining accurate methodologies for studying these interactions.

Chapter 5 expands from group I and II metals to the computationally challenging transition and post-transition metals. Transition and post-transition metals also contaminate the environment[101] and pose a risk to human health,[102] which highlights the need for nucleic acid sensing and extraction technologies to detect and capture these metals. However, transition and post-transition metals can be challenging to describe using computational techniques due to factors such as varying oxidation and spin states, relativistic effects, and ligand-dependent behavior.[103, 104] This chapter aims to take the first step towards designing nucleic acid sensors for transition and post-transition metal detection by assessing the performance of DFT methods for reliably describing transition and post-transition metal–nucleic-acid complexes. This chapter identifies reliable DFT methods for studying the transition and post-transition metal–nucleic-acid complexes, enabling future research on these interactions to develop nucleic acid sensing and extraction platforms.

Chapter 6 moves into applying the methods identified from Chapter 2 to explore Li^+–nucleic acid interactions since Li^+ is a rising environmental contaminant that requires

sensing and extraction strategies for remediation.[1] In Chapter 6, the role of water on the structural and binding information of Li^+–nucleic acid interactions is investigated. Building on the models and methodology from Chapter 2, this chapter investigates hydrated structures of Li^+–nucleic acid complexes, varying the number of water molecules coordinated to Li^+. The structural and thermodynamic information gained from this chapter provides a fundamental understanding of how water affects Li^+–nucleic acid interactions and how to computationally model these interactions in solvent by using a combination of implicit and explicit solvent.

Prior chapters have focused on understanding metal interactions with canonical nucleic acids, while Chapter 7 introduces sugar–backbone modified nucleic acid (PNA). The negatively charged phosphate backbone is replaced with a neutral peptide backbone to investigate how metal binding changes from canonical DNA to PNA, with this modification motivated by findings from previous chapters showing that Li^+ binds to multiple sites in DNA (phosphate moiety and nucleobases) with different geometries. This chapter uses MD simulations on DNA and PNA duplexes in the presence of Li^+ or Na^+ in different salt concentrations to compare how these metals bind in a duplex environment, with the results from this chapter emphasizing the utility of PNA to drive Li^+ binding to more selective nucleic acid regions. This chapter provides insights into the structural stability of PNA in varying environmental conditions, which is important information for future development of nucleic acid metal sensing and extraction technologies.

Chapter 8 uses a combined computational approach of DFT calculations and ab initio molecular dynamics simulations (AIMD) to investigate a di-hydroxy crown-ether-modified cytosine for Li^+ capture. Specifically, DFT calculations performed on nucleobase and base pairs reveal the complexation abilities of Li^+ to the crown ether compared to

canonical G, the strongest nucleobase binding site in canonical DNA. This chapter provides the first insight into how Li^+ binds to a crown-ether-modified nucleobase. Altogether, Chapters 7 and 8 how to design a nucleic acid extraction platform for Li^+ by combining sugar–backbone and nucleobase modifications. Finally, Chapter 9 discusses the key conclusions gained from this thesis and possible future directions for these projects.

1.9. References

(1) Bolan, N.; Hoang, S. A.; Tanveer, M.; Wang, L.; Bolan, S.; Sooriyakumar, P.; Robinson, B.; Wijesekara, H.; Wijesooriya, M.; Keerthanan, S.; et al. From Mine to Mind and Mobiles – Lithium Contamination and Its Risk Management. *Environ. Pollut.* **2021**, *290*, 118067.

(2) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(3) Zimmermann, A. C.; White, I. M.; Kahn, J. D. Nucleic Acid-Cleaving Catalytic DNA for Sensing and Therapeutics. *Talanta* **2020**, *211*, 120709.

(4) Yu, X.; Zhang, S.; Guo, W.; Li, B.; Yang, Y.; Xie, B.; Li, K.; Zhang, L. Recent Advances on Functional Nucleic-Acid Biosensors. *Sensors* **2021**, *21* (21), 7109.

(5) Lake, R. J.; Yang, Z.; Zhang, J.; Lu, Y. DNAzymes as Activity-Based Sensors for Metal Ions: Recent Applications, Demonstrated Advantages, Current Challenges, and Future Directions. *Acc. Chem. Res.* **2019**, *52* (12), 3275-3286.

(6) Liu, L.; Xiong, M.; Rong, Q.; Zhang, M.; Zhang, X.-B. Nucleic Acid Sensors in Vivo: Challenges and Opportunities. *VIEW* **2023**, *4* (3), 20220064.

(7) Yu, H.; Alkhamis, O.; Canoura, J.; Liu, Y.; Xiao, Y. Advances and Challenges in Small-Molecule DNA Aptamer Isolation, Characterization, and Sensor Development. *Angew. Chem. Int. Ed.* **2021**, *60* (31), 16800-16823.

(8) Rodgers, M. T.; Armentrout, P. B. Cationic Noncovalent Interactions: Energetics and Periodic Trends. *Chem. Rev.* **2016**, *116* (9), 5642-5687.

(9) Sigel, R. K. O.; Sigel, H. A Stability Concept for Metal Ion Coordination to Single-Stranded Nucleic Acids and Affinities of Individual Sites. *Acc. Chem. Res.* **2010**, *43* (7), 974-984.

(10) Nakano, S.-i.; Miyoshi, D.; Sugimoto, N. Effects of Molecular Crowding on the Structures, Interactions, and Functions of Nucleic Acids. *Chem. Rev.* **2014**, *114* (5), 2733-2758.

(11) Metal–Nucleic Acid Interactions. *J. Biol. Inorg. Chem* **2014**, *19* (2), 815-832.

(12) Largy, E.; König, A.; Ghosh, A.; Ghosh, D.; Benabou, S.; Rosu, F.; Gabelica, V. Mass Spectrometry of Nucleic Acid Noncovalent Complexes. *Chem. Rev.* **2022**, *122* (8), 7720-7839.

(13) Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M. Targeting Metalloenzymes for Therapeutic Intervention. *Chem. Rev.* **2019**, *119* (2), 1323-1455.

(14) Cassandri, M.; Smirnov, A.; Novelli, F.; Pitolli, C.; Agostini, M.; Malewicz, M.; Melino, G.; Raschellà, G. Zinc-Finger Proteins in Health and Disease. *Cell Death Discov.* **2017**, *3* (1), 17071.

(15) Zhao, H.; Lin, Z.; Lynn, A. Y.; Varnado, B.; Beutler, J. A.; Murelli, R. P.; Le Grice, Stuart F. J.; Tang, L. Two Distinct Modes of Metal Ion Binding in the Nuclease Active Site of a Viral DNA-Packaging Terminase: Insight into the Two-Metal-Ion Catalytic Mechanism. *Nucleic Acids Res.* **2015**, *43* (22), 11003-11016.

(16) Hadian-Dehkordi, L.; Abdi, Z.; Zarei, A.; Aghahosseini, H.; Mohammadi, Z.; Farokhi, A.; Yazdely, T. M.; Nouri, S. H.; Hosseini, M.; Ramazani, A.; et al. DNA-Metal Composites as Bioinspired Catalysts in Chemical Reactions. *Coord. Chem. Rev.* **2024**, *505*, 215635.

(17) Zhang, K.; Sun, H.; Gao, X.; Liang, T.; Gao, X.; Song, H.; Guan, C.; Sun, Z.; Zhang, J. Metal-Coordination-Driven Self-Assembly of DNA Nanohybrids for Robust Intracellular Anti-Tuberculosis Therapy. *New J. Chem.* **2025**, *49* (7), 2651-2655.

(18) Robinson, B. H.; Yalamanchali, R.; Reiser, R.; Dickinson, N. M. Lithium as an Emerging Environmental Contaminant: Mobility in the Soil-Plant System. *Chemosphere* **2018**, *197*, 1-6.

(19) Xu, W.; Jin, Y.; Zeng, G. Introduction of Heavy Metals Contamination in the Water and Soil: A Review on Source, Toxicity and Remediation Methods. *Green Chem. Lett. Rev.* **2024**, *17* (1), 2404235.

(20) Adamczyk-Szabela, D.; Wolf, W. M. The Impact of Soil pH on Heavy Metals Uptake and Photosynthesis Efficiency in Melissa Officinalis, Taraxacum Officinalis, Ocimum Basilicum. *Molecules* **2022**, *27* (15), 4671.

(21) Shahjahan, M.; Taslima, K.; Rahman, M. S.; Al-Emran, M.; Alam, S. I.; Faggio, C. Effects of Heavy Metals on Fish Physiology – a Review. *Chemosphere* **2022**, *300*, 134519.

(22) Shahzad, B.; Mughal, M. N.; Tanveer, M.; Gupta, D.; Abbas, G. Is Lithium Biologically an Important or Toxic Element to Living Organisms? An Overview. *Environ. Sci. Pollut. Res. Int.* **2017**, *24* (1), 103-115.

(23) Mitra, S.; Chakraborty, A. J.; Tareq, A. M.; Emran, T. B.; Nainu, F.; Khusro, A.; Idris, A. M.; Khandaker, M. U.; Osman, H.; Alhumaydhi, F. A.; et al. Impact of Heavy Metals on the Environment and Human Health: Novel Therapeutic Insights to Counter the Toxicity. *J. King Saud Univ. Sci.* **2022**, *34* (3), 101865.

(24) Jomova, K.; Alomar, S. Y.; Nepovimova, E.; Kuca, K.; Valko, M. Heavy Metals: Toxicity and Human Health Effects. *Arch. Toxicol.* **2025**, *99* (1), 153-209.

(25) Young, W. Review of Lithium Effects on Brain and Blood. *Cell Transplant.* **2009**, *18* (9), 951-975.

(26) Nam, K. H. Fluorescent Protein-Based Metal Biosensors. *Chemosensors* **2023**, *11* (4), 216.

(27) Blake, D. A.; Jones, R. M.; Blake, R. C.; Pavlov, A. R.; Darwish, I. A.; Yu, H. Antibody-Based Sensors for Heavy Metal Ions. *Biosens. Bioelectron.* **2001**, *16* (9), 799-809.

(28) Wu, T.; Gao, X.-j.; Ge, F.; Zheng, H.-g. Metal–Organic Frameworks (MOFs) as Fluorescence Sensors: Principles, Development and Prospects. *CrystEngComm* **2022**, *24* (45), 7881-7901.

(29) Ma, L.; Liu, J. Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology. *iScience* **2020**, *23* (1).

(30) Du, Y.; Dong, S. Nucleic Acid Biosensors: Recent Advances and Perspectives. *Anal. Chem.* **2017**, *89* (1), 189-215.

(31) Ahmad, K.; Kumar, P.; Kim, H. Recent Progress in Lead Free Tin-Halide Perovskite Materials Based Solar Cells Via Scaps Based Numerical Simulation. *ChemistrySelect* **2024**, *9* (31), e202402044.

(32) Breaker, R. R.; Joyce, G. F. A DNA Enzyme That Cleaves RNA. *Chem. Biol.* **1994**, *1* (4), 223-229.

(33) Frezza, E.; Laage, D.; Duboué-Dijon, E. Molecular Origin of Distinct Hydration Dynamics in Double Helical DNA and RNA Sequences. *J. Phys. Chem. Lett.* **2024**, *15* (16), 4351-4358.

(34) Lietzke, S. E.; Barnes, C. L.; Berglund, J. A.; Kundrot, C. E. The Structure of an RNA Dodecamer Shows How Tandem U–U Base Pairs Increase the Range of Stable RNA Structures and the Diversity of Recognition Sites. *Structure* **1996**, *4* (8), 917-930.

(35) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA Dodecamer: Conformation and Dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78* (4), 2179-2183.

(36) Yang, R.; Cui, J. Advances and Applications of RNA Vaccines in Tumor Treatment. *Mol. Cancer* **2024**, *23* (1), 226.

(37) Wang, Q.; Wang, Z.; He, Y.; Xiong, B.; Li, Y.; Wang, F. Chemical and Structural Modification of RNA-Cleaving DNAzymes for Efficient Biosensing and Biomedical Applications. *TrAC-Trend Anal. Chem.* **2023**, *159*, 116910.

(38) Li, H.; Lee, T.; Dziubla, T.; Pi, F.; Guo, S.; Xu, J.; Li, C.; Haque, F.; Liang, X.-J.; Guo, P. RNA as a Stable Polymer to Build Controllable and Defined Nanostructures for Material and Biomedical Applications. *Nano Today* **2015**, *10* (5), 631-655.

(39) Lake, A.; Shang, S.; Kolpashchikov, D. M. Molecular Logic Gates Connected through DNA Four-Way Junctions. *Angew. Chem. Int. Ed.* **2010**, *49* (26), 4459-4462.

(40) Xiao, L.; Zhao, Y.; Yang, M.; Luan, G.; Du, T.; Deng, S.; Jia, X. A Promising Nucleic Acid Therapy Drug: DNAzymes and Its Delivery System. *Front. Mol. Biosci.* **2023**, *10*.

(41) Ni, S.; Zhuo, Z.; Pan, Y.; Yu, Y.; Li, F.; Liu, J.; Wang, L.; Wu, X.; Li, D.; Wan, Y.; et al. Recent Progress in Aptamer Discoveries and Modifications for Therapeutic Applications. *ACS Appl. Mater. Interfaces* **2021**, *13* (8), 9500-9519.

(42) Mazumdar, D.; Nagraj, N.; Kim, H.-K.; Meng, X.; Brown, A. K.; Sun, Q.; Li, W.; Lu, Y. Activity, Folding and Z-DNA Formation of the 8-17 DNAzyme in the Presence of Monovalent Ions. *J. Am. Chem. Soc.* **2009**, *131* (15), 5506-5515.

(43) Liu, H.; Gao, Y.; Mathivanan, J.; Armour-Garb, Z.; Shao, Z.; Zhang, Y.; Zhao, X.; Shao, Q.; Zhang, W.; Yang, J.; et al. Crystal Structures and Identification of Novel Cd^{2+}-Specific DNA Aptamer. *Nucleic Acids Res.* **2023**, *51* (9), 4625-4636.

(44) Yang, X.; Li, T.; Li, B.; Wang, E. Potassium-Sensitive G-quadruplex DNA for Sensitive Visible Potassium Detection. *Analyst* **2010**, *135* (1), 71-75.

(45) Takezawa, Y.; Kanemaru, D.; Kudo, N.; Shionoya, M. Phenanthroline-Modified DNA Three-Way Junction Structures Stabilized by Interstrand 3 : 1 Metal Complexation. *Dalton Trans.* **2023**, *52* (32), 11025-11029.

(46) Duffy, K.; Arangundy-Franklin, S.; Holliger, P. Modified Nucleic Acids: Replication, Evolution, and Next-Generation Therapeutics. *BMC Biology* **2020**, *18* (1), 112.

(47) McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. Recent Progress in Non-Native Nucleic Acid Modifications. *Chem. Soc. Rev.* **2021**, *50* (8), 5126-5164.

(48) Berdis, A. Nucleobase-Modified Nucleosides and Nucleotides: Applications in Biochemistry, Synthetic Biology, and Drug Discovery. *Front. Chem.* **2022**, *10*, 1051525.

(49) Saran, R.; Huang, Z.; Liu, J. Phosphorothioate Nucleic Acids for Probing Metal Binding, Biosensing and Nanotechnology. *Coord. Chem. Rev.* **2021**, *428*, 213624.

(50) Punt, Philip M.; Clever, G. H. Imidazole-Modified G-quadruplex DNA as Metal-Triggered Peroxidase. *Chem. Sci.* **2019**, *10* (8), 2513-2518.

(51) Hirva, P.; Nielsen, A.; Bond, A. D.; McKenzie, C. J. Potential Cross-Linking Transition Metal Complexes (M = Ni, Cu, Zn) in the Ligand-Modified LNA Duplexes. *J. Phys. Chem. B* **2010**, *114* (36), 11942-11948.

(52) Kamali, M. J.; Salehi, M.; Fatemi, S.; Moradi, F.; Khoshghiafeh, A.; Ahmadifard, M. Locked Nucleic Acid (LNA): A Modern Approach to Cancer Diagnosis and Treatment. *Exp. Cell Res.* **2023**, *423* (1), 113442.

(53) An, N.; Fleming, A. M.; White, H. S.; Burrows, C. J. Crown Ether–Electrolyte Interactions Permit Nanopore Detection of Individual DNA Abasic Sites in Single Molecules. *Proc. Natl. Acad. Sci.* **2012**, *109* (29), 11504-11509.

(54) Dean, D. A. Peptide Nucleic Acids: Versatile Tools for Gene Therapy Strategies. *Adv. Drug Deliv. Rev.* **2000**, *44* (2), 81-95.

(55) Kotammagari, T. K.; Saleh, L. Y.; Lönnberg, T. Organometallic Modification Confers Oligonucleotides New Functionalities. *Chem. Comm.* **2024**, *60* (23), 3118-3128.

(56) Tereshko, V.; Wilds, C. J.; Minasov, G.; Prakash, T. P.; Maier, M. A.; Howard, A.; Wawrzak, Z.; Manoharan, M.; Egli, M. Detection of Alkali Metal Ions in DNA Crystals Using State-of-the-Art X-Ray Diffraction Experiments. *Nucleic Acids Res.* **2001**, *29* (5), 1208-1215.

(57) Leonarski, F.; D'Ascenzo, L.; Auffinger, P. Binding of Metals to Purine N7 Nitrogen Atoms and Implications for Nucleic Acids: A Csd Survey. *Inorg. Chim. Acta* **2016**, *452*, 82-89.

(58) Hangan, A. C.; Oprean, L. S.; Dican, L.; Procopciuc, L. M.; Sevastre, B.; Lucaciu, R. L. Metal-Based Drug–DNA Interactions and Analytical Determination Methods. *Molecules* **2024**, *29* (18), 4361.

(59) Cerda, B. A.; Wesdemiotis, C. Li^+, Na^+, and K^+ Binding to the DNA and RNA Nucleobases. Bond Energies and Attachment Sites from the Dissociation of Metal Ion-Bound Heterodimers. *J. Am. Chem. Soc.* **1996**, *118* (47), 11884-11892.

(60) Rodgers, M. T.; Armentrout, P. B. Noncovalent Interactions of Nucleic Acid Bases (Uracil, Thymine, and Adenine) with Alkali Metal Ions. Threshold Collision-Induced Dissociation and Theoretical Studies. *J. Am. Chem. Soc.* **2000**, *122* (35), 8548-8558.

(61) Yang, Z.; Rodgers, M. T. Tautomerization in the Formation and Collision-Induced Dissociation of Alkali Metal Cation-Cytosine Complexes. *Phys. Chem. Chem. Phys.* **2012**, *14* (13), 4517-4526.

(62) Ruan, C.; Huang, H.; Rodgers, M. T. A Simple Model for Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase: Alkali Metal Cations and Trimethyl Phosphate. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 305-314.

(63) Ruan, C.; Huang, H.; Rodgers, M. T. Modeling Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase Via Alkali Metal Cation-Triethyl Phosphate Complexes. *J. Phys. Chem. A* **2007**, *111* (51), 13521-13527.

(64) Ward, W. L.; Plakos, K.; DeRose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chem. Rev.* **2014**, *114* (8), 4318-4342.

(65) Anastassopoulou, J. Metal–DNA Interactions. *J. Mol. Struct.* **2003**, *651-653*, 19-26.

(66) Ferré-D'Amaré, A. R.; Winkler, W. C. The Roles of Metal Ions in Regulation by Riboswitches. In *Structural and Catalytic Roles of Metal Ions in RNA*, Sigel, A., Sigel, H., Sigel, R. K. O., Sigel, A., Sigel, H., Sigel, R. K. O., Sigel, A., Sigel, H., Sigel, R. K. O., Sigel, A., et al. Eds.; Vol. 9; The Royal Society of Chemistry, 2011; p 0.

(67) Fingerhut, B. P. The Mutual Interactions of RNA, Counterions and Water – Quantifying the Electrostatics at the Phosphate–Water Interface. *Chem. Comm.* **2021**, *57* (96), 12880-12897.

(68) Schauss, J.; Dahms, F.; Fingerhut, B. P.; Elsaesser, T. Phosphate–Magnesium Ion Interactions in Water Probed by Ultrafast Two-Dimensional Infrared Spectroscopy. *J. Phys. Chem. Lett.* **2019**, *10* (2), 238-243.

(69) Gillis, E. A. L.; Rajabi, K.; Fridgen, T. D. Structures of Hydrated Li^+−Thymine and Li^+−Uracil Complexes by IRMPD Spectroscopy in the N−H/O−H Stretching Region. *J. Phys. Chem. A.* **2009**, *113* (5), 824-832.

(70) Gillis, E. A. L.; Fridgen, T. D. The Hydrated Li^+–Adenine–Thymine Complex by IRMPD Spectroscopy in the N–H/O–H Stretching Region. *Int. J. Mass Spectrom.* **2010**, *297* (1), 2-8.

(71) Rajabi, K.; Gillis, E. A. L.; Fridgen, T. D. Structures of Alkali Metal Ion−Adenine Complexes and Hydrated Complexes by IRMPD Spectroscopy and Electronic Structure Calculations. *J. Phys. Chem. A.* **2010**, *114* (10), 3449-3456.

(72) Power, A. A.; Ali, O. Y.; Burt, M. B.; Fridgen, T. D. IRMPD Spectroscopic and Computational Study of Gas Phase $[M(Ura-H)(Ura)]^+$ and $[M(Ura-H)(H_2O)_N]^+$ (M=Sr, Ba; N=1, 2) Complexes. *Int. J. Mass Spectrom.* **2012**, *330-332*, 233-240.

(73) Cheng, R.; Martens, J.; Fridgen, T. D. A Vibrational Spectroscopic and Computational Study of Gaseous Protonated and Alkali Metal Cationized G–C Base Pairs. *Phys. Chem. Chem. Phys.* **2020**, *22* (20), 11546-11557.

(74) Ochoa, S.; Milam, V. T. Modified Nucleic Acids: Expanding the Capabilities of Functional Oligonucleotides. *Molecules* **2020**, *25* (20), 4659.

(75) Yu, Z.; Cowan, J. A. Metal Complexes Promoting Catalytic Cleavage of Nucleic Acids—Biochemical Tools and Therapeutics. *Curr. Opin. Chem. Biol.* **2018**, *43*, 37-42.

(76) Jash, B.; Scharf, P.; Sandmann, N.; Fonseca Guerra, C.; Megger, D. A.; Müller, J. A Metal-Mediated Base Pair That Discriminates between the Canonical Pyrimidine Nucleobases. *Chem. Sci.* **2017**, *8* (2), 1337-1343.

(77) Jash, B.; Müller, J. Concomitant Site-Specific Incorporation of Silver(I) and Mercury(II) Ions into a DNA Duplex. *Chem. Eur. J.* **2018**, *24* (42), 10636-10640.

(78) Bagchi, S.; Mandal, D.; Ghosh, D.; Das, A. K. Density Functional Theory Study of Interaction, Bonding and Affinity of Group Iib Transition Metal Cations with Nucleic Acid Bases. *Chem. Phys.* **2012**, *400*, 108-117.

(79) Rao, J. S.; Dinadayalane, T. C.; Leszczynski, J.; Sastry, G. N. Comprehensive Study on the Solvation of Mono- and Divalent Metal Cations: Li^+, Na^+, K^+, Be^{2+}, Mg^{2+} and Ca^{2+}. *J. Phys. Chem. A.* **2008**, *112* (50), 12944-12953.

(80) Andrushchenko, V.; Bouř, P. Infrared Absorption Detection of Metal Ion-Deoxyguanosine Monophosphate Binding: Experimental and Theoretical Study. *J. Phys. Chem. B* **2009**, *113* (1), 283-291.

(81) Noguera, M.; Branchadell, V.; Constantino, E.; Ríos-Font, R.; Sodupe, M.; Rodríguez-Santiago, L. On the Bonding of First-Row Transition Metal Cations to Guanine and Adenine Nucleobases. *J. Phys. Chem. A* **2007**, *111* (39), 9823-9829.

(82) Zhu, W.; Luo, X.; Puah, C. M.; Tan, X.; Shen, J.; Gu, J.; Chen, K.; Jiang, H. The Multiplicity, Strength, and Nature of the Interaction of Nucleobases with Alkaline and

Alkaline Earth Metal Cations: A Density Functional Theory Investigation. *J. Phys. Chem. A* **2004**, *108* (18), 4008-4018.

(83) Xing, D.; Tan, X.; Chen, X.; Bu, Y. Theoretical Study on the Gas-Phase Acidity of Multiple Sites of Cu^{+}−Adenine and Cu^{2+}−Adenine Complexes. *J. Phys. Chem.* **2008**, *112* (32), 7418-7425.

(84) Burt, M. B.; Fridgen, T. D. Structures and Physical Properties of Gaseous Metal Cationized Biological Ions. *Eur. J. Mass Spectrom.* **2012**, *18* (2), 235-250.

(85) Marino, T.; Mazzuca, D.; Russo, N.; Toscano, M.; Grand, A. On the Interaction of Rubidium and Cesium Mono-, Strontium and Barium Bi-Cations with DNA and RNA Bases. A Theoretical Study. *Int. J. Quantum Chem.* **2010**, *110*, 138-147.

(86) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112* (1), 289-320.

(87) Stasyuk, O. A.; Solà, M.; Swart, M.; Fonseca Guerra, C.; Krygowski, T. M.; Szatylowicz, H. Effect of Alkali Metal Cations on Length and Strength of Hydrogen Bonds in DNA Base Pairs. *ChemPhysChem.* **2020**, *21* (18), 2112-2126.

(88) Amraoui, N. E.; Messaoudi, A.; Hammoutène, D. Copper Ion Cu(I) Interaction Effect on DNA Nucleotides: DFT Study. *Inorg. Chem. Commun.* **2020**, *119*, 108078.

(89) Burda, J. V.; Šponer, J.; Hobza, P. Ab Initio Study of the Interaction of Guanine and Adenine with Various Mono- and Bivalent Metal Cations (Li^{+}, Na^{+}, K^{+}, Rb^{+}, Cs^{+}; Cu^{+}, Ag^{+}, Au^{+}; Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}; Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem. A.* **1996**, *100* (17), 7250-7255.

(90) Xing, D.; Tan, X.; Jiang, X.; Wang, B. Deprotonation Studies of Cu^{+}–Guanine and Cu^{2+}–Guanine Complexes by Theoretical Investigation. *Comp. Theor. Chem.* **2011**, *963* (2), 490-496.

(91) Ruan, C.; Huang, H.; Rodgers, M. T. Modeling Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase Via Alkali Metal Cation-Triethyl Phosphate Complexes. *J. Phys. Chem. A.* **2007**, *111* (51), 13521-13527.

(92) Long, M. P.; Alland, S.; Martin, M. E.; Isborn, C. M. Molecular Dynamics Simulations of Alkaline Earth Metal Ions Binding to DNA Reveal Ion Size and Hydration Effects. *Phys. Chem. Chem. Phys.* **2020**, *22* (10), 5584-5596.

(93) Yoo, J.; Aksimentiev, A. Competitive Binding of Cations to Duplex DNA Revealed through Molecular Dynamics Simulations. *J. Phys. Chem. B.* **2012**, *116* (43), 12946-12954.

(94) Cruz-León, S.; Schwierz, N. RNA Captures More Cations Than DNA: Insights from Molecular Dynamics Simulations. *J. Phys. Chem. B.* **2022**, *126* (43), 8646-8654.

(95) Cruz-León, S.; Vanderlinden, W.; Müller, P.; Forster, T.; Staudt, G.; Lin, Y.-Y.; Lipfert, J.; Schwierz, N. Twisting DNA by Salt. *Nucleic Acids Res.* **2022**, *50* (10), 5726-5738.

(96) Savelyev, A.; MacKerell, A. D. Differential Deformability of the DNA Minor Groove and Altered Bi/Bii Backbone Conformational Equilibrium by the Monovalent Ions Li^+, Na^+, K^+, and Rb^+ Via Water-Mediated Hydrogen Bonding. *J. Chem. Theory Comput.* **2015**, *11* (9), 4473-4485.

(97) Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion–Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8* (4), 1493-1502.

(98) Cesare Marincola, F.; Denisov, V. P.; Halle, B. Competitive Na^+ and Rb^+ Binding in the Minor Groove of DNA. *J. Am. Chem. Soc.* **2004**, *126* (21), 6739-6750.

(99) Xi, K.; Wang, F.-H.; Xiong, G.; Zhang, Z.-L.; Tan, Z.-J. Competitive Binding of Mg^{2+} and Na^+ Ions to Nucleic Acids: From Helices to Tertiary Structures. *Biophys J* **2018**, *114* (8), 1776-1790.

(100) Li, P.; Merz, K. M., Jr. Metal Ion Modeling Using Classical Mechanics. *Chem. Rev.* **2017**, *117* (3), 1564-1686.

(101) Zhang, Y.; Shen, Z.; Zhou, W.; Liu, C.; Li, Y.; Ding, B.; Zhang, P.; Zhang, X.; Zhang, Z. Environmental Problems of Emerging Toxic Metals and Treatment Technology and Methods. *RSC Adv.* **2024**, *14* (50), 37299-37310.

(102) Rodríguez, J.; Mandalunis, P. M. A Review of Metal Exposure and Its Effects on Bone Health. *J. Toxicol.* **2018**, *2018* (1), 4854152.

(103) Maurer, L. R.; Bursch, M.; Grimme, S.; Hansen, A. Assessing Density Functional Theory for Chemically Relevant Open-Shell Transition Metal Reactions. *J. Chem. Theory Comp.* **2021**, *17* (10), 6134-6151.

(104) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *J. Chem. Theory Comput.* **2019**, *15* (6), 3610-3622.

Chapter 2: An Assessment of the Accuracy of DFT Predicted Li^+–Nucleic Acid Binding Energies

Preface: The contents of this chapter have been published [Boychuk, B. T. A.; Jeong, R. Y. E.; Wetmore, S. D. Assessment of the Accuracy of DFT-Predicted Li^+–Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408].

2.1. Introduction

Metal ions are involved in an abundance of vital biological functions such as enzyme catalysis, nucleic acid and protein stabilization, and signal transduction.[1-3] Nevertheless, human exposure to toxic levels of metals is increasing due to activities such as mining, coal burning and waste dumping.[4] Therefore, many works have strived to understand the interactions between metals and biosystems.[5,6] Uncovering the fundamental details of metal interactions with nucleic acids is critical both in terms of understanding the overall role of metals in biology and in terms of the design of new metal detection platforms. Indeed, many unique single-stranded nucleic acids have been successfully constructed to target specific metals with high binding affinity and specificity.[7,8]

Among metals, alkali metals are of central importance to many cellular processes.[3,9] Na^+ and K^+ are the most abundant alkali metals in biosystems, each playing key roles in the human body including, their involvement in blood regulation, osmo-regulation, and sodium and potassium pumps (Na^+/K^+-ATPase).[9] On the other hand, Rb^+ and Cs^+ are nonessential for cells.[10] Rb^+ and Cs^+ undergo competitive binding with K^+ in several biological processes, and are therefore mildly toxic.[11-13] In addition to their biological roles, alkali metals such as Na^+, K^+ and Cs^+ have been selectively detected by nucleic acid metal sensors.[7]

Although also a nonessential metal, unlike Rb^+ and Cs^+, lithium can exhibit beneficial biochemical effects. For example, lithium is a vital component of drugs used to treat bipolar disorders.[14] Furthermore, Li^+ has been reported to upregulate DNA ligase IV activity, an important DNA repair enzyme, resulting in the shielding of retinal neurocytes from ischemic damage.[15] Indeed, studies of hypoxic–ischemic brain injuries revealed that lithium provides long-term protection against DNA damage.[16,17] Another interesting

example is the reported activity of the naturally-occurring hammerhead ribozyme in the presence of Li^+.[18] Nevertheless, there is a narrow range of lithium concentration that is therapeutic,[19] with prolonged exposure linked to cerebellar and renal dysfunction.[20,21] Furthermore, lithium exposure is increasing due to the growing production and improper disposal of lithium–ion batteries,[19,22] which when coupled with its geochemical mobility in the environment results in Li^+ uptake through the food chain.[23] Unlike other alkali metals, no nucleic acid platform has been designed to detect lithium to the best of our knowledge. A strongly binding nucleic acid platform could also provide a potential solution for the growing interest in lithium extraction from oil sands,[24] which would fill the demand for manufacturing lithium–ion batteries and have the added benefit of aiding environmental remediation.

To better understand how Li^+ interacts with complex biosystems and its biological roles, as well as aid the design of new detection or extraction strategies, fundamental information about metal ion binding to nucleic acids is required. In this light, the kinetic method (KM) approach based on gas-phase mass spectrometry has been used to determine the bond dissociation energies of lithium in complex with each of the 5 nucleobases (adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U)).[25,26] However, due to limitations in the accuracy of the resulting data,[26] threshold collision-induced dissociation (TCID) experiments have been performed to re-evaluate the bond dissociation energies of lithium in complex with 4 nucleobases (A, C, T and U),[27,28] as well as trimethyl phosphate (TMP) and triethyl phosphate (TEP).[29,30] While these experiments have provided key insight into the binding affinity of lithium to the individual subcomponents of nucleic acids, accurate data is missing for G and appropriately charged phosphate models, and no structural information is directly available from these experiments.

Complementarily to experimental studies, computational chemistry can provide detailed structural and energetic information about metal–nucleic acid interactions.[31-36] Indeed, previous TCID data was compared to (MP2 or B3LYP) computed interaction energies of lithium bound to various positions in the nucleobases and (TMP and TEP) phosphate models to aid structural assignments.[27,29,30] Other (MP2 and B3LYP) studies have compared computed Li^+–nucleobase interactions, including cation–π interactions, with those for a broad spectrum of monovalent and divalent metal ions[31,36] or those involving different Li^+ species (Li_3^+ and CH_3Li).[32] Through comparison to infrared multiple photon dissociation (IRMPD) spectra, gas-phase B3LYP structural information has been reported for mono and dihydrated Li^+–thymine and uracil,[34] as well as A,[35] complexes. The maintenance of the preferred DFT nucleobase binding position of Li^+ has also been considered for (BLYP) Watson-Crick DNA base pairs,[37] as well as (B3LYP) cyclic G, C, T and U tetrads.[38,39] Finally, molecular dynamics (MD) simulations have been used to identify Li^+ binding sites in DNA duplexes, as well as study Li^+ competition with other mono and divalent metal ions.[40-43] Despite these computational studies providing critical information about the preferred nucleobase binding positions, the accuracy of the DFT computed binding strengths is uncertain at least in part due to incomplete experimental datasets. Furthermore, the reliability of existing force field parameters for Li^+ has been called into question[41] and studies of larger biomolecular complexes in a range of environments are lacking.

Prior to performing large-scale computational studies to understand the behavior of Li^+ in broader biological and nucleic acid-based systems, methods that can accurately and efficiently describe Li^+–nucleic acid interactions must be identified. With this goal in mind, the present work generates a complete dataset of binding strengths for Li^+–nucleic acid

systems using the gold standard CCSD(T) method at the complete basis set (CBS) limit. Specifically, the gas-phase binding energies of lithium complexed to each of the 5 canonical DNA/RNA nucleobases (A, G, C, T, and U, Figure 2.1) and a model phosphate backbone (dimethyl phosphate) are evaluated. This highly accurate data will be useful for fine-tuning force fields for MD simulations that will further advance our understanding of Li^+–nucleic acid binding. Subsequently, the present work assesses the ability of 54 functionals (Table 2.1) that span many functional families, including those incorporating empirical dispersion corrections, in combination with 3 basis sets to reproduce the CCSD(T)/CBS results. The most accurate DFT methods identified in the present work will be ideal for future investigations of lithium interactions in larger nucleic acid systems, with the goal to provide insight into the biological roles of this metal, and the design of novel biosensing and extraction strategies. Our data complements previous benchmark studies focused on lithium polysulfides[44] and Li^+–benzene cation–π interactions[45] to aid the design of lithium–sulfur and lithium–ion batteries.

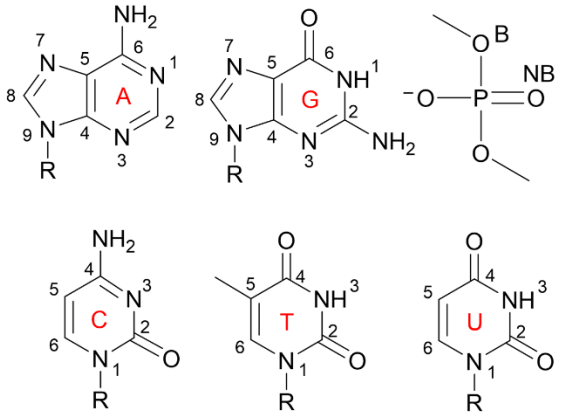


Figure 2.1. Nucleobase and phosphate components considered in this study, with B = bridging and NB = nonbridging oxygens in the dimethyl phosphate model.

2.2. Computational Details

All calculations were performed using Gaussian 16 (Rev. B01).[46] The only exceptions were calculations conducted with revDSD-PBEP86-D3(BJ), ωB97M-V and ωB97X-V, which were carried out using ORCA 4.1.1,[47] and the exchange-hole dipole moment (XDM) dispersion[48-52] corrected functional LC-ωPBE-XDM, which were done using the postG external program.[51,52]

Models were built by initially placing the metal at each potential binding position of the nucleobase (N1 and N6 of A, O6 of G, N3 of A and G, O4 of T and U, N4 of C, and O2 of the pyrimidines) or phosphate (bridging and nonbridging oxygens) nucleic acid component (Figure 2.1). Dimethyl phosphate was used as the phosphate model, which replaces the attached sugars by methyl groups and allows phosphate interactions to be considered in the absence of potential interference from the nucleobase. This model is an improvement over the experimentally implemented TMP[29] and TEP[30] models since dimethyl phosphate accounts for the biologically-relevant charge and permits different (bridging versus non-bridging oxygen) binding locations to be explored. All structures were optimized with MP2(full)/aug-cc-pVTZ and frequency calculations were subsequently performed to confirm stable minima. The geometries of the most stable metal complex for each nucleic acid component were also considered for a selection of DFT functionals.

For each stable minimum, the binding energy was calculated as:

$$E^{BE} = E^{Dimer} - E^{Nucleic\ Acid} - E^{Li^+} \quad (2.1)$$

where E^{BE} is the binding energy of the complex, E^{Dimer} is the dimer energy, and $E^{Nucleic\ Acid}$ and E^{Li^+} are the monomer energies of the nucleic acid component and

Li^+, respectively. The reference binding energies were calculated using CCSD(T) at the CBS limit according to the extrapolation scheme of Helgaker *et al.*[53] Specifically, basis set extrapolation was performed with MP2 using the aug-cc-pVTZ and aug-cc-pVQZ basis sets, followed by CCSD(T)/aug-cc-pVTZ calculations. Subsequently, the total energy was calculated as:

$$E(CCSD(T)/CBS) = E(HF) + E^{corr}(MP2/CBS) + \Delta CCSD(T) \quad (2.2)$$

where $E(HF)$ was evaluated using aug-cc-pVQZ, $E^{corr}(MP2/CBS)$ is the MP2 energy extrapolated to the CBS limit, and $\Delta CCSD(T)$ was calculated as the difference between the MP2 and CCSD(T) energies evaluated with the aug-cc-pVTZ basis set. As done previously in the literature[54-58] and proposed by Helgaker *et al.,*[59] the averages of the counterpoise uncorrected and corrected binding strengths were used as the final benchmark values. This approach addresses convergence concerns when solely using uncorrected or corrected CCSD(T)/CBS binding energies that are known to overestimate or underestimate the true interaction energy, respectively. The counterpoise corrections were evaluated according to the Boys and Bernardi scheme.[60]

In the present study, 39 unique DFT functionals were selected based on the metaphorical 'Jacob's Ladder'[61] (Table 2.1). Furthermore, Grimme's D3[62,63] empirical dispersion correction using Becke–Johnson damping (BJ)[49] was added to 15 functionals. We note that the functionals examined include newer variants by the Head-Gordon group, namely ωB97M-V and ωB97X-V, which contain the Vydrov and Van Voorhis' VV10[64] nonlocal (NL) dispersion correction. Additionally, an exchange-hole dipole moment (XDM) dispersion corrected functional (LC-ωPBE-XDM) was examined. For each functional, three triple-zeta (TZ) basis sets were considered, namely Pople's 6-311+G(3df,2p), Dunning aug-cc-pVTZ[65-67] and Karlsruhe def2-TZVPP.[68] The impact of

basis set expansion to quadruple-zeta (QZ) and truncation to double-zeta (DZ) quality basis sets was also considered for the Dunning and Karlsruhe families. All DFT binding energies were calculated with and without a counterpoise correction using the Boys and Bernardi scheme to determine the impact of the basis set superposition error (BSSE).[60] While we acknowledge that different environments will reduce the magnitude of the predicted binding strengths, all calculations in the present work were performed in the gas phase.

Table 2.1. DFT methods tested in this study for their ability to describe Li^+–nucleic acid interactions.

Family	Functional	D3[a]	%HF[b]	%MP2	Ref.
Double-Hybrid	B2-PLYP	X	53	27	69
	mPW2-PLYP		55	25	70
	PBE-QIDH		69	33.4	71
	PBE0-DH		50	12.5	72
RSH	M11		42.8–100[c]		73
	MN12-SX	X	25–0[c]		74
RSH meta-GGA	ωB97M-V		15		75
RSH GGA	ωB97				76
	ωB97X		16		76
	ωB97X-D		22		77
	ωB97X-V		16.7		78
	HSE06	X	25–0[c]		79-85
	LC-PBE				86,87
	LC-ωPBE	X			88-90
	LC-ωPBE-XDM				52
	CAM-B3LYP	X	19–65[c]		91
GH meta-GGA	M06		27		92
	M06-2X		54		92
	M06-HF		100		93
	BMK		42		94
	MN15		44		95
GH GGA	BH&HLYP		50		96
	SOGGA11-X		40.15		97
	B3PW91	X	20		98
	PBE0	X	25		99
	B3LYP	X	20		98,100
	X3LYP	X	21		101
	O3LYP	X			102
	B97-2		21		103
Local meta-GGA	TPSS	X			104

Table 2.1. Continued.

Family	Functional	D3[a)]	%HF[b)]	%MP2	Ref.
	M11-L				105
	M06-L				106
	MN12-L				107
	MN15-L				108
Local GGA	mPW91				109
	BLYP	X			100,110,111
	BP86	X			100,110
	PBE	X			86
Local LDA	SVWN5				112

[a)]Functionals indicated by an X were considered with and without the D3(BJ) empirical dispersion correction. [b)] Percentage of Hartree–Fock exchange (%HF). [c)] For these RSH functionals, the first value is the %HF at short range, and the second value is the %HF at long range.

2. 3. Results and Discussion

2.3.1. CCSD(T)/CBS Li^+–Nucleic Acid Binding Positions and Strengths

As discussed in the Introduction, both nucleobase (A, G, C, T and U) and phosphate models (Figure 2.1) were considered to determine the preferred lithium–nucleic acid binding sites. Specifically, all potential nucleobase metalation sites were examined, including N1, N3 and N7 of the purines, as well as O6 of G, N6 of A, the N3, N4 and O2 binding sites of C, and the O2 and O4 positions of T and U. The possibility of cation–π interactions was also investigated as previous research indicates their potential important roles in terrestrial environments and protein-ligand interactions.[113,114] For the phosphate model, Li^+ coordination was considered to the two non-bridging oxygen atoms, the two-bridging oxygen atoms or one bridging and one non-bridging oxygen. Minima for all Li^+–nucleic acid complexes characterized in the present work are reported in Figure 2.2, along with the CCSD(T)/CBS binding energies (square brackets) and the relative binding energies for complexes with each nucleic acid component (parentheses).

Among purine metalation sites, lithium preferentially coordinates to the major groove side, with binding being stronger to G (–72.5 kcal/mol) than A (–51.1 kcal/mol). For G, the greatest interaction energy involves bidentate coordination to N7 and O6, which results in a significantly more stable complex than bidentate coordination to N3 and N2 (by 32.0 kcal/mol). This trend correlates with the predicted gas-phase proton affinities (PA) for different G positions,[115] as well as the preference for hard cations like Li^+ for hard bases (oxygen donors) and electron rich sites (i.e., the O6–N7 face of G). For A, four Li^+ complexes were characterized. Two structures involve bidentate coordination with a rotated (nonplanar) N6 amino group and N7 or N3, which are very similar in stability (within 2.3 kcal/mol), which correlates with the similar reported PA for these sites.[115] A

monodentate Li^+–A complex was also identified in which lithium is coordinated to N3 (within 6.9 kcal/mol of the most stable complex). Due to the lack of a strong electron withdrawing group for A compared to G, a cation–π interaction with the 6-membered ring of A is uniquely stabilized (within 19.2 kcal/mol).

Among pyrimidine metalation sites, the strongest binding occurs with C (–68.8 kcal/mol), which arises due to bidentate coordination with N3 and O2, again reflecting the chemical properties of the nucleobase sites and Li^+. A second bidentate coordinated Li^+–C complex involving N3 and a rotated (nonplanar) amino group falls 21.4 kcal/mol higher in energy. Two nearly isoenergetic complexes arise for T and U, with coordination to O4 being slightly preferred over O2 (by 1.2–3.2 kcal/mol) due to the higher proton affinity.[116] An additional cation–π interaction was isolated for T, which is 32.1 kcal/mol less stable than the O4 complex. The isolation of a cation–π complex for T, but not U, agrees with previous studies suggesting that addition of a methyl group to the π ring strengthens cation–π contacts.[117,118]

Overall, when the preferred binding site for each nucleobase is considered, the trend in the Li^+–nucleobase interaction energy decreases as G > C >> A ~ T = U, which follows the trend in the proton affinities of the preferred nucleobase binding sites.[115,116,119] The predicted stronger lithium binding to C compared to A, T or U matches experimental data from TCID.[27,28] Although previous computational literature generally reproduces the CCSD(T)/CBS trend with respect to the binding sites, the difference in the magnitude of the computed DFT (B3LYP) binding strengths can be over 2 kcal/mol, suggesting that there may be room for other functionals to yield improved descriptions of lithium–nucleobase interactions.

In addition to various nucleobase binding sites, lithium forms stable complexes with dimethyl phosphate through the two non-bridging oxygens, two bridging oxygens, or one bridging and one non-bridging oxygen. The strongest binding interaction (–163.6 kcal/mol) occurs when Li^+ coordinates to the highly charged non-bridging oxygen atoms. While the stability of the complex with lithium coordinated to one bridging and one non-bridging oxygen is only 5.5 kcal/mol less stable, coordination to 2 bridging oxygens is far less favorable (by 17.6 kcal/mol). We note that our calculated binding strengths are approximately three times larger than those evaluated using TMP models,[26,29,30] which when coupled with the ability to investigate a wider variety of binding orientations underscores the importance of considering the dimethyl phosphate model. Overall, Li^+ interactions to the phosphate backbone appear to be approximately 2 times greater than the strongest interaction with a nucleobase, which correlates with the increased binding affinity for the phosphate over the nucleobases previously recorded with TCID.[26]

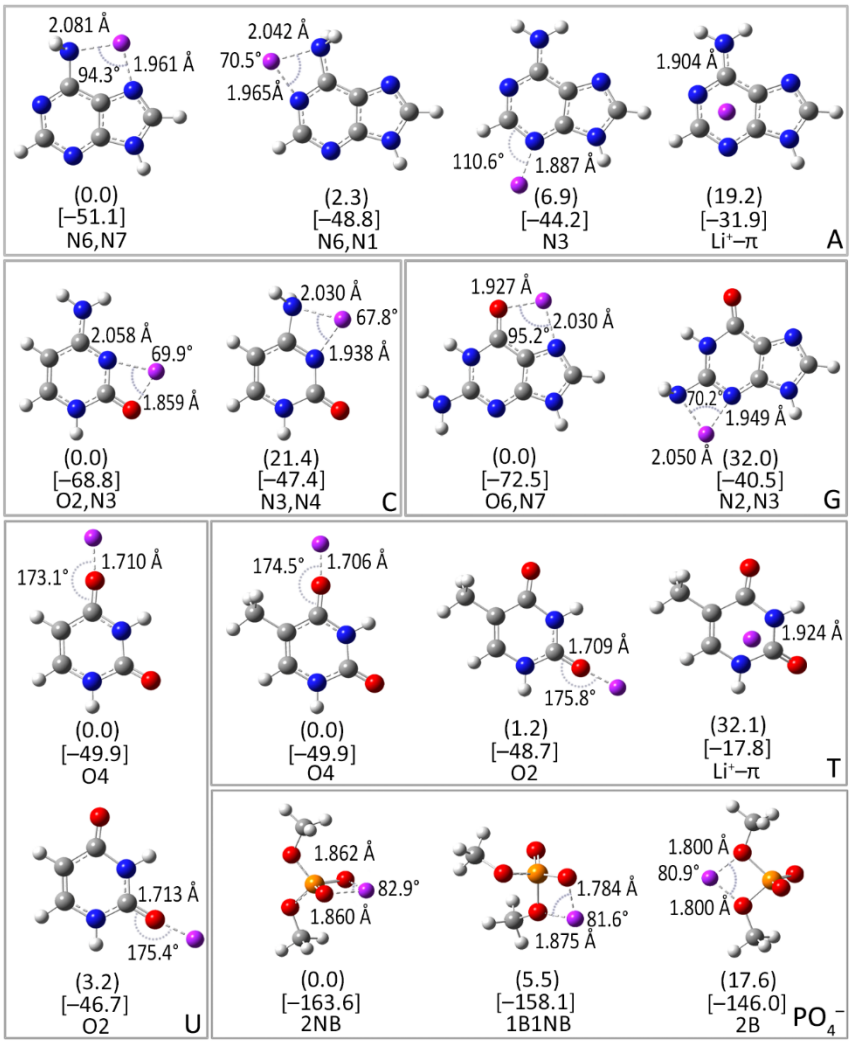


Figure 2.2. CCSD(T)/CBS//MP2/aug-cc-pVTZ characterized complexes between Li^+ and the nucleobases (A, G, C, T and U) or dimethyl phosphate (P). Key structural features (Å and degrees), including bond lengths and angles for direct coordination and metal–centroid distance for cation–π contacts are provided. Half-and-half counterpoise-corrected binding energies [square bracket] and relative energies for a given nucleic acid component (parenthesis) are provided (kcal/mol).

2.3.2. The Accuracy of Counterpoise-Corrected DFT Li^+–Nucleic Acid Binding Strengths

As indicated in the Computational Methods, the ability of 54 functionals (Table 2.1) in combination with 3 basis sets to reproduce the CCSD(T)/CBS Li^+–nucleic acid binding

strengths was evaluated. Signed errors for the DFT binding strengths of the most stable complexes between lithium and nucleic acid components as a function of the functional and basis set are illustrated as a heat map in Figure 2.3. Most combinations of functionals and basis sets overestimate the binding energies of Li^+–nucleic acid complexes as indicated by the prevalence of red shades. Furthermore, there is minor color change vertically, indicating consistency in functional performance across systems and basis sets. In contrast, greater color variation occurs on the horizontal axis, highlighting significant fluctuations in performance between and within different functional families.

To identify the best overall functional(s) for the investigation of Li^+–nucleic acid interactions, the percent errors (PE) for each individual interaction and mean percent errors (MPE) over the nucleobases or the nucleobases and phosphate moiety, as well as the corresponding averages over the three basis sets (AMPE), were evaluated in reference to the CCSD(T)/CBS binding energy (Table 2.2, Figures 2.4, and A.1–A.2, Appendix A). Averaging over basis sets has been utilized in previous benchmark studies to aid identification of functionals that are robust with respect to the basis set.[44] Focusing on these metrics ensures that the descriptions of Li^+–nucleobase and Li^+–phosphate interactions are more equally weighted given the significant difference in the magnitude of the respective binding energies. As a rough guide, a 1% AMPE for the Li^+–nucleobase interactions is equivalent to an ~ 0.5 kcal/mol difference from the CCSD(T)/CBS reference value, while a 1% AMPE is an ~ 1.6 kcal/mol difference for Li^+–phosphate interactions. The unsigned error (UE), mean unsigned error (MUE), and root-mean-square error (RMSE), as well as the corresponding averages (AMUE and ARMSE), are also tabulated and considered in our analysis to identify performance outliers and provide another metric to gauge method accuracy (Table 2.3, Figures A.3–A.4, Appendix A). We note that the trends, and therefore

the major conclusions, based on all metrics are highly similar. We initially focus on the counterpoise-corrected DFT binding energies and organize our discussion in the following subsections according to the metaphorical ‘Jacob’s Ladder’.

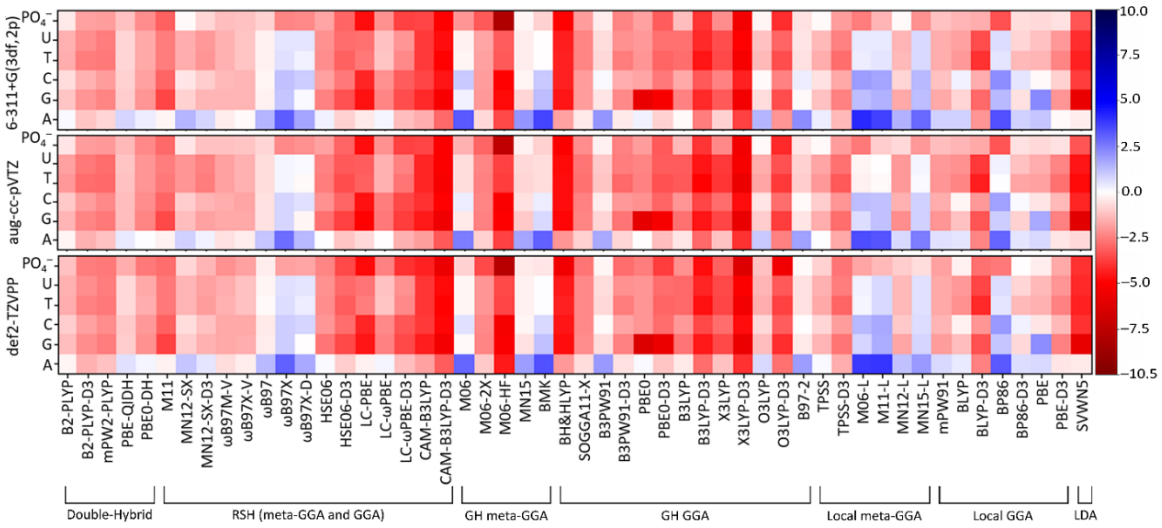


Figure 2.3. Signed errors (kcal/mol) for counterpoise-corrected DFT binding energies of the most stable Li^+–nucleobase and Li^+–phosphate complexes, where negative values (red shades) represent over binding relative to CCSD(T)/CBS reference values

Table 2.2. The percent error (PE), mean percent error (MPE) and average mean percent error (AMPE) across nucleic acid components and/or basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies.[a,b]

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
Double-Hybrid	B2-PLYP	1.82	0.28	1.57	2.32	0.33	1.99	1.97	0.72	1.76	2.04	0.44	1.77
	B2-PLYP-D3	3.38	1.14	3.01	3.99	1.19	3.52	3.63	1.58	3.29	3.67	1.30	3.27
	mPW2-PLYP	3.58	1.15	3.18	4.19	1.20	3.69	3.90	1.61	3.51	3.89	1.32	3.46
	PBE-QIDH	1.49	0.76	1.37	1.85	0.83	1.68	1.40	0.99	1.33	1.58	0.86	1.46
	PBE0-DH	2.56	1.35	2.36	2.96	1.41	2.70	2.56	1.62	2.41	2.70	1.46	2.49
	revDSD-PBEP86-D3	–	–	–	–	–	–	1.89	0.86	1.70	–	–	–
RSH	M11	3.98	1.88	3.63	4.02	1.80	3.65	4.14	1.75	3.74	4.04	1.81	3.67
	MN12-SX	2.24	0.08	1.88	2.65	0.32	2.26	2.10	0.48	1.83	2.33	0.29	1.99
	MN12-SX-D3	2.64	0.50	2.28	3.05	0.74	2.66	2.50	0.90	2.23	2.73	0.71	2.39
RSH meta-GGA	ωB97M-V	2.16	0.81	1.94	2.60	0.76	2.29	2.50	1.30	2.30	2.42	0.96	2.17
RSH GGA	ωB97	0.91	0.55	0.85	1.29	0.51	1.16	0.91	0.24	0.80	1.04	0.43	0.94
	ωB97X	2.32	1.45	2.18	1.70	1.39	1.65	2.25	1.12	2.06	2.09	1.32	1.96
	ωB97X-D	1.32	1.39	1.33	0.85	1.30	0.93	1.21	1.07	1.18	1.13	1.25	1.15
	ωB97X-V	1.79	0.74	1.62	2.31	0.69	2.04	2.03	1.12	1.88	2.04	0.85	1.84
	HSE06	3.06	1.09	2.73	3.44	1.11	3.05	3.14	1.44	2.86	3.21	1.21	2.88
	HSE06-D3	4.44	1.84	4.01	5.05	1.86	4.52	4.64	2.19	4.23	4.71	1.96	4.25
	LC-PBE	4.86	2.88	4.53	5.49	2.88	5.05	4.71	2.93	4.41	5.02	2.89	4.66
	LC-ωPBE	2.30	1.31	2.14	2.73	1.29	2.49	2.37	1.52	2.23	2.47	1.37	2.29
	LC-ωPBE-XDM	–	–	–	3.86	1.71	3.50	2.77	1.64	2.58	3.32	1.68	3.04
	LC-ωPBE-D3	4.13	2.27	3.82	4.72	2.25	4.31	4.20	2.48	3.91	4.35	2.33	4.01
	CAM-B3LYP	5.98	2.22	5.35	6.58	2.24	5.86	6.24	2.63	5.63	6.26	2.36	5.61
	CAM-B3LYP-D3	7.80	3.09	7.01	8.41	3.11	7.52	8.06	3.50	7.30	8.09	3.23	7.28
GH meta-GGA	M06	1.83	1.26	1.73	1.75	0.94	1.62	1.55	0.57	1.38	1.71	0.92	1.58
	M06-2X	2.85	2.00	2.71	3.65	1.84	3.35	3.42	2.15	3.21	3.31	2.00	3.09
	M06-HF	6.74	4.94	6.44	7.32	4.56	6.86	7.63	4.76	7.15	7.23	4.75	6.82

Table 2.2. Continued.

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
	MN15	1.39	0.10	1.17	1.80	0.17	1.53	1.17	0.25	1.02	1.45	0.17	1.24
	BMK	2.12	0.56	1.86	2.05	0.57	1.80	1.89	0.32	1.63	2.02	0.48	1.76
GH GGA	BH&HLYP	6.74	3.05	6.13	7.36	3.11	6.65	6.95	3.35	6.35	7.02	3.17	6.38
	SOGGA11-X	3.17	1.17	2.84	3.94	1.38	3.51	3.38	1.65	3.09	3.50	1.40	3.15
	B3PW91	1.51	0.20	1.29	1.87	0.15	1.58	1.53	0.12	1.30	1.64	0.16	1.39
	B3PW91-D3	3.12	1.44	2.84	3.83	1.49	3.44	3.29	1.76	3.04	3.41	1.56	3.10
	PBE0	3.88	1.14	3.43	4.25	1.19	3.74	4.06	1.54	3.64	4.06	1.29	3.60
	PBE0-D3	4.89	1.94	4.40	5.49	1.98	4.91	5.22	2.33	4.74	5.20	2.08	4.68
	B3LYP	3.49	0.80	3.04	4.14	0.85	3.59	3.83	1.29	3.41	3.82	0.98	3.35
	B3LYP-D3	6.62	2.39	5.91	7.27	2.43	6.46	6.96	2.88	6.28	6.95	2.57	6.22
	X3LYP	4.78	1.36	4.21	5.42	1.40	4.75	5.17	1.88	4.62	5.12	1.55	4.53
	X3LYP-D3	8.36	3.34	7.52	8.99	3.38	8.06	8.74	3.86	7.93	8.70	3.52	7.83
	O3LYP	1.79	0.10	1.50	2.17	0.06	1.81	2.10	0.74	1.87	2.02	0.30	1.73
	O3LYP-D3	5.09	2.53	4.66	5.74	2.57	5.21	5.91	3.37	5.49	5.58	2.82	5.12
	B97-2	1.41	0.37	1.23	1.63	0.29	1.41	1.36	0.08	1.15	1.47	0.24	1.26
Local meta-GGA	TPSS	1.84	0.06	1.54	2.19	0.08	1.84	1.97	0.54	1.73	2.00	0.23	1.70
	TPSS-D3	3.84	1.17	3.40	4.47	1.19	3.93	4.14	1.66	3.73	4.15	1.34	3.68
	M06-L	2.87	1.55	2.65	2.09	1.60	2.01	2.35	1.08	2.14	2.44	1.41	2.27
	M11-L	2.77	0.89	2.46	2.05	0.60	1.81	3.03	0.84	2.67	2.62	0.78	2.31
	MN12-L	2.42	0.09	2.03	2.95	0.45	2.53	2.36	0.59	2.06	2.57	0.38	2.21
	MN15-L	2.46	1.32	2.27	1.64	1.22	1.57	1.63	0.40	1.43	1.91	0.98	1.75
Local GGA	mPW91	2.46	0.74	2.17	2.86	0.80	2.52	2.57	1.12	2.33	2.63	0.89	2.34
	BLYP	2.18	0.89	1.97	2.40	0.88	2.14	2.23	0.18	1.89	2.27	0.65	2.00
	BLYP-D3	5.10	0.99	4.41	5.75	1.00	4.96	5.65	1.71	4.99	5.50	1.23	4.79
	BP86	2.83	2.06	2.70	2.11	2.05	2.10	2.61	1.62	2.44	2.52	1.91	2.41
	BP86-D3	1.52	0.40	1.33	1.67	0.38	1.46	1.45	0.05	1.22	1.55	0.28	1.34

Table 2.2. Continued.

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
	PBE	2.17	0.45	1.89	2.31	0.44	2.00	2.23	0.20	1.89	2.24	0.36	1.93
	PBE-D3	2.64	0.36	2.26	3.31	0.37	2.82	3.09	1.01	2.74	3.01	0.58	2.61
Local LDA	SVWN5	6.22	2.04	5.53	6.94	2.04	6.13	6.38	2.48	5.73	6.52	2.19	5.79
AMPE		3.34	1.32	3.01	3.68	1.32	3.29	3.47	1.49	3.14	–	–	–

[a] Calculated with respect to CCSD(T)/CBS values (Figure 2).
[b] revDSD-PBEP86-D3 is excluded from the reported average.

Table 2.3. The mean unsigned errors (MUE), root-mean squared errors (RMSE), average mean unsigned errors (AMUE) and average root-mean squared errors (ARMSE) across nucleic acid components and/or basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies.[a,b]

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
		Base		**PO_4^-**	**Base**		**PO_4^-**	**Base**		**PO_4^-**	**AMUE**		**ARMSE**
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Base**	**PO_4^-**	**Base**
Double-Hybrid	B2-PLYP	1.02	1.16	0.46	1.31	1.46	0.54	1.11	1.22	1.17	1.15	0.73	1.28
	B2-PLYP-D3	1.91	1.97	1.87	2.25	2.30	1.95	2.06	2.09	2.58	2.07	2.13	2.12
	mPW2-PLYP	2.04	2.14	1.88	2.38	2.47	1.96	2.23	2.29	2.63	2.22	2.16	2.30
	PBE-QIDH	0.86	0.86	1.24	1.08	1.12	1.35	0.82	0.84	1.63	0.92	1.41	0.94
	PBE0-DH	1.51	1.62	2.21	1.75	1.95	2.30	1.52	1.67	2.65	1.59	2.39	1.75
	revDSD-PBEP86-D3	–	–	–	–	–	–	1.19	1.81	1.41	–	–	–
RSH	M11	2.38	2.64	3.08	2.39	2.63	2.94	2.47	2.73	2.87	2.41	2.96	2.67
	MN12-SX	1.21	1.28	0.13	1.46	1.54	0.52	1.15	1.19	0.78	1.27	0.48	1.34
	MN12-SX-D3	1.46	1.56	0.82	1.72	1.90	1.20	1.41	1.50	1.47	1.53	1.16	1.65
RSH meta-GGA	ωB97M-V	1.25	1.32	1.33	1.49	1.56	1.24	1.45	1.51	2.13	1.40	1.56	1.46
RSH GGA	ωB97	0.49	0.68	0.90	0.72	0.75	0.83	0.51	0.68	0.39	0.57	0.71	0.70

Table 2.3. Continued.

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
		Base		PO_4^-	Base		PO_4^-	Base		PO_4^-	AMUE		ARMSE
Family	Functional	MUE	RMSE	UE	MUE	RMSE	UE	MUE	RMSE	UE	Base	PO_4^-	Base
	ωB97X	1.31	1.63	2.37	0.95	1.38	2.27	1.25	1.56	1.83	1.17	2.16	1.52
	ωB97X-D	0.73	0.94	2.28	0.48	0.67	2.12	0.66	0.86	1.74	0.62	2.05	0.82
	ωB97X-V	1.06	1.16	1.21	1.36	1.44	1.13	1.21	1.30	1.83	1.21	1.39	1.30
	HSE06	1.79	1.95	1.78	2.01	2.25	1.82	1.85	2.04	2.36	1.88	1.99	2.08
	HSE06-D3	2.58	2.73	3.00	2.92	3.06	3.04	2.70	2.84	3.58	2.73	3.21	2.88
	LC-PBE	2.95	3.32	4.71	3.31	3.64	4.71	2.87	3.25	4.79	3.04	4.73	3.40
	LC-ωPBE	1.42	1.62	2.14	1.67	1.91	2.11	1.46	1.67	2.48	1.52	2.24	1.73
	LC-ωPBE-XDM	–	–	–	2.32	2.50	2.79	1.71	1.95	2.69	2.02	2.23	2.74
	LC-ωPBE-D3	2.46	2.58	3.71	2.79	2.90	3.68	2.50	2.64	4.05	2.59	3.82	2.71
	CAM-B3LYP	3.45	3.57	3.63	3.80	3.91	3.66	3.61	3.72	4.31	3.62	3.86	3.73
	CAM-B3LYP-D3	4.48	4.55	5.05	4.83	4.90	5.08	4.64	4.71	5.73	4.65	5.28	4.72
GH meta-GGA	M06	1.00	1.55	2.06	0.93	1.24	1.54	0.83	1.38	0.94	0.92	1.51	1.39
	M06-2X	1.65	1.71	3.27	2.11	2.19	3.00	1.99	2.05	3.52	1.92	3.27	1.98
	M06-HF	3.95	4.02	8.08	4.27	4.32	7.46	4.45	4.51	7.78	4.22	7.77	4.28
	MN15	0.76	1.03	0.17	1.01	1.09	0.28	0.64	0.92	0.40	0.80	0.28	1.01
	BMK	1.21	1.79	0.92	1.11	1.49	0.92	1.08	1.63	0.53	1.13	0.79	1.64
GH GGA	BH&HLYP	3.90	4.04	5.00	4.25	4.38	5.09	4.02	4.15	5.48	4.06	5.19	4.19
	SOGGA11-X	1.81	1.95	1.91	2.25	2.35	2.25	1.95	2.05	2.69	2.01	2.28	2.12
	B3PW91	0.80	1.05	0.32	1.01	1.11	0.24	0.82	1.02	0.20	0.88	0.25	1.06
	B3PW91-D3	1.79	1.94	2.35	2.19	2.32	2.44	1.90	2.03	2.88	1.96	2.55	2.10
	PBE0	2.40	3.03	1.87	2.61	3.23	1.94	2.53	3.26	2.51	2.52	2.11	3.17
	PBE0-D3	2.92	3.23	3.17	3.25	3.50	3.24	3.13	3.47	3.81	3.10	3.41	3.40
	B3LYP	1.99	2.23	1.32	2.35	2.57	1.38	2.19	2.39	2.12	2.18	1.61	2.40
	B3LYP-D3	3.75	3.83	3.91	4.12	4.19	3.98	3.96	4.02	4.71	3.95	4.20	4.01
	X3LYP	2.73	2.91	2.22	3.09	3.25	2.29	2.96	3.10	3.08	2.93	2.53	3.09

Table 2.3. Continued.

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
		Base		PO_4^-	Base		PO_4^-	Base		PO_4^-	AMUE		ARMSE
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Base**	**PO_4^-**	**Base**
	X3LYP-D3	4.77	4.81	5.46	5.13	5.17	5.52	5.00	5.03	6.31	4.96	5.76	5.00
	O3LYP	0.94	1.07	0.16	1.18	1.26	0.10	1.17	1.23	1.21	1.10	0.49	1.19
	O3LYP-D3	2.90	2.95	4.14	3.27	3.31	4.20	3.39	3.42	5.51	3.19	4.61	3.23
	B97-2	0.74	1.05	0.60	0.86	1.04	0.48	0.71	0.97	0.13	0.77	0.40	1.02
Local meta-GGA	TPSS	1.03	1.13	0.10	1.25	1.43	0.13	1.12	1.26	0.89	1.13	0.37	1.27
	TPSS-D3	2.17	2.24	1.92	2.53	2.59	1.95	2.35	2.41	2.71	2.35	2.19	2.41
	M06-L	1.64	2.15	2.54	1.17	1.69	2.61	1.32	1.88	1.76	1.38	2.30	1.91
	M11-L	1.59	1.99	1.45	1.18	1.66	0.99	1.72	2.08	1.37	1.50	1.27	1.91
	MN12-L	1.36	1.39	0.15	1.70	1.78	0.74	1.35	1.37	0.97	1.47	0.62	1.51
	MN15-L	1.40	1.61	2.15	0.94	1.26	2.00	0.91	1.06	0.66	1.08	1.60	1.31
Local GGA	mPW91	1.40	1.43	1.21	1.65	1.75	1.31	1.49	1.54	1.83	1.51	1.45	1.57
	BLYP	1.16	1.33	1.46	1.28	1.55	1.44	1.21	1.44	0.29	1.22	1.06	1.44
	BLYP-D3	2.83	2.96	1.62	3.20	3.32	1.64	3.16	3.25	2.80	3.06	2.02	3.18
	BP86	1.64	1.96	3.38	1.23	1.64	3.35	1.50	1.81	2.65	1.45	3.12	1.80
	BP86-D3	0.82	0.88	0.65	0.90	1.08	0.62	0.79	0.87	0.08	0.84	0.45	0.94
	PBE	1.24	1.42	0.74	1.29	1.40	0.72	1.28	1.41	0.33	1.27	0.59	1.41
	PBE-D3	1.50	1.72	0.58	1.88	2.06	0.60	1.77	1.93	1.65	1.72	0.94	1.90
Local LDA	SVWN5	3.68	4.10	3.33	4.08	4.46	3.34	3.78	4.18	4.06	3.85	3.57	4.25
AMUE		1.92	–	2.15	2.11	–	2.16	2.00	–	2.44	–	–	–

[a] Calculated with respect to CCSD(T)/CBS values (Figure 2).
[b] revDSD-PBEP86-D3 is excluded from the reported average.

2.3.2.1. Double-Hybrid Functionals

For any given basis set and double-hybrid functional combination, the AMUE and ARMSE for nucleic acid interactions fluctuate from ~ 0.9 to 2.4 kcal/mol. The best performing double-hybrid functionals across all Li^+–nucleic acid interactions and the three basis sets are PBE-QIDH (1.46% AMPE) and B2-PLYP (1.77% AMPE). Although PBE-QIDH (1.58% nucleobases AMPE, 0.86% phosphate AMPE) is the best functional for Li^+–nucleobase interactions, B2-PLYP (2.04% nucleobases AMPE, 0.44% phosphate AMPE) demonstrates improved performance for Li^+–phosphate interactions. The worst performer of this family is mPW2-PLYP, with the AMPE (3.46%) being more than two times those of the most accurate double-hybrid approaches and the ARMSE being the largest (2.3 kcal/mol). The AMPEs for the remaining double-hybrid functionals tested in the present work range from ~ 2.50 to 3.30%. Although not tested for our full complement of basis sets, we note that revDSD-PBEP86-D3(BJ)[120] combined with the def2-TZVPP basis set results in an MPE of 1.70%, suggesting this is also a reasonably performing double-hybrid functional.

2.3.2.2. Range-Separated Hybrids (RSH), RSH meta-Generalized Gradient Approximation (GGA), and RSH GGAs

The performances of three functional subcategories are combined under the broad RSH classification. Across all Li^+–nucleic acid interactions and basis sets, the best performing RSH functionals include ωB97 (0.94% AMPE and 0.7 kcal/mol ARMSE) and ωB97X-D (1.15% AMPE and 0.82 kcal/mol ARMSE). While ωB97 and ωB97X-D are similar in performance for Li^+–nucleobase interactions, ωB97 results in a lower AMPE for phosphate interactions (0.43%) compared to ωB97X-D (1.25%). Interestingly, the ωB97

series generally outperforms all other functionals within this family. Indeed, the top 4 performers of the ωB97 series surpass the accuracy of other functionals in this family, with only MN12-SX (1.99% AMPE) approaching the accuracy of ωB97. The least accurate members of the broad RSH family are CAM-B3LYP-D3(BJ) and CAM-B3LYP (7.28% and 5.61% AMPE, respectively). In fact, the error is over 6 times greater for CAM-B3LYP-D3(BJ) compared ωB97. Furthermore, both worst performers lead to higher ARMSEs (3.73 – 4.72 kcal/mol). The AMPEs for the remaining functionals fall within ~ 2.4 to 4.7%.

2.3.2.3. Global Hybrid (GH) meta-GGA

In general, the AMPE for GH meta-GGAs ranges from 1.24 to 6.82% across all Li^+–nucleic acid interactions and the ARMSE varies from 1 to 4.28 kcal/mol. Within this family, four of the five tested functionals fall under the Minnesota formalism, two of which performed particularly well. Specifically, the best performer is MN15 (1.24% AMPE and 1.01 kcal/mol ARMSE). The accuracy of MN15 is followed by that of M06 (1.58% AMPE and 1.39 kcal/mol ARMSE) and BMK (1.76% AMPE and 1.64 kcal/mol ARMSE). Although both functionals similarly describe nucleobase interactions, BMK yields slightly improved Li^+–phosphate interaction energies over M06. M06-2X displays mid-range performance (3.09% AMPE) and M06-HF is by far the worst performer of the GH meta-GGA family (6.82% AMPE and 4.28 kcal/mol ARMSE).

2.3.2.4. Global Hybrid (GH) GGA

The GH GGA functional family demonstrates the greatest fluctuation in accuracy across all functional classifications, with the AMPE ranging from 1.26 to 7.83% and the ARMSE varying from ~ 1 to 5 kcal/mol. The best performer of the GH GGA family is

B97-2 (1.26% AMPE), which is closely followed by B3PW91 (1.39%), with both functionals having an ARMSE of ~ 1 kcal/mol. While B97-2 more accurately predicts Li^+–nucleobase binding strengths compared to B3PW91 (1.47% and 1.64% AMPE, respectively), B3PW91 (0.16% AMPE) marginally outperforms B97-2 (0.24%) for phosphate interactions. O3LYP binding strengths are also quite reliable (1.73% AMPE and 1.19 kcal/mol ARMSE). The worst performer of this functional family is X3LYP-D3 (7.83% AMPE), which results in an over 6 times greater AMPE compared to B97-2. The AMPE for the remaining functionals falls within the range of ~ 3.1 to 6.4%.

2.3.2.5. Local meta-GGA

For local meta-GGA functionals, the AMPE ranges from 1.70 to 3.68%, while the ARMSE varies from ~ 1.3 to 2.4 kcal/mol. The most accurate functionals in this family include TPSS (1.70% AMPE; 1.27 kcal/mol ARMSE) and MN15-L (1.75% AMPE; 1.31 kcal/mol ARMSE). While both functionals perform similarly, MN15-L describes Li^+–nucleobase interactions slightly better than TPSS (1.91% versus 2.00% AMPE, respectively), while the trend is reversed and the performance gap larger for Li^+–phosphate interactions (0.98% and 0.23% AMPE). All remaining Minnesota meta-GGA functionals tested (M06-L, M11-L and MN12-L) yield comparable accuracy, with an AMPE ranging from 2.21 to 2.31%. The worst performer in this family is TPSS-D3 (3.68% AMPE and 2.41 kcal/mol ARMSE).

2.3.2.6. Local GGA and Local Density Approximation (LDA)

The most accurate local GGA functional for Li^+–nucleobase and phosphate contacts is BP86-D3 (1.34% AMPE; ARMSE 0.94 kcal/mol), followed by PBE (1.93% AMPE;

1.41 kcal/mol ARMSE). The worst performer is BLYP-D3 (4.79% AMPE; 3.18 kcal/mol ARMSE), while the AMPE for the remaining functionals ranges between ~ 2 and 2.6%. Although BLYP-D3 (5.50% AMPE for nucleobases; 1.23% AMPE for phosphate) is the least reliable method for nucleobase contacts, BP86 leads to larger errors for phosphate interactions (2.52% AMPE for nucleobases; 1.91% AMPE for phosphate). The accuracy of SVWN5 (5.79% AMPE; 4.25 kcal/mol ARMSE) is extremely poor, which is expected from the lowest rung of Jacob's Ladder.

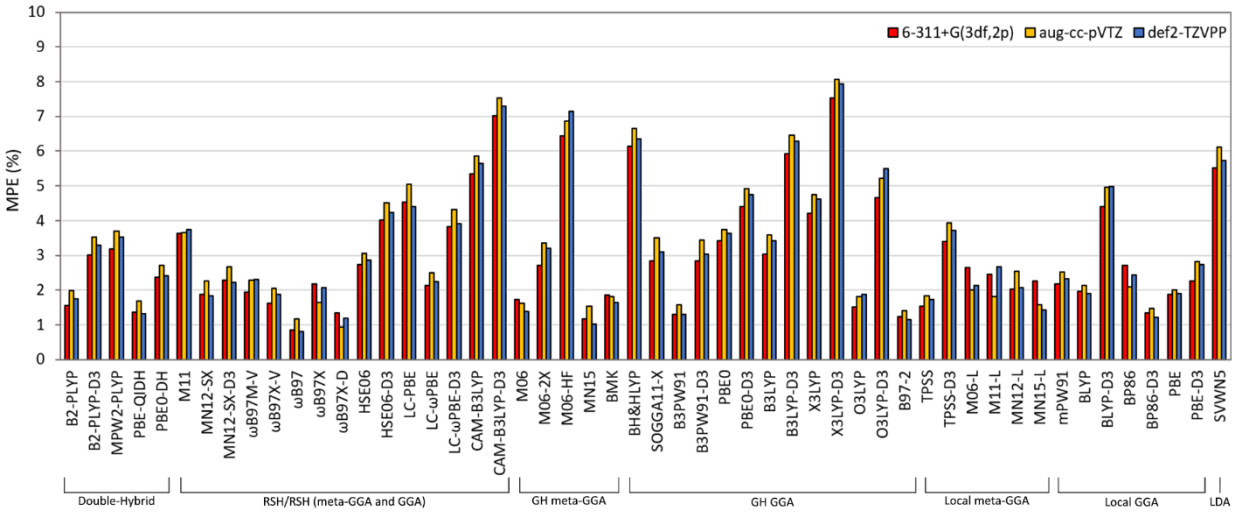


Figure 2.4. MPEs for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies evaluated with the 6-311+G(3df,2p) (red), aug-cc-pVTZ (yellow) and def2-TZVPP (blue) basis sets relative to CCSD(T)/CBS reference values.

2.3.2.7. Dispersion Effects

As mentioned in the Introduction, the impact of an empirical dispersion correction was considered by adding Grimme's D3[62,63] correction with Becke–Johnson damping (BJ)[49] to 15 functionals (denoted D3 for simplicity; Table 2.1, Figures 2.5, and A.5–A.6, Appendix A). The accuracy of functionals that include the empirical dispersion correction

fluctuates considerably and tend to over bind Li^+–nucleic acid interactions. The overall best performing D3 corrected functional is BP86-D3 (1.34% AMPE). In contrast, X3LYP-D3 (7.83% AMPE), CAM-B3LYP-D3 (7.28% AMPE), and B3LYP-D3 (6.22% AMPE) poorly describe both Li^+–nucleobase and Li^+–phosphate interactions. Perhaps most interestingly, with only one exception, the DFT-D3 variants perform worse than the corresponding uncorrected functionals across all nucleobases and phosphate contacts, with the increase in AMPE ranging from ~ 0.4 to 3.4%. A similar trend was reported in a previous study of the ability of DFT methods to reliably describe ionic hydrogen bonding between organic molecules.[121] The greatest decline in performance occurs for X3LYP-D3 and O3LYP-D3 (AMPE increases by > 3% and ARMSE by > 2 kcal/mol). The exception to this trend is BP86, where inclusion of the D3 correction slightly improves the predicted binding strengths (AMPE decreases from 2.41% to 1.34%; ARMSE from 1.80 to 0.94 kcal/mol). In addition to the D3 empirical correction, nonlocal dispersion corrections to ωB97 and ωB97X were considered through the ωB97M-V and ωB97X-V functionals, as well as the D2 correction (ωB97X-D). The only correction in this group that closely maintains the accuracy of ωB97 is ωB97X-D, which increases the AMPE by only 0.21%. Interestingly, ωB97X-V (1.84% AMPE) offers only a modest improvement over ωB97X (1.96% AMPE). Finally, the exchange-hole dipole dispersion moment correction, LC-ωPBE-XDM, leads to a very modest (0.32%) increase in accuracy over the uncorrected LC-ωPBE counterpart (Figure S7). Thus, overall, dispersion corrections most often decrease or lead to very little improvement in the accuracy of lithium–nucleic acid interactions, which has also been previously reported in previous DFT benchmark study that investigated ligand dissociation energies of transition-metal complexes.[122]

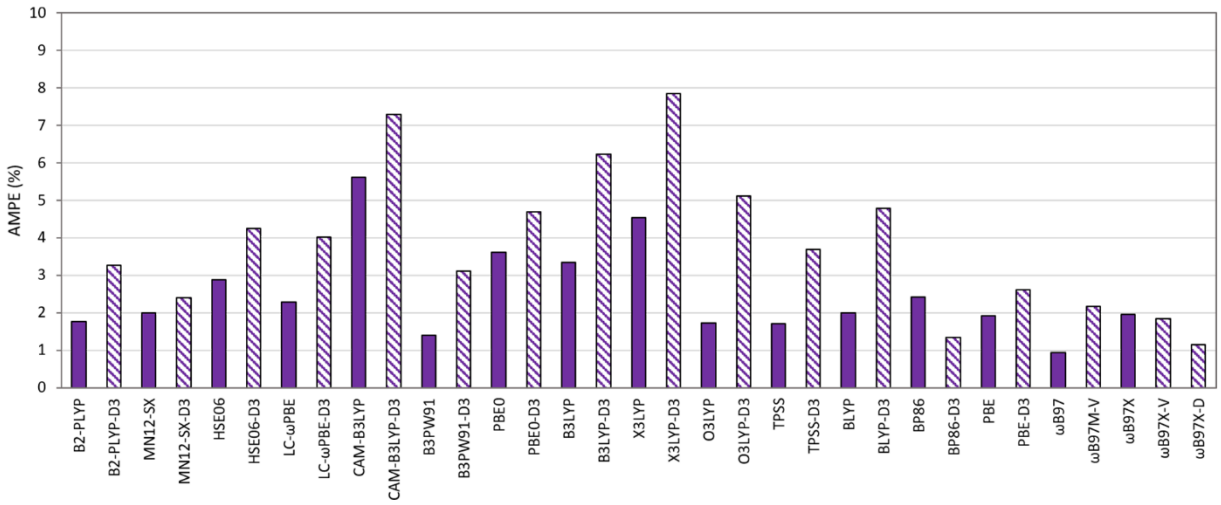


Figure 2.5. AMPEs over all basis sets for Li^+–nucleobase and Li^+–phosphate counterpoise and dispersion-corrected DFT binding energies relative to CCSD(T)/CBS reference values.

2.3.2.8. Comparison Across Families

We find significant variation in the ability of particular functionals to accurately predict Li^+–nucleobase and Li^+–phosphate binding energies across and within families (Figure 2.6). Nevertheless, there are functionals in most rungs of Jacob's ladder with AMPEs below 2%, which represents an ~0.5–1.5 kcal/mol difference from the CCSD(T)/CBS reference values for nucleobase contacts and an ~3 kcal/mol discrepancy for phosphate interactions. These reliable functionals include double-hybrids (the non-empirical PBE-QIDH and the semi-empirical B2-PLYP), semi-empirical RSHs (ωB97, ωB97X-D, ωB97X-V and ωB97X) and the screened exchange MN12-SX, semi-empirical GH meta-GGAs (MN15, M06 and BMK), semi-empirical GH GGAs (B97-2, B3PW91 and O3LYP), local meta-GGAs (the non-empirical TPSS and semi-local MN15-L), and local GGAs (the semi-empirical BP86-D3 and BLYP, and the non-empirical PBE, Figure 2.6). It is interesting to note that there are cost-effective functionals from the lower rungs of Jacob's

Ladder, including meta-GGAs and GGAs, that can be applied to investigate larger Li^+–nucleic acid interactions.

In general, the RSH GGA family includes a number of reliable functionals, with 4 methods (all in the ωB97 series) having an AMPE < 2%. The ωB97 series of functionals has been previously highlighted to accurately predict many properties,[123] including lithium–benzene cation–π interactions.[45] Among the GH meta-GGA functionals, while MN15 is the top performer (1.24% AMPE), M06 is noteworthy and has been similarly highlighted in benchmark studies for a range of other properties.[123] The GH GGA family demonstrates the most significant fluctuation in performance (1.26 to 7.83% AMPE). Interestingly, not all double-hybrid functionals are among the best performers. In fact, mPW2-PLYP has an ARMSE of 2.30 kcal/mol. Nevertheless, PBE-QIDH yields highly accurate descriptions of Li^+–nucleic acid interactions (1.46% AMPE and 0.94 kcal/mol ARMSE), and has been highlighted as a reliable functional for describing lithium polysulfide complexes in a previous benchmark study.[44] Comparing across the 6 families, there is not a clear trend in the accurate description of Li^+–nucleic acid interactions based on the amount of Hartree–Fock (exact) exchange or the fraction of MP2 correlation included. For example, despite B2-PLYP and mPW2-PLYP having a similar amount of exact exchange (53% and 55%) and MP2 correlation (27 and 25%), B2-PLYP is the second-top performer among the double-hybrids, while mPW2-PLYP is the worst functional in this family. Furthermore, although B3PW91, B3LYP, X3LYP and B97-2 in the GH GGA family have similar amounts of exact exchange (~ 20–21%), the corresponding AMPEs vary from 1.26–4.53%.

Despite the good performance of certain methods within each functional family, particular functionals from each class also tend to be unreliable. For example, the AMPE

for B3LYP-D3, BH&HLYP and M06-HF is greater than 6%, while that for CAM-B3YLP-D3 and X3LYP-D3 exceeds 7%. Although B3LYP (3.35% AMPE) has been widely used in other computational studies of Li^+–nucleic acid interactions,[27,29-32,34,35,37] our work highlights that several other functionals provide improved descriptions of Li^+–nucleic acid complexes.

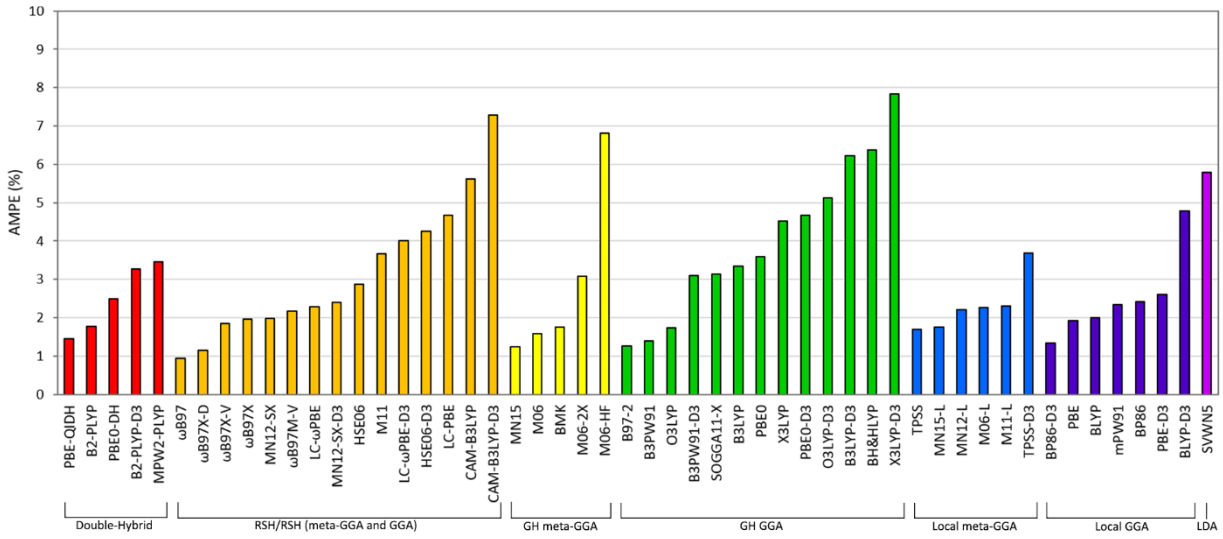


Figure 2.6. AMPEs over all basis sets for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies relative to CCSD(T)/CBS reference values arranged in ascending order of accuracy within families.

2.3.3. The Impact of Basis Sets on DFT Predicted Binding Strengths

As previously illustrated by the heat maps (Figure 2.3), the three triple-zeta basis sets considered in the present work (6-311+G(3df,2p), aug-cc-pVTZ and def2-TZVPP) exhibit similar accuracy with respect to the CCSD(T)/CBS Li^+–nucleobase and Li^+–phosphate interactions (Tables 2.2 and 2.3, Figure 2.4). Specifically, the AMPE across all functionals equals 3.01% for 6-311+G(3df,2p), 3.29% for aug-cc-pVTZ and 3.14% for def2-TZVPP. Furthermore, all three basis sets are equally reliable for nucleobase contacts

(AMUE ~1.9–2.1 kcal/mol) and phosphate interactions (AMUE ~2.2–2.4 kcal/mol). Nevertheless, the AMPE significantly varies with the functional choice for each basis set, ranging from 0.85 to 7.52% for 6-311+G(3df,2p), 0.93 to 8.06% for aug-cc-pVTZ, and 0.80 to 7.93% for def2-TZVPP.

To further investigate the impact of the chosen basis set on Li^+–nucleobase and Li^+–phosphate interactions, we considered the effect of decreasing the TZ basis set to DZ quality and increasing the TZ basis set to a QZ for the Dunning and Karlsruhe families. Overall, the accuracies of the TZ and QZ basis sets are very similar for the vast majority of functionals (Figures 2.7 and A.10, Tables A.3–A.10, Appendix A), with AMPEs across all functionals being 3.29% for aug-cc-pVTZ compared to 3.62% for aug-cc-pVQZ, and 3.15% for def2-TZVPP compared to 3.51% for def2-QZVPP. Furthermore, both TZ and QZ basis sets exhibit similar reliability for the description of Li^+–nucleobase (MUEs 1.99–2.31 kcal/mol) and Li^+–phosphate (MUEs 2.16–2.73 kcal/mol) contacts. Interestingly, a size reduction of the Dunning basis set results in a notable improvement in the DFT description of the Li^+–nucleic acid interactions, with aug-cc-pVDZ (AMPE = 2.54%) outperforming aug-cc-pVTZ and aug-cc-pVQZ for most functionals, suggesting a cancellation of errors for the DZ basis set. In contrast, there is a significant reduction in the accuracy of def2-SVP (4.72% AMPE) compared to def2-TZVPP (3.15% AMPE), likely due to the much smaller size of this DZ basis set. Overall, due to the similar performance of the TZ and QZ basis sets and the sporadic accuracy of the DZ variants, basis sets of at least TZ quality are recommended as the best cost–performance solution for the accurate prediction of Li^+–nucleic acid interactions, while smaller basis sets in the Dunning family can be used for larger nucleic acid systems that prohibit the application of TZ quality basis sets.

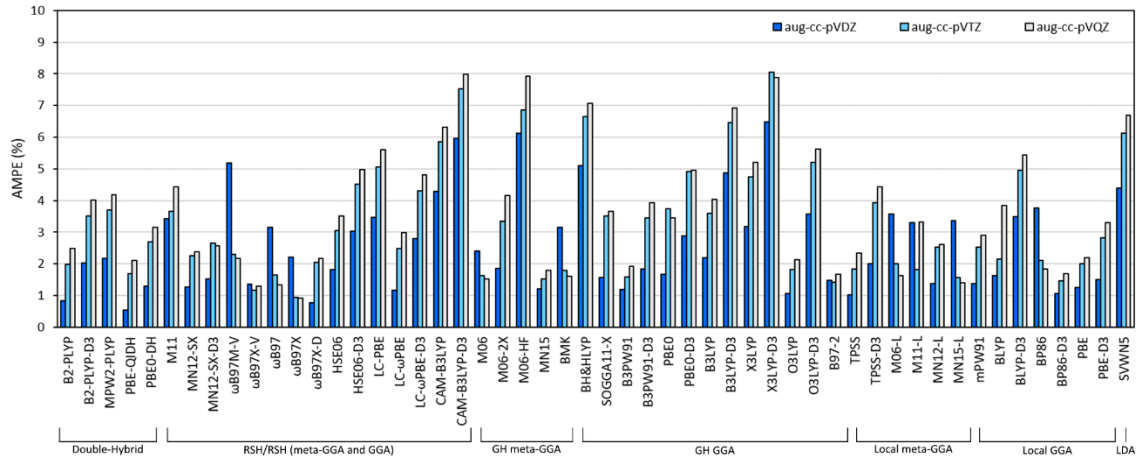


Figure 2.7. AMPEs as a function of the Dunning basis set size for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies relative to CCSD(T)/CBS reference values.

2.3.4. The Impact of Counterpoise Corrections on DFT Predicted Binding Strengths

Although all reported interaction energies discussed thus far include counterpoise corrections, it is intriguing to consider the performance of DFT functionals without these corrections, which are not always computationally feasible when modeling larger biosystems. In general, all trends discussed previously for Li^+–nucleobase and Li^+–phosphate interactions are maintained regardless of whether counterpoise corrections are included in the predicted binding energies (Figures 2.8, A.7–A.9 and A.11–A.13, Tables A.1–A.2, A.5–A.6 and A.9–A.10, Appendix A). As DFT methods typically over bind Li^+–nucleic acid complexes, it is not surprising that removal of the counterpoise correction predominantly decreases the accuracy of the computed binding strengths compared to CCSD(T)/CBS values (Figures 2.3 and A.8, Appendix A). Nevertheless, certain functionals (ωB97X, ωB97X-D, BMK, M06-L, M11-L, MN15-L and BP86) demonstrate slight improvement upon removal of the counterpoise correction (AMPE decreases by

0.02 – 0.82%; ARMSE by 0.1 – 1.9 kcal/mol), most likely due to a cancellation of errors. Furthermore, the trends in functional performance are unchanged (Figures 2.4–2.6, 2.8, and A.7–A.9, A.11, Appendix A). Indeed, empirical dispersion corrections generally do not improve the accuracy of the predicted binding strengths (Figure 2.8) and each functional family contains both good and poor performing methods (Figure A.9). Most importantly, the removal of the counterpoise correction increases the AMPE by less than 1% for all functionals with any of the three basis sets (Figures 2.8, and A.11–A.13, Appendix A), with the exception of only M06-HF (6-311+G(3df,2p) and aug-cc-pVTZ) and O3LYP-D3 (def2-TZVPP). This suggests that the binding energies of Li^+–nucleic acid systems can be reliably evaluated using triple-zeta basis sets in the absence of correcting for the BSSE.

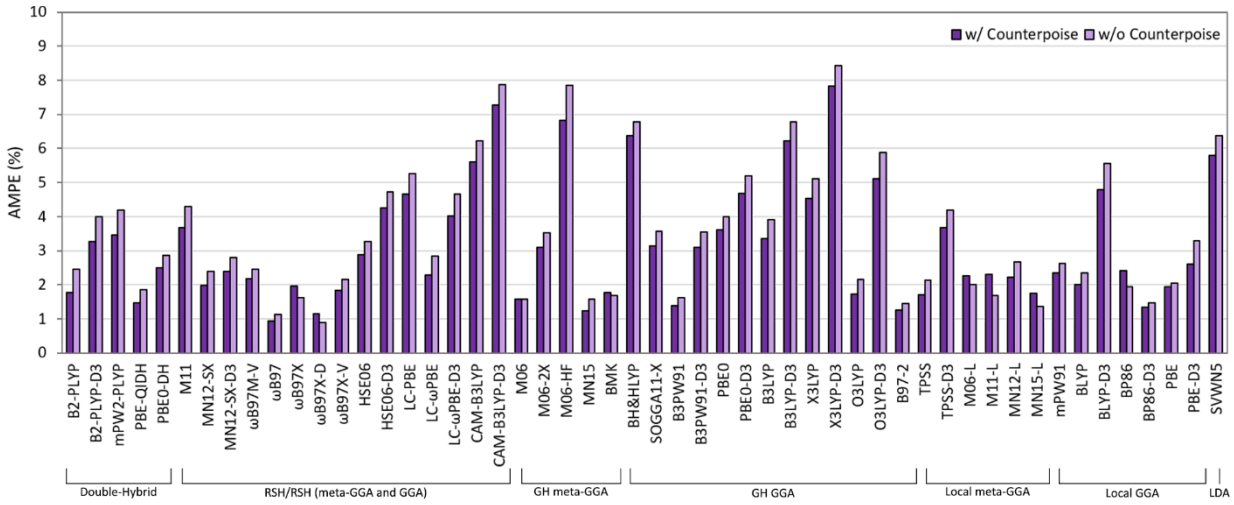


Figure 2.8. AMPEs over all basis sets for Li^+–nucleobase and Li^+–phosphate DFT binding energies with (dark purple) and without (light purple) counterpoise corrections included relative to CCSD(T)/CBS reference values.

2.3.5. The Ability of DFT to Predict the Preferred Nucleobase and Phosphate Binding Locations

In addition to understanding whether DFT methods can broadly reproduce the magnitude of CCSD(T)/CBS binding strengths for Li^+–nucleobase and Li^+–phosphate interactions, it is important to analyze the trends in the computed DFT interaction energies across nucleic acid components and binding sites. Due to the large number of functional and basis set combinations considered in the present work, these trends were investigated by more closely examining the interaction energies calculated using the top two performing functionals in each family (namely B2-PLYP, PBE-QIDH, ωB97, ωB97X-D, M06, MN15, B3PW91, B97-2, TPSS, MN15-L, BP86-D3 and PBE) and the aug-cc-pVTZ basis set, which was arbitrarily chosen based on the comparable performance of all basis sets employed in the present work.

The Li^+–nucleic acid complexes were initially subdivided into those involving direct metal coordination and those involving cation–π interactions (Tables A.11 and A.12, Figures 2.9 and A.14, Appendix A). Although the AMPE in the DFT binding strengths for complexes involving the favored coordination sites are below 2% for all functionals and as low as 0.93% for the top performer (ωB97X-D), greater fluctuation in the AMPE occurs for other (nonpreferred) coordination sites, which ranges from ~ 0.6 to 3.4%. Nevertheless, all functionals except M06 and MN15-L result in an AMPE across all Li^+ coordination sites below 2%, which is equivalent to an error in the computed binding strength of < 1 kcal/mol for nucleobase contacts and ~3 kcal/mol for phosphate interactions. This underscores that there are many reliable functionals from each family for direct Li^+–nucleobase coordination, including more computationally feasible functionals that can be applied to larger Li^+–nucleic acid complexes.

The AMPE for cation–π interactions is larger than that for direct coordination for all functionals. In fact, the AMPE for cation–π contacts is well above 3% for all functionals except B2-PLYP, PBE-QIDH, ωB97X-D and MN15-L, being as large at 8% for M06. Interestingly, there is significant variation between ωB97 (7.80% AMPE) and ωB97X-D (3.26% AMPE) in the description of cation–π interactions, which may be attributed to a combination of the increased amount of short-range exchange (by 22.2%, with 0% for ωB97) and the added atom–atom dispersion corrections. Nevertheless, the addition of empirical dispersion corrections does not necessarily result in improved descriptions of cation–π interactions. This correlates with previous benchmark studies of alkali metal cation–benzene complexes that revealed certain functionals with empirical dispersion corrections overbind cation–π interactions.[62,124] These examples illustrate that care must be taken when selecting a functional to examine nucleic acid systems where a range of interaction types may occur. Nevertheless, PBE-QIDH (1.26% AMPE), B2-PLYP (1.78%) and ωB97X-D (2.06%) have a low AMPE across all types of Li^+–nucleic acid complexes, which complements the greater functional choice when cation–π interactions are not prevalent.

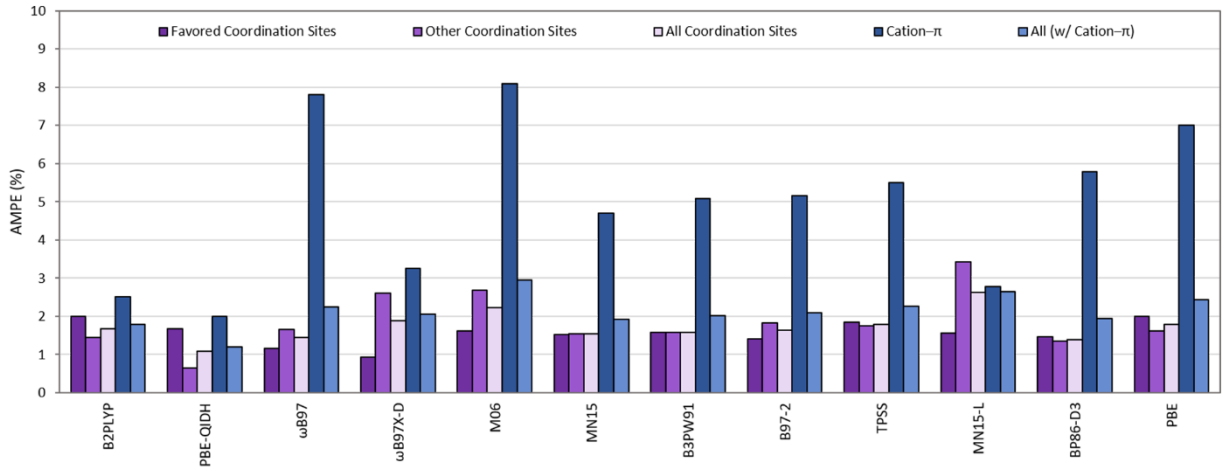


Figure 2.9. AMPEs across the favored nucleic acid coordination sites (dark purple), other (nonpreferred) coordination sites (medium purple), all coordination sites (light purple), only cation–π interactions (dark blue) and all binding locations (including cation–π, medium blue) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies evaluated with the aug-cc-pVTZ basis set relative to CCSD(T)/CBS reference values.

Direct comparison of the predicted binding strengths for all Li^+–nucleic acid interactions reveals a highly similar trend for CCSD(T) and DFT (Figure 2.10). Indeed, when each nucleic acid component is considered, the CCSD(T)/CBS trends in the interaction energies for different coordination sites are well preserved for all 12 DFT methods (Figures 2.10 and A.15–A.21, Appendix A). Indeed, the overall agreement persists across sites with interaction energies falling within the more limited range of ~ 40 – 50 kcal/mol (Figure S20), although DFT functionals predict U and T binding to be (albeit slightly) more favorable than A (Figure A.21, Appendix A). As a result, regardless of the functional considered, Li^+ is predicted to bind most strongly to the phosphate backbone (Figure 2.10) and the interaction energy for the preferred coordination site for each nucleobase decreases as G > C >> A ~ T = U (Figure A.21, Appendix A). Thus, DFT

methods can accurately reproduce the CCSD(T)/CBS predicted trends in the preferred nucleic acid component and coordination site.

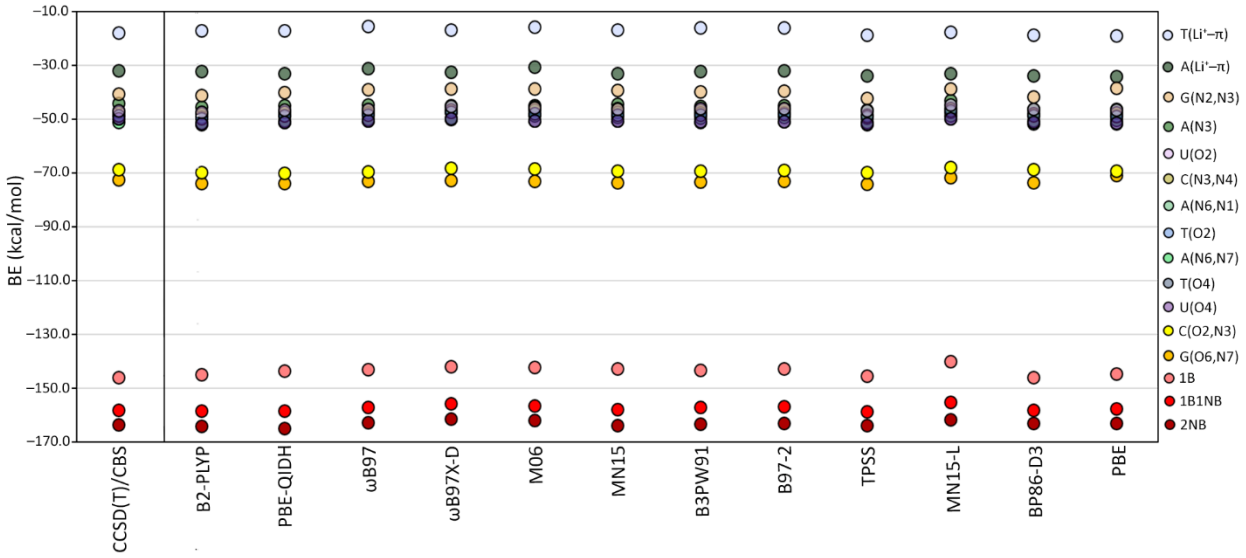


Figure 2.10. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for all Li^+–nucleobase and Li^+–phosphate complexes.

2.3.6. The Ability of DFT to Describe the Structure of Li^+–Nucleic Acid Complexes

Although MP2 geometries were used in the present work to provide a consistent framework for the comparison of the DFT predicted binding energies, MP2 optimizations are not feasible for the larger Li^+–nucleic acid complexes required to understand the impact of the metal in biology or to design metal detection platforms. Therefore, to consider the impact of using DFT predicted structures on Li^+–nucleic acid interactions, the 12 top-performing functionals identified in the previous section were used in combination with aug-cc-pVTZ to optimize the most stable Li^+–nucleobase and Li^+–phosphate complexes. Subsequently, the PE for the DFT predicted interaction energies obtained using the MP2 and DFT optimized geometries were evaluated (Figure A.22 and Table A.13, Appendix A). The resulting highly similar Li^+–nucleic acid interaction energies reflect the negligible

difference between the MP2 and DFT optimized structures. Indeed, most differences in the AMPEs for the binding energies fall between 0.02 and 0.08%, while the greatest difference in the AMPE is only 0.17% for MN15-L, which is followed by 0.12% for BP86-D3 and 0.11% for PBE. Furthermore, for any given complex, the error is less than 0.40% (MN15-L for Li^+–U). Therefore, we are confident that the functionals identified in the present work accurately describe the energetic properties of Li^+–nucleic acid interactions and can be used to obtain structural properties for larger nucleic acid systems.

2.4. Conclusions

In summary, a database of CCSD(T)/CBS Li^+–nucleobase and Li^+–phosphate binding energies has been generated and used to test the performance of 54 DFT functionals including those with empirical dispersion corrections. The results demonstrate that both good and poor performers exist within each functional family. Specifically, an AMPE below 2% is observed within all rungs of Jacob's ladder (except for LDA), while all families (except for the double-hybrids) have at least one functional yielding an AMPE above 5%. Interestingly, the dispersion-corrected functionals typically lead to slightly larger AMPEs (by ~0.4 to 3.4%). In general, all three tested triple-zeta basis sets yield similar performance (AMPEs within ~0.1–0.3%). Furthermore, expansion to quadruple-zeta quality does not significantly impact functional performance, while the performance of the double-zeta basis sets depends on the family. Although the above conclusions are based on binding energies that include counterpoise corrections, the impact of BSSE is minimal (less than 1% change in AMPE). Overall, there are multiple functionals that perform well across all families for direct Li^+ coordination including those from the double-hybrids (B2-PLYP and

PBE-QIDH), RSH variants (ωB97, ωB97M-V, and ωB97X-D), GH meta-GGAs (M06 and MN15), GH GGAs (B3PW91 and B97-2), local meta-GGAs (TPSS and MN15-L) and GGAs (BP86-D3 and PBE). Nevertheless, Li^+–nucleobase cation–π interactions are more difficult to describe than direct nucleic acid coordination. Therefore, functionals must be carefully selected to describe biosystems that include such contacts, with PBE-QIDH, B2-PLYP and ωB97X-D being the best performers across all types of Li^+–nucleic acid interactions. Among the top three performing functionals, PBE-QIDH is the most accurate, albeit computationally more expensive, while ωB97X-D offers an affordable alternative. Importantly, DFT methods also generally reproduce the CCSD(T)/CBS predicted trends in the preferred binding site for a given nucleic acid component as well as the global trend across different components (i.e., phosphate >> G > C >> A ~ T = U).

Overall, this work provides the first highly accurate energetic data for a range of Li^+–nucleic acid interactions, which is critical for the future development of reliable parameters for large-scale MD simulations as existing parameters have been called into question.[41] Furthermore, our new data set has been used to identify dependable DFT methods for the investigation of similar interactions in larger biosystems, which will permit future QM/MM studies to understand the behavior of Li^+ in biology and aid the design of new sensing platforms for lithium detection and extraction.

2.5. References

(1) Role of Metal Ions in Brain Function, Metal Transport, Storage and Homoeostasis. In *Metal-Based Neurodegeneration.*, Ward, R. C. a. R. Ed.; 2013; pp 23-50.

(2) Maret, W. The Metals in the Biological Periodic System of the Elements: Concepts and Conjectures. *Int. J. Mol. Sci.* **2016**, *17* (1), 66.

(3) Dudev, T.; Lim, C. Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, *114* (1), 538-556.

(4) Tchounwou, P. B.; Yedjou, C. G.; Patlolla, A. K.; Sutton, D. J. Heavy Metals Toxicity and the Environment. *EXS* **2012**, *101*, 133-164.

(5) Song, W. J.; Sontz, P. A.; Ambroggio, X. I.; Tezcan, F. A. Metals in Protein–Protein Interfaces. *Ann. Rev. Biophys.* **2014**, *43* (1), 409-431.

(6) Harding, M. M.; Nowicki, M. W.; Walkinshaw, M. D. Metals in Protein Structures: A Review of Their Principal Features. *Crystallogr. Rev.* **2010**, *16* (4), 247-302.

(7) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(8) Zhou, W.; Liu, J. Multi-Metal-Dependent Nucleic Acid Enzymes. *Metallomics* **2018**, *10* (1), 30-48.

(9) Kim, Y.; Nguyen, T.-T. T.; Churchill, D. G. Bioinorganic Chemistry of the Alkali Metal Ions. In *The Alkali Metal Ions: Their Role for Life*, Sigel, A., Sigel, H., Sigel, R. K. O. Eds.; Springer International Publishing, 2016; pp 1-10.

(10) Chellan, P.; Sadler, P. J. The Elements of Life and Medicines. *Philos. Trans. Royal Soc. A.* **2015**, *373* (2037), 20140182.

(11) Burk, P.; Tammiku-Taul, J.; Tamp, S.; Sikk, L.; Sillar, K.; Mayeux, C.; Gal, J.-F.; Maria, P.-C. Computational Study of Cesium Cation Interactions with Neutral and Anionic Compounds Related to Soil Organic Matter. *J. Phys. Chem. A* **2009**, *113* (40), 10734-10744.

(12) Roberts, B. R.; Doecke, J. D.; Rembach, A.; Yévenes, L. F.; Fowler, C. J.; McLean, C. A.; Lind, M.; Volitakis, I.; Masters, C. L.; Bush, A. I.; et al. Rubidium and Potassium Levels Are Altered in Alzheimer's Disease Brain and Blood but Not in Cerebrospinal Fluid. *Acta Neuropathol. Commun.* **2016**, *4* (1), 119.

(13) Meltzer, H. L. A Pharmacokinetic Analysis of Long-Term Administration of Rubidium Chloride. *J. Clin. Pharmacol.* **1991**, *31* (2), 179-184.

(14) Pagani, R.; Gasparini, A.; Ielmini, M.; Caselli, I.; Poloni, N.; Ferrari, M.; Marino, F.; Callegari, C. Twenty Years of Lithium Pharmacogenetics: A Systematic Review. *Psychiatry Res.* **2019**, *278*, 42-50.

(15) Yang, Y.; Wu, N.; Tian, S.; Li, F.; Hu, H.; Chen, P.; Cai, X.; Xu, L.; Zhang, J.; Chen, Z.; et al. Lithium Promotes DNA Stability and Survival of Ischemic Retinal Neurocytes by Upregulating DNA Ligase Iv. *Cell Death Dis.* **2016**, *7* (11), e2473.

(16) Li, Q.; Li, H.; Roughton, K.; Wang, X.; Kroemer, G.; Blomgren, K.; Zhu, C. Lithium Reduces Apoptosis and Autophagy after Neonatal Hypoxia-Ischemia. *Cell Death Dis.* **2010**, *1* (7), e56.

(17) Ren, M.; Senatorov, V. V.; Chen, R.-W.; Chuang, D.-M. Postinsult Treatment with Lithium Reduces Brain Damage and Facilitates Neurological Recovery in a Rat Ischemia/Reperfusion Model. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (10), 6210.

(18) Curtis, E. A.; Bartel, D. P. The Hammerhead Cleavage Reaction in Monovalent Cations. *RNA* **2001**, *7* (4), 546-552.

(19) Aral, H.; Vecchio-Sadus, A. Toxicity of Lithium to Humans and the Environment—a Literature Review. *Ecotoxicol. Environ. Saf.* **2008**, *70* (3), 349-356.

(20) Schrauzer, G. N. Lithium: Occurrence, Dietary Intakes, Nutritional Essentiality. *J. Am. Coll. Nutr.* **2002**, *21* (1), 14-21.

(21) Young, W. Review of Lithium Effects on Brain and Blood. *Cell Transplant.* **2009**, *18* (9), 951-975.

(22) Robinson, B. H.; Yalamanchali, R.; Reiser, R.; Dickinson, N. M. Lithium as an Emerging Environmental Contaminant: Mobility in the Soil-Plant System. *Chemosphere* **2018**, *197*, 1-6.

(23) Tanveer, M.; Hasanuzzaman, M.; Wang, L. Lithium in Environment and Potential Targets to Reduce Lithium Toxicity in Plants. *J. Plant Growth Regul.* **2019**, *38* (4), 1574-1586.

(24) Reiger, S. *New Tech Aims to Extract Lithium for Electric Car Batteries from Oilfield Waste*. 2020. https://www.cbc.ca/news/canada/calgary/lithium-alberta-oilsands-1.5424527 (accessed 2020 03-13).

(25) Cerda, B. A.; Wesdemiotis, C. Li+, Na+, and K+ Binding to the DNA and RNA Nucleobases. Bond Energies and Attachment Sites from the Dissociation of Metal Ion-Bound Heterodimers. *J. Am. Chem. Soc.* **1996**, *118* (47), 11884-11892.

(26) Rodgers, M. T.; Armentrout, P. B. Cationic Noncovalent Interactions: Energetics and Periodic Trends. *Chem. Rev.* **2016**, *116* (9), 5642-5687.

(27) Rodgers, M. T.; Armentrout, P. B. Noncovalent Interactions of Nucleic Acid Bases (Uracil, Thymine, and Adenine) with Alkali Metal Ions. Threshold Collision-Induced Dissociation and Theoretical Studies. *J. Am. Chem. Soc.* **2000**, *122* (35), 8548-8558.

(28) Yang, Z.; Rodgers, M. T. Tautomerization in the Formation and Collision-Induced Dissociation of Alkali Metal Cation-Cytosine Complexes. *Phys. Chem. Chem. Phys.* **2012**, *14* (13), 4517-4526.

(29) Ruan, C.; Huang, H.; Rodgers, M. T. A Simple Model for Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase: Alkali Metal Cations and Trimethyl Phosphate. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 305-314.

(30) Ruan, C.; Huang, H.; Rodgers, M. T. Modeling Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase Via Alkali Metal Cation-Triethyl Phosphate Complexes. *J. Phys. Chem. A* **2007**, *111* (51), 13521-13527.

(31) Zhu, W.; Luo, X.; Puah, C. M.; Tan, X.; Shen, J.; Gu, J.; Chen, K.; Jiang, H. The Multiplicity, Strength, and Nature of the Interaction of Nucleobases with Alkaline and Alkaline Earth Metal Cations: A Density Functional Theory Investigation. *J. Phys. Chem. A* **2004**, *108* (18), 4008-4018.

(32) Pakiari, A. H.; Farrokhnia, M. Nature of Lithium Interactions with DNA Nucleobases: Theoretical Study. *Phys. Chem. Res.* **2014**, *2* (2), 229-243.

(33) Russo, N.; Toscano, M.; Grand, A. Lithium Affinity for DNA and RNA Nucleobases. The Role of Theoretical Information in the Elucidation of the Mass Spectrometry Data. *J. Phys. Chem. B.* **2001**, *105* (20), 4735-4741.

(34) Gillis, E. A. L.; Rajabi, K.; Fridgen, T. D. Structures of Hydrated Li^+−Thymine and Li^+−Uracil Complexes by IRMPD Spectroscopy in the N−H/O−H Stretching Region. *J. Phys. Chem. A* **2009**, *113* (5), 824-832.

(35) Rajabi, K.; Gillis, E. A. L.; Fridgen, T. D. Structures of Alkali Metal Ion−Adenine Complexes and Hydrated Complexes by IRMPD Spectroscopy and Electronic Structure Calculations. *J. Phys. Chem. A* **2010**, *114* (10), 3449-3456.

(36) Burda, J. V.; Šponer, J.; Hobza, P. Ab Initio Study of the Interaction of Guanine and Adenine with Various Mono- and Bivalent Metal Cations (Li^+, Na^+, K^+, Rb^+, Cs^+; Cu^+, Ag^+, Au^+; Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}; Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem. A*. **1996**, *100* (17), 7250-7255.

(37) Stasyuk, O. A.; Solà, M.; Swart, M.; Fonseca Guerra, C.; Krygowski, T. M.; Szatylowicz, H. Effect of Alkali Metal Cations on Length and Strength of Hydrogen Bonds in DNA Base Pairs. *ChemPhysChem.* **2020**, *21* (18), 2112-2126.

(38) Meyer, M.; Sühnel, J. Interaction of Cyclic Cytosine-, Guanine-, Thymine-, Uracil- and Mixed Guanine-Cytosine Base Tetrads with K^+, Na^+ and Li^+ Ions—a Density Functional Study. *J. Biomol. Struct. Dyn.* **2003**, *20* (4), 507-517.

(39) Azargun, M.; Fridgen, T. D. Guanine Tetrads: An IRMPD Spectroscopy, Energy Resolved Sori-Cid, and Computational Study of M(9-Ethylguanine)4+ (M = Li, Na, K, Rb, Cs) in the Gas Phase. *Phys. Chem. Chem. Phys.* **2015**, *17* (39), 25778-25785.

(40) Yoo, J.; Aksimentiev, A. Competitive Binding of Cations to Duplex DNA Revealed through Molecular Dynamics Simulations. *J. Phys. Chem. B.* **2012**, *116* (43), 12946-12954.

(41) Savelyev, A.; MacKerell, A. D. Competition among Li^+, Na^+, K^+, and Rb^+ Monovalent Ions for DNA in Molecular Dynamics Simulations Using the Additive Charmm36 and Drude Polarizable Force Fields. *J. Phys. Chem. B.* **2015**, *119* (12), 4428-4440.

(42) Gebala, M.; Bonilla, S.; Bisaria, N.; Herschlag, D. Does Cation Size Affect Occupancy and Electrostatic Screening of the Nucleic Acid Ion Atmosphere? *J. Am. Chem. Soc.* **2016**, *138* (34), 10925-10934.

(43) Yoo, J.; Aksimentiev, A. Improved Parametrization of Li^+, Na^+, K^+, and Mg^{2+} Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J. Phys. Chem. Lett.* **2012**, *3* (1), 45-50.

(44) He, Q.; Liao, X.; Xia, L.; Li, Z.; Wang, H.; Zhao, Y.; Truhlar, D. G. Accurate Binding Energies for Lithium Polysulfides and Assessment of Density Functionals for Lithium–Sulfur Battery Research. *J. Phys. Chem. C.* **2019**, *123* (34), 20737-20747.

(45) Savarese, M.; Bremond, E.; Adamo, C. Exploring the Limits of Recent Exchange-Correlation Functionals in Modeling Lithium/Benzene Interaction. *Theor. Chem. Acc.* **2016**, *135* (4), 1-11.

(46) *Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16 Rev. B.01*; Wallingford, CT, 2016.

(47) Neese, F. The Orca Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73-78.

(48) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction Revisited. *J. Chem. Phys.* **2007**, *127* (15), 154108.

(49) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124* (17), 174104.

(50) Becke, A. D.; Johnson, E. R. A Unified Density-Functional Treatment of Dynamical, Nondynamical, and Dispersion Correlations. *J. Chem. Phys.* **2007**, *127* (12), 124108.

(51) Kannemann, F. O.; Becke, A. D. Van Der Waals Interactions in Density-Functional Theory: Intermolecular Complexes. *J. Chem. Theory Comput.* **2010**, *6* (4), 1081-1088.

(52) Otero-de-la-Roza, A.; Johnson, E. R. Non-Covalent Interactions and Thermochemistry Using Xdm-Corrected Hybrid and Range-Separated Hybrid Density Functionals. *J. Chem. Phys.* **2013**, *138* (20), 204109.

(53) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, N_2, and H_2O. *Chem. Phys. Lett.* **1998**, *286* (3,4), 243-252.

(54) Kim, K. S.; Mhin, B. J.; Choi, U. S.; Lee, K. Ab Initio Studies of the Water Dimer Using Large Basis Sets: The Structure and Thermodynamic Energies. *J. Chem. Phys.* **1992**, *97* (9), 6649-6662.

(55) Mackie, I. D.; DiLabio, G. A. Approximations to Complete Basis Set-Extrapolated, Highly Correlated Non-Covalent Interaction Energies. *J. Chem. Phys.* **2011**, *135* (13), 134318.

(56) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *J. Chem. Theory Comput.* **2014**, *10* (1), 49-57.

(57) Sherrill, C. D. Chapter 4 - Wavefunction Theory Approaches to Noncovalent Interactions. In *Non-Covalent Interactions in Quantum Chemistry and Physics*, Otero de la Roza, A., DiLabio, G. A. Eds.; Elsevier, 2017; pp 137-168.

(58) García, J. S.; Brémond, É.; Campetella, M.; Ciofini, I.; Adamo, C. Small Basis Set Allowing the Recovery of Dispersion Interactions with Double-Hybrid Functionals. *J. Chem. Theory Comput.* **2019**, *15* (5), 2944-2953.

(59) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P.; Taylor, P. R. Basis Set Convergence of the Interaction Energy of Hydrogen-Bonded Complexes. *J. Chem. Phys.* **1999**, *111* (20), 9157-9167.

(60) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19* (4), 553-566.

(61) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conf. Proc.* **2001**, *577* (1), 1-20.

(62) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.

(63) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456-1465.

(64) Vydrov, O. A.; Van Voorhis, T. Nonlocal Van Der Waals Density Functional: The Simpler the Better. *J. Chem. Phys.* **2010**, *133* (24), 244103.

(65) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007-1023.

(66) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96* (9), 6796-6806.

(67) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98* (2), 1358-1371.

(68) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297-3305.

(69) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124* (3), 034108.

(70) Schwabe, T.; Grimme, S. Towards Chemical Accuracy for the Thermodynamics of Large Molecules: New Hybrid Density Functionals Including Non-Local Correlation Effects. *Phys. Chem. Chem. Phys.* **2006**, *8* (38), 4398-4401.

(71) Brémond, É.; Sancho-García, J. C.; Pérez-Jiménez, Á. J.; Adamo, C. Communication: Double-Hybrid Functionals from Adiabatic-Connection: The QIDH Model. *J. Chem. Phys.* **2014**, *141* (3), 031101.

(72) Brémond, E.; Adamo, C. Seeking for Parameter-Free Double-Hybrid Functionals: The PBE0-Dh Model. *J. Chem. Phys.* **2011**, *135* (2), 024106.

(73) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2* (21), 2810-2817.

(74) Peverati, R.; Truhlar, D. G. Screened-Exchange Density Functionals with Broad Accuracy for Chemistry and Solid-State Physics. *Phys. Chem. Chem. Phys.* **2012**, *14* (47), 16187-16191.

(75) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized, Range-Separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144* (21), 214110.

(76) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128* (8), 084106.

(77) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615-6620.

(78) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 9904-9924.

(79) Heyd, J.; Scuseria, G. E. Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd–Scuseria–Ernzerhof Screened Coulomb Hybrid Functional. *J. Chem. Phys.* **2004**, *121* (3), 1187-1192.

(80) Heyd, J.; Scuseria, G. E. Assessment and Validation of a Screened Coulomb Hybrid Density Functional. *J. Chem. Phys.* **2004**, *120* (16), 7274-7280.

(81) Heyd, J.; Peralta, J. E.; Scuseria, G. E.; Martin, R. L. Energy Band Gaps and Lattice Parameters Evaluated with the Heyd-Scuseria-Ernzerhof Screened Hybrid Functional. *J. Chem. Phys.* **2005**, *123* (17), 174101.

(82) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Erratum: "Hybrid Functionals Based on a Screened Coulomb Potential" [J. Chem. Phys. 118, 8207 (2003)]. *J. Chem. Phys.* **2006**, *124* (21), 219906.

(83) Henderson, T. M.; Izmaylov, A. F.; Scalmani, G.; Scuseria, G. E. Can Short-Range Hybrids Describe Long-Range-Dependent Properties? *J. Chem. Phys.* **2009**, *131* (4), 044108.

(84) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. Efficient Evaluation of Short-Range Hartree-Fock Exchange in Large Molecules and Periodic Systems. *J. Chem. Phys.* **2006**, *125* (10), 104103.

(85) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, *125* (22), 224106.

(86) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.

(87) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A Long-Range Correction Scheme for Generalized-Gradient-Approximation Exchange Functionals. *J. Chem. Phys.* **2001**, *115* (8), 3540-3544.

(88) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. Importance of Short-Range Versus Long-Range Hartree-Fock Exchange for the Performance of Hybrid Density Functionals. *J. Chem. Phys.* **2006**, *125* (7), 074106.

(89) Vydrov, O. A.; Scuseria, G. E. Assessment of a Long-Range Corrected Hybrid Functional. *J. Chem. Phys.* **2006**, *125* (23), 234109.

(90) Vydrov, O. A.; Scuseria, G. E.; Perdew, J. P. Tests of Functionals for Systems with Fractional Electron Number. *J. Chem. Phys.* **2007**, *126* (15), 154109.

(91) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1), 51-57.

(92) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120* (1), 215-241.

(93) Zhao, Y.; Truhlar, D. G. Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average Than B3LYP for Ground States. *J. Phys. Chem. A* **2006**, *110* (49), 13126-13130.

(94) Boese, A. D.; Martin, J. M. L. Development of Density Functionals for Thermochemical Kinetics. *J. Chem. Phys.* **2004**, *121* (8), 3405-3416.

(95) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032-5051.

(96) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372-1377.

(97) Peverati, R.; Truhlar, D. G. Communication: A Global Hybrid Generalized Gradient Approximation to the Exchange-Correlation Functional That Satisfies the Second-Order Density-Gradient Constraint and Has Broad Applicability in Chemistry. *J. Chem. Phys.* **2011**, *135* (19), 191102.

(98) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648-5652.

(99) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158-6170.

(100) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785-789.

(101) Xu, X.; Goddard, W. A. The X3LYP Extended Density Functional for Accurate Descriptions of Nonbond Interactions, Spin States, and Thermochemical Properties. *Proc. Natl. Acad. Sci.* **2004**, *101* (9), 2673.

(102) Cohen, A. J.; Handy, N. C. Dynamic Correlation. *Mol. Phys.* **2001**, *99* (7), 607-615.

(103) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid Exchange-Correlation Functional Determined from Thermochemical Data and Ab Initio Potentials. *J. Chem. Phys.* **2001**, *115* (20), 9233-9242.

(104) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta--Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(105) Peverati, R.; Truhlar, D. G. M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics. *J. Chem. Phys. Lett.* **2012**, *3* (1), 117-124.

(106) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125* (19), 194101.

(107) Peverati, R.; Truhlar, D. G. An Improved and Broadly Accurate Local Approximation to the Exchange–Correlation Density Functional: The MN12-L Functional for Electronic Structure Calculations in Chemistry and Physics. *Phys. Chem. Chem. Phys.* **2012**, *14* (38), 13171-13174.

(108) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional for Kohn–Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. *J. Chem. Theory Comput.* **2016**, *12* (3), 1280-1293.

(109) Adamo, C.; Barone, V. Exchange Functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW Models. *J. Chem. Phys.* **1998**, *108* (2), 664-675.

(110) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098-3100.

(111) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157* (3), 200-206.

(112) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200-1211.

(113) Keiluweit, M.; Kleber, M. Molecular-Level Interactions in Soils and Sediments: The Role of Aromatic π-Systems. *Environ. Sci. Technol.* **2009**, *43* (10), 3421-3429.

(114) Dougherty, D. A. The Cation−π Interaction. *Acc. Chem. Res.* **2013**, *46* (4), 885-893.

(115) McConnell, T. L.; Wheaton, C. A.; Hunter, K. C.; Wetmore, S. D. Effects of Hydrogen Bonding on the Acidity of Adenine, Guanine, and Their 8-Oxo Derivatives. *J. Phys. Chem. A* **2005**, *109* (28), 6351-6362.

(116) Whittleton, S. R.; Hunter, K. C.; Wetmore, S. D. Effects of Hydrogen Bonding on the Acidity of Uracil Derivatives. *J. Phys. Chem. A* **2004**, *108* (38), 7709-7718.

(117) Sayyed, F. B.; Suresh, C. H. Quantitative Assessment of Substituent Effects on Cation-π Interactions Using Molecular Electrostatic Potential Topography. *J. Phys. Chem. A* **2011**, *115* (33), 9300-9307.

(118) Amunugama, R.; Rodgers, M. T. Influence of Substituents on Cation−π Interactions. 1. Absolute Binding Energies of Alkali Metal Cation−Toluene Complexes Determined by Threshold Collision-Induced Dissociation and Theoretical Studies. *J. Phys. Chem. A.* **2002**, *106* (22), 5529-5539.

(119) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. Theoretical Study of the Protonation and Deprotonation of Cytosine. Implications for the Interaction of Cytosine with Water. *J. Mol. Struct.* **2000**, *519* (1), 1-11.

(120) Santra, G.; Sylvetsky, N.; Martin, J. M. L. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: Revdsd-Pbep86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123* (24), 5129-5143.

(121) Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets: Hydrogen Bonding. *J. Chem. Theory Comput.* **2020**, *16* (4), 2355-2368.

(122) Husch, T.; Freitag, L.; Reiher, M. Calculation of Ligand Dissociation Energies in Large Transition-Metal Complexes. *J. Chem. Theory Comput.* **2018**, *14* (5), 2456-2468.

(123) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115* (19), 2315-2372.

(124) Zhao, Y.; Truhlar, D. G. Density Functional Theory for Reaction Energies: Test of Meta and Hybrid Meta Functionals, Range-Separated Functionals, and Other High-Performance Functionals. *J. Chem. Theory Comput.* **2011**, *7* (3), 669-676.

Chapter 3: Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes

Preface: The contents of this chapter have been published [Boychuk, B. T. A.; Meyer, S. P.; Wetmore, S. D. Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* 2023, 11].

3.1. Introduction

Metal ions are utilized in various aspects of life. For example, metals are involved in basic biological functions, useful for the design of new energy storage materials, and are integral to the development of pharmaceutical drugs to treat diseases.[1-3] Group I metals in particular play critical roles and are actively exploited in such applications.[4] For example, Na^+ and K^+ are essential cofactors for maintaining the structure and function of nucleic acids, stabilizing cell membranes, and aiding ATP utilization.[5, 6] Additionally, Na^+ is commonly used in body soaps, while K^+ is used in laundry detergents.[7, 8] On the other hand, Li^+ is used in clean energy initiatives, including the production of lithium-ion batteries,[9] and in drugs to treat bipolar disorders.[10] Rb^+ and Cs^+ have been utilized in perovskite materials to produce enhanced solar cells.[11]

Despite their usefulness in our lives, an overabundance of group I metals can occur in our food sources and drinking water through industrial processes such as mining, waste dumping, and emissions from industrial plants.[12] As a result, group I metals can accumulate in the environment and human body, which causes various detrimental effects. In the human body, high levels of Li^+ can result in cerebellar or renal dysfunction.[13] Alternatively, an imbalance of Na^+ and K^+ can lead to hypertension, which may cause a stroke or heart disease.[14] On the other hand, Rb^+ and Cs^+ are known to be mildly toxic, having the ability to displace K^+ and interfere with cellular functions.[15] As a specific example, increased levels of Rb^+ or Cs^+ in the brain have been correlated with the onset of Parkinson's disease.[16] Since group I metals are involved in different aspects of ecological and human biology, as well as disease progression, it is important to understand how these metals interact with biomolecules (e.g., nucleic acids and proteins).

In addition to answering interesting questions in biology, investigating interactions between group I metals and nucleic acids in particular is key to unlocking many new applications. For example, the development of novel biomaterials holds promise in the energy sector. Specifically, nucleic acids have been explored as a way to enhance lithium–sulfur batteries.[17] Alternatively, there is increasing interest in the design of nucleic acid biosensors to detect toxic metals in the environment or body.[15] Indeed, nucleic acids have advantageous metal binding properties, inherently possessing numerous metal binding sites that exhibit varying binding affinities and having the ability to fold into diverse 3D architectures that afford unique metal-binding pockets (e.g., G-quadruplexes and helical junctions).[15] As a result, multiple nucleic acid sensors have been designed to target metal ions. For example, DNAzymes are currently available in the market to selectively sense Pb^{2+},[18] while an aptamer has been designed to specifically target Cd^{2+}.[19] Among group I metals, the NaA43 DNAzyme has been shown to target Na^{+} in the presence of other mono-, di-, and trivalent metals.[20] However, although nucleic acid sensors have also been designed for K^{+} and Cs^{+},[15, 21] these solutions are not exclusive for the given metal[15, 21, 22] and/or detailed testing of their function in the presence of a wide range of metals has yet to be done, and no sensors have been made to date for other group I metals. Therefore, more work is necessary to establish functional nucleic acid sensors for group I metals, as well as exploit metal–nucleic acid interactions in other applications such as new materials for energy storage.

As a first step to understanding the roles of metals in nature and designing nucleic acid sensors or novel energy storage materials, among other applications, fundamental information about the structure and binding energies of metal–nucleic acid complexes is required. Although early experimental studies used the kinetic method approach to estimate

the gas-phase binding energies between Li^+, Na^+, or K^+ and each canonical DNA/RNA nucleobase,[23] this technique is unable to yield absolute metal ion binding affinities due to the required use of a reference, with unique references employed for different metals also preventing cross-comparisons between metals and/or nucleic acid components.[24, 25] While the absolute gas-phase binding energies of Li^+, Na^+ or K^+ to A, C, T, or U, and Rb^+ or Cs^+ to C were determined using threshold collision-induced dissociation (TCID),[24-26] limitations exist in these studies at least in part due to the lack of structural information regarding the complexes formed and missing metal–nucleobase combinations. TCID measurements can also lack sensitivity, with Li^+ having been deemed particularly challenging due to a low mass-to-charge ratio and high velocity that results in ineffective metal trapping.[27] Furthermore, there is a high risk of nucleobase tautomerization,[25, 26] which may result in complexes that are not relevant to DNA/RNA present in biosystems or nucleic acid-based applications. Although infrared multiple photon dissociation (IRMPD) spectroscopy has also been used to gain information about the complexes formed between Li^+, Na^+, K^+, Rb^+, or Cs^+ bound to A,[28] as well as mono- or dihydrated Li^+ bound to T or U,[29, 30] the method necessitates the use of computational methods (B3LYP) to obtain the corresponding structures and binding energies. Thus, atomic level information that unequivocally correlates the structures and binding energies for the complete set of possible group I metal–nucleic acid complexes is still required.

To fill in knowledge gaps and complement previous experimental work on group I metal–nucleic acid interactions, several computational studies have been performed.[24, 28-37] For example, MP2 has been used to investigate the gas-phase structure and stability of select group 1–nucleobase complexes.[24, 33, 36, 37] However, all possible complexes for each nucleobase and nucleobase binding sites have not been systematically explored in this

capacity. B3LYP has been more recently used to investigate a more complete set of group 1 metal–nucleobase complexes [28-30, 33-35], but the reliability of DFT (especially the popular, nondispersion-corrected B3LYP) for the binding affinities of group I metals to various nucleic acid sites remains ambiguous. Although a computational study has provided CCSD(T) reference binding strengths at the complete basis set limit (CBS) for Li^+ interactions with A, C, G, T, U, or the phosphate moiety, as well as tested the performance of 54 DFT methods (Chapter 2),[31] the remaining group I metal–nucleic acid complexes have yet to be explored. Structural and energetic evaluation of complexes involving the remaining group I metals with a high-level of theory is essential for providing accurate chemical information that is currently missing from experimental data sets. Furthermore, it is necessary to establish reliable functionals for the entire group due to known variations in functional performance across different metals.[38-40]

To gain fundamental insight into the structures and binding energies of group I metal–nucleic acid complexes as well as identify DFT methods that can accurately describe these interactions for future applications, the present work generates a complete data set of gas-phase binding strengths for group I metal–nucleic acid complexes, which has proven challenging for experimental determination. Specifically, the binding energies of each metal to each canonical DNA/RNA nucleobase (A, C, G, T, and U) and a model for the phosphate moiety (dimethyl phosphate; Figure 3.1) are calculated using CCSD(T) extrapolated to the CBS limit. Subsequently, 61 DFT methods (Table 1) across different functional families according to the metaphorical ‘Jacob’s Ladder’ are tested for their ability to reproduce the CCSD(T)/CBS results. The most reliable and robust DFT methods identified from this work can be applied in future studies to uncover the roles of group I metals in biology, guide the rational design of nucleic acid sensors to target metals in the

environment and the human body, and construct new energy storage materials. The highly-accurate data set generated in the present work can also be used to test new functionals and develop new parameters for molecular dynamics (MD) simulations to expand computational studies of group I metal–nucleic acid complexes more broadly.

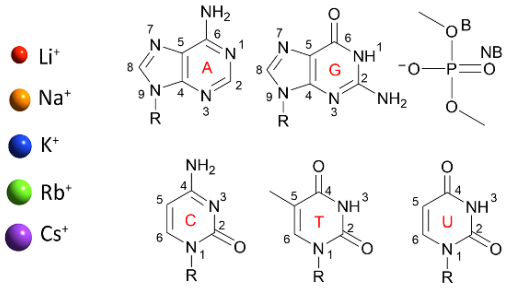


Figure 3.1. Group I metals and nucleic acid components considered in this study. Structure and chemical numbering provided for nucleic acid components (B = bridging and NB = nonbridging oxygen in the dimethyl phosphate model).

Table 3.1. Functionals tested for their ability to describe group I metal–nucleic acid interactions.

Family	Functional	D3[a]	D4[a]	%HF[b]	%MP2	Ref.
Double-Hybrid	B2-PLYP	X		53	27	[41]
	mPW2-PLYP			55	25	[42]
	PBE-QIDH			69	33.4	[43]
	PBE0-DH			50	12.5	[44]
	DSD-PBEP86			68	(51,23)[d]	[45]
RSH	M11			42.8–100[c]		[46]
	MN12-SX	X		25–0[c]		[47]
RSH meta-GGA	ωB97M-V			15		[48]

Table 3.1. Continued.

Family	Functional	D3[a]	D4[a]	%HF[b]	%MP2	Ref.
RSH GGA	ωB97					49
	ωB97X	X	X	16		49
	ωB97X-D			22		50
	ωB97X-V			16.7		51
	HSE06	X		25–0[b]		52-58
	LC-PBE					59, 60
	LC-ωPBE	X				61-63
	CAM-B3LYP	X		19–65[b]		64
GH meta-GGA	M06			27		65
	M06-2X			54		65
	M06-HF			100		66
	BMK			42		67
	MN15			44		68
	PW6B95			28		69
GH GGA	BH&HLYP			50		70
	SOGGA11-X			40.15		71
	B3PW91	X		20		72
	PBE0	X		25		73
	B3LYP	X	X	20		72, 74
	X3LYP	X		21		75
	O3LYP	X				76
	TPSSh			10		77, 78
	B97-2			21		79
Local meta-GGA	TPSS	X	X			80
	revTPSS					81, 82
	M11-L					83
	M06-L					84
	MN12-L					85
	MN15-L					86
Local GGA	mPW91					87
	BLYP	X				74, 88, 89
	BP86	X				74, 88
	PBE	X	X			59
Local LDA	SVWN5					90

[a]Functionals indicated by an X were considered with and without the D3(BJ) or D4 empirical dispersion correction. [b]Percentage of Hartree-Fock exchange (%HF). [c]The first value is the %HF at short range and the second value is the %HF at long range. [d]The MP2 component of this functional is calculated using the spin-component-scaled formulism and the two values represent opposite- and same-spin values, respectively.

3.2. Computational Methodology

Models were constructed by placing a group I metal (Li^+, Na^+, K^+, Rb^+, or Cs^+) at each potential binding site of each nucleobase (A, C, G, T, or U) or the phosphate component (Figure 3.1). Dimethyl phosphate was used to represent the phosphate moiety, which places methyl caps at the location of the connecting sugars to prevent self-interactions between a hydrogen cap and the phosphate moiety, and has been applied in previous work investigating metal–phosphate interactions in the context of nucleic acids.[31, 91] Although previous work used MP2/aug-cc-pVTZ to optimize Li^+–nucleic acid complexes (Chapter 2),[31] all metal–nucleic acid complexes investigated in the present work were optimized with MP2(full)/def2-TZVPP for consistency between the metals, with the def2 series being the best basis sets available for heavier group I metals. Frequency calculations were subsequently performed to confirm stable minima. The MP2 geometries were used for all subsequent calculations.

For each metal–nucleic acid complex, the binding energy (BE) was calculated as

$$E^{BE} = E^{complex} - E^{nucleic\ acid} - E^{metal} \quad (3.1)$$

where E^{BE} is the BE of the complex, $E^{complex}$ is the dimer energy, and $E^{nucleic\ acid}$ and E^{metal} are the isolated monomer energies of the nucleic acid subcomponent and metal, respectively. The reference BEs were evaluated at the CCSD(T)/CBS level of theory using the most common approach for obtaining benchmark values for noncovalent interactions.[92] Specifically, the following equation was used

$$E(CCSD(T)/CBS) = E(HF) + E^{corr}(MP2/CBS) + \Delta CCSD(T) \quad (3.2)$$

where E(HF) was evaluated using def2-QZVPP, E^{corr}(MP2/CBS) is the MP2 correlation energy extrapolated to the CBS limit using def2-TZVPP and def2-QZVPP according to the

Helgaker extrapolation scheme,[93, 94] and ΔCCSD(T) was calculated as the difference between the MP2 and CCSD(T) energies evaluated with the def2-TZVPP basis set. The def2 basis sets applied in the present work were previously shown to be well suited for this extrapolation scheme.[95] All-electron calculations were carried out with CCSD(T) and MP2(full) for the lighter metals (i.e., Li^+ and Na^+), while CCSD(T) and MP2 were combined with the Stuttgart Dresden relativistic (RLC) effective core potential (ECP) for the heavier metals (i.e., K^+, Rb^+, and Cs^+). The large-core ECP was used to reduce computational costs, with minimal differences in ground-state energies and ionization potentials compared to all-electron relativistic calculations found in a previous study for the heavier group I metals.[96] As done in the literature and proposed by Halkier *et al*.,[97] the averages of the counterpoise uncorrected and corrected binding strengths were evaluated to account for the overestimation and underestimation of true CCSD(T)/CBS BE values, respectively.[31, 98-101]

The accurate CCSD(T)/CBS data set was subsequently used to test the performance of 61 functionals (Table 3.1) from different families of the hierarchal 'Jacob's Ladder'. These calculations were performed using the same MP2 geometries employed for the CCSD(T)/CBS calculations. We note that previous work has highlighted little impact in functional performance when the geometry of select Li^+–nucleic acid complexes were re-optimized with DFT.[31] The def2-TZVPP basis set was used throughout, with the small-core Stuttgart Dresden RLC ECP used for the heavier metals (K^+, Rb^+, and Cs^+). We note that previous work on a selection of Li^+–nucleic acid complexes revealed a negligible difference in DFT functional performance upon basis set expansion to def2-QZVPP.[31] Therefore, due to the large number of functional, metal, and nucleic acid component combinations considered in the present work as well as the intended application of the best methods to

larger nucleic acid systems, we employed the computationally more efficient TZ basis set in the present work. The reported binding energies include counterpoise corrections to account for the basis set superposition error (BSSE), which was determined using the Boys and Bernardi scheme, although the impact of neglecting the counterpoise correction is also discussed in the Results and Discussion. For 15 functionals, Grimme's D3[102, 103] empirical dispersion correction with Becke–Johnson damping (BJ)[104] was used, while the D4[105] model was combined with 4 functionals. All calculations were performed in the gas phase using the default settings in Gaussian 16 (B.01),[106] with the exception of the B3LYP-D4, PBE-D4, TPSS-D4, ωB97X-D4, ωB97X-D3(BJ), ωB97X-V, and ωB97M-V calculations, which were performed using the default settings in ORCA 5.0.2 and the convergence criteria set to "VeryTightSCF".[107]

3.3 Results and Discussion

3.3.1. Generating a Highly Accurate CCSD(T)/CBS Data Set of Binding Strengths for Group I Metal–Nucleic Acid Complexes

As noted in the Computational Methodology, each group I metal (Li^+, Na^+, K^+, Rb^+, or Cs^+) was placed at each potential binding site in each nucleic acid component (A, C, G, T, U, or dimethyl phosphate; Figure 3.1). Specifically, for the nucleobases, the N3 and N7 binding sites of the purines were explored, as well as N1 and N6 of A, and O6 of G, while the O2 and O4 binding sites of the pyrimidines were considered, as well as N4 of C. For the nucleic acid backbone, coordination to the bridging (B) and nonbridging (NB) oxygens of dimethyl phosphate was considered. 64 unique complexes were identified between a group I metal and a nucleic acid component (Figure 3.2; Table B.1, Appendix B). The nucleobase binding sites identified are consistent with structures previously reported in the

literature.[24, 28-37] These structures were used to calculate the corresponding gas-phase CCSD(T)/CBS relative energies for each metal–nucleic acid component combination (Figure 3.2) and binding strengths for each complex (Table B.2).

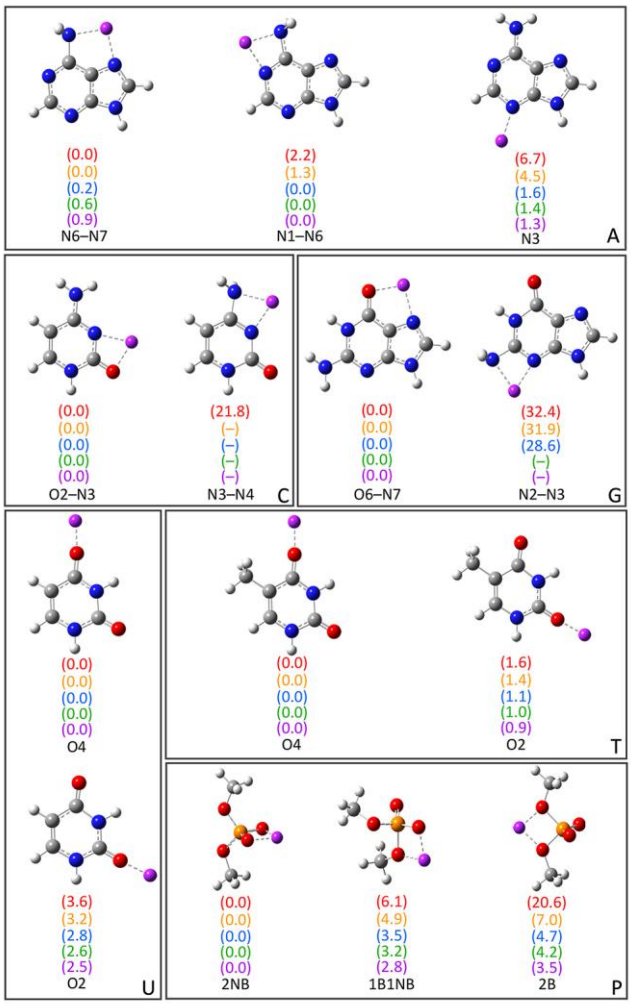


Figure 3.2. CCSD(T)/CBS//MP2/def2-TZVPP characterized complexes between group I metals and the nucleobases (A, G, C, T, and U) or dimethyl phosphate (P). The relative energies for a given metal–nucleic acid component combination (parentheses) are provided in kcal/mol and colored according to metal identity (Li^+ (red), Na^+ (orange), K^+ (blue), Rb^+ (green), and Cs^+ (purple)). The CCSD(T)/CBS binding strengths and MP2/def2-TZVPP coordination distances are provided in Tables B.1 and B.2 in Appendix B.

Bidentate complexes were characterized between each group I metal and the N6–N7, and N1–N6 sites of A, which involve rotation of the N6 amino group compared to isolated A. Although the binding strengths are nearly equal at both binding sites for all metals (within ~2 kcal/mol), Li^+ and Na^+ preferentially coordinate to the N6–N7 position (binding energies of –51.1 and –36.0 kcal/mol, respectively), while K^+, Rb^+, and Cs^+ slightly prefer binding at N1–N6 (binding energies of –19.6, –17.1, and –15.1 kcal/mol, respectively). The monodentate complex formed at N3 of A for all metals is ~4 – 7 kcal/mol less stable for the lighter group I metals, but within 1.6 kcal/mol for K^+, Rb^+, and Cs^+. For G, the most stable complex for each group I metal involves bidentate coordination at the O6–N7 position, with the binding strength decreasing as Li^+ (–73.0 kcal/mol) >> Na^+ >> K^+ > Rb^+ > Cs^+ (–33.4 kcal/mol). Although this is the only G complex characterized for Rb^+ and Cs^+, Li^+, Na^+, and K^+ also coordinate to N3 of G, which results in a significantly less stable complex (by ~29 – 32 kcal/mol).

For C, the most energetically favorable complex for all group I metals involves bidentate coordination at the O2–N3 position, with the binding strength decreasing as Li^+ (–69.3 kcal/mol) >> Na^+ >> K^+ > Rb^+ > Cs^+ (–30.5 kcal/mol). Only Li^+ forms an additional complex at N3–N4 that is afforded by N4 amino group rotation compared to isolated C, which is 21.8 kcal/mol less stable. Within T and U binding sites, group I metals form monodentate complexes at O4 and O2, with the O4 site being preferred for both nucleobases by up to 1.6 kcal/mol for T and 3.6 kcal/mol for U. In general, there are minimal differences (0.1 – 0.3 kcal/mol) between the binding strengths at O4 of T and U regardless of metal identity, while the differences range from ~1.0–2.0 kcal/mol at O2.

Three complexes were isolated for each group I metal that involve coordination to dimethyl phosphate (P), including those with coordination to one bridging and one

nonbridging oxygen (1B1NB), two bridging oxygens (2B), or two nonbridging oxygens (2NB). All metals except Li^+ also coordinate to one nonbridging oxygen in the P(2B) complex (Figure B.1). However, the strongest binding interactions for all metals occur at the 2NB position, followed by 1B1NB and then 2B. Nevertheless, the difference in the binding strength as a function of dimethyl phosphate site decreases from ~20 kcal/mol to ~4 kcal/mol as group I is descended.

Overall, the magnitude of group I metal–nucleic acid interaction energies decreases with an increase in the size of the metal ($Li^+ >> Na^+ >> K^+ > Rb^+ > Cs^+$), which matches the general trends seen in previous experimental[23] and computational[33-37] studies. Indeed, the metal–ligand distances increase with group I metal size (Table B.1, Appendix B), which weakens ion–dipole and ion–induced dipole interactions between the metal and nucleic acid component. As a result, there is less distinction between the strongest and weakest binding sites for a given nucleobase as the size of the metal increases, especially for the largest metal Cs^+. When the preferred binding site for each nucleobase is considered, the group I metal–nucleobase interaction energies decrease as G > C >> A ~ T = U, which matches the trends predicted by TCID for A, C, T, and U.[23] Furthermore, the group 1 metal binding affinity to dimethyl phosphate is ~2 (Li^+ and Na^+) or ~3 (K^+, Rb^+, and Cs^+) times greater than that to the most favorable nucleobase site (G(O6–N7)), which correlates with experimental data for trimethylphosphate.[91]

In addition to reproducing trends in binding strengths, our CCSD(T)/CBS data sometimes agrees well with TCID predicted interaction energies. For example, the differences between the experimental[24] and CCSD(T)/CBS results for T and U are <1 kcal/mol. Nevertheless, deviations exist between experimental and computational data, which may arise at least in part due to limitations in the experimental methodology. For

example, there is a significant difference (2.9 kcal/mol) between our CCSD(T)/CBS and previously published TCID data[24] for the Li^+–A complex, with TCID having a known low sensitivity for Li^+.[27] There is an even larger difference (13 kcal/mol) between the TCID and CCSD(T)/CBS results for Li^+ bound to C, which may arise from the low Li^+ sensitivity coupled with nucleobase tautomerization due to the use of thermal vaporization.[25, 26] Indeed, the experimental[24, 25] and theoretical results for the remaining group I metals (Na^+, K^+, Rb^+, and Cs^+) bound to C or A deviate by ~1–4 kcal/mol, possibly due at least in part to nucleobase tautomerization. In addition to providing accurate data for previously studied complexes, our CCSD(T)/CBS calculations fill in gaps arising from missing experimental thermochemical data for G interactions with all group I metals, and Rb^+ or Cs^+ bound to A, T, or U. Furthermore, accurate metal binding strengths for all group I metals to dimethyl phosphate are provided herein for the first time, which affords a more realistic model than that used in experimental studies (neutral trimethylphosphate).[91] Thus, our work offers the first complete and accurate data set of the thermochemistry related to group I metal binding to each nucleic acid component.

3.3.2. Accuracy of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes Involving Direct Coordination

As discussed in the Computational Methodology, the performance of 61 functionals (Table 3.1) in combination with def2-TZVPP is assessed against our newly generated CCSD(T)/CBS binding strengths for 64 directly coordinated group I metal–nucleic acid complexes (Figure B.2, Appendix B). Initially, we consider the percent errors (PEs) in the DFT predicted binding strengths for each metal–nucleic acid component combination as a function of the functional (Figure 3.3). In general, PEs change both as a function of the

complex and functional considered. Indeed, functional performance is dependent on the identity of the metal, nucleobase, and binding site. For example, interactions with the phosphate moiety tend to give rise to small errors regardless of metal (<5%), while weaker complexes, such as those involving N3 of G, and A binding sites, give rise to larger deviations (>25%). Furthermore, larger errors occur as group 1 is descended, which underscores the importance of testing the reliability of DFT methods for different metals. Nevertheless, some functionals consistently exhibit the poorest performance across many complexes for all group I metals, such as O3LYP-D3(BJ) and X3LYP-D3(BJ).

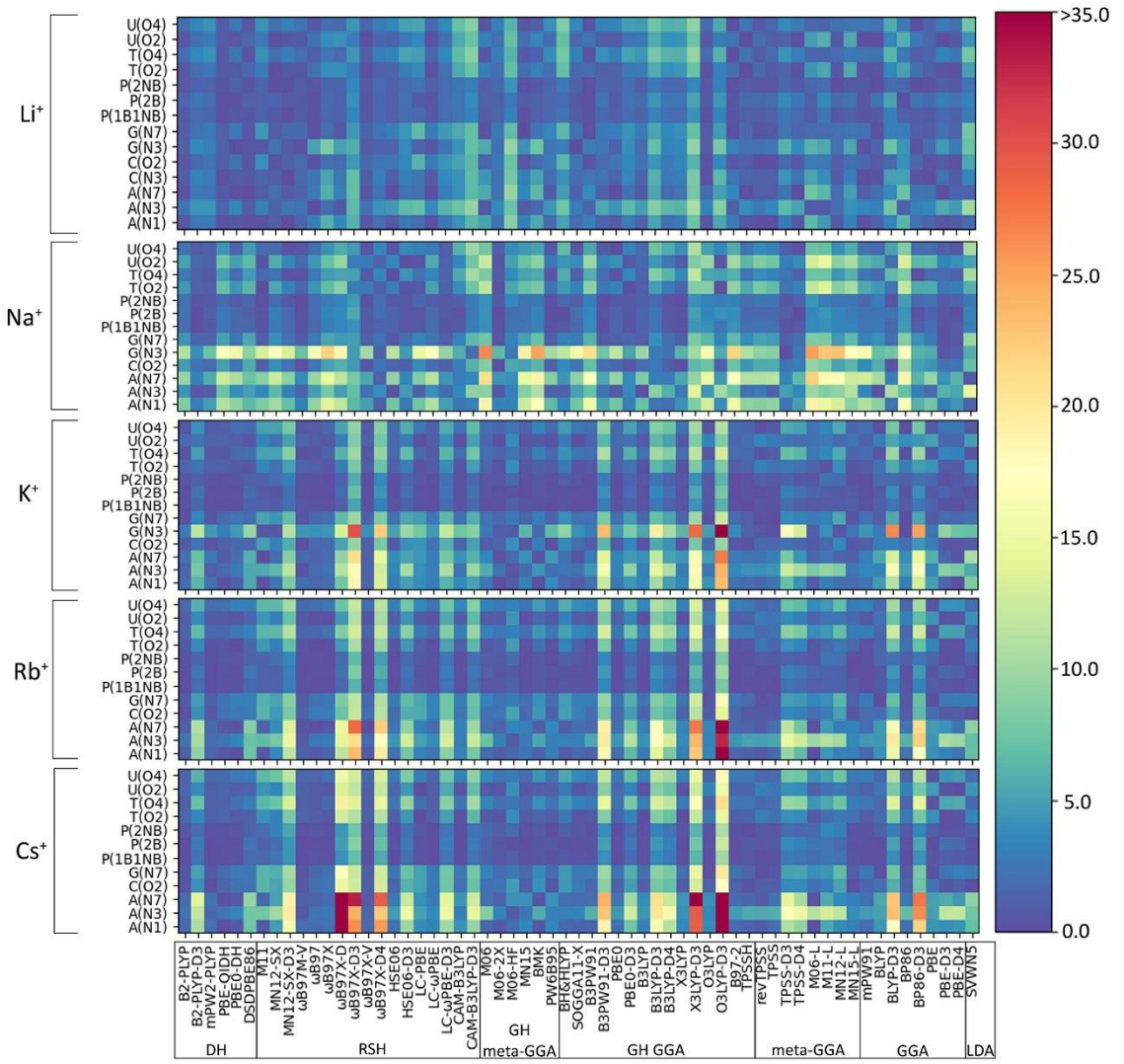


Figure 3.3. Percent errors (%) in DFT binding energies of group I metal–nucleic acid complexes calculated relative to CCSD(T)/CBS reference values, with small errors shown as dark blue shades and large errors highlighted by red shades.

With future applications in mind and our goal to identify a robust functional regardless of group I metal–nucleic acid system considered, we gauge the reliability of functionals using both the PEs and unsigned errors (UEs) evaluated with respect to the CCSD(T)/CBS data across metal–nucleic acid complexes for each DFT method. Given the significant difference in the magnitude of the binding strength for various complexes, evaluation of both PEs and UEs ensures all interactions are more equally weighted.

Functional recommendations were then selected based on the statistics for both metrics. Despite the consideration of several metrics, we only report the mean percent error (MPE) and mean unsigned error (MUE) evaluated over complexes for each group I metal or over the entire data set in the main text for simplicity. We organize our discussion in the following subsections based on functional performance for the smallest (Li^+) to largest (Cs^+) metal, and subsequently consider functional performance over the entire group I, as well as the impact of neglecting counterpoise corrections (Figures 3.4–3.6, B.2–B.6; Tables B.3–B.9, Appendix B).

3.3.2.1. Li^+: The most challenging nucleic acid complex to describe involving direct Li^+ coordination for many functionals is A(N3) (Figure 3.3), which results in the most outliers in the box plot statistics (Table B.3). Regardless, when the statistics are considered over all complexes (Figures 3.4 and B.2; Table 3.2, Appendix B), PBE-QIDH, DSD-PBEP86, and B2-PLYP demonstrate the best performance within the double-hybrid family for Li^+–nucleic acid complexes (<1% MPE; <1 kcal/mol MUE). PBE-QIDH is marginally superior to the other double-hybrid functionals (0.7±0.4% MPE; 0.4±0.2 kcal/mol MUE), with a low spread in the data. The best RSH functional is ωB97X-V (0.9±0.6% MPE; 0.6±0.4 kcal/mol MUE), closely followed by ωB97M-V, MN12-SX-D3(BJ), and MN12-SX (~1% MPEs; 0.7–0.9 kcal/mol MUEs). Within the GH meta-GGA family, the smallest average errors occur for MN15 (1.6±1.4% MPE; 0.8±0.6 kcal/mol MUE), followed by M06-2X and PW6B95, which offer similar performance to each other (~2.4% MPE; ~1.6 kcal/mol MUE), but larger maximum UEs than MN15. Among the 15 GH-GGA functionals explored in this study, O3LYP (1.0±0.9% MPE; 0.6±0.4 kcal/mol MUE) emerges as the top performer for Li^+ complexes, with the smallest MPE, MUE, and standard deviations, as

well as spread in the data, while SOGGA11-X, B3PW91, B97-2, and TPSSh show similar, slightly reduced accuracies (1.4–1.9% MPE; 0.7–1.2 kcal/mol MUE). revTPSS is the best performing meta-GGA (1.1±0.8%; 0.8±0.6 kcal/mol), while TPSS and MN12-L are close runner-ups (1.3% MPE; 0.9 kcal/mol MUE). mPW91 is the only functional of the GGA family that stands out as a solid performer (1.1±1.1% MPE; 0.7±0.6 kcal/mol MUE; maximum errors = 3.5%; 1.6 kcal/mol). Multiple functionals that account for dispersion effects are reliable for Li^+–nucleic acid complexes, including those that incorporate dispersion through corrections (e.g., ωB97M-V, ωB97X-V, and MN12-SX-D3(BJ)) or parameterization (e.g., M06-2X and MN15). When the best performers across different functional families are compared, the most reliable and robust methods recommended for investigating Li^+–nucleic acid interactions are PBE-QIDH and ωB97X-V (**X*,** Table 3.2). Although PBE-QIDH was previously recommended as the overall top performing functional for Li^+ systems based on only the most stable complex for each nucleic acid component [31], other functionals that were previously deemed accurate (ωB97, ωB97X-D, BP86-D3(BJ), PBE) are now replaced with those containing newer dispersion corrections (ωB97X-V), highlighting the importance of considering larger data sets that contain diverse complexes when identifying the top performers for thermochemical data of metal–nucleic acid complexes. For more cost-effective options, O3LYP leads to the lowest MPE and MUE, and the smallest maximum deviation (1.5 kcal/mol) compared to other functionals in the GH GGA, meta-GGA, and GGA families.

Table 3.2. Functionals identified as the top performer(s) for each functional family (X) and the best overall (recommended) functionals (**X*, bold**) for directly coordinated group I metal–nucleic acid complexes.[a]

Family	Functional	Li^+	Na^+	K^+	Rb^+	Cs^+	Group 1
Double-Hybrid	B2-PLYP	X		**X***	X	**X***	
	B2-PLYP-D3		**X***				
	mPW2-PLYP		X		**X***	X	**X***
	PBE-QIDH	**X***		X	X	X	
	PBE0-DH						
	DSD-PBEP86	X					
RSH	M11						
	MN12-SX	X					
	MN12-SX-D3	X					
RSH meta-GGA	ωB97M-V	X	X	X	**X***	**X***	**X***
RSH GGA	ωB97			X	X	X	
	ωB97X			X	X	X	
	ωB97X-D						
	ωB97X-D3						
	ωB97X-V	**X***		**X***	X	**X***	X
	ωB97X-D4		X				
	HSE06			X			
	HSE06-D3		**X***				
	LC-PBE						
	LC-ωPBE						
	LC-ωPBE-D3						
	CAM-B3LYP						
	CAM-B3LYP-D3						
GH meta-GGA	M06						
	M06-2X	X	X	**X***	X	X	X
	M06-HF		X				
	MN15	X		X	X	X	
	BMK			X			
	PW6B95	X	X	X	X	X	X
GH GGA	BH&HLYP						
	SOGGA11-X	X		X		X	
	B3PW91	X					
	B3PW91-D3		X				
	PBE0				X		X
	PBE0-D3		**X***				
	B3LYP				X	X	X
	B3LYP-D3						
	B3LYP-D4		X				
	X3LYP		X		X	X	
	X3LYP-D3						
	O3LYP	**X***			X	X	
	O3LYP-D3						
	B97-2	X		X	X		
	TPSSh	X					
Local meta-GGA	revTPSS	X		X	X	X	**X***
	TPSS	X		X	X	X	**X***
	TPSS-D3		**X***				
	TPSS-D4		X				
	M06-L						
	M11-L						
	MN12-L	X					

Table 3.2. Continued.

Family	Functional	Li^+	Na^+	K^+	Rb^+	Cs^+	Group 1
	MN15-L			**X***	**X***	**X***	
Local GGA	mPW91	X		X	X	X	
	BLYP				X	X	X
	BLYP-D3						
	BP86			X			
	BP86-D3						
	PBE				X		X
	PBE-D3		**X***				
	PBE-D4		X				
Local LDA	SVWN5						

[a]Functional(s) indicated by an X were deemed to be reliable methods for a particular family for a given metal or over all group I metals. The overall top performing functionals that were found to be the most accurate technique(s) regardless of functional family for a given metal or over all group I metals are indicated by **X***.

3.3.2.2. Na^+: In general, larger MPEs arise across functionals for Na^+ relative to Li^+–nucleic acid complexes (Figure 3.4), while the MUEs are more comparable between these lighter group I metals (Figure B.2, Appendix B). For Na^+, the most challenging metal–nucleic acid complexes for the tested functionals to describe include A(N1–N6), A(N6–N7), and G(N3), which result in the most outliers according to the boxplot (Figure 3.3; Table B.4, Appendix B). Among the double-hybrid functionals, B2-PLYP-D3(BJ) is the top performing method for predicting the binding energy of Na^+–nucleic acid complexes (1.1±0.9% MPE; 0.4±0.3 kcal/mol MUE), followed by mPW2-PLYP (1.6±1.6% MPE; 0.7±0.5 kcal/mol MUE). Among the 17 functionals within the RSH family, HSE06-D3(BJ) is the most reliable (1.6±1.3% MPE; 0.7±0.3 kcal/mol MUE), with no outliers for either metric in the boxplot. ωB97M-V and ωB97X-D4 are close runners-ups (1.2–1.5% MPE; 0.4–0.9 kcal/mol), although result in larger maximum errors (≤5.2%; ≤1.9 kcal/mol). M06-2X, which was deemed reliable for Li^+, stands out as the best GH meta-GGA functional for Na^+ interactions, with the smallest magnitude and spread in the deviations (1.1±1.4% MPE; 0.4±0.4 kcal/mol MUE), albeit with several outliers for PE. M06-2X is followed by M06-HF and PW6B95, which have similar errors to each other (2.0–2.6%

MPE; 0.7–1.4 kcal/mol MUE), but larger maximum UEs than M06-2X (2.2–3.7 kcal/mol). Unlike Li^+, PBE0-D3(BJ) and X3LYP from the GH GGA family emerge as the top performing functionals (1.5–1.7% MPE; 0.6 kcal/mol MUE), followed by B3PW91-D3(BJ) and B3LYP-D4, which have slightly reduced accuracies (1.8–1.9% MPE; 0.7–0.8 kcal/mol MUE) and reliability (SDs of 1.5–1.7% MPE; 0.4–0.5 kcal/mol MUE). TPSS-D3(BJ) is the meta-GGA that leads to the least errors (1.4±1.0% MPE; 0.6±0.3 kcal/mol MUE), while TPSS-D4 is a close runner-up (1.8±1.4% MPE; 0.8±0.5 kcal/mol MUE). Of the GGA functionals, only the D3(BJ) and D4-corrected versions of PBE offer solid performance (1.5–1.7% MPE; 0.6–0.8 kcal/mol MUE), with small standard deviations (1.3–1.4% MPE; 0.4–0.5 kcal/mol). Like Li^+, many of the reliable functionals for Na^+ complexes account for dispersion through corrections (ωB97M-V and ωB97X-V) or parameterization (M06-2X), with both ωB97M-V and M06-2X being highly dependable for both metals and ωB97X-V offering reasonable reliability.

Overall, the recommended functionals that consistently offer the highest accuracy for Na^+–nucleic acid complexes are B2PLYP-D3(BJ) and HSE06-D3(BJ) (**X***, Table 3.2). While neither functional is the best option for Li^+, B2PLYP-D3(BJ) still offers reasonable performance for the smallest group I metal, with only moderately larger errors (2.7% MPE; 1.8 kcal/mol MUE) than for Na^+. In contrast, although the top-performing HSE06-D3(BJ) for Na^+ fails to accurately describe Li^+ (3.4% MPE; 2.2 kcal/mol MUE), ωB97M-V works well for both metals (1.2–1.3% MPE; 0.4–0.9 kcal/mol MUE). Despite more inexpensive options not being consistent among the lighter group I metals, the PBE0-D3(BJ) GH GGA functional, TPSS-D3(BJ) from the meta-GGA family, and PBE-D3(BJ) from the GGA family are all reliable for Na^+–nucleic acid complexes, with low MPEs and MUEs, and small maximum deviations (<5%; ≤1.4 kcal/mol).

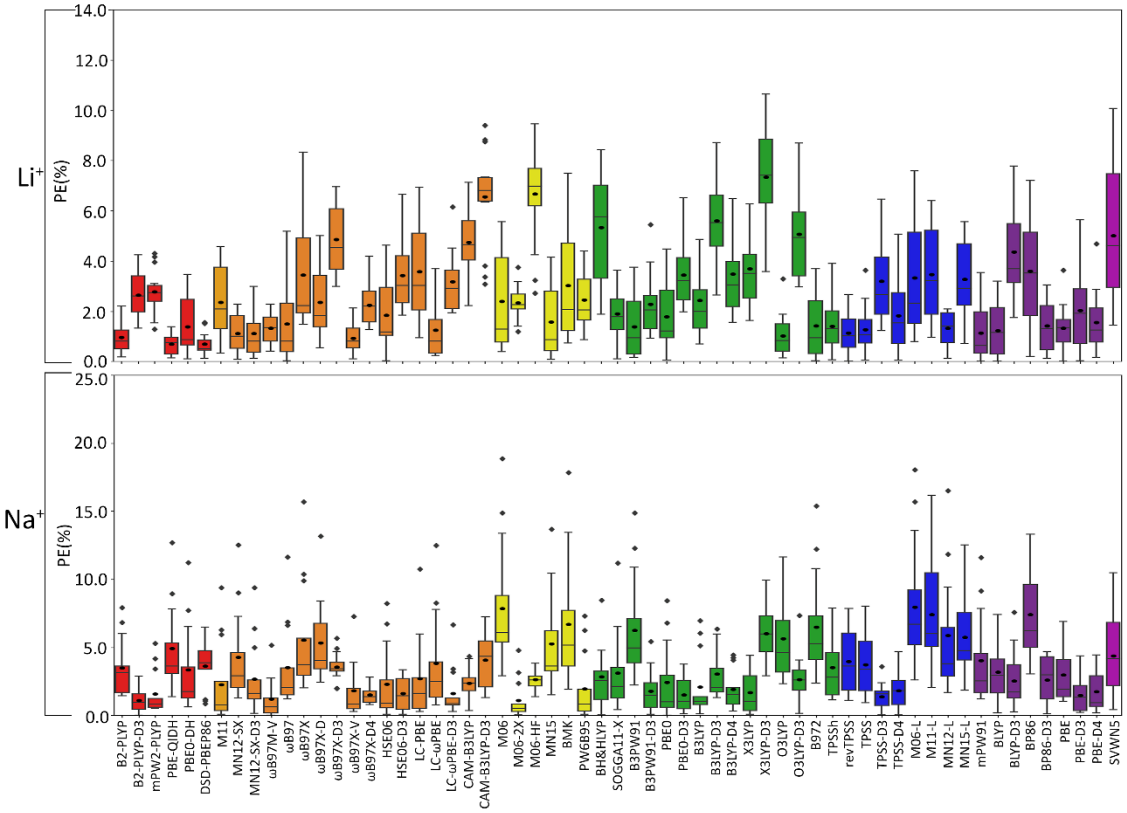


Figure 3.4. Boxplot of the percent errors (%) in Li^+ or Na^+–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

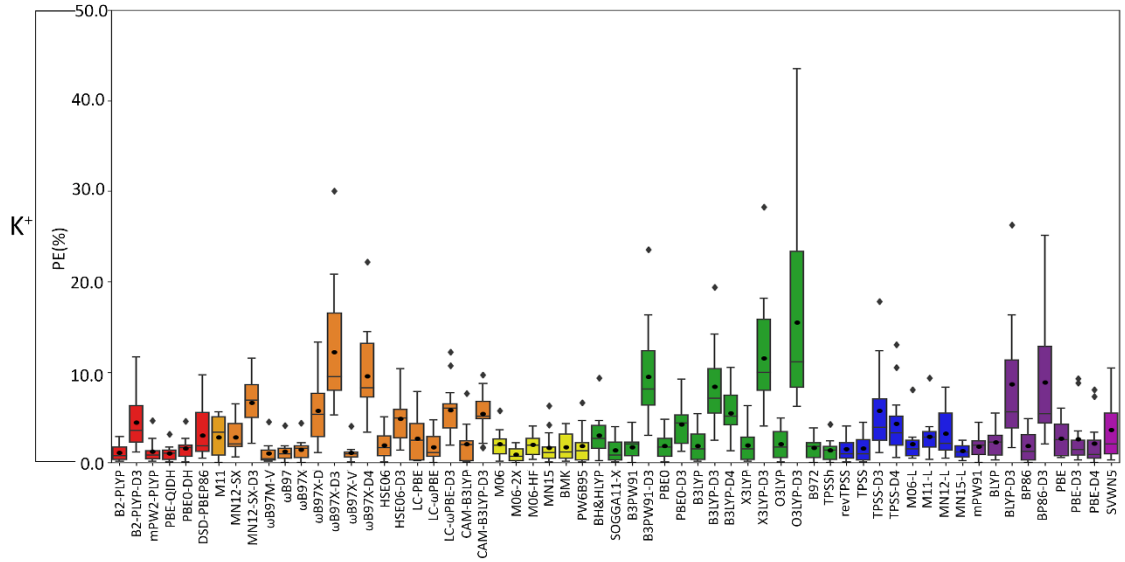


Figure 3.5. Boxplot of the percent errors (%) in K^+–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

3.3.2.3. K^+: In general, the magnitude of error from CCSD(T)/CBS data increases for K^+ compared to the lighter metals (Figure 3.3), with over double the maximum PE (~45% PE K^+; <19% PE Li^+, Na^+; Figures 3.4–3.5; Tables B.3–B.5, Appendix B), which is largely due to G and A complexes formed at the N3 position. These results may be attributed to the weaker binding energies between K^+ and each nucleic acid component (by ~25–49 kcal/mol compared to Li^+ and ~12–24 kcal/mol with respect to Na^+; Table S2). In fact, three functionals (e.g., O3LYP-D3(BJ), X3LYP-D3(BJ), and ωB97X-D3(BJ)) have MPEs greater than 10% for K^+. Among the double-hybrids, B2-PLYP and PBE-QIDH demonstrate similar accuracy (1.1–1.2% MPE; 0.3 kcal/mol MUE), with low maximum PEs (~3%). B2-PLYP provides marginally better descriptions of K^+–nucleic acid interactions, which is also a reliable double-hybrid functional for Li^+ and the dispersion-corrected variant was one of the best functionals for Na^+ complexes. Within the RSH family, four of the six functionals in the ωB97 series (ωB97, ωB97X, ωB97M-V, and

ωB97X-V) offer excellent performance (1.1–1.4% MPE; 0.3 kcal/mol MUE) that is comparable to the double-hybrid functionals, with ωB97X-V demonstrating slightly enhanced accuracy (1.1±1.0% MPE; 0.3±0.1 kcal/mol MUE). HSE06 also performs well, with no outliers appearing in the boxplot, but a larger spread in the data compared to ωB97X-V (1.9±1.5% MPE; 0.4±0.2 kcal/mol MUE). Among the ωB97 series of functionals, ωB97M-V and ωB97X-V both demonstrate reasonable reliability across Li^+, Na^+, and K^+ complexes. The best performing functional of the GH meta-GGA family for Na^+ complexes is M06-2X (0.9±0.7% MPE; 0.3±0.3 kcal/mol MUE). MN15, PW6B95, and BMK are the next best functionals in this family (1.7–1.8% MPE; 0.4–0.5 kcal/mol MUE), which are less reliable than M06-2X due to outliers and larger spreads in the data according to the statistics displayed in the boxplot. SOGGA11-X and B97-2 of the GH GGA family consistently demonstrate accurate performance across all K^+ complexes (1.4–1.6% MPE; 0.4 kcal/mol MUE). Both SOGGA11-X and B97-2 show reasonable accuracy for Li^+ (~1–2% MPE; ~1 kcal/mol MUE), but not Na^+ (~3–7% MPE; ~1–3 kcal/mol MUE). Within the meta-GGA family, revTPSS, MN15-L, and TPSS are the most reliable methods (1.3–1.6% MPE; 0.4 kcal/mol MUE), with MN15-L having slightly better performance mainly due to a smaller spread in the data (1.3±0.8% MPE; 0.4±0.3 kcal/mol MUE). In terms of the GGA functionals, although mPW91 and BP86 have minimal average (1.8% MPE; 0.4–0.6 kcal/mol MUE) and maximum (<5.0%) errors, mPW91 emerges as the most reliable option due to the smallest spread in the data over both metrics. Some functionals identified for K^+ within the meta-GGA (revTPSS) and GGA (mPW91) families are also accurate for Li^+, but less reliable for Na^+ (4.0% MPE; 1.8 kcal/mol MUE).

Based on low MPEs, MUEs, and minor maximum errors, the recommended functionals for investigating K^+–nucleic acid interactions are B2-PLYP, ωB97X-V, and M06-2X (**X***, Table 3.2). ωB97X-V and M06-2X, which each account for dispersion, also offer reasonable reliability for Li^+ and Na^+. While B2-PLYP is also an accurate double-hybrid functional for Li^+, the corresponding dispersion-corrected variant was one of the recommended methods for Na^+ complexes. MN15-L is also available as a computationally affordable option for K^+–nucleic acid complexes.

3.3.2.4. Rb^+: As seen with K^+, certain functionals emerge as having substantially larger errors for Rb^+–nucleic acid interactions (maximum PE of ~22–42%), including O3LYP-D3(BJ), X3LYP-D3(BJ), ωB97X-D3(BJ), and ωB97X-D4, which is mainly attributed to the A complexes (Figures 3.3 and 3.6; Table B.6, Appendix B). Indeed, the number of functionals exceeding an MPE of 10% doubles for Rb^+ (6; Figure 3.6) compared to K^+ (3; Figure 3.5). mPW2-PLYP and B2-PLYP are the top performers within the double-hybrid family (~1% MPE; 0.2 kcal/mol MUE; Figures 3.6 and B.3, Appendix B), closely followed by PBE-QIDH (1.2±0.9% MPE; 0.4±0.2 kcal/mol MUE). mPW2-PLYP emerges as the best double-hybrid functional (maximum errors = 1.9%; 0.4 kcal/mol). The most accurate functional from the RSH family is ωB97M-V (0.8±05% MPE; 0.2±0.1 kcal/mol MUE), followed by ωB97X-V, ωB97, and ωB97X (≤1.3% MPE; 0.2–0.3 kcal/mol MUE). Although multiple functionals from the ωB97 series are recommended for Li^+, Na^+, K^+, and Rb^+, others from this category (ωB97X-D3(BJ) and ωB97X-D4) exhibit large errors for Rb^+ and K^+. MN15 is the most reliable functional of the GH meta-GGA family (1.5±1.2% MPE; 0.4±0.2 kcal/mol MUE), closely followed by M06-2X (1.6±0.9% MPE; 0.5±0.3 kcal/mol MUE) and PW6B95 (1.7±1.3% MPE; 0.4±0.2 kcal/mol MUE). Within

the GH GGA family, O3LYP is the most accurate (1.8±1.2% MPE; 0.5±0.2 kcal/mol; maximum errors = 3.6%; 0.7 kcal/mol), while B3LYP, X3LYP, PBE0, and B97-2 offer comparable performance to each other as viable alternatives (1.6–1.8% MPE; 0.4 kcal/mol MUE). From the meta-GGA family, MN15-L is more reliable (1.8±1.3% MPE; 0.4±0.2 kcal/mol MUE) than TPSS and revTPSS (1.5–1.6% MPE; 0.4 kcal/mol) due to the absence of outliers and lower maximum errors, with MN15-L also being one of the recommended methods for K^+. Like Li^+ and K^+, mPW91 emerges as the top performer of the GGA family (1.7±1.5% MPE; 0.4±0.3 kcal/mol MUE), with no outliers for either metric. For Rb^+ complexes, mPW91 is followed by BLYP and PBE (2.1–2.3% MPE; 0.6–0.7 kcal/mol MUE). Multiple accurate functionals for Rb^+ complexes account for dispersion through corrections (ωB97M-V and ωB97X-V) or parameterization (M06-2X).

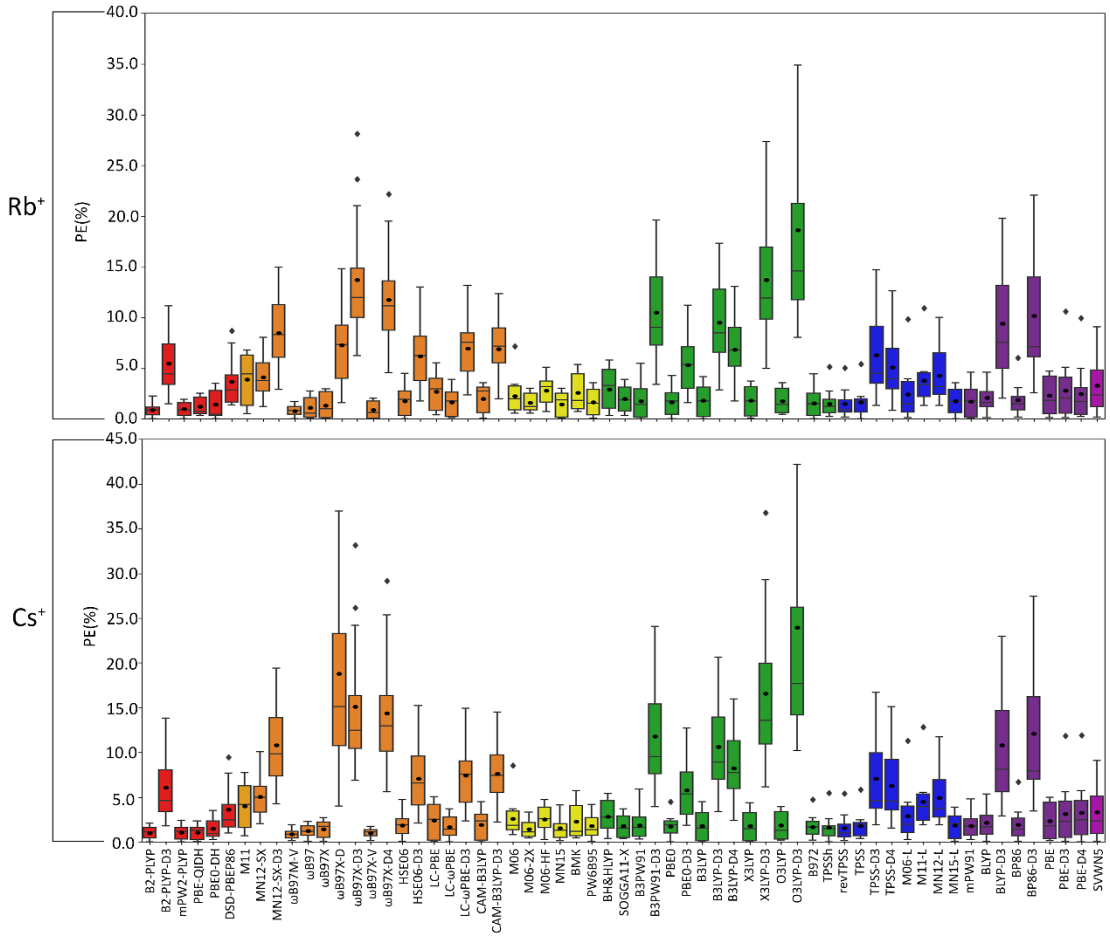


Figure 3.6. Boxplot of the percent errors (%) in Rb^+ or Cs^+–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

Overall, despite being more challenging to describe than the metals discussed thus far, there are many reliable functionals for exploring Rb^+–nucleic acid interactions. Our recommendations include mPW2-PLYP and ωB97M-V (**X***, Table 3.2), which display low MPEs, MUEs, and maximum deviations (≤1.9%; 0.4 kcal/mol). Although not the overall best double-hybrid functional for each metal discussed previously, mPW2-PLYP performs very well for K^+ complexes and offers reasonable agreement to CCSD(T)/CBS data for Li^+ and Na^+. Additionally, ωB97M-V has demonstrated consistently small errors for Li^+, Na^+,

K^+, and Rb^+. As a cost-effective alternative for Rb^+–nucleic acid complexes, MN15-L offers low MPEs, MUEs, and reasonable maximum deviations (3.6%; 0.7 kcal/mol).

3.3.2.5. Cs^+: Cs^+ demonstrates the largest errors among group I metals (maximum PE of 59.0%; Figure 3.6; Table B.7, Appendix B). Indeed, 10 functionals result in significantly large MPEs (>10%). This highlights the increased sensitivity of DFT with increased size of the metal and decreased magnitude of the binding strength. As for K^+ and Rb^+, the A(N3) complex results in the most deviations from CCSD(T)/CBS data (Figure 3.3; Table B.7, Appendix B). Within the double-hybrid family, B2-PLYP is the best performer (1.0±0.7% MPE; 0.2±0.1 kcal/mol MUE; maximum errors = 2.1%; 0.5 kcal/mol; Figure 3.6). Although B2-PLYP was also recommended for K^+ and offers small errors for Li^+ and Rb^+, the performance is not as good as other double-hybrids for Na^+, requiring the addition of a dispersion correction. The next best performers in the double-hybrid family are mPW2-PLYP (1.1% MPE; 0.2 kcal/mol MUE), which is reasonable for all metals, and PBE-QIDH (1.1% MPE; 0.3 kcal/mol MUE), which is recommended for Li^+, reliable for K^+ and Rb^+, but a poor performer for Na^+. The most reliable functionals for Cs^+ among the RSH family are ωB97M-V and ωB97X-V (0.9–1.0% MPE; 0.2 kcal/mol), while ωB97 and ωB97X are the next best performers (1.2–1.4% MPE; 0.3 kcal/mol). ωB97M-V was one of the recommended functionals for Rb^+, while ωB97X-V was recommended for Li^+ and K^+. Nevertheless, the performance of both ωB97M-V and ωB97X-V is excellent for all metals, even though HSE06-D3(BJ) was the best RSH family member for Na^+. Amongst the GH meta-GGA functionals, MN15 demonstrates the smallest deviations from CCSD(T)/CBS (1.5% MPE; 0.3kcal/mol MUE), followed by M06-2X and PW6B95 (1.5–1.8% MPE; 0.4 kcal/mol). Indeed, MN15 demonstrates a slightly smaller spread in the data and the

lowest maximum UE (0.6 kcal/mol). Like Li^+ (the lightest group I metal), O3LYP offers consistent and reliable performance for Cs^+ (the heaviest group I metal) within the GH GGA family (1.9±1.5% MPE; 0.4±0.2 kcal/mol MUE), closely followed by B3LYP, X3LYP, and SOGGA11-X (1.8% MPE; 0.4–0.5 kcal/mol MUE). Within the meta-GGA family, three functionals (revTPSS, TPSS, and MN15-L) offer the best performance (1.6–1.8% MPE; 0.4 kcal/mol MUE), with MN15-L emerging as the top performer due to the smallest standard deviations (1.4% MPE; 0.1 kcal/mol) and maximum errors (3.9%; 0.4 kcal/mol). Within the GGA family, mPW91 is the most accurate functional (1.8±1.3% MPE; 0.4±0.2 kcal/mol MUE), followed by BLYP (2.2±1.9% MPE; 0.5±0.3 kcal/mol MUE). Both methods are also among the most reliable GGA functionals for Rb^+.

Overall, the recommended functionals for investigating Cs^+–nucleic acid interactions are B2-PLYP, ωB97M-V, and ωB97X-V (**X***, Table 3.2), which demonstrate consistent and accurate performance across all complexes, including low maximum errors (≤2.1%; ≤0.5 kcal/mol). As mentioned above, B2-PLYP yields small deviations from CCSD(T)/CBS data for all metals except Na^+, although the MUE across Na^+ complexes is still below 2 kcal/mol. Both ωB97M-V and ωB97X-V, which include the more recently designed VV10 dispersion correction,[108] have emerged as solid performers across all group I metal–nucleic acid complexes. In terms of cost-effective options for Cs^+ complexes, MN15-L is deemed reliable.

3.3.2.6. Although the Top-Performing Functionals Can Vary with Metal, Some Methods Prevail as Being Reliable Across Group I: Thus far, variation in functional performance has been seen for each group I metal as well as between the lighter (Li^+ and

Na^+) and heavier (K^+, Rb^+, and Cs^+) metals (Figures 3.3–3.6). Nevertheless, a robust method is required that offers a reasonable description of all group I metal–nucleic acid complexes. This is important for understanding, for example, problems related to metal competition in biology, diseases, energy storage materials, and nucleic acid sensors. Therefore, in this section, functional performance is compared over all Li^+, Na^+, K^+, Rb^+, and Cs^+–nucleic acid complexes that involve direct metal coordination (Figures 3.2, 3.7 and B.4, Appendix B). Among the six double-hybrid functionals, mPW2-PLYP is the most accurate across all group I metal–nucleic acid binding strengths (1.6±1.3% MPE; 0.7±0.7 kcal/mol MUE; maximum errors = 5.3%; 2.7 kcal/mol). While mPW2-PLYP was not the best performing double-hybrid for the lighter metals, this method still offers a reasonable description across complexes involving Li^+ or Na^+ (≤2.8% MPE; ≤1.8 kcal/mol MUE). ωB97M-V emerges as the best performer over group I within the RSH family (1.1±1.0% MPE; 0.4±0.4 kcal/mol MUE), as well as Rb^+ and Cs^+, closely followed by ωB97X-V (1.1±1.2% MPE; 0.4±0.4 kcal/mol MUE), which is one of the best functionals for Li^+ and K^+. Although slightly less accurate for Na^+ complexes, both ωB97M-V and ωB97X-V offer reasonable binding energies (< 1 kcal/mol MUE deviation from CCSD(T)/CBS). Within the GH meta-GGA family, M06-2X demonstrates the smallest errors (1.5±1.1% MPE; 0.7±0.6 kcal/mol MUE), with PW6B95 being a close runner-up (2.0±1.7% MPE; 0.7±0.7 kcal/mol MUE), albeit having more outliers. M06-2X was one of the best functionals for K^+ and offers a reasonable description for Li^+, Na^+, Rb^+, and Cs^+ (≤2.3% MPE; ≤1.5 kcal/mol MUE). The most reliable functionals of the GH GGA family are B3LYP and PBE0 (≤2.0% MPE; ≤0.8 kcal/mol MUE). These functionals are generally more reliable for the heavier metals (K^+, Rb^+, and Cs^+), but still afford reasonable accuracy for the lighter metals (Li^+ and Na^+; ~2% MPE; ~1–2 kcal/mol MUE). Both revTPSS and TPSS offer the best

performance within the meta-GGA family (1.9–2.0% MPE; 0.8 kcal/mol), while BLYP and PBE emerge as the best GGA functionals (2.2–2.3% MPE; 0.9 kcal/mol; maximum errors ≤7.4%; ≤3.3 kcal/mol).

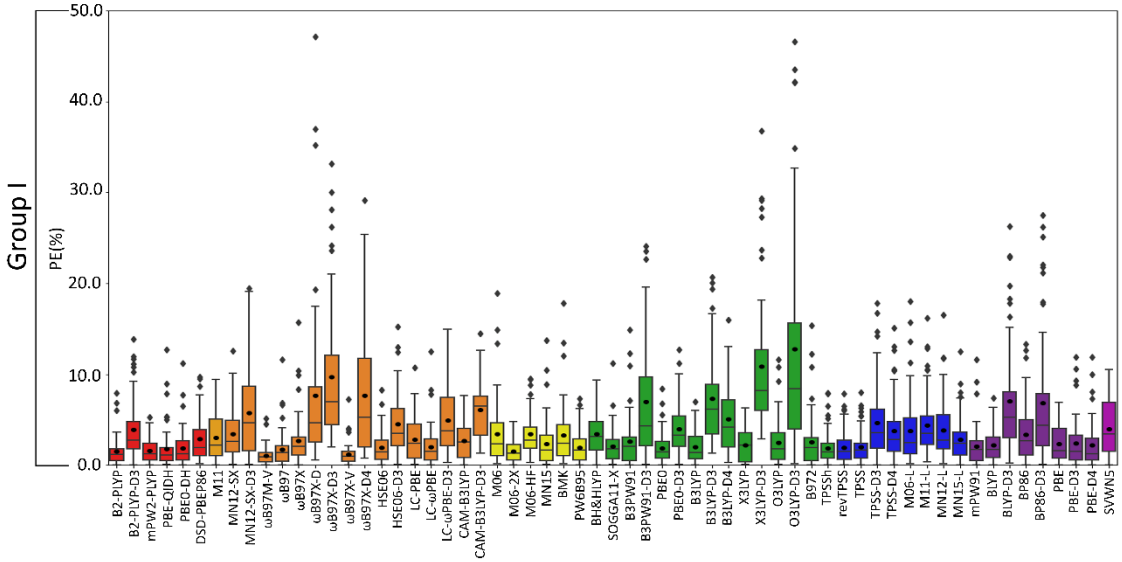


Figure 3.7. Boxplot of the percent errors (%) for group I metal–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

Overall, the recommended functionals for exploring group I metal–nucleic acid interactions are mPW2-PLYP and ωB97M-V (**X***, Table 3.2), which demonstrate the smallest MPEs and MUEs, and low maximum errors (≤5.3%; ≤2.7 kcal/mol). Indeed, ωB97M-V has been identified as the most promising functional for main-group chemistry,[108] and is deemed one of the most accurate functionals for exploring transition metal chemistry.[40] Since both recommended functionals are from more expensive families, cost-effective methods must also be identified to explore group I metal interactions using larger models that are required to accurately study biosystems. Although the most popular

DFT method (B3LYP) offers reasonable computational performance on average for group I metals, B3LYP leads to larger maximum errors (7.0%; 2.7 kcal/mol) compared to the top-performing functionals, as well as a higher spread in the boxplot data. Indeed, cost-effective revTPSS and TPSS are on average the most reliable across group I (**X***, Table 3.2), with MPEs and MUEs comparable to the best performing functionals, albeit having larger spreads and more outliers in the boxplot data than the more computationally intense methods.

3.3.2.7. Impact of Counterpoise Corrections on Group I Metal–Nucleic Acid Binding Energies: Although all results discussed thus far include counterpoise corrections, such corrections are not always computationally feasible, especially when using large models that are necessary to investigate complex biosystems or biomaterials. To ensure that functional performance remains unchanged regardless of whether counterpoise corrections are included, the counterpoise uncorrected and corrected binding energies are compared for the most reliable functionals for each group I metal and over all group I metals (Figures B.5–B.6; Table B.9, Appendix B). In general, there is a slight increase in deviations from the CCSD(T)/CBS binding energies when the counterpoise corrections are not included in the DFT binding strengths. Nevertheless, the differences between the counterpoise-corrected and uncorrected MPEs and MUEs are ≤1.5% and ≤0.6 kcal/mol, respectively, which correlates with previous literature reporting a minimal BSSE impact for similar systems.[31] This suggests that counterpoise corrections can be neglected when implementing larger computational models while minimally impacting the accuracy of the resulting DFT description of group I metal–nucleic acid interactions.

3.4. Conclusions

In summary, a new database of accurate gas-phase CCSD(T)/CBS group I metal–nucleic acid binding energies has been generated, which has otherwise proven to be challenging to compile experimentally due to measurement sensitivities and nucleobase tautomerizations. Moreover, this study provides highly accurate interaction energies between group I metals and G as well as between heavier metals (Rb^+ and Cs^+) and A, T, or U for which no experimental thermochemical data is currently available. Subsequently, the performance of 61 DFT functionals was tested across 64 directly coordinated group I metal–nucleic acid complexes. Although functional performance can vary with the metal, nucleic acid component, and binding site, reliable methods across all group I metal–nucleic acid interactions come to light. Indeed, mPW2-PLYP and ωB97M-V demonstrate remarkably small MPEs and MUEs (≤1.6%; <1.0 kcal/mol) and low spreads in the data based on boxplot visualizations. If more computationally efficient methods are required, revTPSS and TPSS offer reasonable reliability, albeit having larger spreads and more outliers in the boxplots compared to the more computationally intensive functionals. Counterpoise corrections are found to have negligible effects on the accuracy of the top-performing functionals for group I metal–nucleic acid binding energies (≤1.5% and ≤0.6 kcal/mol), which is promising for investigations that require the implementation of larger models that render counterpoise corrections unfeasible. Overall, the accurate methods identified in our work can be used in the future to understand the role of metals in ecological and human biology, to generate new therapeutics, to explore the design of novel chemical technologies (e.g., new materials for energy storage and nucleic acid sensors that target metal contaminants in the environment or human body), and to aid the design of new

computational methods that permit broader investigations of group I metal–nucleic acid interactions.

3.5. References

(1) Shi, P.; Zhang, X.-Q.; Shen, X.; Zhang, R.; Liu, H.; Zhang, Q. A Review of Composite Lithium Metal Anode for Practical Applications. *Adv. Mater. Technol.* **2020**, *5* (1), 1900806.

(2) Buccella, D.; Lim, M. H.; Morrow, J. R. Metals in Biology: From Metallomics to Trafficking. *Inorg. Chem.* **2019**, *58* (20), 13505-13508.

(3) Franz, K. J.; Metzler-Nolte, N. Introduction: Metals in Medicine. *Chem. Rev.* **2019**, *119* (2), 727-729.

(4) Gentner, T. X.; Mulvey, R. E. Alkali-Metal Mediation: Diversity of Applications in Main-Group Organometallic Chemistry. *Angew. Chem. Int. Ed.* **2021**, *60* (17), 9247-9262.

(5) Auffinger, P.; D'Ascenzo, L.; Ennifar, E. Sodium and Potassium Interactions with Nucleic Acids. In *The Alkali Metal Ions: Their Role for Life*, Sigel, A., Sigel, H., Sigel, R. K. O. Eds.; Springer International Publishing, 2016; pp 167-201.

(6) Tomita, A.; Zhang, M.; Jin, F.; Zhuang, W.; Takeda, H.; Maruyama, T.; Osawa, M.; Hashimoto, K.-i.; Kawasaki, H.; Ito, K.; et al. Atp-Dependent Modulation of Mgte in Mg^{2+} Homeostasis. *Nat. Commun.* **2017**, *8* (1), 148.

(7) Lin, B.; McCormick, A. V.; Davis, H. T.; Strey, R. Solubility of Sodium Soaps in Aqueous Salt Solutions. *J. Colloid Interface Sci.* **2005**, *291* (2), 543-549.

(8) Khan, R. A. Chapter 7 - Detergents. In *Environmental Micropollutants*, Hashmi, M. Z., Wang, S., Ahmed, Z. Eds.; Elsevier, 2022; pp 117-130.

(9) Raabe, D. The Materials Science Behind Sustainable Metals and Alloys. *Chem. Rev.* **2023**, *123* (5), 2436-2608.

(10) Machado-Vieira, R.; Manji, H. K.; Zarate Jr, C. A. The Role of Lithium in the Treatment of Bipolar Disorder: Convergent Evidence for Neurotrophic Effects as a Unifying Hypothesis. *Bipolar Disord.* **2009**, *11* (s2), 92-109.

(11) Hu, Y.; Hutter, E. M.; Rieder, P.; Grill, I.; Hanisch, J.; Aygüler, M. F.; Hufnagel, A. G.; Handloser, M.; Bein, T.; Hartschuh, A.; et al. Understanding the Role of Cesium and Rubidium Additives in Perovskite Solar Cells: Trap States, Charge Transport, and Recombination. *Adv. Energy Mater.* **2018**, *8* (16), 1703057.

(12) Gomes, H. I.; Mayes, W. M.; Rogerson, M.; Stewart, D. I.; Burke, I. T. Alkaline Residues and the Environment: A Review of Impacts, Management Practices and Opportunities. *J. Clean. Prod.* **2016**, *112*, 3571-3582.

(13) Schrauzer, G. N. Lithium: Occurrence, Dietary Intakes, Nutritional Essentiality. *J. Am. Coll. Nutr.* **2002**, *21* (1), 14-21.

(14) Levings, J. L.; Gunn, J. P. The Imbalance of Sodium and Potassium Intake: Implications for Dietetic Practice. *J. Acad. Nutr. Diet.* **2014**, *114* (6), 838-841.

(15) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(16) Ramos, P.; Santos, A.; Pinto, E.; Pinto, N. R.; Mendes, R.; Magalhães, T.; Almeida, A. Alkali Metals Levels in the Human Brain Tissue: Anatomical Region Differences and Age-Related Changes. *J. Trace Elem. Med. Biol.* **2016**, *38*, 174-182.

(17) Li, Q.; Zhou, C.; Ji, Z.; Han, B.; Feng, L.; Wu, J. High-Performance Lithium/Sulfur Batteries by Decorating Cmk-3/S Cathodes with DNA. *J. Mater. Chem. A.* **2015**, *3* (14), 7241-7247.

(18) McKernan, J. A. B. Y. Andalyze Lead 100 Test Kit and And1000 Fluorimeter Environmental Technology Verification Report and Statement. *U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-14/052* **2014**.

(19) Liu, H.; Gao, Y.; Mathivanan, J.; Armour-Garb, Z.; Shao, Z.; Zhang, Y.; Zhao, X.; Shao, Q.; Zhang, W.; Yang, J.; et al. Crystal Structures and Identification of Novel Cd2+-Specific DNA Aptamer. *Nucleic Acids Res.* **2023**, *51* (9), 4625-4636.

(20) Torabi, S. F.; Wu, P.; McGhee, C. E.; Chen, L.; Hwang, K.; Zheng, N.; Cheng, J.; Lu, Y. In Vitro Selection of a Sodium-Specific DNAzyme and Its Application in Intracellular Sensing. *Proc Natl Acad Sci U S A* **2015**, *112* (19), 5903-5908.

(21) Lin, S.; Yang, C.; Mao, Z.; He, B.; Wang, Y. T.; Leung, C. H.; Ma, D. L. A G-Pentaplex-Based Assay for Cs(+) Ions in Aqueous Solution Using a Luminescent Ir(III) Complex. *Biosens. Bioelectron* **2016**, *77*, 609-612.

(22) Zhao, L.; Ahmed, F.; Zeng, Y.; Xu, W.; Xiong, H. Recent Developments in G-Quadruplex Binding Ligands and Specific Beacons on Smart Fluorescent Sensor for Targeting Metal Ions and Biological Analytes. *ACS Sens.* **2022**, *7* (10), 2833-2856.

(23) Rodgers, M. T.; Armentrout, P. B. Cationic Noncovalent Interactions: Energetics and Periodic Trends. *Chem. Rev.* **2016**, *116* (9), 5642-5687.

(24) Rodgers, M. T.; Armentrout, P. B. Noncovalent Interactions of Nucleic Acid Bases (Uracil, Thymine, and Adenine) with Alkali Metal Ions. Threshold Collision-Induced Dissociation and Theoretical Studies. *J. Am. Chem. Soc.* **2000**, *122* (35), 8548-8558.

(25) Yang, B.; Rodgers, M. T. Alkali Metal Cation Binding Affinities of Cytosine in the Gas Phase: Revisited. *Phys. Chem. Chem. Phys.* **2014**, *16* (30), 16110-16120.

(26) Yang, Z.; Rodgers, M. T. Tautomerization in the Formation and Collision-Induced Dissociation of Alkali Metal Cation-Cytosine Complexes. *Phys. Chem. Chem. Phys.* **2012**, *14* (13), 4517-4526.

(27) Rodgers, M. T.; Armentrout, P. B. A Critical Evaluation of the Experimental and Theoretical Determination of Lithium Cation Affinities. *Int. J. Mass Spectrom.* **2007**, *267* (1), 167-182.

(28) Rajabi, K.; Gillis, E. A. L.; Fridgen, T. D. Structures of Alkali Metal Ion−Adenine Complexes and Hydrated Complexes by IRMPD Spectroscopy and Electronic Structure Calculations. *J. Phys. Chem. A.* **2010**, *114* (10), 3449-3456.

(29) Burt, M. B.; Fridgen, T. D. Structures and Physical Properties of Gaseous Metal Cationized Biological Ions. *Eur. J. Mass Spectrom.* **2012**, *18* (2), 235-250.

(30) Gillis, E. A. L.; Rajabi, K.; Fridgen, T. D. Structures of Hydrated Li^+−Thymine and Li^+−Uracil Complexes by IRMPD Spectroscopy in the N−H/O−H Stretching Region. *J. Phys. Chem. A.* **2009**, *113* (5), 824-832.

(31) Boychuk, B. T. A.; Jeong, R. Y. E.; Wetmore, S. D. Assessment of the Accuracy of DFT-Predicted Li^+–Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408.

(32) Stasyuk, O. A.; Solà, M.; Swart, M.; Fonseca Guerra, C.; Krygowski, T. M.; Szatylowicz, H. Effect of Alkali Metal Cations on Length and Strength of Hydrogen Bonds in DNA Base Pairs. *ChemPhysChem.* **2020**, *21* (18), 2112-2126.

(33) Monajjemi, M.; Ghiasi, R.; Sadjadi, M. A. S. Metal-Stabilized Rare Tautomers: N4 Metalated Cytosine (M = Li^+, Na^+, K^+, Rb^+ and Cs^+), Theoretical Views. *Appl. Organomet. Chem.* **2003**, *17* (8), 635-640.

(34) Marino, T.; Mazzuca, D.; Russo, N.; Toscano, M.; Grand, A. On the Interaction of Rubidium and Cesium Mono-, Strontium and Barium Bi-Cations with DNA and RNA Bases. A Theoretical Study. *Int. J. Quantum Chem.* **2010**, *110*, 138-147.

(35) Zhu, W.; Luo, X.; Puah, C. M.; Tan, X.; Shen, J.; Gu, J.; Chen, K.; Jiang, H. The Multiplicity, Strength, and Nature of the Interaction of Nucleobases with Alkaline and Alkaline Earth Metal Cations: A Density Functional Theory Investigation. *J. Phys. Chem. A* **2004**, *108* (18), 4008-4018.

(36) Burda, J. V.; Šponer, J.; Hobza, P. Ab Initio Study of the Interaction of Guanine and Adenine with Various Mono- and Bivalent Metal Cations (Li^+, Na^+, K^+, Rb^+, Cs^+; Cu^+, Ag^+, Au^+; Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}; Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem. A.* **1996**, *100* (17), 7250-7255.

(37) Hashemianzadeh, S. M.; Faraji, S.; Amin, A. H.; Ketabi, S. Theoretical Study of the Interactions between Isolated DNA Bases and Various Groups Ia and Iia Metal Ions by Ab Initio Calculations. *Monatsh. Chem.* **2008**, *139* (2), 89-100.

(38) Boychuk, B. T. A.; Meyer, S. P.; Wetmore, S. D. Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* **2023**, *11*.

(39) Grauffel, C.; Dudev, T.; Lim, C. Metal Affinity/Selectivity of Monophosphate-Containing Signaling/Lipid Molecules. *J. Chem. Theory Comput.* **2021**, *17* (4), 2444-2456.

(40) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *J. Chem. Theory Comput.* **2019**, *15* (6), 3610-3622.

(41) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124* (3), 034108.

(42) Schwabe, T.; Grimme, S. Towards Chemical Accuracy for the Thermodynamics of Large Molecules: New Hybrid Density Functionals Including Non-Local Correlation Effects. *Phys. Chem. Chem. Phys.* **2006**, *8* (38), 4398-4401.

(43) Brémond, É.; Sancho-García, J. C.; Pérez-Jiménez, Á. J.; Adamo, C. Communication: Double-Hybrid Functionals from Adiabatic-Connection: The QIDH Model. *J. Chem. Phys.* **2014**, *141* (3), 031101.

(44) Brémond, E.; Adamo, C. Seeking for Parameter-Free Double-Hybrid Functionals: The PBE0-DH Model. *J. Chem. Phys.* **2011**, *135* (2), 024106.

(45) Kozuch, S.; Martin, J. M. L. Dsd-Pbep86: In Search of the Best Double-Hybrid DFT with Spin-Component Scaled MP2 and Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20104-20107.

(46) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2* (21), 2810-2817.

(47) Peverati, R.; Truhlar, D. G. Screened-Exchange Density Functionals with Broad Accuracy for Chemistry and Solid-State Physics. *Phys. Chem. Chem. Phys.* **2012**, *14* (47), 16187-16191.

(48) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized, Range-Separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144* (21), 214110.

(49) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128* (8), 084106.

(50) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615-6620.

(51) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal

Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 9904-9924.

(52) Heyd, J.; Scuseria, G. E. Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd–Scuseria–Ernzerhof Screened Coulomb Hybrid Functional. *J. Chem. Phys.* **2004**, *121* (3), 1187-1192.

(53) Heyd, J.; Scuseria, G. E. Assessment and Validation of a Screened Coulomb Hybrid Density Functional. *J. Chem. Phys.* **2004**, *120* (16), 7274-7280.

(54) Heyd, J.; Peralta, J. E.; Scuseria, G. E.; Martin, R. L. Energy Band Gaps and Lattice Parameters Evaluated with the Heyd-Scuseria-Ernzerhof Screened Hybrid Functional. *J. Chem. Phys.* **2005**, *123* (17), 174101.

(55) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Erratum: "Hybrid Functionals Based on a Screened Coulomb Potential" [J. Chem. Phys. 118, 8207 (2003)]. *J. Chem. Phys.* **2006**, *124* (21), 219906.

(56) Henderson, T. M.; Izmaylov, A. F.; Scalmani, G.; Scuseria, G. E. Can Short-Range Hybrids Describe Long-Range-Dependent Properties? *J. Chem. Phys.* **2009**, *131* (4), 044108.

(57) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. Efficient Evaluation of Short-Range Hartree-Fock Exchange in Large Molecules and Periodic Systems. *J. Chem. Phys.* **2006**, *125* (10), 104103.

(58) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, *125* (22), 224106.

(59) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.

(60) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A Long-Range Correction Scheme for Generalized-Gradient-Approximation Exchange Functionals. *J. Chem. Phys.* **2001**, *115* (8), 3540-3544.

(61) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. Importance of Short-Range Versus Long-Range Hartree-Fock Exchange for the Performance of Hybrid Density Functionals. *J. Chem. Phys.* **2006**, *125* (7), 074106.

(62) Vydrov, O. A.; Scuseria, G. E. Assessment of a Long-Range Corrected Hybrid Functional. *J. Chem. Phys.* **2006**, *125* (23), 234109.

(63) Vydrov, O. A.; Scuseria, G. E.; Perdew, J. P. Tests of Functionals for Systems with Fractional Electron Number. *J. Chem. Phys.* **2007**, *126* (15), 154109.

(64) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1), 51-57.

(65) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120* (1), 215-241.

(66) Zhao, Y.; Truhlar, D. G. Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average Than B3LYP for Ground States. *J. Phys. Chem. A* **2006**, *110* (49), 13126-13130.

(67) Boese, A. D.; Martin, J. M. L. Development of Density Functionals for Thermochemical Kinetics. *J. Chem. Phys.* **2004**, *121* (8), 3405-3416.

(68) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032-5051.

(69) Zhao, Y.; Truhlar, D. G. Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A.* **2005**, *109* (25), 5656-5667.

(70) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372-1377.

(71) Peverati, R.; Truhlar, D. G. Communication: A Global Hybrid Generalized Gradient Approximation to the Exchange-Correlation Functional That Satisfies the Second-Order Density-Gradient Constraint and Has Broad Applicability in Chemistry. *J. Chem. Phys.* **2011**, *135* (19), 191102.

(72) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648-5652.

(73) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158-6170.

(74) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785-789.

(75) Xu, X.; Goddard, W. A. The X3LYP Extended Density Functional for Accurate Descriptions of Nonbond Interactions, Spin States, and Thermochemical Properties. *Proc. Natl. Acad. Sci.* **2004**, *101* (9), 2673.

(76) Cohen, A. J.; Handy, N. C. Dynamic Correlation. *Mol. Phys.* **2001**, *99* (7), 607-615.

(77) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(78) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119* (23), 12129-12137.

(79) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid Exchange-Correlation Functional Determined from Thermochemical Data and Ab Initio Potentials. *J. Chem. Phys.* **2001**, *115* (20), 9233-9242.

(80) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta--Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(81) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry. *Phys. Rev. Lett.* **2009**, *103* (2), 026403.

(82) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Erratum: Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry [Phys. Rev. Lett. 103, 026403 (2009)]. *Phys. Rev. Lett.* **2011**, *106* (17), 179902.

(83) Peverati, R.; Truhlar, D. G. M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics. *J. Phys. Chem. Lett.* **2012**, *3* (1), 117-124.

(84) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125* (19), 194101.

(85) Peverati, R.; Truhlar, D. G. An Improved and Broadly Accurate Local Approximation to the Exchange–Correlation Density Functional: The MN12-L Functional for Electronic Structure Calculations in Chemistry and Physics. *Phys. Chem. Chem. Phys.* **2012**, *14* (38), 13171-13174.

(86) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional for Kohn–Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. *J. Chem. Theory Comput.* **2016**, *12* (3), 1280-1293.

(87) Adamo, C.; Barone, V. Exchange Functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW Models. *J. Chem. Phys.* **1998**, *108* (2), 664-675.

(88) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098-3100.

(89) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157* (3), 200-206.

(90) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200-1211.

(91) Ruan, C.; Huang, H.; Rodgers, M. T. A Simple Model for Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase: Alkali Metal Cations and Trimethyl Phosphate. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 305-314.

(92) Řezáč, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chemical Reviews* **2016**, *116* (9), 5038-5071.

(93) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106* (23), 9639-9646.

(94) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, N_2, and H_2O. *Chem. Phys. Lett.* **1998**, *286* (3,4), 243-252.

(95) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated Ab Initio Methods? *J. Chem. Theory Comput.* **2011**, *7* (1), 33-43.

(96) Lim, I. S.; Schwerdtfeger, P.; Metz, B.; Stoll, H. All-Electron and Relativistic Pseudopotential Studies for the Group 1 Element Polarizabilities from K to Element 119. *J. Chem. Phys.* **2005**, *122* (10).

(97) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P.; Taylor, P. R. Basis Set Convergence of the Interaction Energy of Hydrogen-Bonded Complexes. *J. Chem. Phys.* **1999**, *111* (20), 9157-9167.

(98) Kim, K. S.; Mhin, B. J.; Choi, U. S.; Lee, K. Ab Initio Studies of the Water Dimer Using Large Basis Sets: The Structure and Thermodynamic Energies. *J. Chem. Phys.* **1992**, *97* (9), 6649-6662.

(99) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *J. Chem. Theory Comput.* **2014**, *10* (1), 49-57.

(100) Sherrill, C. D. Chapter 4 - Wavefunction Theory Approaches to Noncovalent Interactions. In *Non-Covalent Interactions in Quantum Chemistry and Physics*, Otero de la Roza, A., DiLabio, G. A. Eds.; Elsevier, 2017; pp 137-168.

(101) García, J. S.; Brémond, É.; Campetella, M.; Ciofini, I.; Adamo, C. Small Basis Set Allowing the Recovery of Dispersion Interactions with Double-Hybrid Functionals. *J. Chem. Theory Comput.* **2019**, *15* (5), 2944-2953.

(102) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.

(103) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456-1465.

(104) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124* (17), 174104.

(105) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147* (3), 034112.

(106) *Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16 Rev. B.01*; Wallingford, CT, 2016.

(107) Neese, F. The Orca Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73-78.

(108) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115* (19), 2315-2372.

Chapter 4: Building a Data Set of Group II Metal–Nucleic Acid Interactions and Evaluating the Performance of Density Functional Theory Methods

4.1. Introduction

Group II metals play important roles in diverse fields, contributing to biological processes, and the development of materials and therapeutics.[1] In biology, Mg^{2+} ions are essential cofactors for maintaining the structure and function of nucleic acids, providing stabilization to cell membranes, and aiding ATP utilization.[2] Ca^{2+} is the most abundant metal in the human body, playing a vital role in cellular signaling machinery and immune system responses.[3] In materials applications, Mg^{2+} and Ca^{2+} are key components in commercially approved bioactive glasses that are employed in soft tissue regeneration. [4] While Be^{2+}, Sr^{2+}, and Ba^{2+} are not biologically essential, they have proven to be useful in diverse industrial and technological applications. For example, Be^{2+} is used as a reflector in nuclear reactors, electronics, and aerospace materials.[5] Sr^{2+} is utilized in photocatalysts as well as a therapeutic treatment for osteoporosis. [6, 7] Ba^{2+} has been utilized in photovoltaic cells to improve solar cell performance, in addition to Ba^{2+} being applied in nonlinear optical materials for UV lasers and multiphoton generators, as well as in drilling fluids for use in the oil sands.[8, 9]

Although group II metals are important in biology, materials, and therapeutics, industrial processes, such as mining, manufacturing, and waste dumping can lead to excess levels of group II metals in the environment.[10, 11] These metals can disrupt the soil environment by altering pH levels, interfering with ion homeostasis, and reducing microbial activity, which diminishes plant growth and crop yields.[12, 13] Additionally, group II metals can leach into ground water, thereby contaminating our drinking water.[14] Excess

levels of group II metals in our food sources and drinking water can lead to their accumulation in the human body, which can cause negative effects on human health. For example, exposure to Be^{2+} in industrial workers has been correlated with the development of chronic beryllium disease and lung cancer.[15] Serum Mg^{2+} levels between 1.74 and 2.61 mM cause magnesium toxicity in the body, which can lead to cardiac arrest. [16] Additionally, patients with impaired kidneys are more susceptible to magnesium toxicity due to inefficient regulation of magnesium levels in the body.[16] An overabundance of Ca^{2+} in the human body causes dysfunction of Ca^{2+} channels, which leads to hypertension, atherosclerosis, and chronic fatigue syndrome.[3] Sr^{2+} mimics Ca^{2+} and therefore can displace Ca^{2+} in the body, with excess levels of Sr^{2+} leading to defective bone mineralization.[7] Ba^{2+} present in high levels in the human body can cause tremors, paralysis, and hypokalemia.[17] Given that elevated levels of group II metals negatively affect human and environmental health, it is important to design detection and extraction technologies for monitoring and remediation of group II metals.

Biosensors are powerful analytical tools made from biological components (e.g., antibodies, enzymes, nucleic acids) that can provide early detection of metal contaminants.[18] Among biological components, nucleic acids are a particularly promising solution for designing metal sensing platforms to detect metal contaminants in the environment and human body due to their diverse functionality.[19] Specifically, nucleic acids contain distinct sites including the negatively charged phosphate moieties, sugars, and nitrogenous bases that can bind metals with varying affinity.[19] Furthermore, nucleic acids are structurally versatile, demonstrating the ability to fold into diverse 3D structures (e.g., tetrahedrons, G-quadruplexes, pseudo-knots, hairpins) that can bind specific metals.[20] For

this reason, there have been several nucleic acid sensing platforms developed to detect metals.[19] For example, nucleic acid sensors have been designed to selectively detect toxic metals such as Cd^{2+}, Hg^{2+}, and Pb^{2+}.[21, 22] While nucleic acids have proven to be successful in detecting heavy metals, designing nucleic acid sensors for group II metals remains challenging largely due to metal competition.[19] For example, current nucleic acid probes are not specific enough to detect Mg^{2+} without interference from Ca^{2+}.[19] However, the EtNa DNAzyme can selectively detect Ca^{2+} in aqueous solutions, being the first nucleic acid sensor that can selectively distinguish between Mg^{2+} and Ca^{2+}.[23] Another DNA-based sensor was designed for Ba^{2+}; however, a limited selection of metals were tested for competition with Ba^{2+}.[24] Aside from limitations in nucleic acid sensors for Mg^{2+} and Ba^{2+}, no nucleic acid sensors have been designed to detect Be^{2+} or Sr^{2+} to date.[19, 24] To develop nucleic acid sensors specific for group II metals, it is important to first understand the structure and binding strengths of group II metal–nucleic acid interactions. Aside from developing metal sensing platforms, fundamental information about group II metal–nucleic acid interactions can aid the development of materials like DNA nanowires for electrical devices,[25] and the design of tunable therapeutics that enable controlled-release mechanisms for drugs.[26]

There have been multiple experimental studies performed to elucidate how group II metals interact with nucleic acids.[27-34] Melting temperature measurements indicate that Mg^{2+} enhances nucleic acid stability,[27] while Raman spectroscopy shows that there are minimal structural deviations to B-DNA in the presence of group II metals.[28] Another experimental study using CD spectroscopy showed that Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+} cause minimal structural impact, with heavier metals inducing less distortion to the structure of

DNA.[29] X-ray crystallography has provided structural insight into how Mg^{2+} and Ca^{2+} bind to nucleic acids, with Mg^{2+} interacting with the phosphate backbone, N7 or O6 of G, and O4 of U, and Ca^{2+} with the phosphate, O6 of G, and O4 of U.[30] Additional work that analyzed crystallographic structures of Mg^{2+}–nucleic acid interactions identified that Mg^{2+} does not generally bind to N7 sites but rather to the phosphate backbone.[31] Other NMR and X-ray crystallographic studies have obtained structural information of group II metals interacting with G-quadruplexes, with group II metals identified to either interact in the G-quadruplex channel or the phosphate backbone, thereby providing greater structural stability to G-quadruplexes.[32] For group II metal binding with nucleic acids, NMR experiments identified that Mg^{2+} and Ca^{2+} primarily interact with the phosphate backbone.[33] Ion exchange measurements identified Ca^{2+} to bind more strongly to DNA than Mg^{2+}.[34] Ultrasonic velocity experiments determined binding constants between group II metals and DNA in solution, with the order of ($Ca^{2+} > Sr^{2+} > Ba^{2+} > Mg^{2+}$) suggesting that heavier group II metals bind directly to DNA while Mg^{2+} undergoes outer-sphere coordination.[29] However, this method does not directly measure metal binding, which results in uncertainties on the interactions between group II metals and nucleic acids. While these experimental approaches have provided valuable insights, gaps remain in our understanding of the structural and binding properties for group II metal–nucleic acid interactions.

Computational studies have also been carried out using density functional theory (DFT) to gain atomic-level insights into the structure and binding of group II metal–nucleic acid interactions.[35-41] Specifically, the structure and stability of complexes of Be^{2+}, Mg^{2+}, or Ca^{2+} metals bound to each nucleobase binding site was investigated using B3LYP in the

gas phase, with the binding strengths decreasing as $Be^{2+} > Mg^{2+} > Ca^{2+}$.[35] A similar study investigated Mg^{2+} and Ca^{2+} binding to each nucleobase using CCSD(T) and found the binding strength decreases as $Mg^{2+} > Ca^{2+}$ and G > C > A.[36] Another study using B3LYP explored Be^{2+}, Mg^{2+}, Ca^{2+}, or Sr^{2+} interacting with select nucleobase binding sites.[37] Only one metal binding position in the phosphate backbone was explored, with Be^{2+}, Mg^{2+} or Ca^{2+} bound to the two non-bridging oxygens in the backbone.[38] A separate study investigated A, C, G, T, or U complexes of canonical and tautomeric nucleobases with Sr^{2+} or Ba^{2+} using B3LYP.[39] Although there have been computational works geared towards understanding how group II metals interact with nucleic acids, not all possible nucleobase sites were explored for Sr^{2+} or Ba^{2+} binding, and the phosphate moiety sites interacting with each group II metal has yet to be investigated. Additionally, the accuracy of the DFT methods used in these previous studies remain in question.

The present work provides a fundamental and complete understanding of how group II metals (i.e., Be^{2+}, Mg^{2+}, Ca^{2+}, Sr^{2+}, or Ba^{2+}) interact with each binding site in nucleic acid components (i.e., phosphate moiety and nucleobases; Figure 4.1). Specifically, the gas-phase structures of all possible directly bound group II metal–nucleic acid complexes are uncovered using ab initio methods (MP2/def2-TZVPP). Subsequently, a comprehensive data set of all group II metal–nucleic acid binding strengths is generated using CCSD(T)/CBS. Finally, the accuracy of 61 DFT methods for group II metal–nucleic acid binding strengths is assessed. For the first time, recommended DFT methods are identified for future research on the roles of these interactions in materials, therapeutics, and biotechnologies.

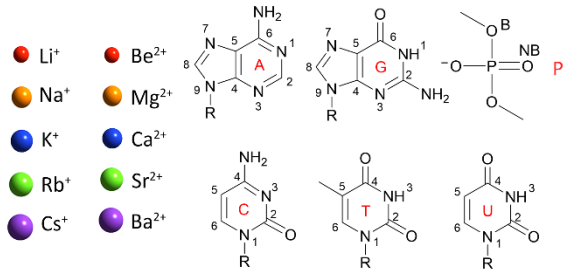


Figure 4.1. Metals and nucleic acid components investigated. Atom numbering and abbreviations for the binding sites (B = bridging and NB = non-bridging within the phosphate moiety) are included for clarity.

4.2. Computational Methodology

The model construction and methodology discussed throughout this section are based on previous computational work that established an accurate protocol for exploring the structure and energetics of group I metal–nucleic acid complexes.[42, 43] Initial models of group I metals interacting with each binding site within the nucleobases (i.e., A, C, G, T, or U) or the phosphate moiety were taken from previous work and each group I metal was subsequently replaced with its adjacent group II metal counterpart (i.e., Be^{2+}, Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}; Figure 4.1). All group II metal–nucleic acid complexes were optimized using MP2 with full electronic correlation in combination with the def2-TZVPP basis set. Subsequently, vibrational frequencies were evaluated at the same level of theory to confirm minima. The binding energy was calculated as the difference between the total energy of the dimer complex and the sum of the monomer energies (i.e., the metal and nucleic acid component). Our reference binding strengths were calculated using CCSD(T) extrapolated to the complete basis set limit (CBS) using the Helgaker scheme,[44, 45] with the extrapolation performed using MP2 calculations with the def2-TZVPP and the def2-QZVPP basis sets, and CCSD(T)/def2-TZVPP. The def2 basis sets have been confirmed to be compatible with

this approach.[46] The large-core Stuttgart Dresden (SDD) scalar relativistic effective core potential (ECP) was used for Ca^{2+}, Sr^{2+}, and Ba^{2+}. The reported CCSD(T)/CBS values included averaged counterpoise-corrected and uncorrected MP2 binding strengths, which has been reported to minimize errors arising from incomplete basis sets.[47-50]

The ability of 61 DFT methods (Table 4.1) to reliably describe the CCSD(T)/CBS binding strengths of group II metal–nucleic acid complexes was tested in combination with def2-TZVPP, as previous work has showed that different basis sets negligibly affect metal–nucleic acid binding strengths.[43] The scalar relativistic SDD ECP was used to describe Ca^{2+}, Sr^{2+}, and Ba^{2+}. We assess 15 functionals that include Grimme's D3 correction with Becke-Johnson (BJ) damping.[51-53] Four functionals tested contain the D4 correction.[54] We note that previous work investigated the impact of counterpoise corrections on group I metal–nucleic acid interactions and found that differences between corrected and uncorrected binding strengths are minimal (≤1.5% or ≤0.6 kcal/mol).[42] Therefore, only DFT counterpoise-corrected values are considered in the present work.[42] All calculations employing D4-corrected functionals, in addition to ωB97M-V, ωB97X-V, and ωB97X-D3(BJ), were carried out in ORCA 5.0.2 using "VeryTight" SCF convergence criteria.[55] The remaining 54 functionals were implemented using Gaussian 16 (B.01) with the default convergence criteria.[56]

Table 4.1. Functionals tested in the present work for their ability to reproduce CCSD(T)/CBS group II metal–nucleic acid binding strengths.

Family	Functional	D3[a]	D4[a]	%HF[b]	%MP2	Ref.
Double-Hybrid	B2-PLYP	X		53	27	57
	mPW2-PLYP			55	25	58
	PBE-QIDH			69	33.4	59
	PBE0-DH			50	12.5	60
	DSD-PBEP86			68	(51,23)[d]	61
RSH	M11			42.8–100[c]		62
	MN12-SX	X		25–0[c]		63
RSH meta-GGA	ωB97M-V			15		64
RSH GGA	ωB97					65
	ωB97X	X	X	16		65
	ωB97X-D			22		66
	ωB97X-V			16.7		67
	HSE06	X		25–0[b]		68-74
	LC-PBE					75, 76
	LC-ωPBE	X				77-79
	CAM-B3LYP	X		19–65[b]		80
GH meta-GGA	M06			27		81
	M06-2X			54		81
	M06-HF			100		82
	BMK			42		83
	MN15			44		84
	PW6B95			28		85
GH GGA	BH&HLYP			50		86
	SOGGA11-X			40.15		87
	B3PW91	X		20		88
	PBE0	X		25		89
	B3LYP	X	X	20		88, 90
	X3LYP	X		21		91
	O3LYP	X				92
	TPSSh			10		93, 94
	B97-2			21		95
Local meta-GGA	TPSS	X	X			96
	revTPSS					97, 98
	M11-L					99
	M06-L					100
	MN12-L					101
	MN15-L					102
Local GGA	mPW91					103
	BLYP	X				90, 104, 105
	BP86	X				90, 104
	PBE	X	X			75

Table 4.1. Continued.

Family	Functional	D3[a]	D4[a]	%HF[b]	%MP2	Ref.
Local LDA	SVWN5					106

[a]X signifies the functional was used with and without the D3(BJ) or D4 empirical dispersion correction. [b]Hartree-Fock exchange percentage (%HF) contained in the functional. [c]%HF at short range (first value) and long range (second value). [d]The spin-component-scaled formalism is used to determine the amount of MP2 in this functional, with the numbers representing opposite- and same-spin values, respectively.

4.3. Results and Discussion

4.3.1. Establishing the CCSD(T)/CBS data set consisting of group II metal–nucleic acid complexes

From the CCSD(T)/CBS data set on group I metal–nucleic acid interactions,[42] each group I metal generally preferred the 2NB > 1B1NB > 2B sites in the phosphate moiety, followed by nucleobase binding sites in the order of O6–N7 >> N2–N3 of G, O2–N3 >> N3–N4 of C, N6–N7 > N1–N6 > N3 of A, and O4 > O2 of T/U (Figure 4.2). The binding affinity decreased according to group I metal as Li^+ >> Na^+ >> K^+ > Rb^+ > Cs^+ and according to nucleic acid component as P >> G > C >> A ~ T = U. In A, N6–N7 is dominant for Li^+ and Na^+, while N6–N7 and N1–N6 are equivalent for K^+, Rb^+, and Cs^+. In C, N3–N4 exclusively forms with Li^+, while N2–N3 in G only occurs with Li^+, Na^+, and K^+.

Group II metals bind to A sites in a similar fashion to group I metals, forming bidentate interactions at the N6–N7 and N1–N6 sites of A, while monodentate interactions occur at the N3 site. Like group I metals, all group II metals form bidentate interactions with O6–N7 in G. However, unlike group I metals, all group II metals consistently form bidentate interactions with N2–N3 of G, with this complex not formed for Rb^+ and Cs^+. All group II metals bind to the O2–N3 sites in C. Additionally, all group II metals except

for Ba^{2+} form complexes with N3–N4 in C, unlike Na^{+}, K^{+}, Rb^{+}, and Cs^{+}. All group II metals bind to O4 or O2 in T/U and the 2NB sites in the phosphate moiety and form same structures as observed for group I metals. In the phosphate moiety, Be^{2+} and Mg^{2+} form bidentate complexes with the 1B1NB sites, like Li^{+}. All remaining group II metals form the same structural motif for the 1B1NB site in the phosphate moiety, forming a 3-coordinate complex with 2B and 1NB oxygen sites (Figure C.1, Appendix C). Overall, there were more complexes formed between group II metals (69) and nucleic acid components than for group I (64), with the 5 additional complexes observed due to more group II metals binding to N3–N4 of C and N2–N3 of G.

The nucleic acid binding decreases as P >> G >> C >> A >> T~U (Figure 4.2), with greater differences in binding strengths between metals for each nucleic acid component. The binding strengths decrease as P(2NB > 1B1NB > 2B), followed by G(O6–N7 >> N2–N3), C(O2–N3 >> N3–N4), A(N6–N7 > N1–N6 > N3), and T/U(O4 > O2). Notably, there are alterations in metal binding behavior between the groups, with all group II metals preferentially binding to N6–N7 in A, followed by N1–N3, unlike group I metals which showed near equivalency between N1–N3 and N6–N7 with slight preference for N1–N3 in A. Nevertheless, the remaining metal–nucleic acid binding strength trends are consistent between group I and II metals.

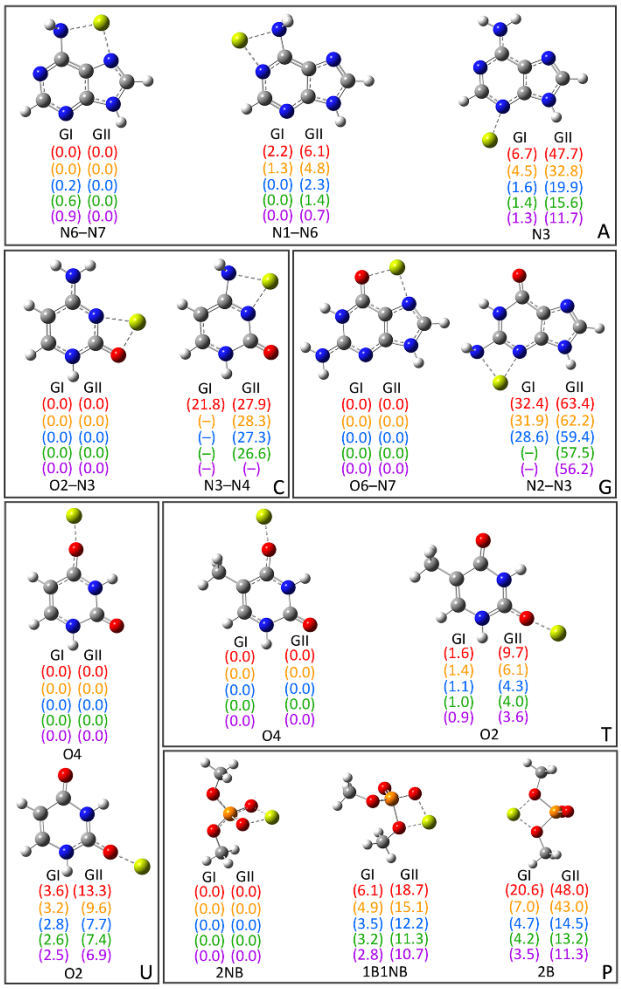


Figure 4.2. Combined CCSD(T)/CBS data set of group I and II metal–nucleic acid complexes. All relative energies (kcal/mol) are reported with group I (GI) values in the first column and group II (GII) values in the second column, which are listed in increasing order of atomic size. Coordination distances for group I metal–nucleic acid interactions are available in previous work,[42] while coordination distances for group II–metal nucleic acid interactions are available in Table C.1 in Appendix C.

The magnitude of group II metal–nucleic acid binding strengths (Tables C.1–C.2, Appendix C) is significantly greater than for group I (Table C.3, Appendix C).[42] Comparing the smallest metal from each group (Be^{2+} and Li^{+}), which have the strongest binding affinities, Be^{2+} binds to the phosphate moiety (2NB) and O6–N7 in G, respectively, ~3 and ~4 times stronger than Li^{+}. As the size of the metal increases, a drastic increase in the

magnitude in the binding strength occurs in group II relative to group I metal–nucleic acid complexes. As a specific example, Ba^{2+} is bound to the 2NB sites of the phosphate moiety ~13 times more strongly than Cs^{+}, while there is ~3 times enhanced Ba^{2+} binding to O6–N7 in G compared to Cs^{+}. The greater binding affinities are supported by shorter coordination distances in group II than group I metal–nucleic acid complexes, with the largest decrease in coordination distance being 0.479 Å from Cs^{+} to Ba^{2+} (Table C.1, Appendix C).[42] Additionally, the relative energies between complexes for a given metal–nucleic acid binding site are greater for group II than group I metals, implying stronger preference for group II metals at each nucleic acid site. Overall, we have provided the first comprehensive data set of gas-phase interactions between group II metals and nucleic acid components.

4.3.2 Evaluation of functional performance for reproducing CCSD(T)/CBS group II metal–nucleic acid binding strengths

Building on our previous work that generated a CCSD(T)/CBS data set on group I metal–nucleic acid interactions,[42] we have expanded our scope to include group II metals. By combining both CCSD(T)/CBS data sets, we can compare structural and binding strength differences between group I and II metal–nucleic acid interactions. Computational methods that can reliably describe both groups are required to understand the roles of metals and metal competition in different applications that utilize group I and II metal–nucleic acid interactions. We initially compared functional performance using percent errors (PE) and unsigned errors (UE), metrics that will provide a balanced description of

metal–nucleic acid interactions given the variation in binding strengths across metals of different sizes and identify reliable methods for these interactions.

4.3.3. ωB97M-V is the most reliable functional for describing group I and II metal–nucleic acid binding strengths

Combining data sets for both group I and II metal–nucleic acid interactions results in significant outliers and deviations in the statistics for both metrics (Figures 4.3 and C.2; Table C.3, Appendix C). Although there is no single functional that achieves highly accurate descriptions of group I and II metal–nucleic acid interactions with maximum errors within chemical accuracy for main group elements (1.0 kcal/mol),[107] there are a few robust methods that offer consistently reliable performance over both groups of metals. In particular, the functional that comes the closest to achieving chemical accuracy with the smallest average errors (0.8% MPE; 0.6 kcal/mol MUE) and a remarkably low maximum UE of 1.9 kcal/mol is ωB97M-V. Therefore, ωB97M-V is recommended as the most suitable functional for studying group I and II metal–nucleic acid interactions, especially in the context of metal competition. Other robust methods include mPW2-PLYP, M06-2X, and mPW91, with similarly small average deviations (1.2–1.5% MPE; 1.0–1.2 kcal/mol MUE), albeit larger maximum errors (~5–12%; ~4–6 kcal/mol).

According to our previous study, there was significant fluctuation in functional performance across group I metal–nucleic acid interactions, with the largest discrepancies arising from the PEs rather than UEs.[42] Despite variation in functional performance, ωB97M-V and mPW2-PLYP emerged as the most reliable methods for describing group I

metal–nucleic acid interactions, with the smallest MPEs and MUEs (≤1.6%; <1.0 kcal/mol; Figure 4.3 and C.2, Appendix C).[42]

The PEs for group II metal–nucleic acid interactions are more consistent than those of group I metals, with lower PEs and fewer outliers in the boxplots (Figure 4.3 and C.2; Table C.5, Appendix C). However, there are more highly dispersed errors across functional families when comparing the UEs of group I and II metals. This is likely due to group II metals having higher binding energies than group I metals, leading to smaller PEs. Nevertheless, it is important to consider both PEs and UEs to provide a complete understanding of functional performance. The recommended functionals for group II metal–nucleic acid interactions are ωB97M-V and ωB97X-V, which demonstrate the smallest average errors (<1.0% MPE; ≤1.0 kcal/mol MUE) and maximum errors (<3%; <3 kcal/mol), closely followed by PBE0-DH, with slightly greater average (0.9% MPE; 1.3 kcal/mol MUE) and maximum UE (3.8 kcal/mol; Figures 4.3 and C.2; Tables 4.2 and C.4, Appendix C). Additionally, mPW2-PLYP is recommended for describing group II metal–nucleic acid complexes (1.2% MPE; 1.0 kcal/mol MUE), albeit with larger maximum UEs (5.3%; 5.7 kcal/mol).

Overall, group I metal–nucleic acid interactions are more challenging to describe by functionals compared to group II metals. Indeed, 19 functionals are deemed reliable over group II metals, while 11 are recommended for group I metals. This result is highlighted in the combined group I and II metal–nucleic acid dataset (Table C.4. Appendix C), which demonstrates a greater number of outliers in the PEs that is attributed to the weaker binding strengths of group I metal–nucleic acid complexes and thereby group I metals are more sensitive to errors. In contrast, the outliers in UEs are dominated by group II metals, however, the gaps in binding strengths between each group II metal is greater than for group I, which demonstrates less risk of altering binding strength trends of metal–nucleic acid interactions based on functional choice. Given that key differences in functional performance over group I and II metals and between group I or II metals, functional performance is subsequently explored within group II metals.

Table 4.2. Recommended functionals from each family according to the rungs of Jacob's Ladder (X) and the most accurate functionals overall (**X*, bold**) for metal–nucleic acid complexes.[a]

Family	Functional	Group I&II	Group I	Group II	Be^{2+}	Mg^{2+}	Ca^{2+}	Sr^{2+}	Ba^{2+}
Double-Hybrid	B2-PLYP		X	X		X	X	**X***	X
	B2-PLYP-D3					X	X		
	mPW2-PLYP	X*	**X***	X		X	X	**X***	**X***
	PBE-QIDH				X			X	X
	PBE0-DH			**X***	X		X	X	
	DSD-PBEP86			X	X		X	X	
RSH	M11								
	MN12-SX								
	MN12-SX-D3								
RSH meta-GGA	ωB97M-V	**X***	**X***	**X***	**X***	**X***	X	**X***	**X***
RSH GGA	ωB97								X
	ωB97X								**X***
	ωB97X-D						X		
	ωB97X-D3								
	ωB97X-V		X	**X***	**X***	X	X	**X***	**X***
	ωB97X-D4				X	X			
	HSE06			X			**X***	X	
	HSE06-D3								

Table 4.2. Continued.

Family	Functional	Group I&II	Group I	Group II	Be^{2+}	Mg^{2+}	Ca^{2+}	Sr^{2+}	Ba^{2+}
	LC-PBE								
	LC-ωPBE						X	X	X
	LC-ωPBE-D3					**X***	X	X	
	CAM-B3LYP								X
	CAM-B3LYP-D3								
GH meta-GGA	M06				X				
	M06-2X	**X***	X*	X	**X***	X	**X**	X	
	M06-HF						X	X	
	MN15				X				**X***
	BMK						**X***	**X***	X
	PW6B95			X		X	**X**	**X**	X
GH GGA	BH&HLYP					X			X
	SOGGA11-X			X		X		**X***	X
	B3PW91			X	X				
	B3PW91-D3								
	PBE0		X	X			X	X	X
	PBE0-D3								
	B3LYP		X				**X***	X	X
	B3LYP-D3								
	B3LYP-D4								
	X3LYP						X	X	X
	X3LYP-D3								
	O3LYP						X		
	O3LYP-D3								
	B97-2			X	X		X		
	TPSSh								
Local meta-GGA	revTPSS		**X**	X		X	X	X	X
	TPSS		**X**	X		X	X	X	X
	TPSS-D3								
	TPSS-D4								
	M06-L				X				
	M11-L								
	MN12-L				X				
	MN15-L			X			X	X	X
Local GGA	mPW91	**X***		X	X	**X***	**X***	X	**X***
	BLYP		X				X	X	X
	BLYP-D3								
	BP86			X	X	X			
	BP86-D3								
	PBE		X	X			X	X	X
	PBE-D3								
	PBE-D4								
Local LDA	SVWN5								

[a]X signifies functionals that offer reasonable accuracy according to functional family for each group II metal or across groups I and II of metals. **X*** represents the most accurate functionals among the 61 tested for describing metal–nucleic acid complexes.

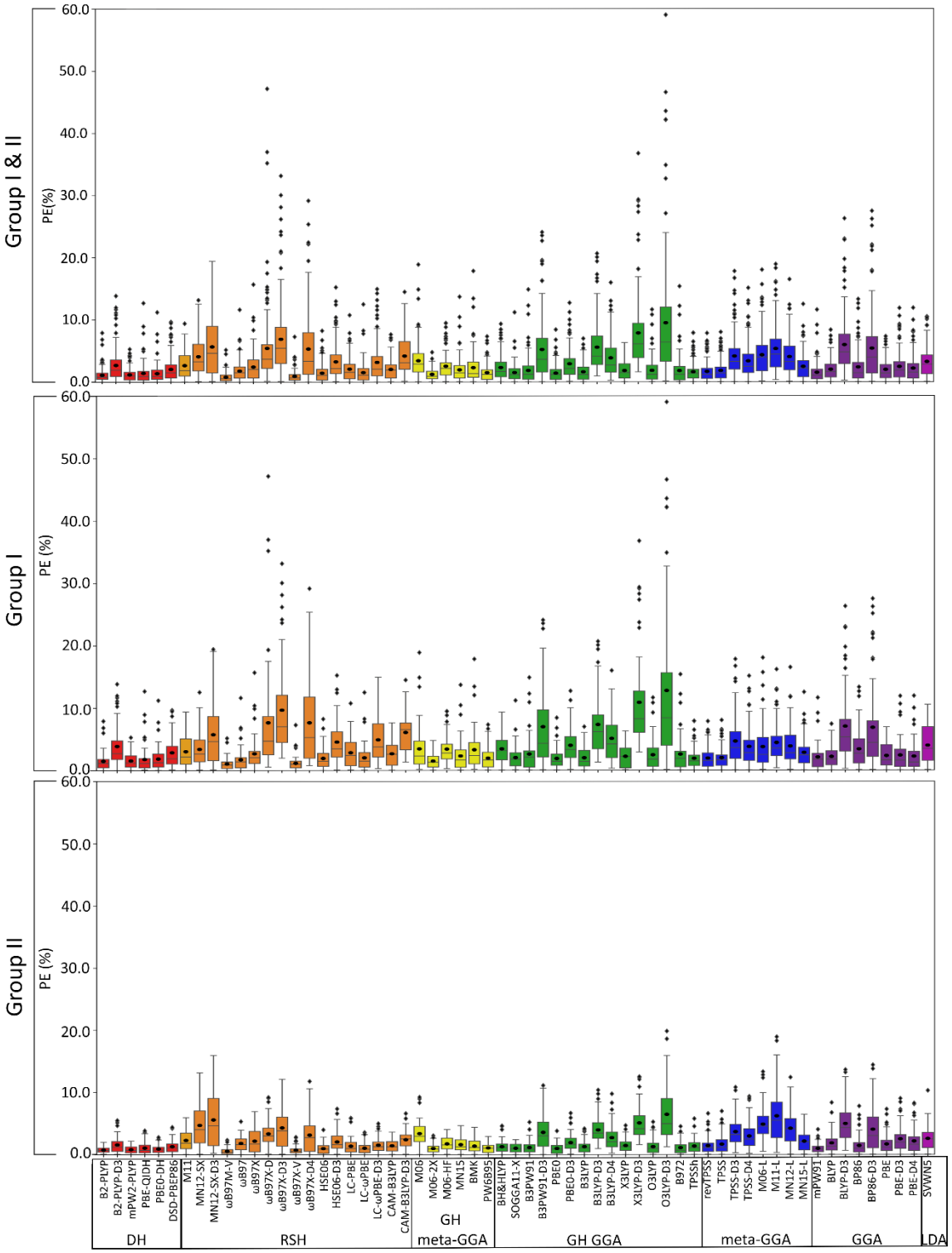


Figure 4. 3. Boxplot of the percent errors (%) for group I metals (top), group II metals (middle), and group I and II metals (bottom). Each color represents a functional family according to the rungs of Jacob's Ladder. The boxplot for group I metal–nucleic acid interactions was obtained from our previous study.[42]

4.3.4. Exploring functional performance for group II metals interacting with nucleic acids according to functional family

To further explore functional variation observed in the boxplots for group II metal–nucleic acid interactions, we focus our analysis on how functionals describe individual interactions between group II metals and each nucleic acid component (Figure 4.4). The disparities in the PEs highlight the dependency of functional performance on metal identity and nucleic acid component considered. To understand how functionals describe each group II metal–nucleic acid complex, the following sections investigate functional performance across different families. Specifically, the following sections are ordered according to the rungs in Jacob's Ladder, with double-hybrid, range-separated hybrid, global-hybrid meta-GGA, global-hybrid GGA, meta-GGA, GGA, and LDA.

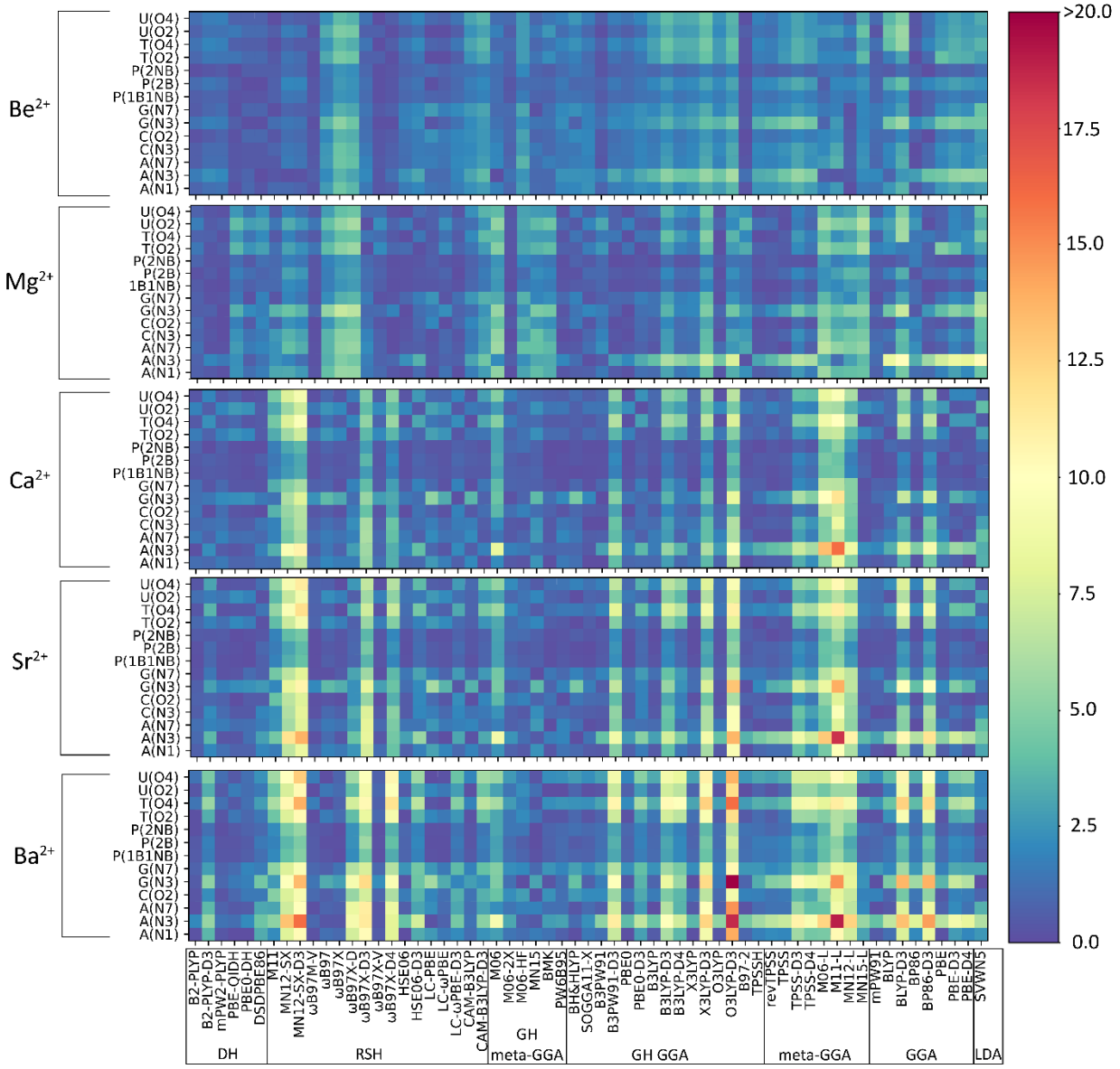


Figure 4.4. Performance of each functional for describing each group II metal–nucleic acid binding strength compared to CCSD(T)/CBS reference values, reported as percent errors (%).

4.3.4.1. Double hybrids

In general, the double-hybrid family provides a reliable description of each group II metal interacting with nucleic acid components, regardless of metal, with minimal discrepancies in the average MPEs (0.3–3.4%) and slightly greater fluctuations in MUEs

(0.6–4.1 kcal/mol; Figures 4.5 and C.3; Table C.6, Appendix C). Nevertheless, there are different functionals that emerge as top performers for each metal. Specifically, PBE0-DH and DSD-PBEP86 are the best functionals for describing Be^{2+} complexes, with the smallest average errors (0.5–0.6% MPE; 1.5–1.7 kcal/mol MUE), standard deviations (0.3%; 0.9 kcal/mol), and maximum errors (1.1%; 3.0 kcal/mol), as well as no outliers regardless of metric considered. Following, PBE-QIDH (0.5% MPE; 1.3 kcal/mol MUE) is a reliable option for studying Be^{2+} complexes, albeit with slightly larger maximum errors (1.5%; 3.3 kcal/mol). While B2-PLYP, B2-PLYP-D3, and mPW2-PLYP are not recommended for complexes involving Be^{2+}, these functionals provide an excellent description (0.3–0.8% MPE; 0.5–1.5 kcal/mol MUE) of Mg^{2+} complexes, with mPW2-PLYP emerging as the most reliable option for Mg^{2+}, with the smallest average (0.3% MPE; 0.5 kcal/mol MUE) and maximum deviations (1.5%; 2.0 kcal/mol). For the Ca^{2+} complexes, there are many functionals in the double-hybrid family that yield accurate results and no outliers, including B2-PLYP, B2-PLYP-D3, mPW2-PLYP, PBE0-DH, and DSD-PBEP86 (0.7–1.1% MPE; 0.8–1.3 kcal/mol MUE). Sr^{2+} and Ba^{2+} complexes exhibit similar functional performance, with B2-PLYP, mPW2-PLYP, and PBE-QIDH all deemed reliable (0.6–0.8% MPE; 0.6–0.8 kcal/mol MUE) for describing both metals. PBE0-DH and DSD-PBEP86 are only recommended for Sr^{2+} complexes (0.9–1.2% MPE; 0.7–1.2 kcal/mol), with low spreads in the data based on boxplot visualizations. Altogether, these results highlight that while the best-performing functionals in the double-hybrid family may vary for different metals, the double-hybrid family provides reliable descriptions of group II metal–nucleic acid binding strengths.

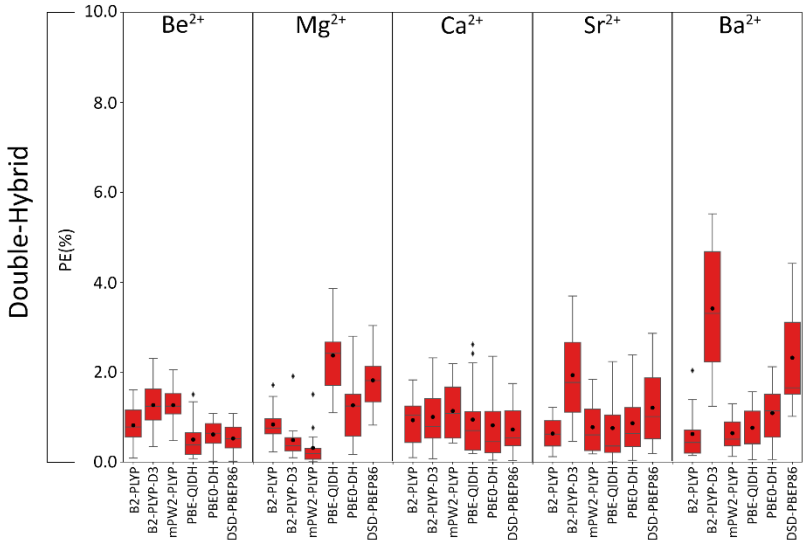


Figure 4.5. Boxplot of the percent errors (%) for the double-hybrid family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

4.3.4.2. Range-separated hybrids

Overall, the range-separated hybrid family exhibits greater fluctuation than the double-hybrids in functional description of each group II metal (Figures 4.6 and C.4; Table C.7, Appendix C). Despite these fluctuations, two functionals within this family consistently yield accurate results across all group II metals, specifically, ωB97M-V and ωB97X-V with remarkably small average (0.3–1.0% MPE; 0.4–1.6 kcal/mol MUE) and maximum deviations (2.8%; 2.9 kcal/mol). Aside from these two functionals, ωB97X-D4 is also deemed reliable for Be^{2+} and Mg^{2+} complexes (0.4–0.5% MPE; 0.7–1.4 kcal/mol MUE), while LC-ωPBE-D3 is also accurate for describing Mg^{2+} complexes (0.5% MPE; 0.9 kcal/mol). For the heavier metals (Ca^{2+}, Sr^{2+}, and Ba^{2+}), a greater number of functionals offer reliable descriptions of these complexes compared to Be^{2+} and Mg^{2+}. HSE06 is the top performing functional from the range-separated hybrid family for complexes containing Ca^{2+}, with minimal average deviations (0.8% MPE; 0.8 kcal/mol MUE) and the smallest

maximum errors (2.1%; 1.7 kcal/mol), followed by ωB97X-D (1.3% MPE; 1.5 kcal/mol MUE). Both LC-ωPBE, and LC-ωPBE-D3 are deemed reliable for complexes involving Ca^{2+} and Sr^{2+} (1.1–1.7% MPE; 1.0–1.6 kcal/mol MUE). For Ba^{2+} complexes, ωB97, LC-ωPBE, and CAM-B3LYP are recommended, as they exhibit exceptionally small average (0.8–1.1% MPE; 0.6–1.0 kcal/mol MUE) and maximum (3.3%; 2.0 kcal/mol) errors. While functionals in the range-separated-hybrid family have more widely dispersed errors than the double-hybrid family, the top performers (ωB97M-V and ωB97X-V) offer reliable and computationally cheaper alternatives to the double-hybrid functionals for describing group II metal–nucleic acid interactions.

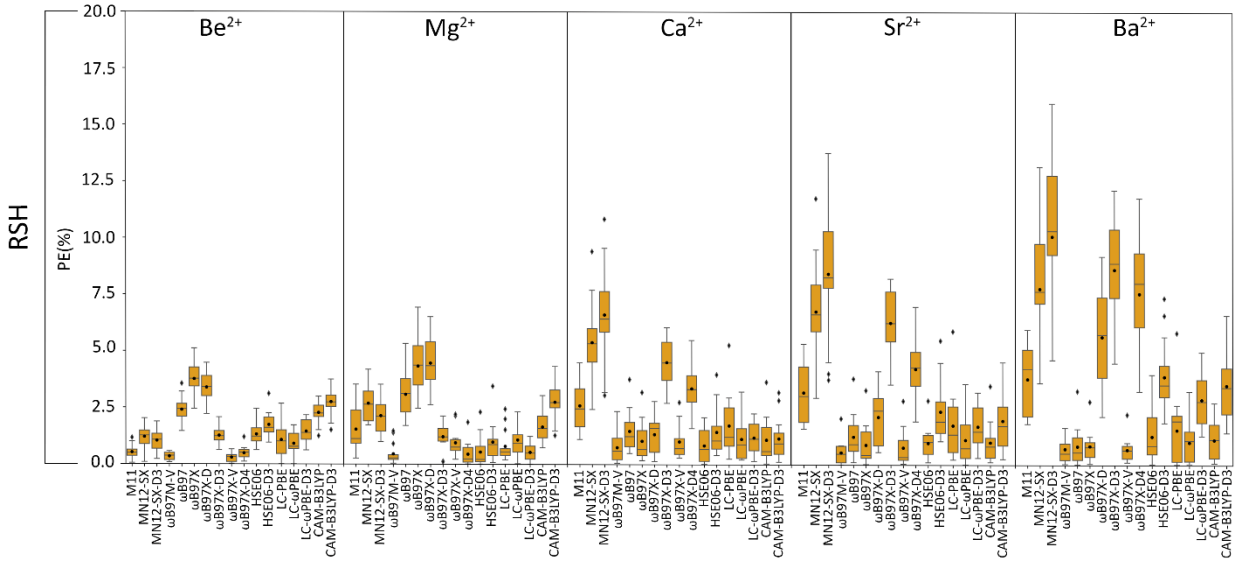


Figure 4.6. Boxplot of the percent errors (%) for the range-separated hybrid family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

4.3.4.3. Global-hybrid meta-GGA

In general, functionals in the global-hybrid meta-GGA family exhibit less spread than those in the range-separated hybrid family. Within the global-hybrid meta-GGA

family, M06-2X provides the best description of Be^{2+}–nucleic acid complexes, with minimal average (0.8% MPE; 2.4 kcal/mol MUE) and maximum (1.3%; 3.3 kcal/mol) errors (Figures 4.7 and C.5; Table C.8, Appendix C). Otherwise, M06 and MN15 offer reasonable performance for Be^{2+} complexes (1.0–1.4% MPE; 3.0–4.1 kcal/mol MUE). Only M06-2X and PW6B95 are deemed reliable from this family to study Mg^{2+}–nucleic acid complexes, with exceptionally low average deviations (0.3–0.5% MPE; 0.5–0.9 kcal/mol MUE). BMK, M06-2X, and PW6B95 are the most accurate functionals for Ca^{2+} and Sr^{2+} complexes, with small errors (0.6–1.2% MPE; 0.5–1.4 kcal/mol MUE) and no outliers in the boxplot, followed by M06-HF (0.9–1.0% MPE; 1.2 kcal/mol). MN15 emerges as the top performing functional for Ba^{2+} complexes with minimal deviations (0.9% MPE; 0.6 kcal/mol MUE), followed by BMK and PW6B95 (1.3% MPE; 1.4 kcal/mol MUE). Overall, M06-2X and PW6B95 serve as excellent cost-effective alternatives to functionals in the range-separated hybrid family for studying group II metal–nucleic acid interactions.

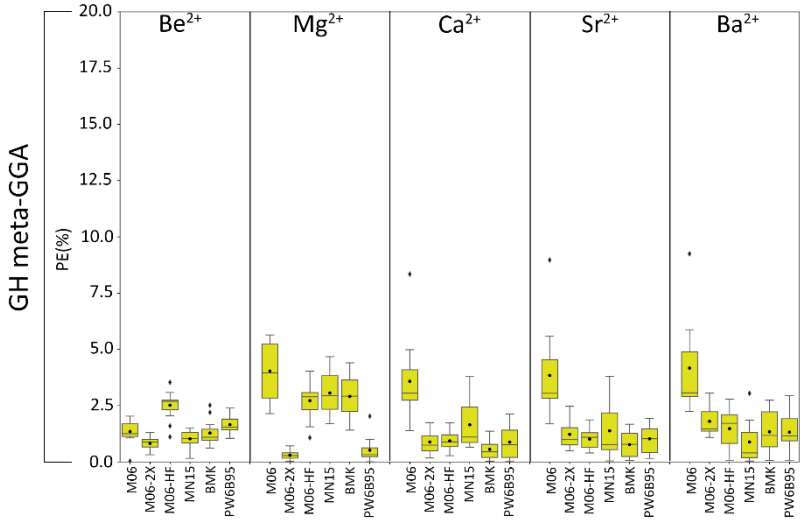


Figure 4.7. Boxplot of the percent errors (%) for the global hybrid meta-GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

4.3.4.4. Global-hybrid GGA

The global-hybrid GGA family demonstrates considerable variation in functional performance, with deviations becoming more pronounced as the size of the metal increases (Figures 4.8 and C.6; Table C.9, Appendix C). Only two (B3PW91 and B97-2) of the 15 functionals within this family offer a reliable description of Be^{2+} complexes (0.3–0.6% MPE; 1.0–1.9 kcal/mol MUE). Like Be^{2+}, only two functionals (BH&HLYP and SOGGA11-X) produce reliable results for complexes involving Mg^{2+}. These functionals lead to the smallest errors (0.9% MPE; 1.6 kcal/mol MUE) and no outliers in the boxplots. There are more options of reliable functionals for describing Ca^{2+} complexes, with five functionals (PBE0, B3LYP, X3LYP, O3LYP, and B97-2) recommended based on their accuracy (0.7–0.9% MPE; 0.7–0.9 kcal/mol MUE). With the smallest maximum errors regardless of metric (1.6%; 1.6 kcal/mol), B3LYP emerges as the most accurate method of the five functionals. Multiple functionals produce similarly reliable results for Sr^{2+} and Ba^{2+} complexes. Specifically, X3LYP, B3LYP, PBE0, and SOGGA11-X perform well for Sr^{2+} (0.8–0.9% MPE; 0.7–1.0 kcal/mol MUE) and Ba^{2+} (1.2–1.6% MPE; 1.3–1.7 kcal/mol MUE), with SOGGA11-X demonstrating the highest accuracy for Sr^{2+} (0.8% MPE; 0.7 kcal/mol MUE). BH&HLYP is also recommended for Ba^{2+} complexes (1.6% MPE; 1.5 kcal/mol MUE). Despite significant variation across functionals in the global-hybrid GGA for describing group II metal–nucleic acid interactions, SOGGA11-X and PBE0 emerge as reliable cost-effective alternatives to characterize the binding strengths of group II metal–nucleic acid complexes.

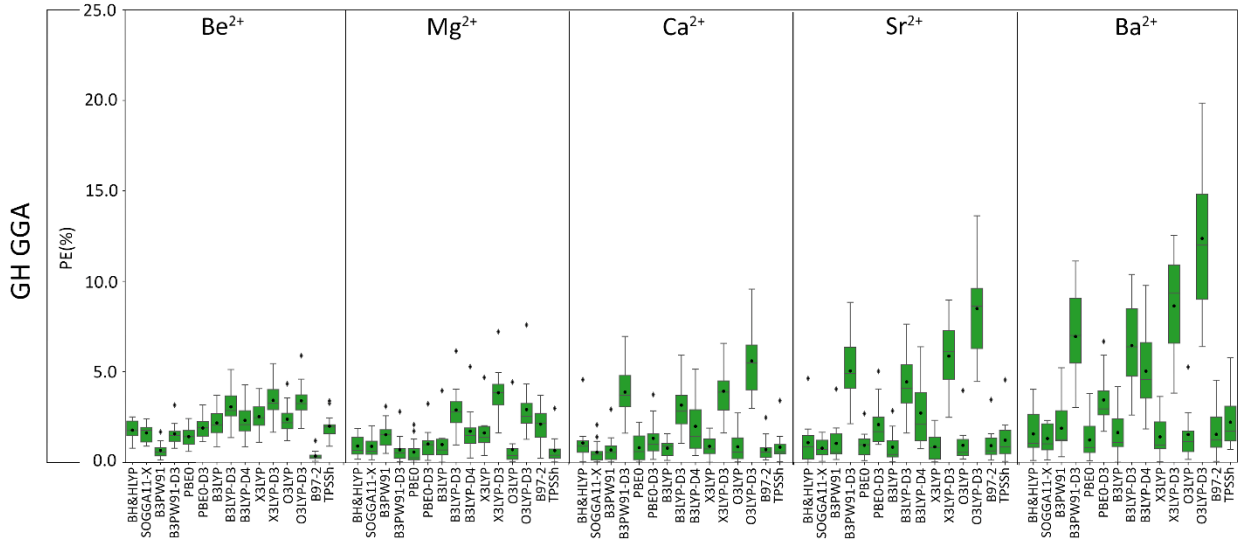


Figure 4.8. Boxplot of the percent errors (%) for the global hybrid GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

4.3.4.5. Meta-GGA, GGA, and LDA

The meta-GGA family shows disparities in functional performance between group II metals (Figures 4.9 and C.7; Table C.10, Appendix C). Two functionals in this family (revTPSS and TPSS) offer reasonable performance in describing complexes involving Mg^{2+}, Ca^{2+}, Sr^{2+}, or Ba^{2+} (0.8–2.6% MPE; 1.0–2.5 kcal/mol MUE). MN15-L is only deemed reliable for Ca^{2+}, Sr^{2+}, and Ba^{2+} complexes (0.8–2.0% MPE; 0.9–1.9 kcal/mol MUE). In contrast, no meta-GGA family members demonstrate excellent performance for Be^{2+}, although M06-L and MN12-L (0.9–1.6% MPE; 2.9–4.5 kcal/mol MUE) offer reasonable accuracy but come with notably large maximum UEs (7.2–9.4 kcal/mol).

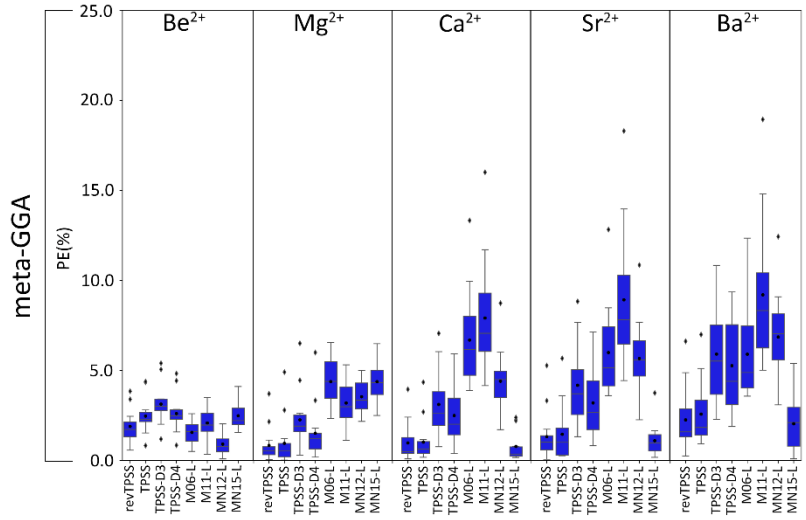


Figure 4.9. Boxplot of the percent errors (%) for the meta-GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

For the Be^{2+} complexes, mPW91 and BP86 are the most dependable functionals in the GGA family, with the smallest average errors (0.7–1.1% MPE; 2.0–3.1 kcal/mol MUE; Figures 4.10 and C.8; Table C.11, Appendix C). These functionals are also recommended for Mg^{2+} (0.9–1.6% MPE; 1.6–2.9 kcal/mol MUE), with mPW91 providing the most accurate description with the smallest average (0.9% MPE; 1.6 kcal/mol MUE) and maximum deviations (2.0–2.8%; 4.7–6.3 kcal/mol). The best-performing GGA functional for Ca^{2+}, Sr^{2+}, and Ba^{2+} complexes is mPW91 (0.7–1.3% MPE; 0.7–1.5 kcal/mol MUE). Additionally, BLYP and PBE yield similarly reliable results (1.1–2.4% MPE; 1.1–2.1 kcal/mol MUE). While the meta-GGA and GGA family show significant variability in functional performance, MN15-L (meta-GGA) and mPW91 (GGA) both provide reasonably accurate descriptions of group II metal–nucleic acid interactions.

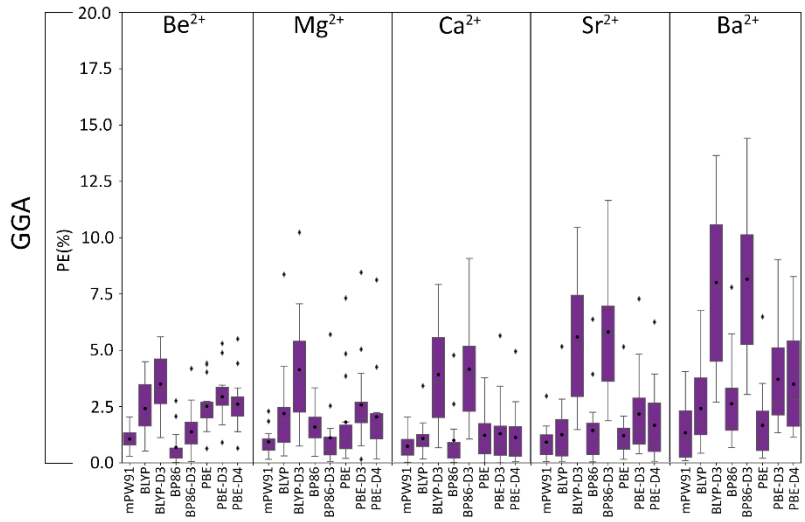


Figure 4.10. Boxplot of the percent errors (%) for the GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

Finally, the LDA family is included for completeness within the hierarchical framework of Jacob's Ladder. SVWN5 yields large average errors (3.0–4.0% MPE; 7.2–7.5 kcal/mol MUE) for smaller metals (Be^{2+} and Mg^{2+}), while SVWN5 is more consistent with small average errors (1.8–2.0% MPE; 1.5–2.2 kcal/mol MUE) for the larger metals (Ca^{2+}, Sr^{2+}, and Ba^{2+}; Figures 4.11 and C.9; Table C.12, Appendix C).

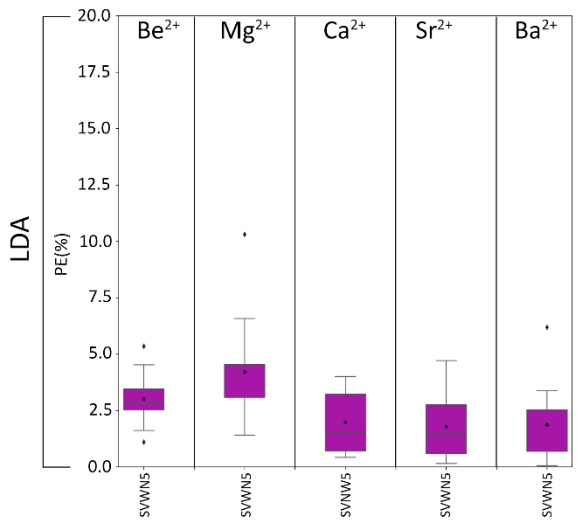


Figure 4.11. Boxplot of the percent errors (%) for the LDA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

4.4. Conclusions

Overall, this study expands the existing data set on group I metal–nucleic acid interactions to include a comprehensive set of group II metal–nucleic acid complexes. For the first time, the structures of group II metals coordinating to each binding site within each nucleic acids (i.e., A, C, G, T, U, and the phosphate moiety) have been determined using a high level of theory (MP2/def2-TZVPP), in addition to gold-standard binding energies calculated using CCSD(T) at the complete basis set limit. Subsequently, 61 DFT methods were assessed on their ability to reproduce CCSD(T)/CBS binding strengths of 69 group II metal–nucleic acid interactions. Overall, the best performing functional for describing both group I and II metal data sets is ωB97M-V. For investigating group II metal–nucleic acid complexes, ωB97M-V, ωB97X-V, and PBE0-DH yield highly accurate results. Within each functional family, the top-performing functionals vary for each metal. When comparing functional performance between group I and II metals, group I demonstrates significant

variation in PEs, but lower UEs than group II metals, while group II demonstrates lower PEs, but larger UEs than group I metals. This result is due to group II metal–nucleic acid binding strengths larger in magnitude than group I metals. Nevertheless, M06-2X and mPW91 are cost-effective alternatives from the lower rungs of Jacob's Ladder compared to ωB97M-V to study both group I and II metal–nucleic acid complexes. Although this chapter discusses gas-phase results, the inclusion of implicit solvent will decrease the magnitude of the binding strengths but the relative preference of metals for each binding site and the relative performance of functionals is likely to be the same in different solvent environments. The impact of both explicit and implicit solvent on metal–nucleic acid interactions is examined in Chapter 6 for the representative example of Li^+.

4.5. References

(1) Halka, M.; Nordstrom, B. *Alkali and Alkaline Earth Metals*; Infobase Publishing, 2010.

(2) Tomita, A.; Zhang, M.; Jin, F.; Zhuang, W.; Takeda, H.; Maruyama, T.; Osawa, M.; Hashimoto, K.-i.; Kawasaki, H.; Ito, K.; et al. ATP-Dependent Modulation of Mgte in Mg^{2+} Homeostasis. *Nat. Commun.* **2017**, *8* (1), 148.

(3) Froghi, S.; Grant, C. R.; Tandon, R.; Quaglia, A.; Davidson, B.; Fuller, B. New Insights on the Role of Trp Channels in Calcium Signalling and Immunomodulation: Review of Pathways and Implications for Clinical Practice. *Clin. Rev. Allergy Immunol.* **2021**, *60* (2), 271-292.

(4) Shearer, A.; Molinaro, M.; Montazerian, M.; Sly, J. J.; Miola, M.; Baino, F.; Mauro, J. C. The Unexplored Role of Alkali and Alkaline Earth Elements (Alaes) on the Structure, Processing, and Biological Effects of Bioactive Glasses. *Biomater. Sci.* **2024**, *12* (10), 2521-2560.

(5) Kaczynski, D. J. Beryllium and Beryllium Alloys and Composities. In *Kirk-Othmer Encyclopedia of Chemical Technology*, 2011; pp 1-25.

(6) Soontornchaiyakul, W.; Yoshino, S.; Kanazawa, T.; Haruki, R.; Fan, D.; Nozawa, S.; Yamaguchi, Y.; Kudo, A. CH_4 Synthesis from CO_2 and H_2O of an Electron Source over Rh–Ru Cocatalysts Loaded on $NaTaO_3$:Sr Photocatalysts. *J. Am. Chem. Soc.* **2023**.

(7) Curtis, E. M.; Cooper, C.; Harvey, N. C. Cardiovascular Safety of Calcium, Magnesium and Strontium: What Does the Evidence Say? *Aging Clin. Exp. Res.* **2021**, *33* (3), 479-494.

(8) Jayathissa, R.; Burns, C. A. A Study of Structural and Dielectric Properties of Ba^{2+} Doped $CH_3NH_3PbI_3$ Crystals. *SN Appl. Sci.* **2020**, *2* (3), 349.

(9) Yang, G.; Wu, K. Atomically Thick Barium Borate for Entangled Multiphoton Generators and Ultraviolet Laser Sources. *ACS Appl. Opt. Mater.* **2023**, *1* (5), 939-944.

(10) Chandra, R.; Anjan Kumar Prusty, B.; Azeez, P. A. Pre-Mining Baseline Characterization of Soils: Alkali and Alkaline Earth Metals. *Acta Ecologica Sinica* **2011**, *31* (6), 283-290.

(11) Zhao, J.; Li, B.; Wei, X.; Zhang, Y.; Li, T. Slagging Characteristics Caused by Alkali and Alkaline Earth Metals During Municipal Solid Waste and Sewage Sludge Co-Incineration. *Energy* **2020**, *202*, 117773.

(12) Leiva Soto, A.; Culman, S. W.; Herms, C.; Sprunger, C.; Doohan, D. Managing Soil Acidity Vs. Soil Ca:Mg Ratio: What Is More Important for Crop Productivity? *Crop, Forage Turfgrass Manag.* **2023**, *9* (1), e20210.

(13) Naz, M.; Dai, Z.; Hussain, S.; Tariq, M.; Danish, S.; Khan, I. U.; Qi, S.; Du, D. The Soil pH and Heavy Metals Revealed Their Impact on Soil Microbial Community. *J. Environ. Manage.* **2022**, *321*, 115770.

(14) Azzam, M. A.; Rizwan Khan, M.; Moustafa Youssef, H. Drinking Water as a Substantial Source of Toxic Alkali, Alkaline and Heavy Metals: Toxicity and Their Implications on Human Health. *J. King Saud Univ. Sci.* **2023**, *35* (6), 102761.

(15) Shankar, A.; Dubey, A.; Saini, D.; Singh, M.; Prasad, C. P.; Roy, S.; Bharati, S. J.; Rinki, M.; Singh, N.; Seth, T.; et al. Environmental and Occupational Determinants of Lung Cancer. *Transl. Lung Cancer Res.* **2019**, *8* (Suppl 1), S31-s49.

(16) Musso, C. G. Magnesium Metabolism in Health and Disease. *Int Urol Nephrol.* **2009**, *41* (2), 357-362.

(17) McNeill, I. R.; Isoardi, K. Z. Barium Poisoning: An Uncommon Cause of Severe Hypokalemia. *Toxicol. Commun.* **2019**, *3* (1), 88-90.

(18) Williams, A.; Aguilar, M. R.; Pattiya Arachchillage, K. G. G.; Chandra, S.; Rangan, S.; Ghosal Gupta, S.; Artes Vivancos, J. M. Biosensors for Public Health and Environmental Monitoring: The Case for Sustainable Biosensing. *ACS Sustain. Chem. Eng.* **2024**, *12* (28), 10296-10312.

(19) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(20) Zhang, R.; Zuo, X.; Yin, F. Nucleic Acid Framework-Enabled Spatial Organization for Biological Applications. *Chem & Bio Eng.* **2025**, *2* (2), 71-86.

(21) Billet, B.; Chovelon, B.; Fiore, E.; Oukacine, F.; Petrillo, M.-A.; Faure, P.; Ravelet, C.; Peyrin, E. Aptamer Switches Regulated by Post-Transition/Transition Metal Ions. *Angew. Chem., Int. Ed.* **2021**, *60* (22), 12346-12350.

(22) Liu, H.; Gao, Y.; Mathivanan, J.; Armour-Garb, Z.; Shao, Z.; Zhang, Y.; Zhao, X.; Shao, Q.; Zhang, W.; Yang, J.; et al. Crystal Structures and Identification of Novel Cd^{2+}-Specific DNA Aptamer. *Nucleic Acids Res.* **2023**, *51* (9), 4625-4636.

(23) Zhou, W.; Saran, R.; Huang, P.-J. J.; Ding, J.; Liu, J. An Exceptionally Selective DNA Cooperatively Binding Two Ca^{2+} Ions. *ChemBioChem* **2017**, *18* (6), 518-522.

(24) Ebrahimi, N.; Raoof, J. B.; Ojani, R.; Ebrahimi, M. Designing a Novel DNA-Based Electrochemical Biosensor to Determine of Ba^{2+} Ions Both Selectively and Sensitively. *Anal. Biochem.* **2022**, *642*, 114563.

(25) Zhang, L.; Ma, X.; Wang, G.; Liang, X.; Mitomo, H.; Pike, A.; Houlton, A.; Ijiro, K. Non-Origami DNA for Functional Nanostructures: From Structural Control to Advanced Applications. *Nano Today* **2021**, *39*, 101154.

(26) Zhang, T.; Nong, J.; Alzahrani, N.; Wang, Z.; Oh, S. W.; Meier, T.; Yang, D. G.; Ke, Y.; Zhong, Y.; Fu, J. Self-Assembly of DNA–Minocycline Complexes by Metal Ions with Controlled Drug Release. *ACS Appl. Mater. Interfaces.* **2019**, *11* (33), 29512-29521.

(27) Izatt, R. M.; Christensen, J. J.; Rytting, J. H. Sites and Thermodynamic Quantities Associated with Proton and Metal Ion Interaction with Ribonucleic Acid, Deoxyribonucleic Acid, and Their Constituent Bases, Nucleosides, and and Nucleotides. *Chem. Rev.* **1971**, *71* (5), 439-481.

(28) Duguid, J.; Bloomfield, V. A.; Benevides, J.; Thomas, G. J., Jr. Raman Spectroscopy of DNA-Metal Complexes. I. Interactions and Conformational Effects of the Divalent Cations: Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd. *Biophys J* **1993**, *65* (5), 1916-1928.

(29) Kankia, B. I. Interaction of Alkaline–Earth Metal Ions with Calf Thymus DNA. Volume and Compressibility Effects in Diluted Aqueous Solutions. *Biophys. Chem.* **2000**, *84* (3), 227-237.

(30) Kolev, S. K.; Petkov, P. S.; Rangelov, M. A.; Trifonov, D. V.; Milenov, T. I.; Vayssilov, G. N. Interaction of Na^{+}, K^{+}, Mg^{2+} and Ca^{2+} Counter Cations with RNA. *Metallomics* **2018**, *10* (5), 659-678.

(31) Leonarski, F.; D'Ascenzo, L.; Auffinger, P. Mg^{2+} Ions: Do They Bind to Nucleobase Nitrogens? *Nucleic Acids Res.* **2016**, *45* (2), 987-1004.

(32) Bhattacharyya, D. M. A., G.; Basu, S.; . Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, *4* (38).

(33) Erat, M. C.; Sigel, R. K. O. Methods to Detect and Characterize Metal Ion Binding Sites in RNA. In *Structural and Catalytic Roles of Metal Ions in RNA*, Sigel, A., Sigel, H., Sigel, R. K. O., Sigel, A., Sigel, H., Sigel, R. K. O., Sigel, A., Sigel, H., Sigel, R. K. O., Sigel, A., et al. Eds.; Vol. 9; The Royal Society of Chemistry, 2011; p 0.

(34) Korolev, N.; Lyubartsev, A. P.; Rupprecht, A.; Nordenskiöld, L. Competitive Binding of Mg^{2+}, Ca^{2+}, Na^{+}, and K^{+} Ions to DNA in Oriented DNA Fibers: Experimental and Monte Carlo Simulation Results. *Biophys J* **1999**, *77* (5), 2736-2749.

(35) Zhu, W.; Luo, X.; Puah, C. M.; Tan, X.; Shen, J.; Gu, J.; Chen, K.; Jiang, H. The Multiplicity, Strength, and Nature of the Interaction of Nucleobases with Alkaline and Alkaline Earth Metal Cations: A Density Functional Theory Investigation. *J. Phys. Chem. A* **2004**, *108* (18), 4008-4018.

(36) Yu, C.-Y.; Yu, Y.; Gong, L.-D.; Yang, Z.-Z. Mg^{2+}/Ca^{2+} Binding to DNA Bases: A Quantum Chemical Method and ABEEMσπ/MM Fluctuating Charge Model Study. *Theor. Chem. Acc.* **2012**, *131* (3), 1098.

(37) Hashemianzadeh, S. M.; Faraji, S.; Amin, A. H.; Ketabi, S. Theoretical Study of the Interactions between Isolated DNA Bases and Various Groups Ia and Iia Metal Ions by Ab Initio Calculations. *Monatsh. Chem.* **2008**, *139* (2), 89-100.

(38) Marynick, D. S.; Schaefer, H. F., 3rd. Theoretical Studies of Metal-Phosphate Interactions: Interaction of Li^{+}, Na^{+}, K^{+}, Be^{++}, Mg^{++}, and Ca^{++} with $H_2po_4^{-}$ and $(Ch_3o)_2po_2^{-}$: Implications for Nucleic Acid Solvation. *Proc Natl Acad Sci U S A* **1975**, *72* (10), 3794-3798.

(39) Marino, T.; Mazzuca, D.; Russo, N.; Toscano, M.; Grand, A. On the Interaction of Rubidium and Cesium Mono-, Strontium and Barium Bi-Cations with DNA and RNA Bases. A Theoretical Study. *Int. J. Quantum Chem.* **2010**, *110*, 138-147.

(40) Grauffel, C.; Dudev, T.; Lim, C. Metal Affinity/Selectivity of Monophosphate-Containing Signaling/Lipid Molecules. *J. Chem. Theory Comput.* **2021**, *17* (4), 2444-2456.

(41) Walden, K.; Martin, M. E.; LaBee, L.; Provorse Long, M. Hydration and Charge-Transfer Effects of Alkaline Earth Metal Ions Binding to a Carboxylate Anion, Phosphate Anion, and Guanine Nucleobase. *J. Phys. Chem. B* **2021**, *125* (44), 12135-12146.

(42) Boychuk, B. T. A.; Meyer, S. P.; Wetmore, S. D. Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* **2023**, *11*.

(43) Boychuk, B. T. A.; Jeong, R. Y. E.; Wetmore, S. D. Assessment of the Accuracy of DFT-Predicted Li^{+}–Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408.

(44) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, N_2, and H_2O. *Chem. Phys. Lett.* **1998**, *286* (3,4), 243-252.

(45) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106* (23), 9639-9646.

(46) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated Ab Initio Methods? *J. Chem. Theory Comput.* **2011**, *7* (1), 33-43.

(47) Kim, K. S.; Mhin, B. J.; Choi, U. S.; Lee, K. Ab Initio Studies of the Water Dimer Using Large Basis Sets: The Structure and Thermodynamic Energies. *J. Chem. Phys.* **1992**, *97* (9), 6649-6662.

(48) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *J. Chem. Theory Comput.* **2014**, *10* (1), 49-57.

(49) Sherrill, C. D. Chapter 4 - Wavefunction Theory Approaches to Noncovalent Interactions. In *Non-Covalent Interactions in Quantum Chemistry and Physics*, Otero de la Roza, A., DiLabio, G. A. Eds.; Elsevier, 2017; pp 137-168.

(50) García, J. S.; Brémond, É.; Campetella, M.; Ciofini, I.; Adamo, C. Small Basis Set Allowing the Recovery of Dispersion Interactions with Double-Hybrid Functionals. *J. Chem. Theory Comput.* **2019**, *15* (5), 2944-2953.

(51) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.

(52) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456-1465.

(53) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124* (17), 174104.

(54) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147* (3).

(55) Neese, F. The Orca Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73-78.

(56) *Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16 Rev. B.01*; Wallingford, CT, 2016.

(57) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124* (3), 034108.

(58) Schwabe, T.; Grimme, S. Towards Chemical Accuracy for the Thermodynamics of Large Molecules: New Hybrid Density Functionals Including Non-Local Correlation Effects. *Phys. Chem. Chem. Phys.* **2006**, *8* (38), 4398-4401.

(59) Brémond, É.; Sancho-García, J. C.; Pérez-Jiménez, Á. J.; Adamo, C. Communication: Double-Hybrid Functionals from Adiabatic-Connection: The QIDH Model. *J. Chem. Phys.* **2014**, *141* (3), 031101.

(60) Brémond, E.; Adamo, C. Seeking for Parameter-Free Double-Hybrid Functionals: The PBE0-Dh Model. *J. Chem. Phys.* **2011**, *135* (2), 024106.

(61) Kozuch, S.; Martin, J. M. L. Dsd-Pbep86: In Search of the Best Double-Hybrid DFT with Spin-Component Scaled MP2 and Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20104-20107.

(62) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2* (21), 2810-2817.

(63) Peverati, R.; Truhlar, D. G. Screened-Exchange Density Functionals with Broad Accuracy for Chemistry and Solid-State Physics. *Phys. Chem. Chem. Phys.* **2012**, *14* (47), 16187-16191.

(64) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized, Range-Separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144* (21), 214110.

(65) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128* (8), 084106.

(66) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615-6620.

(67) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 9904-9924.

(68) Heyd, J.; Scuseria, G. E. Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd–Scuseria–Ernzerhof Screened Coulomb Hybrid Functional. *J. Chem. Phys.* **2004**, *121* (3), 1187-1192.

(69) Heyd, J.; Scuseria, G. E. Assessment and Validation of a Screened Coulomb Hybrid Density Functional. *J. Chem. Phys.* **2004**, *120* (16), 7274-7280.

(70) Heyd, J.; Peralta, J. E.; Scuseria, G. E.; Martin, R. L. Energy Band Gaps and Lattice Parameters Evaluated with the Heyd-Scuseria-Ernzerhof Screened Hybrid Functional. *J. Chem. Phys.* **2005**, *123* (17), 174101.

(71) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Erratum: "Hybrid Functionals Based on a Screened Coulomb Potential" [J. Chem. Phys. 118, 8207 (2003)]. *J. Chem. Phys.* **2006**, *124* (21), 219906.

(72) Henderson, T. M.; Izmaylov, A. F.; Scalmani, G.; Scuseria, G. E. Can Short-Range Hybrids Describe Long-Range-Dependent Properties? *J. Chem. Phys.* **2009**, *131* (4), 044108.

(73) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. Efficient Evaluation of Short-Range Hartree-Fock Exchange in Large Molecules and Periodic Systems. *J. Chem. Phys.* **2006**, *125* (10), 104103.

(74) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, *125* (22), 224106.

(75) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.

(76) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A Long-Range Correction Scheme for Generalized-Gradient-Approximation Exchange Functionals. *J. Chem. Phys.* **2001**, *115* (8), 3540-3544.

(77) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. Importance of Short-Range Versus Long-Range Hartree-Fock Exchange for the Performance of Hybrid Density Functionals. *J. Chem. Phys.* **2006**, *125* (7), 074106.

(78) Vydrov, O. A.; Scuseria, G. E. Assessment of a Long-Range Corrected Hybrid Functional. *J. Chem. Phys.* **2006**, *125* (23), 234109.

(79) Vydrov, O. A.; Scuseria, G. E.; Perdew, J. P. Tests of Functionals for Systems with Fractional Electron Number. *J. Chem. Phys.* **2007**, *126* (15), 154109.

(80) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1), 51-57.

(81) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120* (1), 215-241.

(82) Zhao, Y.; Truhlar, D. G. Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better

Performance on Average Than B3LYP for Ground States. *J. Phys. Chem. A* **2006**, *110* (49), 13126-13130.

(83) Boese, A. D.; Martin, J. M. L. Development of Density Functionals for Thermochemical Kinetics. *J. Chem. Phys.* **2004**, *121* (8), 3405-3416.

(84) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032-5051.

(85) Zhao, Y.; Truhlar, D. G. Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A.* **2005**, *109* (25), 5656-5667.

(86) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372-1377.

(87) Peverati, R.; Truhlar, D. G. Communication: A Global Hybrid Generalized Gradient Approximation to the Exchange-Correlation Functional That Satisfies the Second-Order Density-Gradient Constraint and Has Broad Applicability in Chemistry. *J. Chem. Phys.* **2011**, *135* (19), 191102.

(88) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648-5652.

(89) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158-6170.

(90) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785-789.

(91) Xu, X.; Goddard, W. A. The X3LYP Extended Density Functional for Accurate Descriptions of Nonbond Interactions, Spin States, and Thermochemical Properties. *PNAS* **2004**, *101* (9), 2673.

(92) Cohen, A. J.; Handy, N. C. Dynamic Correlation. *Mol. Phys.* **2001**, *99* (7), 607-615.

(93) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(94) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119* (23), 12129-12137.

(95) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid Exchange-Correlation Functional Determined from Thermochemical Data and Ab Initio Potentials. *J. Chem. Phys.* **2001**, *115* (20), 9233-9242.

(96) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta--Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(97) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry. *Phys. Rev. Lett.* **2009**, *103* (2), 026403.

(98) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Erratum: Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry [Phys. Rev. Lett. 103, 026403 (2009)]. *Phys. Rev. Lett.* **2011**, *106* (17), 179902.

(99) Peverati, R.; Truhlar, D. G. M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics. *J. Phys. Chem. Lett.* **2012**, *3* (1), 117-124.

(100) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125* (19), 194101.

(101) Peverati, R.; Truhlar, D. G. An Improved and Broadly Accurate Local Approximation to the Exchange–Correlation Density Functional: The MN12-L Functional for Electronic Structure Calculations in Chemistry and Physics. *Phys. Chem. Chem. Phys.* **2012**, *14* (38), 13171-13174.

(102) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional for Kohn–Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. *J. Chem. Theory Comput.* **2016**, *12* (3), 1280-1293.

(103) Adamo, C.; Barone, V. Exchange Functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW Models. *J. Chem. Phys.* **1998**, *108* (2), 664-675.

(104) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098-3100.

(105) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157* (3), 200-206.

(106) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200-1211.

(107) Jiang, W.; DeYonker, N. J.; Determan, J. J.; Wilson, A. K. Toward Accurate Theoretical Thermochemistry of First Row Transition Metal Complexes. *J. Phys. Chem. A.* **2012**, *116* (2), 870-885.

Chapter 5: Assessment of DFT Methods for the Structural Prediction of Transition and Post-Transition Metal–Nucleic Acid Complexes

Preface: The contents of this chapter have been published [Boychuk, B. T. A.; Wetmore, S. D. Assessment of Density Functional Theory Methods for the Structural Prediction of Transition and Post-Transition Metal–Nucleic Acid Complexes. *J. Chem. Theory Comput.* 2023, 19 (15), 5273-5288].

5.1. Introduction

Metal ions play vital roles in nucleic acid chemistry.[1-4] For example, metals provide charge and structural stabilization that aid the formation of higher-order structural motifs with distinct functions such as DNA double or triple helices, G-quadruplexes, loops, and helical junctions.[3,5,6] Beyond structural stability, metal ions can act as catalytic co-factors in nucleic acid systems. Indeed, ribozymes and DNAzymes often use metals to aid phosphodiester bond cleavage or general organic reactions (e.g., Diels-Alder reactions).[4] Metal–nucleic acid interactions have been shown to play important roles in medicine, with perhaps the best known example being cisplatin–DNA cross-links formed as a common cancer treatment strategy.[7] Alternatively, nucleic acid platforms can be designed to detect toxic metals for medicinal and environmental applications due to their cost-effectiveness, green properties, and anticipated high binding affinity and specificity.[8]

Among all metals, transition and post-transition metal ions are of particular interest due to their unique physiochemical properties. For example, transition metals are commonly used in pharmaceuticals, with potential uses in anti-inflammatory, anti-infective, and chemotherapeutic agents.[9,10] In fact, Co, Fe, Cu, Pt, Au, Ag, Ru, Os, and Ir-containing complexes have all shown anticancer properties.[9-11] Nevertheless, excess concentrations of transition metals can generate reactive oxygen species, which can lead to cellular and tissue damage.[12] Indeed, human exposure to high levels of Cu results in lipid peroxidation, DNA damage, and mitochondrial damage, which cause liver disease, central nervous system dysfunction, and death.[12-14] In terms of post-transition metals, Pb^{2+} is one of the most toxic heavy metal pollutants transmitted through the food chain via pesticides, herbicides, plumbing, automobile exhaust, and industrial waste.[15] Pb^{2+} is carcinogenic to

humans and exposure during early development is particularly damaging to the nervous system, impacting cognitive ability and inducing seizures.[15] Thus, in addition to metal-containing pharmaceuticals that interact with nucleic acids, there is a need for advanced nucleic acid platforms that can detect toxic metal contaminants. While nucleic acids have been developed to selectively bind certain transition and post-transition metals (e.g., Cu^{2+}, Pb^{2+}, Cd^{2+}, Mn^{2+}, Hg^{2+}, and Fe^{2+}), there are ongoing design challenges at least in part to ensure specificity for each of these metal and detection systems for other important targets have yet to be generated.[8]

To further our understanding of metal behavior in biology and exploit nucleic acids in medicinal and biotechnological applications, the fundamental structural and chemical roles of metal ions in nucleic acid chemistry must be uncovered. In this light, several experimental and computational studies have investigated the physicochemical features of different metal–nucleic acid complexes, including those involving transition or post-transition metals.[2,4,8,16-45] Computational chemistry in particular can provide valuable structural and chemical information about metal–nucleic acid interactions. For example, previous investigations have explored the structures and energetics of complexes between 3d transition metals in the +1 oxidation state and each coordination position of guanine[33] and adenine.[32,33,40] Additional studies have investigated interactions between Zn^{2+}, Cd^{2+}, or Hg^{2+} and the canonical DNA/RNA nucleobases,[24,34,37] the phosphate moiety,[37] or modified (thioguanine and thiocytosine) base pairs.[36] $Au^{-1/0/+}$ or $Ag^{0/+}$ complexes with the DNA/RNA nucleobase, deoxyguanine monophosphate, and GC/AT base pairs have been explored,[35] as well as Ag^{+} or Au^{+} metal-mediated base pairs.[45] Platinum–DNA interactions have also been investigated with the goal to assess the binding affinities of different

cisplatin drugs with DNA.[41] Aside from transition metals, Pb^{2+} binding to thymine or uracil and their thio analogues has been considered.[38,39] Since data only exist for select metal–nucleic acid interactions, gaps remain in the literature[8,31] and a systematic investigation is required to gain fundamental structural and chemical information for a full range of metal–nucleic acid complexes.

As a requirement for studies of metal–nucleic acid interactions, an accurate computational methodology must be selected. While improved MP2 (correlation consistent Composite Approach, ccCA) and full configuration interaction (Full-CI) approaches have been deemed accurate for thermochemical properties of transition metal complexes,[46,47] such methods are not feasible for biosystems due to high computational costs. As a result, the majority of metal–nucleic acid studies to date employed DFT, specifically B3LYP,[32-34,36,37,40] which has been reported to inaccurately describe transition metal complexes.[48] Although BH&HLYP has been used for Pt^{2+} and *ab initio* methods (MP2) for Zn^{2+}, Cd^{2+}, and Hg^{2+} complexes,[36,37,41] there are also uncertainties regarding the accuracy of these approaches for metal–nucleic acid interactions and limited systems (deoxyguanine monophosphate or GC base pairs with select metals) have been explored. Indeed, previous literature reports varying reliability of DFT methods for structure, energetics, and other properties of transition and post-transition metal complexes.[46,49-58] For example, while mPW2-PLYP is the top performer for the ccCA-TM/11 test set of 193 experimental enthalpies,[51] τ-HCTHhyb is recommended for a subset of 70 complexes that contain specific ligands (e.g., hydrides, halides, or oxides).[52] Additionally, a previous computational investigation of ruthenium-based Grubbs catalysts revealed that DFT functionals result in diverse geometries.[59] Although this literature highlights the importance

of testing DFT method performance for specific metal complexes,[46,53-57] to the best of our knowledge no benchmark study has systematically investigated transition and post-transition metal–nucleic acid systems. Since nucleic acid ligands (purines, pyrimidines, and phosphate moieties) contain chemically diverse binding sites compared to complexes previously explored,[46,53-57] it is important to analyze the ability of DFT to describe metal–nucleic acid complexes. Furthermore, there is limited data available in general for larger metal complexes (30–100 atoms) and open-shell systems,[60] which are particularly relevant to metal–nucleic acid chemistry. Thus, the performance of different DFT functionals for metal–nucleic acid systems must be critically assessed.

To identify reliable DFT methods for the accurate structural prediction of transition and post-transition metal–nucleic acid complexes, the present work initially investigates the ability of 20 DFT functionals to reproduce crystal structure geometries. Initially, the PDB and CSD are surveyed to identify representative complexes between purine or pyrimidine-type ligands and transition or post-transition metal, that vary in the metal, oxidation state, ligand (i.e., nucleobase or phosphate mimic), and coordination geometry (Figures 5.1 and D.1–D.2, Appendix D). Subsequently, computed geometries from all-electron DFT optimizations in the environmental extremes of the gas phase and implicit water, as well as with an effective core potential (ECP), are compared to the corresponding experimental reference structure. The most accurate DFT methods identified herein can be reliably used in future computational investigations of transition and post-transition metal–nucleic acid systems geared towards understanding the roles of metal ions in biology as well as the design of novel nucleic acid–based pharmaceuticals and metal-sensing platforms.

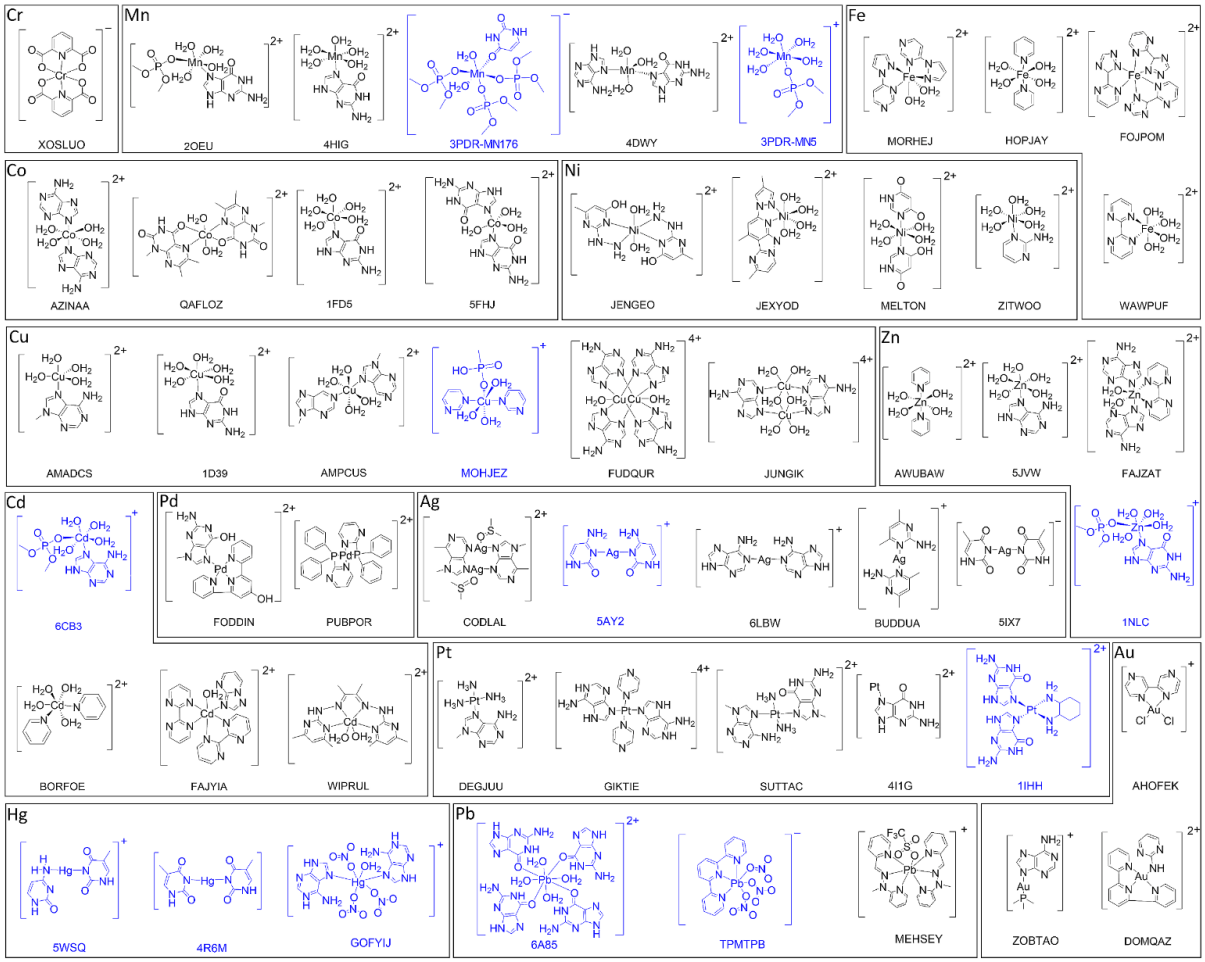


Figure 5.1. Models of the 53 representative transition and post-transition metal–nucleic acid complexes from the PDB and CSD investigated in the present work. The 12 crystal structures highlighted in blue are analyzed separately due to the inability of any functional to reproduce the experimental geometry. Enlarged versions of the models shown in Figure 1 are available in Appendix D (Figures D.1–D.2).

5.2. Computational Methodology

5.2.1. Database Survey

To identify representative transition and post-transition metal–nucleic acid complexes, the PDB was searched for systems containing a nucleic acid and the following metals and oxidation states: $Sc^{+/3+}$, $Ti^{2+/3+/4+}$, $V^{2+/3+/4+/5+}$, $Cr^{2+/3+/4+}$, Mn^{2+}, $Fe^{2+/3+}$, $Co^{2+/3+}$, Ni^{2+}, $Cu^{+/2+}$, Zn^{2+}, Pd^{2+}, $Ag^{+/2+}$, Cd^{2+}, Pt^{2+}, $Au^{+/3+}$, Hg^{2+}, and Pb^{2+}. This selection of metals

covers those previously reported to interact with nucleic acids.[8,16,33,61] All X-ray crystal structures of DNA/RNA and DNA/RNA–protein complexes available in the PDB prior to July 2019 were surveyed. However, only crystal structures with a resolution better than 2.5 Å were used to generate our test set. The electron density map for each crystal structure was visually inspected to confirm the metal location. The metal coordination sphere was carefully analyzed to identify metal–nucleic acid contacts that involve a nucleobase and/or phosphate component. Only structures with coordination distances between the metal and a nitrogen or oxygen atom of the nucleic acid component < 2.6 Å for transition metals or 3.0 Å for post-transition metals were included in the test set.

While PDB structures provide excellent examples of metal–nucleic acid complexes, the quality of the structural data is limited by the resolution, which affects the description of metal–ligand coordination,[62] and there is a lack of suitable high-resolution crystal structures of complexes between nucleic acids and metals in the PDB. Therefore, we expanded our search to the CSD (Version 5.40). Using the substructure feature of the CSD elemental sketcher, purine and pyrimidine rings (i.e., nucleobase subcomponents) were constructed and all nitrogenous and oxygenous metal binding sites were surveyed. The metal–pyrimidine or purine complexes were then visually inspected to confirm at least one metal–nitrogen/oxygen contact with the purine or pyrimidine subcomponent.

Among the representative nucleic acid–metal interactions identified in the PDB and CSD, 53 were selected to form our test set (Figure 5.1), which span all nucleic acid components and many metals. Aside from metal–nucleic acid subcomponent interactions, complexes with the simultaneous coordination of water to satisfy the coordination sphere were selected. Structures containing other ligands (e.g., nitrate or triflate ions) were also

included to increase the size of the test set. The structures extracted from both databases involve the preferred coordination geometries of the metals, including octahedral (i.e., 3d transition metals), trigonal bipyramidal, square-planar (i.e., d^8 metals), linear, and tetrahedral, while no geometrical preference was applied to Pb^{2+} due to its inherently flexible coordination sphere. Additional details of the structures found and those included in our test set are provided in the Results and Discussion, and the general workflow used to design the test set is provided in Appendix D (Figure D.3).

5.2.2. Model Building

Models of the metal–nucleic acid complexes were built based on the 53 crystal structures taken from the PDB and CSD. Only ligands directly coordinated to the metal center were included in the models. In complexes with DNA/RNA strands, nucleobase ligands were generated by truncating at the N-glycosidic bond and adding a hydrogen atom to N1 of the pyrimidines or N9 of the purines using GaussView 6.0, while phosphate ligands were obtained by cleaving the DNA backbone at the neighboring nucleotides such that the ligand maintains the corresponding C5′ and C3′ atoms. Although this approach ensures the models are computationally feasible, the impact of the surrounding crystalline environment and crystal packing effects are ignored. The final models used in the test set are shown in Figures 5.1 and D.1–D.2 in Appendix D, with structures extracted from the same PDB ID differentiated by metal atom ID.

5.2.3. Electronic Structure Calculations

Each model generated from PDB and CSD structures was subjected to geometry optimization using a range of functionals. Specifically, B3LYP,[63,64] BP86,[64,65] BLYP,[64-66] PBE,[67] and TPSS[68] were selected for testing, both in an uncorrected form and a form that includes Grimme's D3[69,70] empirical dispersion correction with Becke–Johnson (BJ) damping.[71] Functionals from the Minnesota family (MN15[72] and MN15-L[73]), and ωB97X,[74] ωB97X-D,[75] and B97D3[70,76] were also considered. These 15 functionals are commonly used to study transition metal complexes,[33,50,72,77-81] and were used in preliminary calculations on the entire 53-model test set. Complexes with geometries that could not be reproduced by any of the 15 functionals were removed from the test set and analyzed separately (12 test set), which reduced the test set from 53 to 41 complexes (Figure D.3, Appendix D). An additional 5 functionals, namely PW6B95,[82] TPSSh,[83] ωB97X-D3(BJ), ωB97X-V,[84] and the composite method PBEh-3c,[85] were also explored and analyzed in detail using the test set composed of the remaining 41 complexes.

The def2-SVP[86] double-zeta Karlsruhe basis set was used due to its coverage of a wide range of the periodic table and its computational efficiency, while remembering the size and number of diverse complexes included in the test set and the goal to identify methods that can be applied in the future to larger nucleic acid systems. Nevertheless, to investigate the impact of basis set size on the predicted geometries, the def2-TZVP triple-zeta Karlsruhe basis set was considered in combination with MN15, one of our recommended functionals, for a subset of 12 complexes that vary in size (2 with 1 nucleic acid ligand, 7 with 2 nucleic acid ligands, and 3 with 3 or more nucleic acid ligands) and metal identity (Table D.1, Appendix D). There are minimal differences

(0.1–1.7% inner-shell MPE) in the inner coordination sphere predicted using double-zeta and triple-zeta basis sets regardless of the complex. In addition, there are small differences in the all-heavy-atom RMSDs (≤ 0.020 Å) for 11 out of 12 complexes, with def2-TZVP affording an improved global geometry for the final complex (BORFOE; Table D.1, Appendix D). Additionally, the effects of diffuse functions were considered on all negatively charged complexes in our 41-system test set (XOSLUO and 5IX7) using MN15 in conjunction with def2-SVPD. Upon comparison to def2-SVP results, there are negligible differences upon inclusion of diffuse functions in the inner coordination sphere (0.2% inner-shell MPE) and the global structure (≤ 0.002 Å all-heavy-atom RMSD) for these complexes (Table D.2, Appendix D). Furthermore, the representative TPMTPB complex was re-optimized with diffuse functions to consider the impact of basis set expansion on a negatively charged system for which no functional could reproduce the experimental structure. However, the experimental structure could still not be reproduced upon inclusion of diffuse functions (Figure D.4, Appendix D). When these minimal BSSE effects on the geometries of the metal–nucleic acid complexes are coupled with the large number of structure and functional combinations considered in the present work, the smaller def2-SVP basis set was confidently used throughout this study.

Both all-electron and ECP calculations (using the Stuttgart-Dresden ECP[87,88]) were performed for each functional. The impact of a relativistic ECP on the predicted metal–nucleic acid geometries was investigated for a subset of 10 complexes containing a heavy metal with $Z > 36$ using one of our recommended DFT methods (MN15) in combination with zeroth-order regular approximation (ZORA), a robust relativistic basis set identified to consistently describe molecular properties[89] that has been applied in QM/MM studies of

metalloenzymes.[90] There are minimal differences in the inner coordination geometries evaluated with a relativistic basis set and scalar-relativistic ECP (≤ 0.8% inner-shell MPE; Table D.3, Appendix D). Additionally, there are negligible differences between the all-heavy-atom RMSDs evaluated with a relativistic basis set and scalar-relativistic ECP for 8 of the 10 complexes (< 0.04 Å), with the remaining two complexes (BORFOE and MEHSEY) having larger deviations (> 0.3 Å) due to ligand rotation. Therefore, the scalar-relativistic ECP was used in the present work to optimize metal–nucleic acid complexes.

For metals with multiple plausible spin states, the lowest spin state was considered, and the unrestricted formulism was used for systems with unpaired electrons. Fractional occupation number weighted density (FOD) analysis[91-93] was used as a qualitative diagnostic tool to explore static electron correlation effects. Within our 41-system test set, 19 complexes were identified to have high multi-reference character and 22 complexes had greater single-reference character (Figures D.5–D.7, Appendix D). Although we noted varying multi-reference effects between different metal complexes, there is no direct correlation between the magnitude of multi-reference character and functional performance across the 41-system test set (Figure D.8; Table D.4, Appendix D). For the 12-system test set that contains challenging complexes that could not be reproduced with any DFT method, the number of fractional occupation differences (NFOD) values range between 0.076 and 1.652, which further underscores that there is no correlation between the ability to optimize nucleic acid–metal complexes and the degree of multi-reference character (Figure D.7, Appendix D). Therefore, we include analysis of all 53 complexes in the present work. Full discussion of the FOD analysis is provided in Appendix D (Figures D.5–D.8; Table D.4, Appendix D).

All complexes were optimized in the environmental extremes of the gas phase and implicit water (ε = 78.3553) using the integral equation formalism polarizable continuum model (IEF-PCM) as implemented in Gaussian 16 and the conductor-like polarizable continuum model (CPCM) in ORCA 5.0.3. All calculations were performed using Gaussian 16 (Rev. B01)[94] and ORCA 5.0.3 (ωB97X-D3(BJ), ωB97X-V, and PBEh-3c),[95] with default grid and convergence criteria used for both programs. All analysis metrics were evaluated using the cpptraj[96] tool in Ambertools18,[97] with the general equations used to calculate each metric provided in the Appendix (Section D.1). Although there are limitations in comparing equilibrium geometries at 0 K with finite-temperature X-ray structures and techniques exist to include anharmonicity effects for DFT geometries (e.g., VPT2 corrections),[98] these are not explored in present work due to the larger number of functional and complex combinations considered.

5.3. Results and Discussion

5.3.1. Building a Structural Test Set of Metal–Nucleic Acid Complexes: Zn^{2+} and Mn^{2+}–Nucleic Acid Interactions Dominate the PDB, While $Cu^{+/2+}$ Interactions are Prominent in the CSD

In total, 2090 X-ray crystal structures were identified in the PDB that contain a transition or post-transition metal along with DNA or RNA, or a DNA or RNA–protein complex. Overall, the most common transition or post-transition metal found in the PDB structures without invoking a limit on the resolution is Zn^{2+} (62%), followed by Mn^{2+} (25%), Cd^{2+} (4%), and $Co^{2+/3+}$ (3%), with the remaining metals found in much lower

percentages (≤ 2%), and no structures were identified for Pd^{2+}, $Sc^{+/3+}$, or $Ti^{2+/3+/4+}$ (Figure D.9, Appendix D). When a 2.5 Å resolution limit was imposed on the PDB structures used to build the test set, the number of nucleic acid systems containing metals decreased by ~2.5 times to yield a total of 824 structures. Among this reduced set (Figure 5.2A), Zn^{2+} (52%) and Mn^{2+} (33%) are still the most common metals, $Co^{2+/3+}$ (5%) is the next most abundant metal, and the remaining metals are present in a similar distribution trend compared to the total set (≤ 2%).

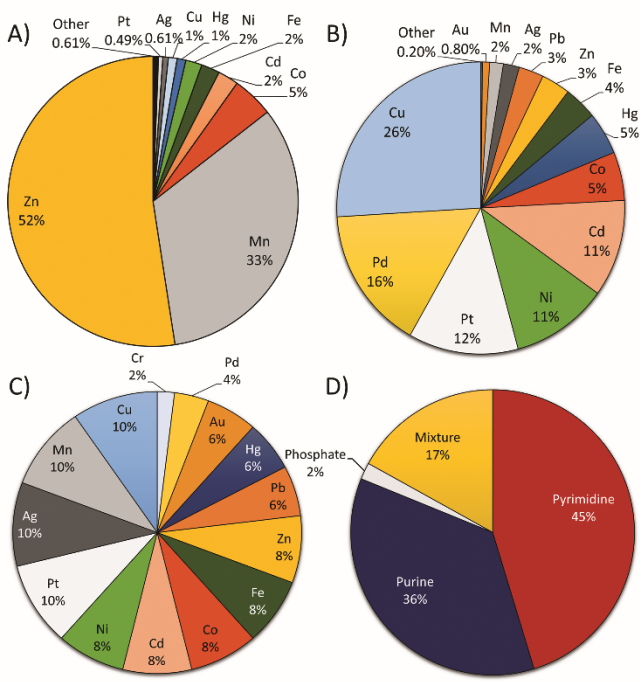


Figure 5.2. A) Metal distribution across all nucleic acid complexes in the PDB with resolution < 2.5 Å (824 structures). 'Other' represents the combined total of $V^{2+/3+/4+/5+}$, $Cr^{2+/3+/4+}$, $Au^{+/3+}$, and Pb^{2+}. B) Metal distribution across all complexes involving purine/pyrimidine mimics in the CSD (1506 structures). 'Other' represents the combined total of $V^{2+/3+/4+/5+}$ and $Cr^{2+/3+/4+}$. Distribution of C) metals and D) nucleic acid ligands across the 53 complexes in our test set. Mixture represents a combination of purine(s), pyrimidine(s), and/or phosphate moieties.

Despite the large number of systems in the PDB that contain a nucleic acid and metal, not all structures contain a metal–nucleic acid interaction. For instance, Zn^{2+} coordinates to the protein rather than the nucleic acid in the crystal structure of the Cas9 endonuclease in complex with the guide RNA and target DNA (5B2O, 1.70 Å resolution). Similarly, Mn^{2+} acts as a spectator ion instead of directly coordinating to the nucleic acid in the crystal structure of NExo, a DNA 3′-phosphatase (6FKE, 2.15 Å resolution). Therefore, visual inspection is necessary to confirm metal coordination, which is supported by previous studies that have challenged reported Mg^{2+} coordination to nucleic acids in PDB structures.[99,100] To ensure a minimum of one metal–nucleobase or phosphate contact is present in the structures included in our test set, the metal coordination spheres in the PDB structures were visually inspected and 19 structures were selected. These structures include metal-mediated base pairs, hydrated metal–nucleobase or phosphate contacts, or simultaneous metal coordination to nucleobase and phosphate components, and cover a range of metals [Mn^{2+} (3PDR-MN5, 3PDR-MN176, 2OEU, 4HIG, 4DWY), Co^{2+} (5FHJ, 1FD5), Cu^{2+} (1D39), Zn^{2+} (5JVW, 1NLC), Ag^{+} (5IX7, 6LBW, 5AY2), Cd^{2+} (6CB3), Hg^{2+} (5WSQ, 4R6M), or Pt^{2+} (4I1G), Figure 5.1]. An additional Pt^{2+} complex (1IHH) is included in the test set, which provides an interesting example of an oxaliplatin compound that forms a DNA intrastrand cross-link. Finally, a G-quadruplex in the presence of Pb^{2+} (6A85) is also incorporated into the test set in which Pb^{2+} adopts a square antiprismatic coordination geometry.

The PDB structures of metal–nucleic acid systems are often limited by resolution and are dominated by Zn^{2+} and Mn^{2+}–nucleic acid complexes, which emphasizes that more X-ray crystal structures of metals interacting with nucleic acids are required. Therefore,

structures available in the CSD were also considered to expand the diversity of the metal–nucleic acid contacts included in our test set. Overall, 1506 structures containing at least one metal contact to a nitrogen/oxygen atom in a representative purine or pyrimidine ring were identified in the CSD. Unlike the PDB, the most common metals found in these complexes are $Cu^{+/2+}$ (26%), Pd^{2+} (16%), Pt^{2+} (12%), Ni^{2+} (11%), and Cd^{2+} (11%), while Zn^{2+} and Mn^{2+} (the most abundant metals in the PDB structures) comprise approximately 3% and 2% of the metal–purine/pyrimidine structures, respectively (Figure 5.2B). Only a handful of complexes (~0.1–5%) were identified for $Co^{2+/3+}$, Hg^{2+}, $Fe^{2+/3+}$, Pb^{2+}, $Ag^{+/2+}$, $Au^{+/3+}$, $Cr^{2+/3+/4+}$, and $V^{2+/3+/4+/5+}$, while no structures are available for $Sc^{+/3+}$ and $Ti^{2+/3+/4+}$. Among the 1506 structures identified in the CSD, there are over 8 times more metal–pyrimidine complexes (1349) than metal–purine (157) complexes. Additionally, the most abundant metals differ between the pyrimidine and purine complexes (Figure D.10, Appendix D). Specifically, the most frequently occurring metal coordinated to a pyrimidine is $Cu^{+/2+}$ (27%), followed by Pd^{2+} (18%), Pt^{2+} (14%), and Cd^{2+} (12%), while no structures exist for $Sc^{+/3+}$ and $Ti^{2+/3+/4+}$ and the remaining metals are present in frequencies below 7%. In contrast, the most frequent metal bound to a purine fragment is Ni^{2+} (52%), followed by $Cu^{+/2+}$ (17%), $Co^{2+/3+}$ (17%), and Hg^{2+} (8%), with Cd^{2+}, Pd^{2+}, Zn^{2+}, Pt^{2+}, $Ag^{+/2+}$, $Fe^{2+/3+}$, and Mn^{2+} falling below 2%, and the remaining metals ($V^{2+/3+/4+/5+}$, $Cr^{2+/3+/4+}$, and $Au^{+/3+}$) having no structures available.

In total, 34 structures were incorporated into our test set from the CSD. Although the occupation of the remaining (non-nucleic acid component) binding sites in the coordination sphere by water was preferred to best mimic a biological environment, structures with alternative ligands, such as azoles ($Fe^{2+/3+}$, Ni^{2+}), NH_3 (Pt^{2+}), nitrate ions

(Pb^{2+}, Hg^{2+}), triflate ions (Pb^{2+}), chloride ions (Au^{3+}), PMe_3 (Au^{+}), Schiff base (Cd^{2+}), DMSO (Ag^{+}), and pyramidal phosphine (Pd^{2+}), were included in the test set due to the lack of other crystal structures for some metals (Figure 5.1). The representative metal–nucleic acid complexes from the CSD include a hydrated metal coordinated to a nitrogenous base (i.e., Fe^{2+} (MORHEJ, WAWPUF, HOPJAY), Co^{2+} (QAFLOZ, AZINAA), Ni^{2+} (ZITWOO, JENGEO, JEXYOD, MELTON), Cu^{2+} (AMPCUS, AMADCS), Zn^{2+} (FAJZAT, AWUBAW), Cd^{2+} (BORFOE, FAJYIA, WIPRUL), Pt^{2+} (DEGJUU), and Hg^{2+} (GOFYIJ)), a hydrated metal simultaneously coordinated to a nitrogenous base and a phosphate component (i.e., Cu^{2+} (MOHJEZ)), an anhydrous metal coordinated to multiple nitrogenous bases (i.e., Cr^{3+} (XOSLUO), Fe^{2+} (FOJPOM), $Au^{2+/3+}$ (DOMQAZ, AHOFEK), Pt^{2+} (GIKTIE, SUTTAC), Pd^{2+} (FODDIN), and Pb^{2+} (MEHSEY, TPMTPB)), an anhydrous metal simultaneously coordinated to nitrogenous base(s) and/or phosphorous components (i.e., Au^{+} (ZOBTAO) and Pd^{2+} (PUBPOR)), or a dinuclear metal complex (i.e., Cu^{2+} (JUNGIK, FUDQUR) and Ag^{+} (CODLAL, BUDDUA)).

Overall, our test set contains models from 53 crystal structures that represent a diverse range of metal–nucleic acid complexes (Figure 5.1). 3–5 PDB and/or CSD structures are included for each metal (6–10%), with the exception of Pd^{2+} and $Cr^{+/3+}$ for which less than 3 structures met our imposed criteria (2–4%; Figure 5.2C). In addition to sampling a variety of metals, our test set considers different nucleic acid components including a pyrimidine (45%), a purine (36%), a phosphate (2%), or a mixture of these nucleic acid components (17%; Figure 5.2D). The greater number of structures from the CSD (34) compared to the PDB (19) aligns with our goal to test the ability of DFT methods

to reproduce accurate experimental data for metal coordination geometries, including metal–ligand distances.

5.3.2. The Tested DFT Functionals Cannot Reproduce the Experimental Structure of Some Representative Metal–Nucleic Acid Complexes

All 53 metal–nucleic acid complexes in our test set were optimized with 15 commonly used functionals and the resulting structures compared to the crystal structure references. Our initial analysis of functional performance is based on the gas-phase all-heavy-atom RMSDs, which describe the global structure (Figure 5.3; Table D.5, Appendix D). One system (5WSQ) is excluded from the analysis since the (Hg^{2+}) metal ion consistently occupies a different coordination site compared to the reference structure (N3 rather than N4 of C; Figure D.11, Appendix D), a problem that has been previously reported for other Hg^{2+} complexes.[101] Dark bands occur in Figure 5.3 for an additional 11 complexes (3PDR-MN5, 3PDR-MN176, MOHJEZ, 1NLC, 5AY2, 6CB3, 1IHH, GOFYIJ, 4R6M, TPMTPB, and 6A85) for all functionals, with similar deviations from the reference structure occurring for each system–functional combination. The corresponding RMSDs are up to 2.9 Å, suggesting the experimental structure for these 11 complexes could not be properly described by gas-phase DFT calculations regardless of the functional considered. To understand the underlying reason for the inability to predict the experimental structures, a series of additional calculations were performed on these 11 systems (Tables D.6–D.10, Appendix D).

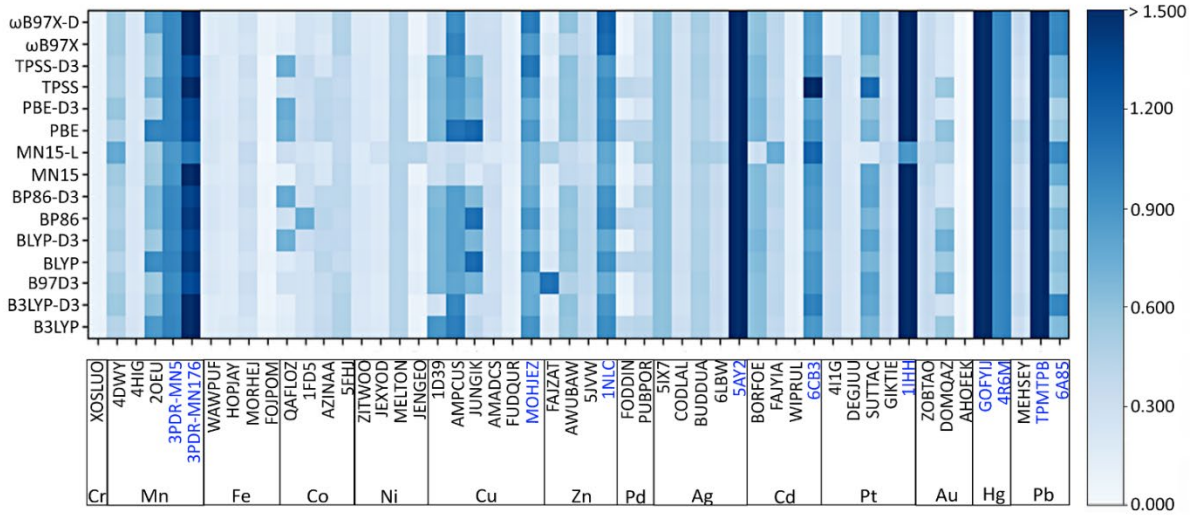


Figure 5.3. All-heavy-atom root-mean-square-deviations (RMSDs, Å) from all-electron gas-phase DFT optimizations relative to crystal structure references for the 52-representative metal–nucleic acid complexes considered in the present work. White indicates no deviation (RMSD = 0.000 Å) from the crystallographic reference, while the shade of blue deepens with increased deviation (dark blue corresponds to RMSD > 1.5 Å). 5WSQ is excluded due to a significant change in the metal coordination site across all functionals. The crystal structures highlighted in blue are analyzed separately.

First, to capture at least a portion of the electrostatic influence of the surrounding environment, the 11 complexes were optimized in implicit water. Although this does not accurately reflect the crystalline environment, consideration of the environmental extremes of the gas phase and water permits analysis of the maximum structural deviation due to electrostatic effects. Overall, implicit water reduces the errors for 4 systems (5AY2, MOHJEZ, 4R6M, and 6CB3; Figures 5.4 and D.12–D.14; Tables D.6–D.7, Appendix D). Specifically, the gas-phase average RMSD (ARMSD) over these 4 complexes is consistently greater than 1 Å for each functional due to reorientation of the ligands (Figures 5.4, D.12–D.14), which can be accompanied by the loss of the inner sphere coordination

(MOHJEZ; Figure D.14, Appendix D). In contrast, there is substantial improvement in the predicted structures in implicit water, with ARMSDs over the 4 complexes for each functional reduced to ~0.5 Å (Table D.6, Appendix D). For example, the large gas-phase all-heavy-atom RMSDs for the Ag^{+} (> 1.7 Å, 5AY2) and Hg^{2+} (> 0.9 Å, 4R6M) complexes for each functional arise due to twisting about the metal center that results in a loss of ligand co-planarity, while all functionals maintain the crystalline ligand co-planarity in water (ARMSD < 0.1 Å for 5AY2 and < 0.5 Å for 4R6M; Figures 5.4 and D.12, Appendix D). Similar twisting was reported in a previous gas-phase DFT (B3LYP-D3 and M06-L) study of mismatched T–Hg–T and U–Hg–U metallo-base pairs.[102]

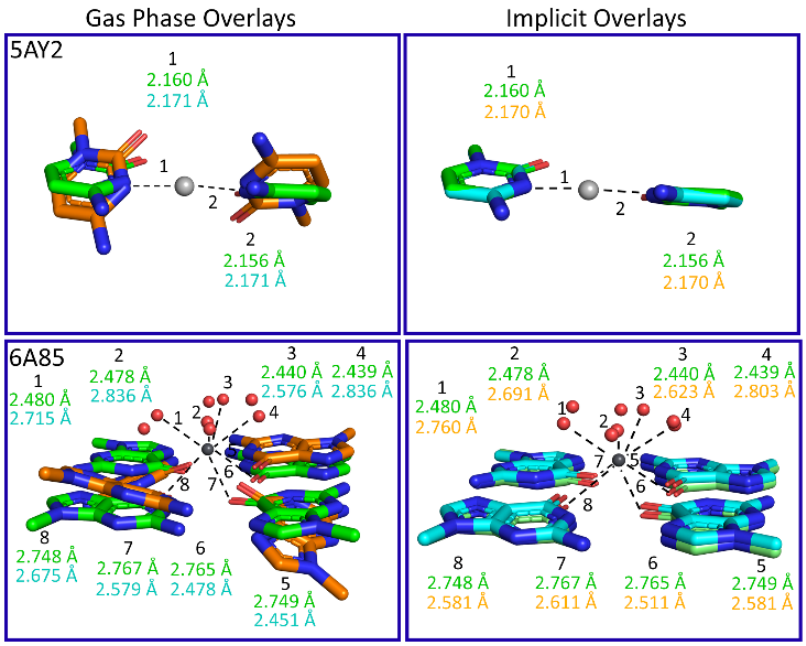


Figure 5.4. Overlays of DFT geometries from all-electron gas-phase (orange) and implicit water (cyan) calculations relative to the reference crystal structure (green) for Ag^{+} (5AY2) and Pb^{2+} (6A85) nucleic acid complexes (constraints were used for the Pb^{2+} complex in both environments), highlighting that implicit water and/or constraints improve the structural description of metal–nucleic acid complexes. Coordination distances are provided in Å.

As the next step to potentially improving agreement between DFT and the experimental structures, the model truncation points were constrained to the crystallographic coordinates and the systems re-optimized in both the gas phase and implicit water. Specifically, nucleobase ligands were extended to include C1′, which was the constraint point, while C5′ and C3′ of phosphate models were fixed to the experimental geometries. This approach improved the overall agreement between the computed and experimental structures for an additional 5 systems (3PDR-MN5, 3PDR-MN176, 1NLC, 6A85, and 1IHH; Table D.8–D.9, Appendix D). Specifically, truncation point constraints reduce the gas-phase ARMSD for 3 complexes for each functional (3PDR-MN176, 1NLC, and 6A85), which is mostly due to reorientation of the bulky ligands about the metal center (Figures 5.4 and D.15–D.16; Table D.8–D9, Appendix D). The most striking example is the octahedral Mn^{2+}–RNA complex (3PDR-MN176) (Figure D.15, Appendix D) for which the RMSDs over all functionals is reduced from ~1.5–2.0 to ~0.4 Å (Table D.9, Appendix D). While the RMSDs for the Pt^{2+} complex (1IHH) are reduced from ~1.0–2.0 Å to ~0.4–0.5 Å for two functionals (MN15 and MN15-L), the errors for the remaining functionals are large (~1.3–1.4 Å). For the Mn^{2+} complex (3PDR-MN5), the RMSDs remain large (~0.7–1.2 Å) with constraints in the gas phase.

Although the implementation of constraints does not universally yield reliable structural data in the gas phase, the experimental structures of all 5 complexes are more accurately reproduced using constraints in a water environment (~0.5 Å ARMSD) than in the gas phase (~0.7 Å ARMSD; Figures 5.4 and D.17–D.18; Tables D.8–D.9, Appendix D). For example, even though the gas-phase DFT structure of Pb^{2+} bound in a G-quadruplex (6A85) is improved upon the implementation of constraints, the experimentally-observed

co-planarity of the purine rings and metal coordination are better maintained by using constraints in implicit water (Figures D.17–D.18, Appendix D). Furthermore, the inclusion of both constraints and implicit water reduces the ARMSDs for complexes for which gas-phase optimizations with constraints failed (1IHH and 3PDR-MN5; Figures D.17–D.18, Appendix D). Thus, constraints address structural features such as reorientation of bulky ligands, while the surrounding (implicit) environment accounts for polarization effects that further finetune the predicted geometry.

All functionals failed to describe the GOFYIJ complex regardless of the environment or constraints (RMSD > 0.8 Å; Table D.10, Appendix D). Indeed, although constraints at the model truncation points aid the correct relative alignment of the nucleic acid ligands, large RMSDs arise for each functional because of a continued shift in the positions of the nitrate ligands, mainly due to the ability of either oxygen to coordinate to the metal (Figure D.19, Appendix D). This suggests that the models for the GOFYIJ complex are unable to capture important environmental effects such as crystal packing.[101] Although unconstrained DFT optimizations of the only other complex in our test set that contains nitrate ligands (TPMTPB) are plagued with the same changes in the coordination sphere as GOFYIJ in both environments (Figures D.19–D.20; Table D.10, Appendix D), the nucleic acid ligand was not truncated in the TPMTPB model and therefore the impact of constraints could not be investigated. In addition to complications arising from the nitrate ligands, the inability of DFT to reproduce these crystal structure geometries may be due to the inherent flexibility of the Hg^{2+} and Pb^{2+} coordination spheres, which are known to adopt irregular geometries.[103–104]

In summary, the 12 complexes that result in large heavy atom RMSDs upon gas-phase optimization regardless of functional (Figure 5.3) can be subcategorized according to: (1) those with improved structural features upon optimization in water, (2) those with improved structural features upon constraining the model truncation points to the crystallographic coordinates, and (3) those that cannot be reproduced regardless of the environment and constraints. The challenges faced reproducing the ligand orientation and/or inner sphere coordination highlight the importance of the surrounding environment in stabilizing the structures of these metal–nucleic acid complexes. Since poor functional performance for these 11 systems along with 5WSQ is likely due to the models rather than the methods, these complexes are removed from further analysis.

5.3.3. Accuracy of All-electron Gas-Phase DFT Predicted Structures of Metal–Nucleic Acid Complexes Varies with Metal and Functional

To further test the ability of DFT to reproduce observed metal–nucleic acid coordination geometries, the all-heavy-atom RMSDs for each functional–metal complex combination obtained from all-electron gas-phase calculations on the remaining 41 representative structures were considered for the same 15 functionals as well as 5 additional methods (Figure D.21; Table D.11, Appendix D). The RMSDs reveal a larger color change horizontally than vertically, indicating a greater dependence of DFT performance on the system studied than the selected functional. Furthermore, the prevalence of lighter bands for the 41 systems demonstrates the broad reliability of the examined functionals across a diverse set of metal–nucleic acid complexes (80% of functional–complex combinations demonstrate ARMSDs < 0.500 Å). Nevertheless, darker bands appear for many functional–

system combinations, highlighting that not all functionals are equally capable of reliably describing the geometries of metal–nucleic acid complexes and a potential dependence of functional performance on metal identity. As a result, the following sections consider the ARMSDs as a function of metal or functional to assess the global quality of the predicted structure. In addition, the inner-shell RMSDs are investigated, which account for various structural features near the metal binding site (e.g., angles between ligands and coordination distances). We also compare the mean percent error calculated over all metal–ligand coordination distances in all complexes (inner-shell MPE), as well as the maximum percent error for a single coordination distance in a given complex (Max PE). Averaging over all complexes for a given metal or over all metals for a given functional allows for the identification of a robust DFT method that can be applied to a broad range of metal–nucleic acid complexes to understand the structural roles metal ions play in biology and various applications.

5.3.4. The Accuracy of Computed Structures Varies with the Metal Identity

To identify variations in functional performance with metal, we consider the all-heavy-atom ARMSD across all functionals and complexes for each metal (Figure 5.5; Table D.12, Appendix D). A clear fluctuation with metal identity is evident in the ability of DFT to predict the experimental structure. The largest errors occur for Cu^{2+} (0.485 Å) and Mn^{2+} (0.484 Å). The significant errors encountered for Cu^{2+} arise due to frequent disruption of the coordination sphere for most functionals, where at least one coordinated water is lost (see, for example, Figure D.22A, Appendix D), which has been reported in the literature for other Cu^{2+} complexes.[37] Nevertheless, the coordination sphere is consistently

maintained by ωB97X-V, ωB97X-D3(BJ), MN15, MN15-L, and PBEh-3c for all Cu^{2+}–nucleic acid complexes (Figure D.22B, Appendix D). In contrast, the large errors for Mn^{2+} complexes arise due to ligand reorientation about the metal center (e.g., relative shift in the phosphate and nucleobase ligands; see, for example, Figure D.23A, Appendix D). Although this ligand reorientation is seen for all functionals, a less pronounced shift occurs for 8 functionals (i.e., MN15, MN15-L, BLYP-D3(BJ), B97D3, TPSS-D3(BJ), ωB97X, ωB97X-D3(BJ), and ωB97X-V). The observed variation in the ARMSD with the metal identity correlates with computational literature highlighting that some metals (i.e., Cu^{2+} and Mn^{2+}) are more difficult to describe using DFT.[44,50,52,105-107] Following Mn^{2+}, Ag^{+} complexes (0.458 Å) result in the next largest ARMSD. However, the structural deviations for Ag^{+} complexes primarily occur due to a change in the ligand co-planarity of the pyrimidine rings in one complex (5IX7; Figure D.24, Appendix D). The smallest ARMSD occurs for Cr^{3+} (0.116 Å) followed by Fe^{2+} (0.194 Å) and Ni^{2+} (0.249 Å), while the ARMSDs for the remaining metals fall between ~0.28 and 0.41 Å.

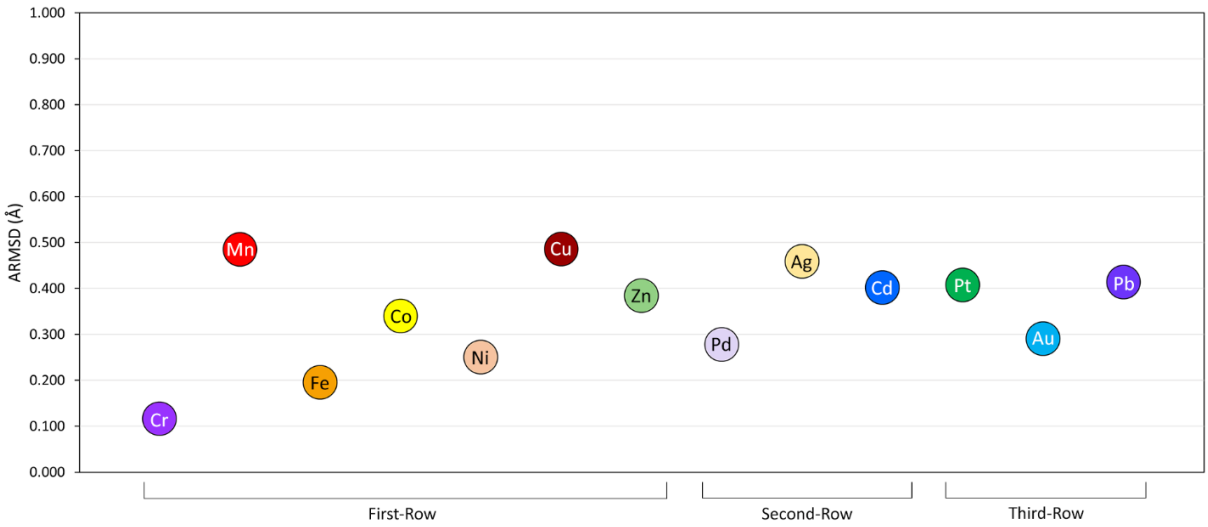


Figure 5.5. All-heavy-atom average root-mean-square-deviations (ARMSDs, Å) across 20 functionals and all complexes relative to crystal structure references for each metal from all-electron gas-phase optimizations.

Although the all-heavy-atom ARMSD accounts for the structural accuracy of the entire metal–nucleic acid complex, the systems considered in our test set contain bulky ligands that are often highly flexible and can readily change orientations in the absence of the surrounding crystalline environment. Therefore, the ability of the functionals to describe the inner coordination sphere was also evaluated using the inner-shell ARMSDs across the 41 complexes (Figure D.25; Table D.12, Appendix D). The inner-shell ARMSDs underscore that the coordination geometry of certain metals is more difficult to describe using DFT. Cu^{2+} complexes (0.498 Å) stand out as the most challenging due to the loss of the inner coordination for most functionals, which is followed by Mn^{2+} complexes (0.336 Å). Among the remaining first-row transition metals, the trends in the description of the global structure (Figure 5.5) and inner coordination sphere (Figure D.25, Appendix D) as a function of metal are consistent, with the exception of Zn^{2+}. In contrast, the trends in

the all-heavy-atom and inner-shell ARMSDs deviate more significantly for the second and third-row metals. Nevertheless, the inner coordination sphere is accurately described for all second and third-row metals (ARMSD ≤ 0.203 Å), emphasizing that errors in the predicted global structures arise from reorientation of bulky ligands rather than the inability of the functional to describe metal coordination. Overall, this analysis highlights that the functional performance can vary with metal identity, but the inner coordination sphere is generally well described for the vast majority of metals over all DFT functionals, with the notable exception of Cu^{2+}–containing complexes.

5.3.5. The Most Reliable Functional for Metal–Nucleic Acid Complexes Can Vary with the Metal

As the all-heavy-atom RMSDs for each complex vary with the DFT method (Figure 5.3) and ARMSDs vary with metal identity (Figure 5.5), it is important to understand the relative accuracy of individual functionals for each metal. This section considers the all-heavy atom and inner-shell ARMSDs for each metal–functional combination (Figures 5.6 and D.26; Tables D.13–D.14, Appendix D). The largest range in the ARMSDs for each functional occurs for Cu^{2+} complexes, with PBE being the worst performer (ARMSD all-heavy-atom = 0.684 Å; inner-shell = 0.778 Å). Nevertheless, MN15 and MN15-L reproduce the crystallographic geometries the most accurately and are recommended for studying Cu^{2+}–nucleic acid complexes (ARMSD all-heavy-atom = 0.265–0.276 Å; inner-shell = 0.223–0.281 Å), closely followed by ωB97X-V, ωB97X-D3(BJ), and PBEh-3c (ARMSD all-heavy-atom = 0.313–0.390 Å; inner-shell = 0.260–0.302 Å). Among the remaining first-row transition metals, ARMSDs for the Co^{2+} complexes

fluctuate the most significantly (ARMSD all-heavy-atom = 0.243–0.473 Å; inner-shell = 0.186–0.414 Å). PBEh-3c most reliably reproduces the structures of Co^{2+}–nucleic acid complexes (ARMSD all-heavy-atom = 0.243 Å; inner-shell = 0.191 Å), closely followed by MN15, ωB97X, ωB97X-D3(BJ), ωB97X-V, and B3LYP-D3(BJ). Although the inner-shell ARMSD are ≤ 0.206 Å for the Zn^{2+} complexes, the all-heavy-atom ARMSDs suggest that B97D3 should be avoided for Zn^{2+} (0.654 Å), while the remaining functionals perform similarly (~0.3–0.4 Å). For Mn^{2+} complexes, the inner-shell ARMSDs fall within a small range (0.320–0.355 Å), while PBE (0.566 Å), BLYP (0.542 Å), and PBEh-3c (0.541 Å) demonstrate the largest all-heavy-atom ARMSDs, and the remaining functionals yield comparable overall structures (ARMSD all-heavy atom = 0.420–0.535 Å). All functionals provide a similar description of complexes involving one of the remaining three first-row transition metals (Fe^{2+}, Ni^{2+}, and Cr^{3+}), with all-heavy-atom and inner-shell ARMSDs for each functional deviating by < 0.041 Å for a given metal (except MN15-L for Ni^{2+}).

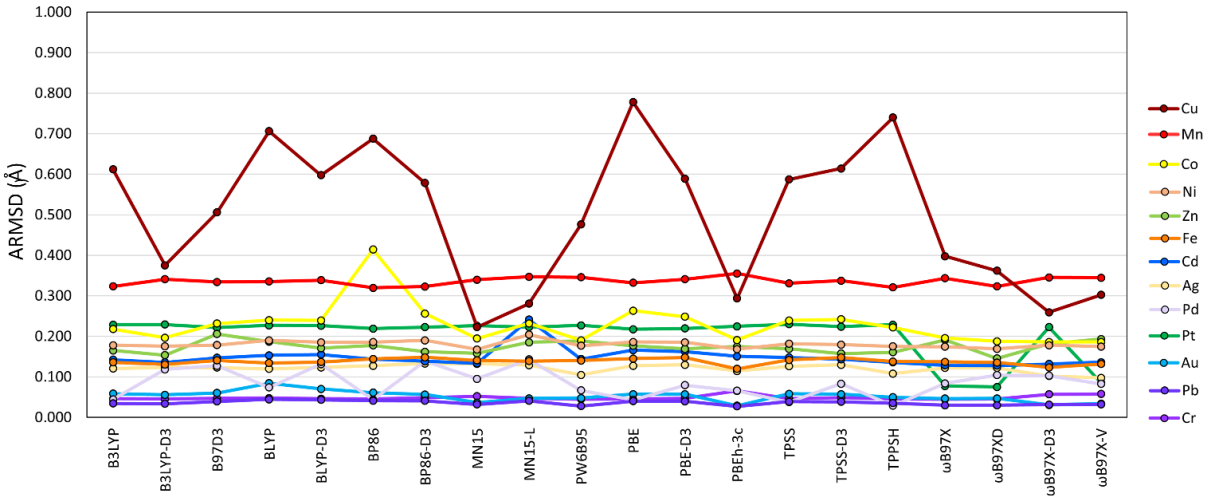


Figure 5.6. Inner-shell average root-mean-square-deviations (ARMSDs, Å) across all complexes relative to crystal structure references for each metal and functional from all-electron gas-phase optimizations.

For the second-row metals, the greatest fluctuation in the geometries across functionals occurs for the Pd^{2+} complexes, although the trend in the all-heavy-atom and inner-shell ARMSDs as a function of the DFT method is not consistent (Figures 5.6 and D.26; Tables D.13–D.14, Appendix D). When both metrics are considered, the functionals that best reproduce the crystallographic geometries of Pd^{2+} complexes include ωB97X-V, ωB97X-D, MN15, PW6B95, PBEh-3c, and ωB97X (ARMSD all-heavy-atom = 0.171–0.200 Å; inner-shell = 0.066–0.104 Å). The remaining second-row transition metals (Cd^{2+} and Ag^{+}) each demonstrate similar all-heavy-atom and inner-shell ARMSDs regardless of functional (within ~0.1 Å for each metal), which indicates that all functionals offer similar descriptions of complexes involving each of these metals. While systems involving third-row metals $Au^{2+/3+}$ or Pb^{2+} consistently result in small inner-shell ARMSDs regardless of functional (~0.03–0.08 Å), a greater variation in the inner-shell ARMSD occurs for Pt^{2+} (0.075–0.230 Å). Additionally, the trend in all-heavy-atom ARMSDs with respect to functional is different for each metal. The greatest variation in the all-heavy-atom ARMSDs for the third-row metals occurs for the Pb^{2+} complexes (0.291–0.773 Å). While ωB97X-D3(BJ) exhibits a large all-heavy-atom ARMSD (0.773 Å) for Pb^{2+} complexes, the inner-shell metal coordination sphere was maintained (inner-shell ARMSD 0.032 Å). The most reliable method for Pb^{2+} complexes is TPSS (0.291 Å). For the $Au^{2+/3+}$ complexes, all functionals yield small all-heavy-atom ARMSDs (≤ 0.386 Å). Finally, MN15-L demonstrates the smallest all-heavy-atom ARMSD for the Pt^{2+} complexes (0.297 Å), with the all-heavy-atom ARMSD for the remaining functionals ranging between 0.358 and 0.515 Å. Overall, although the precise functional for accurate structural description of metal–nucleic acid complexes can vary based on metal identity, a trend in the greater

accuracy of some functionals is starting to emerge, which is more closely explored in the next section.

5.3.6. Multiple DFT Methods are Recommended for Accurate Structural Depiction of Diverse Metal–Nucleic Acid Complexes

To assess functional reliability across nucleic acid complexes containing a broad range of metals and thereby identify a robust method for structure determination, the all-heavy-atom ARMSDs and Max RMSDs for each functional relative to crystallographic data are compared (Figures 5.7A and B; Table D.15, Appendix D). The all-heavy-atom ARMSDs are similar across all functionals (0.310–0.415 Å). However, the all-heavy-atom Max RMSDs vary more substantially, ranging from 0.712 to 1.149 Å and reflecting significant deviations from the experimental structure for some functional–system combinations. In fact, only PBEh-3c (2OEU), ωB97X-V (SUTTAC), MN15 (SUTTAC), ωB97X-D3(BJ) (MEHSEY), and MN15-L (4DWY) have a Max RMSD below 0.8 Å, which arises because of a small change in the ligand orientation about the metal center (e.g., slight rotation of a purine/pyrimidine ring; see, for example, Figure D.27, Appendix D). Nevertheless, ligand reorientation about the metal center can result in large errors for other functionals. For example, the Max RMSD for TPSS arises for the Pt^{2+} complex (SUTTAC) due to significant purine rotation (1.139 Å; Figure D.28, Appendix D). Similarly, the purine ligands in the Zn^{2+} complex (FAJZAT) undergo a large rotation when optimized with B97D3 (Figure D.29, Appendix D), resulting in the Max RMSD of 1.092 Å. Aside from ligand reorientation, functionals such as B3LYP demonstrate a large Max RMSD (1.038 Å) due to expulsion of ligands from the coordination sphere, mainly for Cu^{2+}

complexes as discussed previously. These outcomes highlight the greater reliability of ωB97X-V, ωB97X-D3(BJ), MN15, MN15-L, and PBEh-3c compared to other functionals for the structural description of diverse metal–nucleic acid complexes.

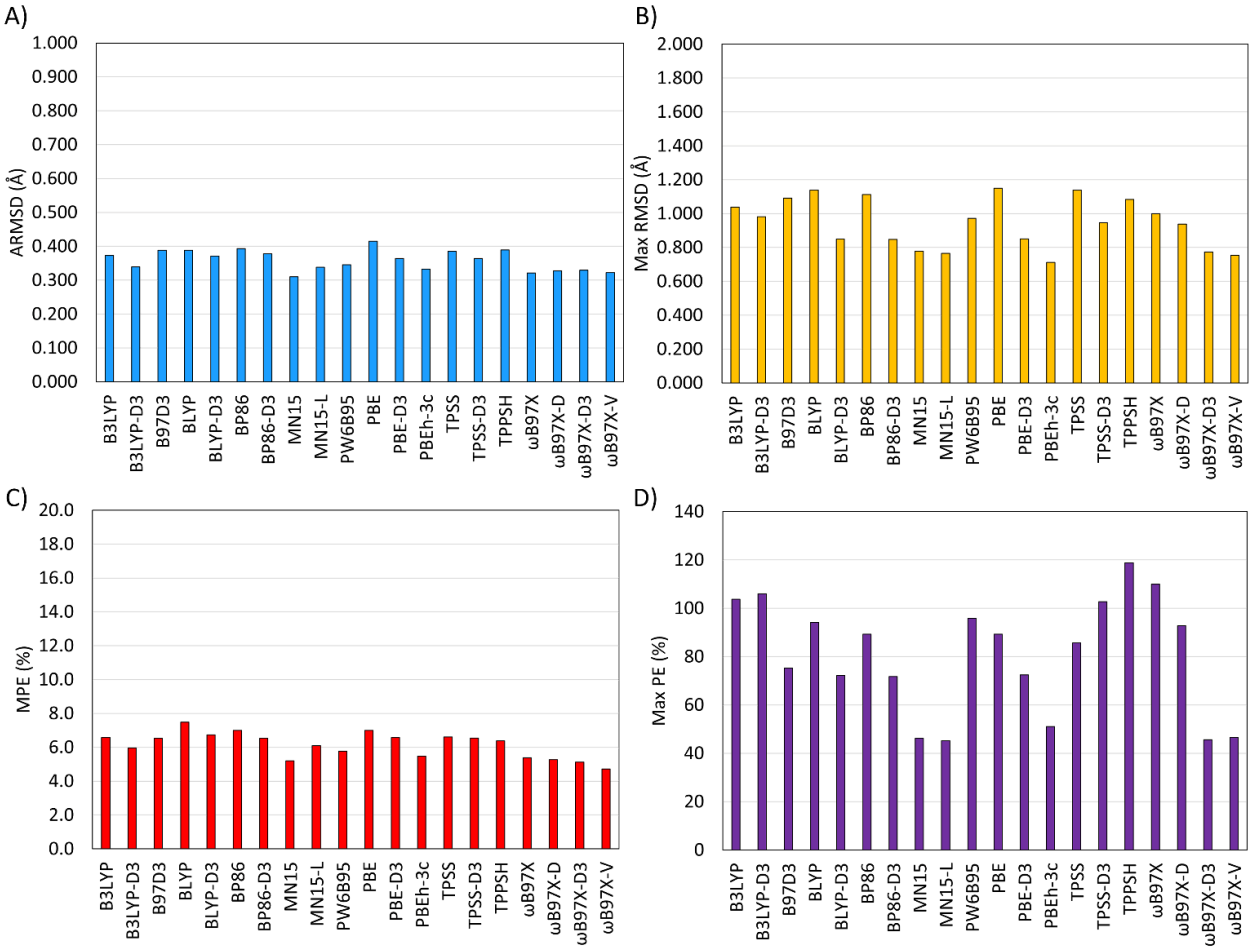


Figure 5.7. All-heavy-atom A) average root-mean-square-deviations (ARMSDs, Å) and B) max RMSDs (Å), and inner-shell C) mean percent errors (MPE, %) and D) max PEs (%) across all complexes relative to crystal structure references for each functional from all-electron gas-phase optimizations.

The relative performance of the 20 DFT functionals is further assessed by comparing the inner-shell MPEs and Max PEs across the 41 systems (Tables D.15–D.17,

Appendix D). In general, the inner-shell MPEs (4.7–7.5%) are similar for all functionals (Figure 5.7C). While this suggests comparable average functional performance, the Max PE in a single coordination distance ranges from ~45 to 120% (Figure 5.7D). Similar to the conclusions from the all-heavy-atom Max ARMSD, all functionals except for ωB97X-V, ωB97X-D3(BJ), MN15, MN15-L, and PBEh-3c yield Max PEs greater than 70%, while these functionals exhibit Max PEs of ~45–51%. These large Max PEs correspond to the inability of functionals to reproduce the inner coordination sphere of Cu^{2+} complexes. Indeed, upon removal of Cu^{2+} complexes, the Max PEs are reduced to ~34–46% for all functionals (Figure D.30A; Table D.18, Appendix D). This correlates with previous literature suggesting that water is expelled from octahedral Cu^{2+} coordinated to N7 of G and 5 water molecules when optimized with B3LYP in the absence of constraints.[37] In a separate study, octahedral $[Cu(H_2O)_6]^{2+}$ complexes similarly expel water from the inner coordination sphere when optimized with PBE, but not other functionals (TPSSh, RevTPSS, B3LYP, and BHLYP).[108] These select examples emphasize that one must cautiously choose the functional when exploring hydrated Cu^{2+} complexes, likely due to increased repulsion in octahedral Cu^{2+} complexes (Figure D.30B; Table D.18, Appendix D).[109] Nevertheless, the Max PEs across only Cu^{2+} complexes are considerably lower for ωB97X-D3(BJ) (28.3%), MN15 (27.2%), and MN15-L (35.0%) than the remaining functionals, with no loss of ligand coordination at the metal center. The Max PE for ωB97X-D3(BJ) (45.7%), MN15 (46.2%), and MN15-L (45.2%) arise for the Pt^{2+}–guanine complex (4I1G) due to a change from monodentate (N7) to bidentate (O6 and N7) coordination in the gas phase (Figure D.31, Appendix D), which occurs for all functionals except ωB97X and ωB97X-D (Max PE equals 11.7% for 4I1G for both functionals), and ωB97X-V (12.2%). Although the Max PEs for ωB97X-V (46.5%) and PBEh-3c (51.0%)

occur for the Cu^{2+} complex (JUNGIK), the resulting change is due to a reorientation of the water ligands about the dinuclear metal centers, with no loss of coordination to either metal center (Figure D.32, Appendix D).

Overall, ωB97X-V, ωB97X-D3(BJ), and MN15 emerge as the top three methods for the optimization of metal–nucleic acid complexes across multiple rows of the periodic table. Indeed, ωB97X-V, ωB97X-D3(BJ), and MN15 are recommended based on the smallest all-heavy-atom ARMSDs (0.310–0.329 Å), Max RMSDs (0.754–0.777 Å), inner-shell MPEs (4.7–5.2%), and Max PEs (45.7–46.5%). It is important to highlight that these top performing methods all include dispersion effects through either parameterization or additional corrections (D3 or VV10). Indeed, previous work has highlighted that dispersion effects can be significant in transition metal complexes.[110] In addition to reliably describing metal–nucleic acid geometries, MN15 has been identified as a top performing functional for 3d transition metal bond dissociation energies involving a range of ligands (i.e., H, F, Cl, Br, O, and S).[111] The PBEh-3c composite method, which is commonly used for QM/MM calculations on biomolecules,[112,113] also accurately describes the structural features of metal–nucleic acid geometries, with only a slightly larger all-heavy-atom ARMSD (0.333 Å), inner-shell MPEs (5.5%), and Max PEs (51.0%) compared to the top three functionals. MN15-L also performs well as a cheaper variation of MN15, with a slightly larger all-heavy-atom ARMSD (0.338 Å) and inner-shell MPE (6.1%) compared to the best three methods. Although ωB97X and ωB97X-D demonstrate low inner-shell MPEs (~5%), these functionals can lead to large Max PEs (109.9% and 92.8%) and Max RMSDs (1.000 Å and 0.938 Å) for Cu^{2+} complexes and therefore must be used with

caution. Nevertheless, both ωB97X and ωB97X-D are deemed reliable for reproducing the geometries of metal–nucleic acid complexes that do not contain Cu^{2+}.

5.3.7. Inclusion of Implicit Water Minimally Refines Metal–Nucleic Acid Geometries

Given that implicit hydration reduces deviations from experimental structures for systems where gas-phase calculations were unsuccessful regardless of functional (discussed above), the effects of implicit water on the DFT predicted geometries are explored for the other 41 complexes in our test set (Figures 5.8 and D.33; Tables D.15 and D.19–D.20, Appendix D). The all-heavy-atom ARMSDs calculated across all complexes for each functional are comparable in the gas phase and implicit water (0.273–0.400 Å; Figure D.33A, Appendix D). The Max RMSDs from gas-phase and implicit water optimizations are also highly similar for most functionals (~0.6–1.3 Å; Figure D.33B, Appendix D). Interestingly, either the gas phase or implicit water environment can more accurately reproduce the global experimental structure, including the relative orientation of the bulky ligands, depending on the functional–system combination considered. The inner-shell MPEs over all complexes suggest that the experimental coordination geometry is equally reproduced in both environmental extremes (Figure D.33C, Appendix D). Indeed, there is generally only a slight change in the coordination geometry upon inclusion of implicit water (< 11%; Figure 5.8), with the deviations being higher (up to ~28%) for only 5 complexes (1D39, AMPCUS, JUNGIK, BORFOE, and 4I1G). Nevertheless, the change in the inner-coordination sphere predicted in implicit solvent in reference to the crystallographic data is minimal, with most functionals demonstrating an on average ~0.5–1.5% improvement upon comparison to the corresponding gas-phase inner-shell MPE

(Figure D.33C, Appendix D). The Max PE for a given distance in the inner coordination sphere decreases in water compared to the gas phase for most functionals (Figure D.33D, Appendix D), with 7 methods (B3LYP, B3LYP-D3(BJ), B97D3, TPSS-D3(BJ), TPSSh, ωB97X, and ωB97X-D) exhibiting ~37–70% reductions. As discussed for the global structure, the inclusion of implicit solvent can improve or deteriorate the structural description of the inner coordination sphere. For example, there is a reduction in errors for many Cu^{2+} complexes, but an overall increase in errors of the Cd^{2+} complexes in implicit water (Figure 5.8). Most importantly, despite the ability of implicit water to improve the description of select metal–nucleic acid complexes by providing charge stabilization, the magnitude of the structural improvement and therefore the benefits of optimizing in implicit water are generally minimal, with the trends in functional performance prevailing regardless of the surrounding environment (Figures 5.9A and D.33, Appendix D).

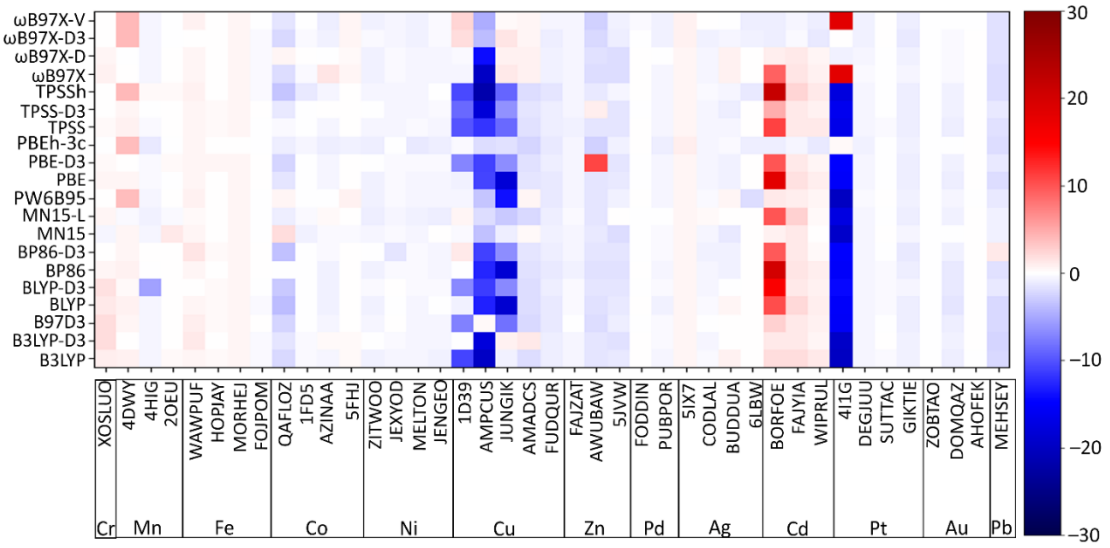


Figure 5.8. Differences in the inner-shell mean percent errors (MPE, %) from all-electron gas-phase and implicit water optimizations relative to crystal structure references. Values represent implicit water minus gas-phase data, with white indicating no change between

the gas-phase and implicit water results, blue being a reduction in errors, red representing an increase in errors upon inclusion of water, and deeper shades of color representing greater deviations.

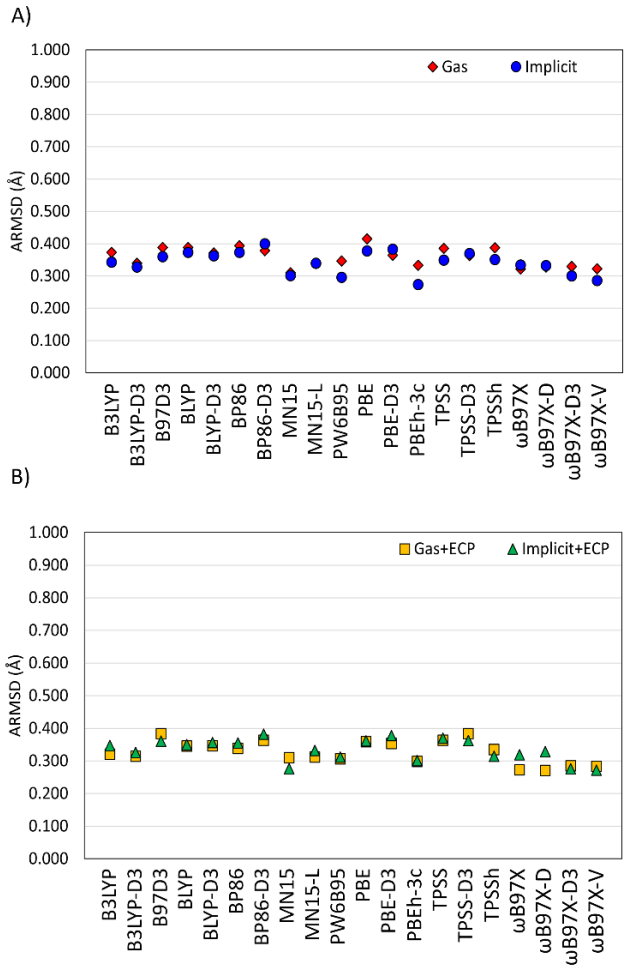


Figure 5.9. All-heavy-atom average root-mean-square-deviations (ARMSDs, Å) from gas-phase and implicit water A) all-electron and B) ECP optimizations relative to crystal structure references.

5.3.8. An ECP Negligibly Affects Metal–Nucleic Acid Geometries Regardless of Environment

While all data discussed to this point have been generated using all-electron calculations, it is important to consider whether the use of an ECP for the metal center

impacts the accuracy of the predicted geometry of metal–nucleic acid complexes, especially when such interactions occur in much larger biosystems that would otherwise require significant computational time. The impact of an ECP on metal–nucleic acid geometries is assessed for the first-row transition metals, which includes 24 complexes (Figures 5.1 and D.1, Appendix D). The gas-phase all-heavy-atom ARMSDs evaluated across all complexes for each functional are similar for all-electron and ECP calculations (within 0.064 Å; Figure D.34A; Table D.20–D.21, Appendix D). However, the effect of an ECP on the Max RMSD is varied. Specifically, the Max RMSD evaluated using an ECP increases by ~0.2–0.3 Å compared to all-electron results for two functionals (MN15 and TPSS-D3(BJ)) and decreases by ~0.2–0.4 Å for six functionals (BLYP, BP86, PBE, PW6B95, ωB97X and ωB97X-D), while the Max RMSD for all-electron and ECP calculations are similar for all remaining functionals (within 0.09 Å; Figure D.34B, Appendix D). Upon consideration of the inner coordination distances, the inner-shell MPE for each functional–system combination calculated with and without an ECP are nearly identical across our test set (Figures 5.10 and D.34C; Tables D.21–D.22, Appendix D). In fact, larger deviations (up to 74% PE) occur for certain functional–system combinations, which are due to the notoriously difficult to describe Cu^{2+} complexes. As discussed for the Max RMSD, implementation of an ECP can significantly decrease (ωB97X, ωB97X-D, PW6B95, and TPSSh) or increase (B97D3 and TPSS) the Max PE corresponding to the all-electron calculations for a given functional. However, the majority of DFT methods yield similar Max PEs with or without the use of an ECP (within ~10%; Figure D.34D, Appendix D). Most importantly, the same trends in all of these important metrics hold true when all-electron and ECP calculations are done in implicit water (Figure D.35; Tables D.21 and D.23, Appendix D). Furthermore, comparison of ECP calculations in the gas phase and

implicit water reaffirm our previous conclusion that the water environment does not significantly impact the predicted structure of metal–nucleic acid complexes (Figure 5.9B). Overall, while structural deviations can arise upon incorporation of an ECP for a few functional–complex combinations regardless of the environment, there are minimal differences compared to the all-electron data and the trends in functional performance are maintained (Figure 5.9).

5.4. Conclusions

In summary, the ability of 20 DFT functionals to reproduce experimental geometries of representative transition and post-transition metal–nucleic acid complexes was investigated using a diverse test set generated from the PDB and CSD. Although the structures of 12 out of the 53 complexes in our test set could not be reproduced using gas-phase optimizations regardless of functional, 8 of these complexes were accurately described using implicit solvent, constrained models, or a combination of both techniques, which underscores the importance of the surrounding environment in stabilizing these structures. The accuracy of DFT gas-phase optimizations across the remaining 41 complexes is metal dependent, with the greatest deviations from experiment occurring for Cu^{2+}-containing complexes. Although DFT optimizations in implicit water can improve the description of the inner coordination sphere for select systems among the 41 complexes in the test set, implicit solvent generally results in a minimal refinement of metal–nucleic acid geometries. Similarly, including an ECP has a negligible impact on the structural description of metal–nucleic acid complexes regardless of the environment. The top three performing functionals across the 41 complexes are ωB97X-V, ωB97X-D3(BJ), and

MN15. We emphasize that these top performing methods all account for dispersion effects, which highlights the importance of accounting for dispersion when using DFT to optimize metal complexes. MN15-L and PBEh-3c are also deemed to be reliable for diverse metal–nucleic acid coordination geometries, while ωB97X and ωB97X-D are suitable for metal–nucleic acid complexes that do not involve Cu^{2+}. Overall, this work serves as the first systematic study to explore DFT performance for the geometries of transition and post-transition metal–nucleic acid complexes. The reliable methods identified herein can be used to develop improved parameters for MD simulations, to perform QM/MM calculations on large metal–nucleic acid systems, and to investigate and design metal–nucleic acid-based pharmaceuticals, nanomaterials, and biosensors.

5.5. References

(1) Algar, J. L.; Findlay, J. A.; Preston, D. Roles of Metal Ions in Foldamers and Other Conformationally Flexible Supramolecular Systems. *ACS Org. Inorg. Au* **2022**, *2* (6), 464-476.

(2) Zhou, W.; Liu, J. Multi-Metal-Dependent Nucleic Acid Enzymes. *Metallomics* **2018**, *10* (1), 30-48.

(3) Müller, J. Functional Metal Ions in Nucleic Acids. *Metallomics* **2010**, *2* (5), 318-327.

(4) Ward, W. L.; Plakos, K.; DeRose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chem. Rev.* **2014**, *114* (8), 4318-4342.

(5) Biver, T. Stabilisation of Non-Canonical Structures of Nucleic Acids by Metal Ions and Small Molecules. *Coord. Chem. Rev.* **2013**, *257* (19), 2765-2783.

(6) Zhao, L.; Ahmed, F.; Zeng, Y.; Xu, W.; Xiong, H. Recent Developments in G-Quadruplex Binding Ligands and Specific Beacons on Smart Fluorescent Sensor for Targeting Metal Ions and Biological Analytes. *ACS Sens.* **2022**, *7* (10), 2833-2856.

(7) Johnstone, T. C.; Suntharalingam, K.; Lippard, S. J. The Next Generation of Platinum Drugs: Targeted Pt(II) Agents, Nanoparticle Delivery, and Pt(Iv) Prodrugs. *Chem. Rev.* **2016**, *116* (5), 3436-3486.

(8) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(9) Ndagi, U.; Mhlongo, N.; Soliman, M. E. Metal Complexes in Cancer Therapy - an Update from Drug Design Perspective. *Drug Des. Devel. Ther.* **2017**, *11*, 599-616.

(10) Lemaître, T. A.; Burgoyne, A. R.; Ooms, M.; Parac-Vogt, T. N.; Cardinaels, T. Inorganic Radiolabeled Nanomaterials in Cancer Therapy: A Review. *ACS Appl. Nano Mater.* **2022**, *5* (7), 8680-8709.

(11) Sun, R. W.; Ma, D. L.; Wong, E. L.; Che, C. M. Some Uses of Transition Metal Complexes as Anti-Cancer and Anti-HIV Agents. *Dalton Trans.* **2007**, (43), 4884-4892.

(12) Tchounwou, P. B.; Yedjou, C. G.; Patlolla, A. K.; Sutton, D. J. Heavy Metals Toxicity and the Environment. In *Molecular, Clinical and Environmental Toxicology: Volume 3: Environmental Toxicology*, Luch, A. Ed.; Vol. 101; Springer Basel, 2012; pp 133-164.

(13) Bian, K.; Chen, F.; Humulock, Z. T.; Tang, Q.; Li, D. Copper Inhibits the Alkb Family DNA Repair Enzymes under Wilson's Disease Condition. *Chem. Res. Toxicol.* **2017**, *30* (10), 1794-1796.

(14) Merker, K.; Hapke, D.; Reckzeh, K.; Schmidt, H.; Lochs, H.; Grune, T. Copper Related Toxic Effects on Cellular Protein Metabolism in Human Astrocytes. *BioFactors* **2005**, *24* (1-4), 255-261.

(15) Kumar, A.; Kumar, A.; M, M. S. C.; Chaturvedi, A. K.; Shabnam, A. A.; Subrahmanyam, G.; Mondal, R.; Gupta, D. K.; Malyan, S. K.; S, S. K.; et al. Lead Toxicity: Health Hazards, Influence on Food Chain, and Sustainable Remediation Approaches. *Int. J. Environ. Res. Public Health* **2020**, *17* (7), 2179.

(16) Sigel, H.; Sigel, R. Metal-Ion Interactions with Nucleic Acids and Their Constituents. In *Comprehensive Inorganic Chemistry II (Second Edition): From Elements to Applications*, 2013; pp 623-660.

(17) Palou-Mir, J.; Barceló-Oliver, M.; Sigel, R. K. O. Lead: Its Effects on Environment and Health. In *The Role of Lead(II) in Nucleic Acids:*, Astrid, S., Helmut, S., Roland, K. O. S. Eds.; Vol. 17; De Gruyter, 2017; pp 403-434.

(18) Nucleic Acid–Metal Ion Interactions. In *Encyclopedia of Inorganic and Bioinorganic Chemistry*, Hud, N. V. Ed.; 2008; pp 1-433.

(19) Pechlaner, M.; Sigel, R. K. O. Characterization of Metal Ion-Nucleic Acid Interactions in Solution. In *Interplay between Metal Ions and Nucleic Acids*, Sigel, A., Sigel, H., Sigel, R. K. O. Eds.; Springer Netherlands, 2012; pp 1-42.

(20) Łęczkowska, A.; Vilar, R. Interaction of Metal Complexes with Nucleic Acids. *Annu. Rep. Sect. A: Inorg. Chem.* **2013**, *109* (0), 299-316.

(21) Kolev, S. K.; Petkov, P. S.; Rangelov, M. A.; Trifonov, D. V.; Milenov, T. I.; Vayssilov, G. N. Interaction of Na^+, K^+, Mg^{2+} and Ca^{2+} Counter Cations with RNA. *Metallomics* **2018**, *10* (5), 659-678.

(22) Liu, H.-K.; Sadler, P. J. Metal Complexes as DNA Intercalators. *Acc. Chem. Res.* **2011**, *44* (5), 349-359.

(23) Cowan, J. A. Transition Metals as Probes of Metal Cofactors in Nucleic Acid Biochemistry. *Comments Inorg. Chem.* **1992**, *13* (5), 293-312.

(24) Aoki, K. Interaction between Transition Metal Ions and Nucleic Acid Bases, Nucleosides, and Nucleotides. *Nihon Kessho Gakkaishi* **1981**, *23* (6), 309-327.

(25) Bagchi, S.; Mandal, D.; Ghosh, D.; Das, A. K. Density Functional Theory Study of Interaction, Bonding and Affinity of Group Iib Transition Metal Cations with Nucleic Acid Bases. *Chem. Phys.* **2012**, *400*, 108-117.

(26) Ennifar, E.; Walter, P.; Dumas, P. A Crystallographic Study of the Binding of 13 Metal Ions to Two Related RNA Duplexes. *Nucleic Acids Res.* **2003**, *31* (10), 2671-2682.

(27) Marzilli, L. G.; Kistenmacher, T. J. Stereoselectivity in the Binding of Transition-Metal Chelate Complexes to Nucleic Acid Constituents: Bonding and Nonbonding Effects. *Acc. Chem. Res.* **1977**, *10* (4), 146-152.

(28) Zhou, J.; Han, H.; Liu, J. Nucleobase, Nucleoside, Nucleotide, and Oligonucleotide Coordinated Metal Ions for Sensing and Biomedicine Applications. *Nano Research* **2022**, *15* (1), 71-84.

(29) Duguid, J.; Bloomfield, V. A.; Benevides, J.; Thomas, G. J., Jr. Raman Spectroscopy of DNA-Metal Complexes. I. Interactions and Conformational Effects of the Divalent Cations: Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd. *Biophys J* **1993**, *65* (5), 1916-1928.

(30) Barone, G.; Terenzi, A.; Lauria, A.; Almerico, A. M.; Leal, J. M.; Busto, N.; García, B. DNA-Binding of Nickel(II), Copper(II) and Zinc(II) Complexes: Structure–Affinity Relationships. *Coord. Chem. Rev.* **2013**, *257* (19), 2848-2862.

(31) Rodgers, M. T.; Armentrout, P. B. Cationic Noncovalent Interactions: Energetics and Periodic Trends. *Chem. Rev.* **2016**, *116* (9), 5642-5687.

(32) Xing, D.; Tan, X.; Chen, X.; Bu, Y. Theoretical Study on the Gas-Phase Acidity of Multiple Sites of Cu^{+}−Adenine and Cu^{2+}−Adenine Complexes. *J. Phys. Chem.* **2008**, *112* (32), 7418-7425.

(33) Noguera, M.; Branchadell, V.; Constantino, E.; Ríos-Font, R.; Sodupe, M.; Rodríguez-Santiago, L. On the Bonding of First-Row Transition Metal Cations to Guanine and Adenine Nucleobases. *J. Phys. Chem. A* **2007**, *111* (39), 9823-9829.

(34) Vazquez, M. V.; Martínez, A. Ca, Cd, Zn, and Their Ions Interacting with Cytosine: A Theoretical Study. *J. Phys. Chem.* **2007**, *111* (39), 9931-9939.

(35) Leal, L. A. E.; Lopez-Acevedo, O. On the Interaction between Gold and Silver Metal Atoms and DNA/RNA Nucleobases – a Comprehensive Computational Study of Ground State Properties. *Nanotechnol. Rev.* **2015**, *4* (2), 173-191.

(36) Sponer, J.; Burda, J. V.; Leszczynksi, J.; Hobza, P. Interactions of Hydrated Iia and Iib Group Metal Cations with Thioguanine-Cytosine DNA Base Pair: Ab Initio and Density Functional Theory Investigation of Polarization Effects, Differences among Cations, and Flexibility of the Cation Hydration Shell. *J. Biomol. Struct. Dyn.* **1999**, *17* (1), 61-77.

(37) Rulíšek, L.; Šponer, J. Outer-Shell and Inner-Shell Coordination of Phosphate Group to Hydrated Metal Ions (Mg^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}) in the Presence and Absence of Nucleobase. The Role of Nonelectrostatic Effects. *J. Phys. Chem. B.* **2003**, *107* (8), 1913-1923.

(38) Salpin, J. Y.; Guillaumont, S.; Tortajada, J.; Lamsabhi, A. M. Gas-Phase Interactions between Lead(II) Ions and Thiouracil Nucleobases: A Combined Experimental and Theoretical Study. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (3), 359-369.

(39) Guillaumont, S.; Tortajada, J.; Salpin, J.-Y.; Lamsabhi, A. M. Experimental and Computational Study of the Gas-Phase Interactions between Lead(II) Ions and Two Pyrimidic Nucleobases: Uracil and Thymine. *Int. J. Mass. Spectrom.* **2005**, *243* (3), 279-293.

(40) Rodgers, M. T.; Armentrout, P. B. Influence of D-Orbital Occupation on the Binding of Metal Ions to Adenine. *J. Am. Chem. Soc.* **2002**, *124* (11), 2678-2691.

(41) Gkionis, K.; Mutter, S. T.; Platts, J. A. QM/Mm Description of Platinum–DNA Interactions: Comparison of Binding and DNA Distortion of Five Drugs. *RSC Adv.* **2013**, *3* (12), 4066-4073.

(42) Kumbhar, S.; Johannsen, S.; Sigel, R. K. O.; Waller, M. P.; Müller, J. A QM/Mm Refinement of an Experimental DNA Structure with Metal-Mediated Base Pairs. *J. Inorg. Biochem.* **2013**, *127*, 203-210.

(43) Gkionis, K.; Kruse, H.; Šponer, J. Derivation of Reliable Geometries in QM Calculations of DNA Structures: Explicit Solvent QM/Mm and Restrained Implicit Solvent QM Optimizations of G-Quadruplexes. *J. Chem. Theory Comput.* **2016**, *12* (4), 2000-2016.

(44) Cheng, L.; Gauss, J.; Ruscic, B.; Armentrout, P. B.; Stanton, J. F. Bond Dissociation Energies for Diatomic Molecules Containing 3d Transition Metals: Benchmark Scalar-Relativistic Coupled-Cluster Calculations for 20 Molecules. *J. Chem. Theory Comput.* **2017**, *13* (3), 1044-1056.

(45) Cao, G.-J. Dinuclear Metal-Mediated Homo Base Pairs with Metallophilic Interactions: Theoretical Studies of $G_2m_2^{2+}$ (M = Cu, Ag, and Au) Ions. *Sci. Rep.* **2017**, *7* (1), 14896.

(46) DeYonker, N. J.; Williams, T. G.; Imel, A. E.; Cundari, T. R.; Wilson, A. K. Accurate Thermochemistry for Transition Metal Complexes from First-Principles Calculations. *Chem. Phys.* **2009**, *131* (2), 024106.

(47) Khedkar, A.; Roemelt, M. Modern Multireference Methods and Their Application in Transition Metal Chemistry. *Phys. Chem. Chem. Phys.* **2021**, *23* (32), 17097-17112.

(48) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme, S.; Checinski, M. P. Comprehensive Thermochemical Benchmark Set of Realistic Closed-Shell Metal Organic Reactions. *J. Chem. Theory Comput.* **2018**, *14* (5), 2596-2608.

(49) Hirao, H. Applications of Computational Chemistry to Selected Problems of Transition-Metal Catalysis in Biological and Nonbiological Systems. In *Transition Metals in Coordination Environments*, Broclawik, E., Borowski, T., Radoń, M. Eds.; Challenges and Advances in Computational Chemistry and Physics, 2019; pp 463-486.

(50) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *J. Chem. Theory Comput.* **2019**, *15* (6), 3610-3622.

(51) Jiang, W.; Laury, M. L.; Powell, M.; Wilson, A. K. Comparative Study of Single and Double Hybrid Density Functionals for the Prediction of 3d Transition Metal Thermochemistry. *J. Chem. Theory Comp.* **2012**, *8* (11), 4102-4111.

(52) Zhang, W.; Truhlar, D. G.; Tang, M. Tests of Exchange-Correlation Functional Approximations against Reliable Experimental Data for Average Bond Energies of 3d Transition Metal Compounds. *J. Chem. Theory Comput.* **2013**, *9* (9), 3965-3977.

(53) Minenkov, Y.; Singstad, Å.; Occhipinti, G.; Jensen, V. R. The Accuracy of DFT-Optimized Geometries of Functional Transition Metal Compounds: A Validation Study of Catalysts for Olefin Metathesis and Other Reactions in the Homogeneous Phase. *Dalton Trans.* **2012**, *41* (18), 5526-5541.

(54) Bühl, M.; Kabrede, H. Geometries of Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comp.* **2006**, *2* (5), 1282-1290.

(55) Waller, M. P.; Braun, H.; Hojdis, N.; Bühl, M. Geometries of Second-Row Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comput.* **2007**, *3* (6), 2234-2242.

(56) Bühl, M.; Reimann, C.; Pantazis, D. A.; Bredow, T.; Neese, F. Geometries of Third-Row Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comp.* **2008**, *4* (9), 1449-1459.

(57) Amin, E. A.; Truhlar, D. G. Zn Coordination Chemistry: Development of Benchmark Suites for Geometries, Dipole Moments, and Bond Dissociation Energies and Their Use to

Test and Validate Density Functionals and Molecular Orbital Theory. *J. Chem. Theory Comput.* **2008**, *4* (1), 75-85.

(58) Tchakoua, T.; Gerrits, N.; Smeets, E. W. F.; Kroes, G. J. SBH17: Benchmark Database of Barrier Heights for Dissociative Chemisorption on Transition Metal Surfaces. *J. Chem. Theory Comput.* **2023**, *19* (1), 245-270.

(59) Minenkov, Y.; Occhipinti, G.; Heyndrickx, W.; Jensen, V. R. The Nature of the Barrier to Phosphane Dissociation from Grubbs Olefin Metathesis Catalysts. *Eur. J. Inorg. Chem.* **2012**, *2012* (9), 1507-1516.

(60) Maurer, L. R.; Bursch, M.; Grimme, S.; Hansen, A. Assessing Density Functional Theory for Chemically Relevant Open-Shell Transition Metal Reactions. *J. Chem. Theory Comput.* **2021**, *17* (10), 6134-6151.

(61) Hannon, M. J.; Reedijk, J. Metal Interactions with Nucleic Acids. *Dalton Trans.* **2015**, *44* (8), 3503-3504.

(62) Zheng, H.; Hou, J.; Zimmerman, M. D.; Wlodawer, A.; Minor, W. The Future of Crystallography in Drug Discovery. *Expert Opin. Drug Disc.* **2014**, *9* (2), 125-137.

(63) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *Chem. Phys.* **1993**, *98* (7), 5648-5652.

(64) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785-789.

(65) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098-3100.

(66) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157* (3), 200-206.

(67) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.

(68) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(69) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.

(70) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456-1465.

(71) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124* (17), 174104.

(72) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032-5051.

(73) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional for Kohn–Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. *J. Chem. Theory Comput.* **2016**, *12* (3), 1280-1293.

(74) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

(75) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615-6620.

(76) Becke, A. Density-Functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *Chem. Phys.* **1997**, *107*, 8554-8560.

(77) Noguera, M.; Bertran, J.; Sodupe, M. A Quantum Chemical Study of Cu^{2+} Interacting with Guanine−Cytosine Base Pair. Electrostatic and Oxidative Effects on Intermolecular Proton-Transfer Processes. *J. Phys. Chem.* **2004**, *108* (2), 333-341.

(78) Zhang, D.; Truhlar, D. G. Spin Splitting Energy of Transition Metals: A New, More Affordable Wave Function Benchmark Method and Its Use to Test Density Functional Theory. *J. Chem. Theory Comput.* **2020**, *16* (7), 4416-4428.

(79) Nandy, A.; Chu, D. B. K.; Harper, D. R.; Duan, C.; Arunachalam, N.; Cytter, Y.; Kulik, H. J. Large-Scale Comparison of 3d and 4d Transition Metal Complexes Illuminates the Reduced Effect of Exchange on Second-Row Spin-State Energetics. *Phys. Chem. Chem. Phys.* **2020**, *22* (34), 19326-19341.

(80) Cramer, C. J.; Truhlar, D. G. Density Functional Theory for Transition Metals and Transition Metal Chemistry. *Phys. Chem. Chem. Phys.* **2009**, *11* (46), 10757-10816.

(81) Iron, M. A.; Janes, T. Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database. *J. Phys. Chem. A.* **2019**, *123* (17), 3761-3781.

(82) Zhao, Y.; Truhlar, D. G. Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A.* **2005**, *109* (25), 5656-5667.

(83) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J Chem Phys* **2003**, *119* (23), 12129-12137.

(84) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16* (21), 9904-9924.

(85) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent Structures and Interactions by Density Functional Theory with Small Atomic Orbital Basis Sets. *J. Chem. Phys.* **2015**, *143* (5), 054107.

(86) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297-3305.

(87) Metz, B.; Stoll, H.; Dolg, M. Small-Core Multiconfiguration-Dirac–Hartree–Fock-Adjusted Pseudopotentials for Post-D Main Group Elements: Application to PbH and PbO. *J. Chem. Phys.* **2000**, *113* (7), 2563-2569.

(88) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-Adjusted Ab Initio Pseudopotentials for the First Row Transition Elements. *Chem. Phys.* **1987**, *86* (2), 866-872.

(89) Rolfes, J. D.; Neese, F.; Pantazis, D. A. All-Electron Scalar Relativistic Basis Sets for the Elements Rb–Xe. *J. Comput. Chem.* **2020**, *41* (20), 1842-1849.

(90) Pang, Y.; Bjornsson, R. Understanding the Electronic Structure Basis for N_2 Binding to Femoco: A Systematic Quantum Mechanics/Molecular Mechanics Investigation. *Inorg. Chem.* **2023**, *62* (14), 5357-5375.

(91) Grimme, S.; Hansen, A. A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects. *Angew. Chem., Int. Edit. Engl.* **2015**, *54* (42), 12308-12313.

(92) Bauer, C. A.; Hansen, A.; Grimme, S. The Fractional Occupation Number Weighted Density as a Versatile Analysis Tool for Molecules with a Complicated Electronic Structure. *Chem. Eur. J.* **2017**, *23* (25), 6150-6164.

(93) Bursch, M.; Mewes, J.-M.; Hansen, A.; Grimme, S. Best-Practice DFT Protocols for Basic Molecular Computational Chemistry**. *Angew. Chem., Int. Edit. Engl.* **2022**, *61* (42), e202205735.

(94) *Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16 Rev. B.01*; Wallingford, CT, 2016.

(95) Neese, F. Software Update: The Orca Program System—Version 5.0. *WIREs Comp. Mol. Sci.* **2022**, *12* (5), e1606.

(96) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084-3095.

(97) D.A. Case, I. Y. B.-S., S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. Amber. **2018**.

(98) Franke, P. R.; Stanton, J. F.; Douberly, G. E. How to VPT2: Accurate and Intuitive Simulations of Ch Stretching Infrared Spectra Using VPT2+K with Large Effective Hamiltonian Resonance Treatments. *J. Phys. Chem. A* **2021**, *125* (6), 1301-1324.

(99) Leonarski, F.; D'Ascenzo, L.; Auffinger, P. Binding of Metals to Purine N7 Nitrogen Atoms and Implications for Nucleic Acids: A Csd Survey. *Inorg. Chim. Acta* **2016**, *452*, 82-89.

(100) Leonarski, F.; D'Ascenzo, L.; Auffinger, P. Nucleobase Carbonyl Groups Are Poor Mg^{2+} Inner-Sphere Binders but Excellent Monovalent Ion Binders-a Critical PDB Survey. *RNA* **2019**, *25* (2), 173-192.

(101) Nehzati, S.; Summers, A. O.; Dolgova, N. V.; Zhu, J.; Sokaras, D.; Kroll, T.; Pickering, I. J.; George, G. N. Hg(II) Binding to Thymine Bases in DNA. *Inorg. Chem.* **2021**, *60* (10), 7442-7452.

(102) Marino, T. DFT Investigation of the Mismatched Base Pairs $(T\text{-}Hg\text{-}T)_3$, $(U\text{-}Hg\text{-}U)_3$, $D(T\text{-}Hg\text{-}T)_2$, and $D(U\text{-}Hg\text{-}U)_2$. *J. Mol. Model.* **2014**, *20* (6), 2303.

(103) Lerner, E. I.; Lippard, S. J. Heavy-Metal Complexes of 2,4,6-Tris(2-Pyrimidyl)-1,3,5-Triazine. Structure of a Dilead Derivative. *Inorg. Chem.* **1977**, *16* (6), 1537-1546.

(104) Zamora, F.; Sabat, M.; Lippert, B. A Bis(9-Methyladeninium) Complex of Hg(II) with a Highly Irregular Coordination Geometry: $[Hg(9\text{-}MeAH\text{-}N7)_2(H_2O)(NO_3)_3]ClO_4$. *Inorgan. Chim. Acta* **1998**, *267* (1), 87-91.

(105) Jiang, W.; DeYonker, N. J.; Determan, J. J.; Wilson, A. K. Toward Accurate Theoretical Thermochemistry of First Row Transition Metal Complexes. *J. Phys. Chem.* **2012**, *116* (2), 870-885.

(106) Varadwaj, P. R.; Varadwaj, A.; Marques, H. M. DFT-B3LYP, NPA-, and QTAIM-Based Study of the Physical Properties of $[M(II)(H_2O)_2(15\text{-}Crown\text{-}5)]$ (M = Mn, Fe, Co, Ni, Cu, Zn) Complexes. *J. Phys. Chem.* **2011**, *115* (22), 5592-5601.

(107) Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2018**, *119* (4), 2453-2523.

(108) Galván-García, E. A.; Agacino-Valdés, E.; Franco-Pérez, M.; Gómez-Balderas, R. $[Cu(H_2O)N]^{2+}$ (N = 1–6) Complexes in Solution Phase: A DFT Hierarchical Study. *Theor. Chem. Acc.* **2017**, *136* (3), 29.

(109) Burda, J. V.; Pavelka, M.; Šimánek, M. Theoretical Model of Copper Cu(I)/Cu(II) Hydration. DFT and Ab Initio Quantum Chemical Study. *J. Mol. Struct.* **2004**, *683* (1), 183-193.

(110) Bursch, M.; Caldeweyher, E.; Hansen, A.; Neugebauer, H.; Ehlert, S.; Grimme, S. Understanding and Quantifying London Dispersion Effects in Organometallic Complexes. *Acc. Chem. Res.* **2019**, *52* (1), 258-266.

(111) Moltved, K. A.; Kepp, K. P. Chemical Bond Energies of 3d Transition Metals Studied by Density Functional Theory. *J. Chem. Theory Comput.* **2018**, *14* (7), 3479-3492.

(112) Zhang, Z.; Vögele, J.; Mráziková, K.; Kruse, H.; Cang, X.; Wöhnert, J.; Krepl, M.; Šponer, J. Phosphorothioate Substitutions in RNA Structure Studied by Molecular Dynamics Simulations, QM/Mm Calculations, and NMR Experiments. *J. Phys. Chem. B.* **2021**, *125* (3), 825-840.

(113) Mráziková, K.; Kruse, H.; Mlýnský, V.; Auffinger, P.; Šponer, J. Multiscale Modeling of Phosphate···π Contacts in RNA U-Turns Exposes Differences between Quantum-Chemical and Amber Force Field Descriptions. *J. Chem. Inf. Model.* **2022**, *62* (23), 6182-6200.

Chapter 6: A DFT Investigation into the Structure and Strength of Hydrated Li^+–Nucleic Acid Interactions

6.1. Introduction

Li^+ has emerged as an advantageous metal for the development of important materials including rechargeable batteries and therapeutics to treat diseases.[1-4] Indeed, lithium-ion batteries demonstrate the highest energy and power density, and longest cycle life to date, and remain the preferred battery technology despite research efforts into alternatives.[5] As low concentrations of Li^+ are beneficial to the central nervous system, particularly exhibiting antimanic properties,[6] lithium therapies have been developed to treat bipolar disorders.[1,2,7] In addition to mood stabilizing effects, Li^+ exhibits anticancer properties, inhibiting tumor growth[2,3] and regulating tumor proliferation, tumor invasion, and metastasis.[3] Li^+ also demonstrates antiviral capabilities against, for example, hepatitis C and adenoviruses.[8]

Despite the importance of Li^+ for materials and medicine, lithium refinery residues produced from mining, battery and ceramic disposal, and fly ash generated by coal power plants contaminate our environment, food sources, and drinking water.[9,10] In fact, the mining and production of lithium has grown exponentially due to the development of lithium–ion batteries for electric vehicles as part of the global clean energy transition from fossil fuels and coal.[11] As a result, Li^+ environmental contamination is on the rise,[10,12] which has harmful effects on biological systems. Accumulation of Li^+ in the body can result in kidney disease, cerebellar dysfunction, and arrhythmia.[2,10] Li^+ can also negatively affect aquatic and plant life, having been shown to disrupt ionic homeostasis in fish and alter plant

metabolism.[10] Since an excess of Li^+ in the human body and environment can have harmful effects, it is important to develop effective technologies for Li^+ detection and extraction.

Nucleic acid metal sensors have proven to be a successful strategy for detecting metal contaminants, with sensors having been developed to target metals such as Pb^{2+}, Zn^{2+}, Na^+, K^+, and Cs^+.[13,14] However, to the best of our knowledge, no nucleic acid sensor has been designed to date to target Li^+. Aside from metal detection, a nucleic acid platform designed to capture Li^+ can be utilized in the energy sector to extract lithium from oil sands,[15] which is of central importance for expanding alternative energy initiatives. To design nucleic acid platforms specific for Li^+, fundamental structural and binding information is required for Li^+–nucleic acid interactions. Aside from Li^+ detection and remediation, investigating Li^+–nucleic acid interactions can also provide key insights into Li^+ behavior in biological systems and contribute to the development of new pharmaceuticals.

Several experimental and computational investigations have been done to uncover the physicochemical properties of Li^+–nucleic acid complexes.[16-32] Specifically, the kinetic method approach based on gas-phase mass spectrometry and threshold collision induced dissociation (TCID) experiments has been used to determine the gas-phase dissociation energies of Li^+ bound to each of the five nucleobases (A, G, C, T, and U; Figure 6.1A)[17] as well as trimethyl (TMP)[28] and triethyl (TEP)[27] phosphate as models of Li^+–phosphate interactions. While these experiments have provided key information regarding the binding affinity of Li^+ to the individual nucleic acid components, no structural information was gained. In direct complement, CCSD(T)/CBS, MP2, and DFT calculations have been carried out on complexes of Li^+ bound to A, C, G, T, U, or a model phosphate moiety

(dimethyl phosphate).[18,24,25,29-31] These studies have characterized the structure and stability of complexes between Li^+ and each binding site in each nucleic acid component (N3 and N7 of the purines, and N4 of C, O2 and O4 of the pyrimidines, and the bridging and non-bridging oxygens of the phosphate moiety, as well as N1 and N6 of A, O6 of G). Both experimental and computational techniques suggest the metal binding affinity decreases as P >> G > C >> A ~ T = U.[17,28,29]

While valuable intrinsic chemical information was gained from these gas-phase experimental and computational studies, solvent can significantly affect metal complexes.[33,34] Furthermore, biological processes occur in, and materials and biosensors are used in, aqueous environments, with previous work highlighting the importance of modelling solvation when studying biomolecules.[35] Therefore, there have been an abundance of experimental[36-39] and computational[38,40,41] studies on the structure and stability of Li^+–water complexes, which confirm the preferred Li^+–water coordination numbers (i.e., 4 and 6) and distances. However, limited work has explored hydrated Li^+–nucleic acid interactions. Previous gas-phase studies have compared experimental and B3LYP calculated infrared multiple photon dissociation (IRMPD) spectra for explicitly hydrated Li^+ complexed with A, T, U or dimethyl phosphate.[22,42,43] A separate study calculated the IR, absorption spectra, and relative electronic energies of Li^+-mediated Watson-Crick-Franklin (WCF) base pairs (G–Li^+–C and A–Li^+–T) in the gas phase and implicit water with models including up to 3 metal-coordinated waters.[44] While molecular dynamics (MD) simulations have been used to investigate Li^+ interactions with nucleic acids,[45-51] these studies focused on either the impact of Li^+ binding on the nucleic acid structure or the distribution of Li^+ binding locations (e.g., major versus minor groove,

nucleobases versus phosphate backbone). Thus, the impact of water on the structure and binding affinity of Li^+–nucleic acid complexes has not been fully explored.

To address gaps in our understanding of the impact of water on Li^+–nucleic acid interactions, the present work systematically investigates how changes in the surrounding environment affect Li^+ binding to nucleic acids using density functional theory (DFT). Initially, the effects of different non-polar and polar environments on the binding of Li^+ to G or the phosphate moiety (P) are explored using a dielectric medium ranging from gas phase to water. Subsequently, the impact of explicit metal hydration in Li^+–G or P complexes is considered using 1 to 5 water molecules and the implicit environmental extremes (i.e., gas phase and water). Based on these results, the preferred coordination numbers of Li^+ (i.e., 4 and 6) are then used to investigate how hydrated Li^+ binds to each canonical nucleobase (A, G, C, T, and U) and WCF base pairs (Figure 6.1B). Overall, our results provide a more complete understanding of how the structure and binding strengths of Li^+–nucleic acid complexes change in the presence of water. This information can be used to aid the understanding of the behavior of Li^+ in biosystems, and to develop technologies such as nucleic acid sensing and extraction platforms, and therapeutics.

6.2. Computational Methodology

All structures were optimized and frequency calculations performed using MN15/aug-cc-pVDZ, while single-point calculations were performed with MN15/aug-cc-pVTZ. This combination was previously identified to reliably describe Li^+–nucleic acid complexes as compared against CCSD(T)/CBS//MP2/aug-cc-pVTZ data.[25,29]

Initial models of Li^+ coordinated to each binding site in G (N2 and N3 or O6 and N7; Figure 6.1) and the phosphate moiety (P; 1 bridging 1 non-bridging (1B1NB), 2 bridging (2B), or 2 non-bridging (2NB) oxygens; Figure 6.1) were adapted from our previous work.[29] As done in previous studies on Li^+–P interactions,[25,28,29,52] dimethyl phosphate was used to represent the phosphate moiety, which introduces methyl caps at the sites where the connecting sugars are positioned in nucleic acids. All calculations were carried out in different environments (gas-phase ($\varepsilon = 1.000$), benzene ($\varepsilon = 2.2706$), diethyl ether ($\varepsilon = 4.2400$), 1-bromopropane ($\varepsilon = 8.0496$),[53] acetone ($\varepsilon = 20.493$), 1,2-ethanediol ($\varepsilon = 40.245$), and water ($\varepsilon = 78.3553$)) using the integral equation formulism–polarizable continuum model (IEF-PCM). To further understand how the surrounding environment affects Li^+–nucleic acid interactions, the impact of explicit water was investigated. Specifically, models of explicitly hydrated Li^+ complexes ($[Li(H_2O)_n]^+$), with n = 1–6 were built. The Li^+ coordination geometries considered were linear (1–2 ligands), trigonal planar (3 ligands), tetrahedral (4 ligands), trigonal bipyramidal (5 ligands), and octahedral (6 ligands). [54] Using the optimized $[Li(H_2O)_n]^+$ complexes, a water molecule was substituted with G (Figure 6.1) or P, bound via the preferred binding site of the nucleic acid component (i.e., O6–N7 of G and 2NB of P). A hydrogen cap was placed at N1 of pyrimidines and N9 of purines to replace the connecting sugars located in nucleic acids in the G models discussed above.

Based on the results for G and P, models were built by placing hydrated Li^+ at each potential binding site of all other nucleobases (A, C, T, or U), while considering the preferred coordination numbers of Li^+ (i.e., 4 and 6). Since binding sites in the sugar edge of the nucleobases are considered, a methyl cap was used at N9 of the purines and N1 of

the pyrimidines to ensure that N9–H or N1–H does not form hydrogen bonds with the Li^+-coordinated waters, which could artificially increase Li^+ binding to nucleic acids. Two solvents were considered (i.e., 1-bromopropane and water), with 1-bromopropane (ε = 8.0496) selected to approximate the electrostatic environment of DNA .[53] Previous experimental work measured the effective dielectric constant of DNA (ε~8) and validated it using MD simulations, yielding more accurate DNA interaction free energies.[53] For this reason, a dielectric constant of ~8 is recommended for computer modeling of systems containing DNA.[53] Base-pair models were built using the optimized hydrated Li^+–nucleobase complexes wherein WCF hydrogen-bonding sites are not occupied by hydrated Li^+. Optimization, frequency, and single-point calculations were carried out in 1-bromopropane and water.

The binding energy (BE) of bare Li^+ interacting to each binding site of G or the phosphate moiety was calculated as:

$$BE_{Li^+\text{-}NA} = E_{Li^+\text{-}NA} - E_{NA} - E_{Li^+} \quad (6.1)$$

where $E_{Li^+\text{-}NA}$ is the energy of the Li^+–nucleic acid complex, E_{NA} is the energy of the G or phosphate component, and E_{Li^+} is the energy of Li^+. For the explicitly hydrated Li^+–nucleic acid models, the BE was calculated between hydrated Li^+ and the nucleic acid component as either an isolated nucleobase or a base pair. Monomer deformation energy was not included in the BEs. There was a notable structural distortion in the G and P monomer when bare Li^+ interacts with G(N2–N3) and P(1B1NB), with the associated monomer deformation energy in the gas phase being 12.8 kcal/mol and 4.1 kcal/mol, respectively. Additionally, the impact of Li^+ binding on the WCF base pair energy was

evaluated by comparing the BE for the canonical base pairs (G:C, A:T, or A:U) to that between a nucleobase and the pairing nucleobase–Li^+ complex (BE_{Li^+B-P}).

All calculations were carried out using Gaussian 16 (rev B.01).[55]

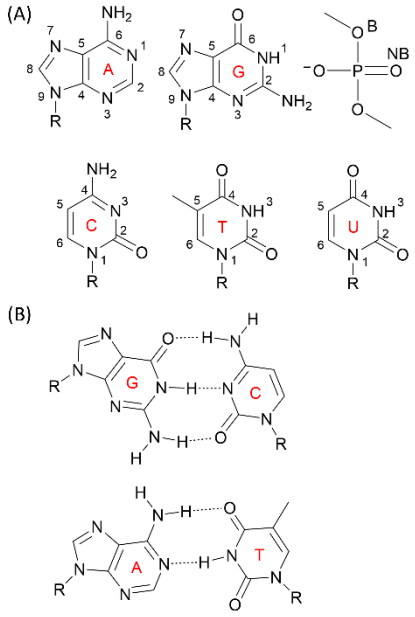


Figure 6.1. (A) Structure and chemical numbering of nucleic acid components considered in this study with R representing the ribose in DNA/RNA, (B = bridging and NB = non-bridging oxygen in the dimethyl phosphate model). (B) The Watson-Crick-Franklin (G:C and A:T) base pairs.

6.3. Results and Discussion

6.3.1. Although Implicit Solvent Can Induce Significant Structural Changes to Li^+–G and Li^+–P Complexes, the General Trends in Binding Strengths as a Function of Binding Site are Consistent in all Environments

To gain initial insight into how different environments affect Li^+ interactions with nucleic acids, the nucleic acid components that result in the strongest Li^+ binding are considered (namely G and P). With a systematic increase in the dielectric constant from the

gas phase to water, coordination distances between Li^+ and each G or P site progressively increase (Figure E.1, Appendix E). For example, the largest change in a G complex occurs for the Li^+–G(N3) coordination distance, which lengthens by 0.248 Å in water compared to the gas phase, while the largest change for P occurs for the Li^+–bridging oxygen distance in the Li^+–P(2B) complex, which increases by 0.285 Å. However, large global structural changes occur in the complexes as the dielectric constant of the implicit environment increases. Specifically, although the Li^+ coordination geometry at the most favorable G(O6–N7) and P(2NB and 2B) sites are maintained regardless of the surrounding medium (Figure E.1, Appendix E), Li^+ coordination to the weaker G(N2–N3) and P(1B1NB) binding sites changes from bidentate to monodentate in diethyl ether, favoring N3 of G or 1NB of P (Figure E.1, Appendix E). To provide adequate metal stabilization, bidentate coordination when the dielectric constant is ≤ 4.24 is coupled with monomer distortion (i.e., rotation of the G amino group ($\angle$N2C2N3) by ~45° or an increase in the $\angle$(OPO) in P(1B1NB) by ~10° for $\varepsilon \leq 4.24$). In contrast, the larger dielectric constants can sufficiently stabilize the metal cation, permitting monodentate coordination and optimal monomer geometries.

Changes in the structure of the Li^+–G and Li^+–P complexes are coupled with weakening of Li^+ binding to G and P (Figure 6.2 and Table E.1, Appendix E). Indeed, there is a sharp decrease in the Li^+–G BEs from the gas phase to 1-bromopropane and a continued decrease to water, resulting in a 12–19 times reduction in water compared to the gas phase. The lack of external stabilization of the metal at low dielectric constants coupled with monomer deformation for the G(N2–N3) complex results in large differences in the BE for the two G sites (26.4 kcal/mol in the gas phase), while environments with larger dielectric

constants results in much smaller differences (3.9 kcal/mol in water). Nevertheless, O6–N7 is the preferred G binding site in all environments. Regardless of binding site, the Li^+–P BEs are reduced by ~14–22 times in water, and the gaps between the BE for different complexes decreasing with increasing dielectric medium, with the Li^+–P(2NB) and Li^+–P(1B1NB) complexes becoming nearly equal stability (within 1.6 kcal/mol) in 1-bromopropane (ε = 8.0496). As a result, when the strongest G and P binding sites are considered, Li^+ binds to P twice as strong as G in the gas phase and water. Nevertheless, the gas phase trend[17,28,29] in the Li^+ binding affinities for G and P sites remains in water, namely P(2NB) > P(1B1NB) > P(2B) >> G(O6–N7) > G(N2–N3). Thus, despite implicit environments influencing the structure and binding strengths for a given Li^+–nucleic acid complex, the general trends in the preference for G and P binding sites remains in all implicit environments.

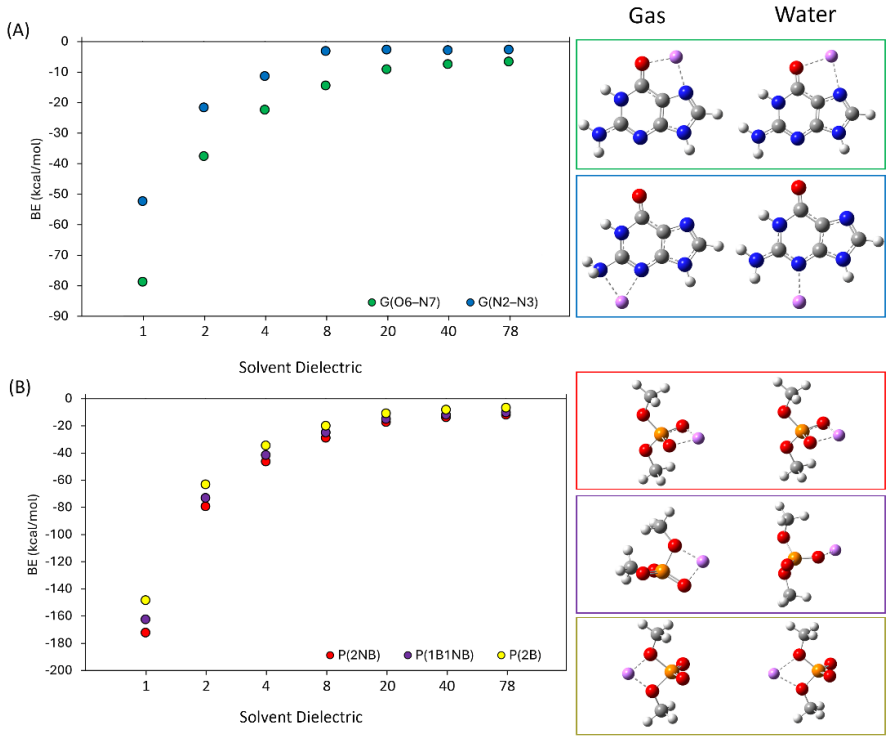


Figure 6.2. BEs (kcal/mol) for Li^+ binding to (A) G or (B) P in different dielectric mediums.

6.3.2. Explicitly Hydrated Lithium Forms Direct and Hydrogen-Bond-Mediated Contacts with Nucleic Acid Components, Impacting Both the Structures and Binding Strengths of Li^+–Nucleic Acid Complexes

Since implicit solvent influences both the structure and strength of Li^+ interactions with G and P, the impact of explicitly coordinated solvent molecules must be considered. Therefore, 0 to 5 explicit water molecules (n) were added around the metal in the gas-phase Li^+–G(O6–N7) and Li^+–P(2NB) complexes, which exhibited the strongest metal binding affinity for each nucleic acid component (Figures 6.3–6.4 and E.2–E.3, Appendix E). As observed for implicit solvent, increasing the number of water molecules directly

coordinated to Li^+ increases the metal–nucleic acid coordination distances. Furthermore, changes in the metal–nucleic acid coordination geometry occur upon increasing the number of explicit water molecules. Specifically, for n = 1 or 2, bidentate coordination between Li^+ and O6–N7 of G observed in the gas phase is maintained; however, for n = 3 – 5, monodentate Li^+ coordination occurs to either O6 or N7 of G, with hydrogen bonds formed between select metal coordinated waters and the unoccupied G site (i.e., O6 or N7). Monodentate coordination to the O6–N7 site of G was not observed in implicit water, underscoring the importance of including explicit solvent molecules. Similarly, although bidentate coordination exists between Li^+ and the 2NB oxygens for n = 0 or 1, with an additional hydrogen bond between the explicit water and a non-bridging oxygen as reported previously,[56] model expansion to n = 2 – 5 results in monodentate Li^+ coordination to a single NB oxygen, with multiple hydrogen-bond-mediated contacts between the explicit water and the other NB oxygen.

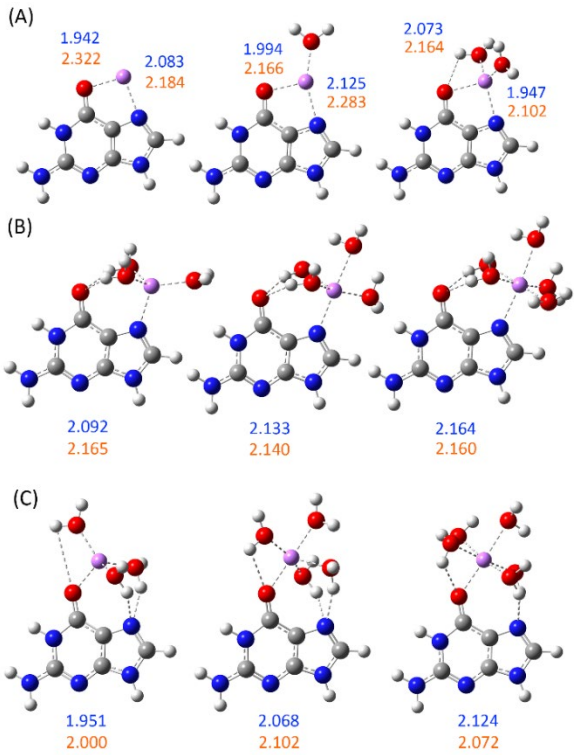


Figure 6.3. Key coordination distances (Å) in explicitly hydrated Li^+–G complexes in the gas phase (blue) and implicit water (orange) involving (A) 0 to 2 waters with bidentate Li^+–G(O6–N7) coordination, (B) 3 to 5 waters with monodentate Li^+–G(N7) coordination, and (C) 3 to 5 waters with monodentate Li^+–G(O6) coordination. Full structural information on hydrated Li^+–G complexes are provided in Appendix E (Figure E.2).

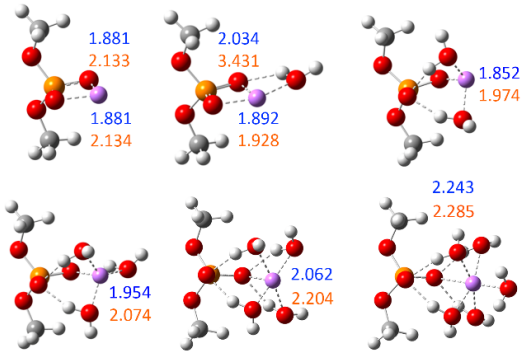


Figure 6.4. Key coordination distances (Å) in explicitly hydrated Li^+–P complexes in the gas phase (blue) and implicit water (orange) involving 0 to 5 waters with a transition from bidentate (Li^+–P(2NB)) to monodentate (Li^+–P(1NB)) coordination. Full structural information on hydrated Li^+–P complexes are provided in Appendix E (Figure E.3).

In general, inclusion of an increased number of explicit water molecules in the Li^+ coordination sphere decreases the gas-phase BE of Li^+ to the nucleic acid component (Figure 6.5 and Table E.2, Appendix E). For Li^+–G complexes, the BE decreases most significantly when the first (by 12.4 kcal/mol for n = 0 to 1) and then the second (by 11.3 kcal/mol for n = 1 to 2) explicit water is added. In contrast, there are much smaller changes in the BE as n increases from 2 to 3 – 5 (~1–4 kcal/mol), regardless of Li^+ binding site (i.e., O6 or N7). For n = 3 – 5, Li^+ shows a slight (~1–2 kcal/mol) preference for the N7 over the O6 site of G. For P, including a single water decreases the BE (8.1 kcal/mol), while larger decreases are observed moving from n = 1 to 2 (4.1 kcal/mol) and n = 2 to 3 (6.0 kcal/mol). The BE for the Li^+–P complex with n = 3 to 4 is increased by 6.0 kcal/mol, while n = 4 to 5 reduces the BE by 2.8 kcal/mol, which is attributed to the fifth water causing minor disruption in the Li^+–P interaction. Thus, explicit solvent can significantly weaken the gas-phase Li^+–nucleic acid binding energy by stabilizing Li^+ and thereby reducing the energetic incentive to bind to nucleobase donor sites.

Overall, as the number of water molecules increases in Li^+–G and Li^+–P complexes, the coordination geometry between Li^+ and G/P is altered, more hydrogen-bond-mediated contacts between the coordinated water and the nucleic acid component occur, and the gas-phase BE decreases. This highlights the importance of including explicit water for an accurate description of Li^+–nucleic acid complexes. Nevertheless, Li^+ binding is energetically preferred to P over G by ~2–3.3 times in the gas phase and ~2–4 times in water.

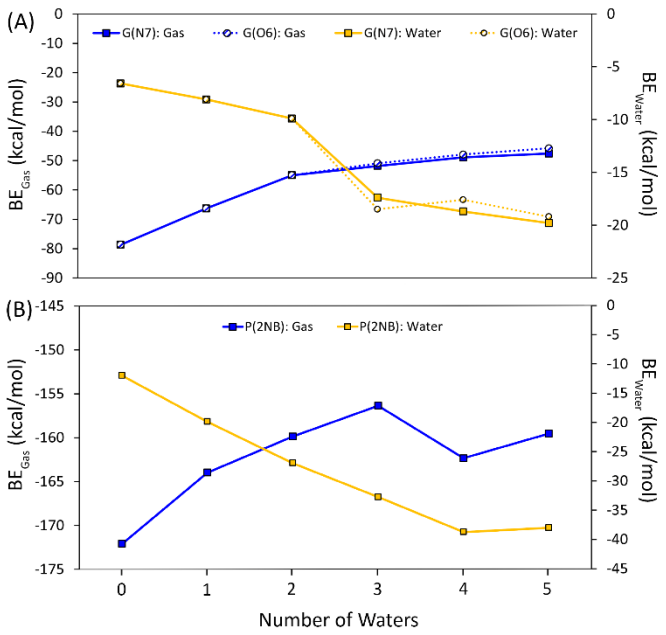


Figure 6.5. BEs (kcal/mol) for Li^+ binding to G(N7) (A), G(O6) (B), or P(2NB) (C) in the gas phase (blue) and implicit water (orange) involving 0 to 5 waters.

6.3.3. While Implicit Solvent Minimally Impacts the Structure of Explicitly Hydrated Li^+–Nucleic Acid Complexes, the Broader Environment Significantly Affects the Thermodynamic Properties

As discussed in the previous sections, the magnitude of the binding strength of bare Li^+ interacting with G or P was significantly affected by increasing the polarity of the dielectric medium and inclusion of explicit water coordinated to the metal. Therefore, it is important to assess how simultaneous incorporation of both explicit and implicit solvent affect the structure and strength of Li^+–G/P interactions. For the $[Li(H_2O)_n(G)]^+$ complexes, the global geometry of the complex (Figures 6.3 and E.2, Appendix E), including the coordination geometry of the metal is preserved upon optimization in water. Nevertheless, all Li^+ coordination distances are lengthened, with the effect being smaller as a greater

number of explicit waters are included in the model. As a result, hydrogen-bond-mediated metal–nucleic acid interactions are frequently strengthened as reflected in shorter hydrogen-bonding distances. For example, in the G(O6) complex with 5 waters, there is a 0.1 Angstrom decrease in the hydrogen-bond length between a water and N7. Although the metal coordination distances increase with the inclusion of the implicit hydration for Li^+–P complexes, with the effects diminishing in magnitude with an increase in the number of water molecules (Figures 6.4 and E.3, Appendix E), the implicit water environment results in an increase in hydrogen-bonding distances between the metal coordinated waters and the nucleic acid, reflecting increased stabilization of the charged P moiety. Furthermore, in the presence of one explicit water molecule, metal coordination changes from bidentate in the gas phase to monodentate in water. Nevertheless, the monodentate coordination for n = 2 – 5 observed for the gas phase is preserved in implicit water.

Unlike in the gas phase, the binding energy of explicit hydrated Li to G decreases with an increase in the number of water molecules in implicit water environment (Figure 6.5 and Table E.2, Appendix E). This arises at least in part because the additional water molecules shield the Li^+ from the effects of the implicit solvent and in part because of the strengthened hydrogen bonding between G and the metal coordinated waters. The largest decrease upon adding water occurs between n = 2 and 3 for both G binding sites (~7.5–8.6 kcal/mol), where the lithium changes from monodentate to bidentate coordination. Once the preferred coordination numbers of Li^+ are achieved (n = 3 and 5), there is near convergence in the binding energies, resulting in a difference between n = 3 and 5 of 0.7–2.4 kcal/mol. In the case of P complexes, the binding strength is reduced by a factor of approximately 2.5 because of the change from bidentate to monodentate Li^+–P coordination

for n = 0 to 1. Subsequently, the inclusion of additional water molecules results in a slight decrease in the binding energy as the number of explicit waters increases as discussed for G complexes (by ~6–7 kcal/mol for each additional water), with a plateau formed between n = 4 and 5 (within 0.7 kcal/mol). Overall, this data highlights that both implicit and explicit solvent molecules can impact the structure and strength of Li^+–nucleic acid molecules, with the dielectric medium providing global screening of the complex and the explicit solvent molecules offering stabilization through formation of discrete hydrogen-bonding interactions.

6.3.4. Solvent Environment Significantly Affects the Structure and Binding Strengths of Li^+–Nucleobase Interactions

To understand how Li^+ interacts with nucleic acids in a hydrated environment, models consisting of hydrated Li^+ interacting with each nucleobase (A, C, G, T, or U) binding site are considered. Due to a significant decrease in the BEs of Li^+–G complexes in implicit and explicit solvent compared to gas-phase results, as well as the small differences (within 3.9 kcal/mol) between G binding sites for Li^+–G(N2–N3) and Li^+–G(O6–N7) in implicit water (Table E.1, Appendix E), which indicates these binding sites are closer in energy and more accessible in implicit solvent. Therefore, how hydrated Li^+ binds to the other purine (A) and pyrimidines (A, C, T, and U) is compared to G. Given that explicitly hydrated Li^+–G complexes achieve structural stabilization with 3 to 5 Li^+ bound waters and the dominant coordination numbers of Li^+ are 4 and 6, the Li^+–nucleobase models considered consist of Li^+ coordination numbers of 4 and 6 (Figures 6.6–6.7 and E.4–E.6; Table E.5, Appendix E) in two implicit solvents (1-bromopropane,

which mimics a DNA environment, and water).[53] Our analysis of the structures and binding strengths of Li^+–nucleobase complexes in implicit water are organized according to Li^+ interactions with the purines, pyrimidines, and comparison between purines and pyrimidines. Finally, the implicit water results are compared to a reduced-dielectric environment (1-bromopropane).

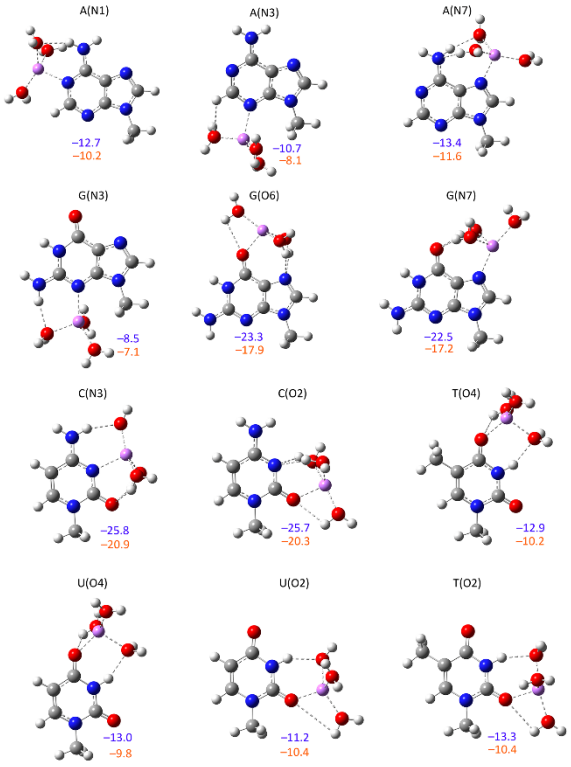


Figure 6.6. Structures and BEs (kcal/mol) of explicitly hydrated Li^+–nucleobase interactions involving 3 waters. Full structural information on hydrated Li^+–nucleobase complexes are provided in Appendix E (Figure E.5).

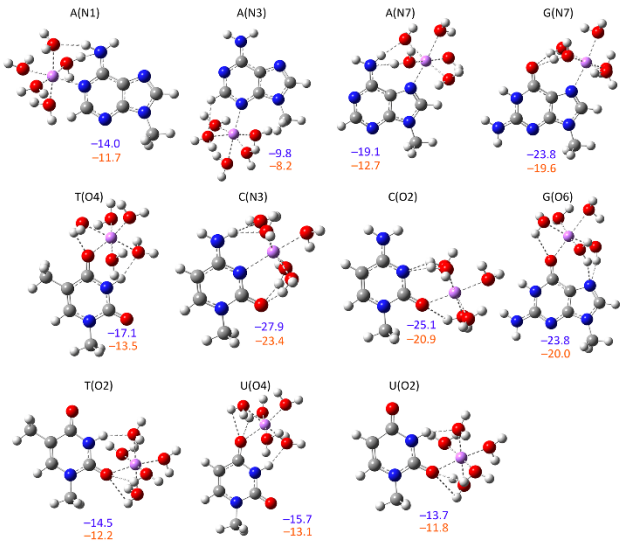


Figure 6.7. Structures and BEs (kcal/mol) of explicitly hydrated Li^+–nucleobase interactions involving 5 waters. Full structural information on hydrated Li^+–nucleobase complexes are provided in Appendix E (Figure E.6).

6.3.4.1. Hydrated Li^+–purine interactions are influenced by water-mediated contacts and steric accessibility, with consistent binding strength trends occurring regardless of number of Li^+-bound waters in 1-bromopropane or water

In general, hydrated Li^+ binds to purines and pyrimidines in a similar fashion, with a monodentate interaction between hydrated Li^+ and each purine or pyrimidine site. Additionally, Li^+-bound waters reorient to form hydrogen bonds with adjacent binding sites. Among the purines, the O6 and N7 sites in G are consistently favored for binding hydrated Li^+, while the N7 site is the strongest in A but weaker than G. Like G, C supports the most hydrogen bonds between Li^+-bound waters and the adjacent O2 and N3 sites. The N3 and O2 sites in C bind hydrated Li^+ the strongest, followed by T and U. Overall, the hydrated Li^+–purine/pyrimidine binding strengths decrease as C > G > A ~ T ~ U (Figure E.6 and Table E.5, Appendix E) regardless of n = 3 or 5 waters in the inner-shell of Li^+,

which contrasts previous gas-phase literature on bare Li^+–nucleobase interactions that identified G > C ≫ A ∼ T = U.[29] Indeed, G demonstrates stronger affinity for binding Li^+ than C in the gas phase by 3.7 kcal/mol; however, Li^+–C interactions show the most hydrogen bonds with well-orientated waters near neighboring nucleobase binding sites, leading to the largest overall BE by ~2.5–3.2 kcal/mol in 1-bromopropane and ~0.6–3.1 kcal/mol in water.

6.3.4.2. WCF Base-Pairing Alters the Number of Available Nucleic Acid Binding Sites and Affects Structural and Binding Strengths of Hydrated Li^+–Nucleic Acid Complexes, with Solvent Effects Varying by Li^+-Coordination Site

Given that nucleobases generally exist as base pairs in nucleic acids as well as the fact that base pairing alters the availability of Li^+ binding sites, it is important to expand the nucleobase models to include the complementary WCF base. By exploring implicit and explicit solvent effects on Li^+–nucleic acid interactions in a base-pair context, key insights into the structures, stabilities, and binding strengths of Li^+–nucleic acid interactions can be obtained. Therefore, WCF base-pair models are considered (Figures 6.8–6.9 and E.7–E.8; Table 6.1). Hydrated Li^+ binding to the WCF base pair is found to either increase (within 1.7 kcal/mol) or decrease (within 1.3 kcal/mol) base pair stability (Table E.6, Appendix E). These results demonstrate that hydrated Li^+ does not significantly influence WCF base pairing, therefore, our following results focus on how hydrated Li^+ interacts with the base pair itself.

Table 6.1. BEs (kcal/mol) evaluated in 1-bromopropane and water for explicitly hydrated Li^+–nucleobase (Li^+–NA) and Li^+–WCF nucleobase-pair (Li^+–BP) complexes with 3 or 5 waters.[a,b]

	n = 3						n = 5					
	1-bromopropane			Water			1-bromopropane			Water		
Complex	BE_{Li^+-NA}	BE_{Li^+-BP}	Diff.[c]	BE_{Li^+-NA}	BE_{Li^+-BP}	Diff.[c]	BE_{Li^+-NA}	BE_{Li^+-BP}	Diff.[c]	BE_{Li^+-NA}	BE_{Li^+-BP}	Diff.[c]
A(N1):T	–	–	–	–	–	–	–	–	–	–	–	–
A(N7):T	–13.4	–13.7	0.3	–11.6	–11.2	–0.4	–19.1	–14.6	–4.5	–12.7	–12.3	–0.4
A(N3):T	–10.7	–11.3	0.6	–8.1	–8.0	–0.1	–9.8	–	–	–8.2	–	–
A:T(O2)	–13.3	–11.9	–1.4	–10.4	–8.4	–2.0	–14.5	–15.9	1.4	–12.2	–15.2	3.0
A:T(O4)	–12.9	–13.0	0.1	–10.2	–8.7	–1.5	–17.1	–17.0	–0.1	–13.5	–13.1	–0.4
A(N7):U	–13.4	–13.8	0.4	–11.6	–11.2	–0.4	–19.1	–14.7	–4.4	–12.7	–12.6	–0.1
A(N3):U	–10.7	–11.3	0.6	–8.1	–8.2	0.1	–9.8	–11.8	2.0	–8.2	–	–
A:U(O2)	–11.2	–12.0	0.8	–10.4	–8.0	–2.4	–13.7	–14.8	1.1	–11.8	–11.7	–0.1
A:U(O4)	–13.0	–13.0	0.0	–9.8	–9.3	–0.5	–15.7	–13.8	–1.9	–13.1	–11.3	–1.8
G:C(N3)	–	–	–	–	–	–	–	–	–	–	–	–
G(N7):C	–22.5	–25.1	2.6	–17.2	–18.2	1.0	–23.8	–25.6	1.8	–20.0	–20.5	0.5
G(O6):C	–23.3	–24.9	1.6	–17.9	–19.1	1.2	–23.6	–26.0	2.4	–19.6	–21.0	1.4
G(N3):C	–8.5	–10.2	1.7	–7.1	–7.5	0.4	–	–	–	–	–	–
G:C(O2)	–25.7	–14.8	–10.9	–20.3	–11.3	–9.0	–25.1	–	–	–20.9	–	–

[a]All geometries were optimized with IEF-PCM-MN15/aug-cc-pVDZ and single-point calculations were carried out using IEF-PCM-MN15/aug-cc-pVTZ in each respective solvent. The data reported does not include deformation energy. All nucleobase models are capped with a methyl group at the N9 and N1 position in purines and pyrimidines, respectively. [b]The hydrated Li^+–nucleobase BEs are taken from Appendix E (Table E.5). [c]The differences are calculated as the hydrated Li^+–base pair BE subtracted by the hydrated Li^+–nucleobase BE in each solvent for 3 or 5 waters.

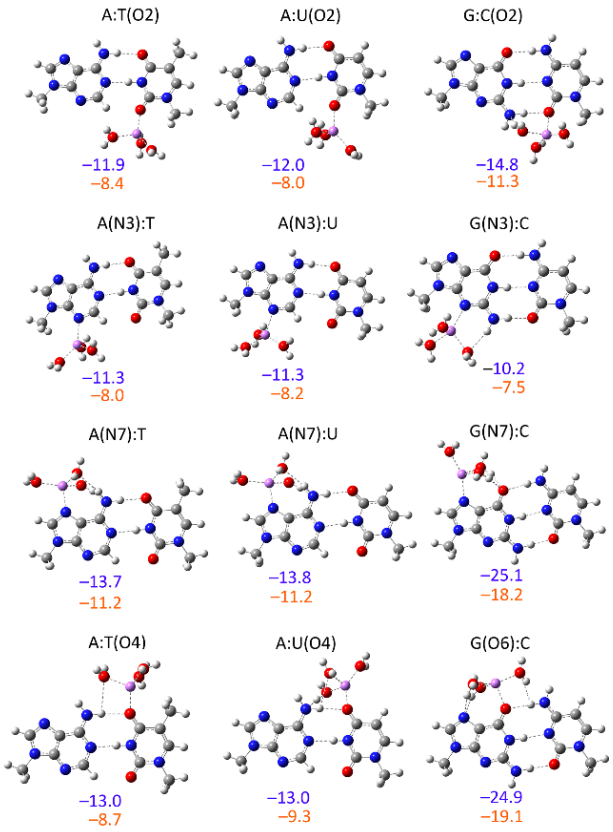


Figure 6.8. Structures and BEs (kcal/mol) of explicitly hydrated Li^+–base pair interactions involving 3 waters. Full structural information on hydrated Li^+–base pair complexes are provided in Appendix E (Figure E.7).

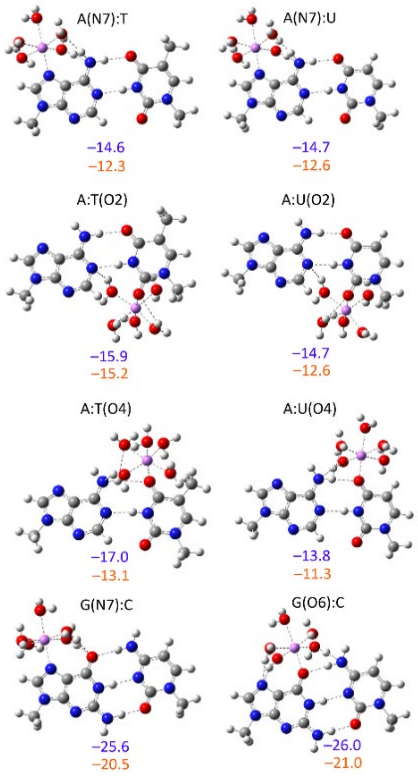


Figure 6.9. Structures and BEs (kcal/mol) of explicitly hydrated Li^+–base pair interactions involving 5 waters. Full structural information on hydrated Li^+–base pair complexes are provided in Appendix E (Figure E.8).

6.3.4.3. The Limited Accessibility to Binding Sites in GC Pairs Alters Hydrated Li^+ Binding Preferences for G over C Compared to Isolated Nucleobases

For hydrated Li^+ binding to a GC pair where n = 3, there are 4 distinct complexes, with 2 formed in the major groove and 2 in the minor groove. Specifically, hydrated Li^+ binds to the O6, N7, or N3 sites in G, or the O2 site in C. Like the nucleobase models, hydrated Li^+–GC interactions are monodentate with additional stabilization provided by hydrogen bonds between Li^+-bound waters and adjacent GC sites. However, unlike the nucleobase models, there are less available binding sites for hydrated Li^+ to bind to, which is due to hydrogen bonding in the WCF face between G and C. Nevertheless, the strongest

GC binding sites for attracting hydrated Li^+ are the O6 and N7 sites in G, followed by the O2 site in C and the N3 site in G. Specifically, the binding strengths of Li^+–G(O6):C and Li^+–G(N7):C are –19.1 and –18.2 kcal/mol in water, respectively, which is ~1.7 and ~1.9 times stronger than hydrated Li^+ interacting with G:C(O2) or G(N3):C. The slightly stronger binding of Li^+ to the G(O6):C site compared to G(N7):C is likely due to the formation of 3 hydrogen bonds between Li^+-bound waters and O6 and N7 sites in Li^+–G(O6):C complex, whereas only 2 hydrogen bonds occur in the Li^+–G(N7):C complex. In the Li^+–G:C(O2) complex, 2 Li^+-bound waters form hydrogen bonds with the O2 site in C and N2 site in G. The water orientation in these models differs from the nucleobase model, which is due to the absence of hydrogen bonding in the WCF face. The Li^+–G(N3):C complex demonstrates the weakest binding strength (–7.5 kcal/mol in water), with only a single hydrogen bond between a Li^+-bound water and N2 in G. Finally, lowering the dielectric from water to 1-bromopropane shortens Li^+–GC distances, reducing the distances by 0.009–0.028 Å. Additionally, the hydrated Li^+–GC binding strengths are strengthened by ~3–7 kcal/mol.

Expanding the hydration sphere of Li^+ from 3 to 5 waters reduces the number of complexes from 4 to 2. The larger hydration sphere around Li^+ leads to only the O6 and N7 sites in the GC base pair remain accessible. The inability of hydrated Li^+ to bind to the N3 site in G or O2 site in C is a result of steric crowding in the minor groove. Hydrated Li^+ binding at the O6 or N7 site in GC shows 3 Li^+-bound waters each forming hydrogen bonds with O6 and N7 in both complexes. As a result, these complexes lead to nearly identical binding strengths (within 0.5 kcal/mol in water). As observed for the hydrated Li^+–GC complexes where n = 3, reducing the dielectric to 1-bromopropane for n = 5 has a minimal

effect on the Li^+–GC distances by ~0.005–0.006 Å, but larger effect on strengthening the hydrated Li^+–GC binding interactions by ~5 kcal/mol.

Overall, hydrated Li^+ binding to the GC pair differs compared to each isolated G or C nucleobase. In the isolated nucleobases, hydrated Li^+ binds to C complexes the strongest, with these interactions occurring in WCF face and forming multiple stabilizing hydrogen bonds. However, base pairing between G and C prevents this interaction, redirecting hydrated Li_+ to bind to the O2 site in C in the minor groove, which deviates significantly from the isolated C. As a result, the hydrated Li^+–G and Li^+–C binding strengths in the base pair return to G(O6 ~ N7) > C(O2) > G(N3), aligning with gas-phase trends. Like the nucleobase models, the larger hydration sphere of Li^+ where n = 5 prevents Li^+ from accessing the N3 site in G in the minor groove, which is due to steric constraints.

6.3.4.4. The Binding Strength Trends for AT and AU are Sensitive to the Implicit and Explicit Solvent Environment

For hydrated Li^+ binding to AT and AU pairs where n = 3, 4 distinct complexes are formed for each base pair, with 2 formed in the major groove and 2 in the minor groove. Within the AT and AU pairs, hydrated Li^+ binds to the N7 or N3 sites in A, and the O4 or O2 sites in T or U. Like the nucleobase models, hydrated Li^+ interactions at each binding site are monodentate with water-mediated contacts formed with adjacent base pair sites. Among base pair binding sites, hydrated Li^+ interacts with the N7 site in AT and AU pairs the strongest, with binding strengths of –11.2 kcal/mol for both complexes. The hydrated Li^+–AT/AU binding strengths are ~1.2 times stronger than A:T/U(O4) and ~1.3 times

stronger than A:T/U(O2) and A(N3):T/U. Li+-bound waters form 2 hydrogen bonds with N7 in AT/AU and O4 of AU, 1 for O4 in AT, and no hydrogen-bond-mediated contacts for N3 or O2 sites in the AT or AU pairs. Following, hydrated Li^+ interacts with the O4 site in AT (–8.7 kcal/mol in water) and AU (–9.3 kcal/mol in water) are the strongest, followed by the O2 and N3 sites in AT and AU base pairs which are nearly equivalent to each other (within 0.4 kcal/mol in water). The Li^+–AT/AU distances are reduced by 0.008–0.031 Å upon lowering the dielectric from water to 1-bromopropane. Additionally, the hydrated Li^+–AT/AU binding strengths are strengthened by ~1–4 kcal/mol from water to 1-bromopropane.

Increasing the hydration sphere of Li^+ from n = 3 to 5 reduces the number of distinct complexes formed for each AT or AU pair from 4 to 3. All hydrated Li^+–AT/AU interactions observed for n = 3 persist for n = 5, except for hydrated Li^+ binding at the A(N3):T/U site. This result is due to steric crowding in the minor groove, which is observed in the isolated nucleobase models. The hydrated Li^+–AT/AU binding strength trends shift where Li^+–A:T(O4) is the strongest (–17.0 kcal/mol) in 1-bromopropane, while the strongest interaction in water occurs at A:T(O2) (–15.2 kcal/mol). The larger hydration sphere generally leads to more water-mediated contacts, with Li^+-bound waters forming 2 hydrogen bonds with O4 and O2 sites in the AT or AU pairs, while only 1 or 0 occur for n = 3. Reducing the dielectric to 1-bromopropane for n = 5 has a minimal effect on the Li^+–AT/AU distances by ~0.000–0.051 Å, but more variable effect on strengthening the hydrated Li^+–AT/AU binding interactions by ~1–4 kcal/mol.

Overall, the relative binding strengths of hydrated Li^+–A and Li^+–T/U differ between hydration spheres of n = 3 or 5 waters. For n = 3, the binding strength trends

decrease as AT/AU(N7) > AT/AU(O4) > AT/AU(O2 ~ N3), which is consistent with gas-phase trends. In contrast, for n = 5, the binding strengths decrease as AT(O2) > AT(O4) > AT/AU(N7) > AU(O2 ~ O4), demonstrating subtle differences in site preferences for hydrated Li^+ with a larger hydration sphere. These results underscore how solvent affects Li^+ binding behavior in nucleic acids.

6.3.4.5. While Hydrated Li^+ Binding to G in a Base Pair Remains Consistently the Strongest, Hydrated Li^+ Interactions with AT or AU Reveal Subtle Differences in Structure and Binding Strength

Overall, when considering the dominant Li^+ binding sites in each WCF base pair, the BE trends decrease as G > C > A > T ~ U in 1-bromopropane and G > C ~ A > T ~ U in water for complexes where n = 3. For n = 5, the hydrated Li^+–BP binding strength trends decrease as G > T > A > U, regardless of 1-bromopropane or water, with C no longer accessible for Li^+ with a larger hydration sphere due to steric constraints. The GC pair demonstrates the largest number of hydrogen-bond-mediated contacts (3) between Li^+-bound waters and adjacent base pair sites, regardless of n = 3 or 5. These results differ from hydrated Li^+ binding to isolated nucleobases, with the BEs following the trend of C > G. For the hydrated Li^+–base pair complexes with 5 waters, the BE trends decrease as G > T > A > U, regardless of implicit solvent.

While incorporation of an opposing WCF nucleobase influences the structure and binding strength of hydrated Li^+–nucleic acid interactions, such effects are generally modest. The most significant change from nucleobase to base-pair models is the loss of

contact between Li^+-coordinated waters and neighboring binding nucleobase sites upon forming the WCF pair, reducing the binding strengths. This highlights the importance of examining the opposing base when extrapolation of the conclusions to large nucleic acid systems is desired.

6.4. Conclusions

In summary, the impacts of implicit and explicit solvent on the structure and binding strength of Li^+–nucleic acid interactions were investigated in single nucleobase/phosphate and WCF base pair models. The results demonstrate that optimizing bare Li^+ interacting with nucleic acids in implicit solvent can lead to different geometries and significantly reduce the magnitude of binding strengths, effectively screening electrostatic interactions. When explicit water molecules are included in the models, the structures of Li^+–nucleic acid complexes diverge from prior gas-phase studies on bare Li^+–nucleic acid interactions. Specifically, as the number of water molecules increases, monodentate coordination between Li^+ and G/P is preferred, with bidentate coordination between Li^+ and G or P no longer observed once 3 or 2 water molecules coordinate to Li^+, respectively. When exploring the preferred coordination numbers of Li^+ (i.e., 4 and 6) at each nucleobase, the trends in the binding strength as a function of nucleobase alter such that C is preferred over G regardless of the number of explicit waters. This C > G preference is lost upon expanding the nucleobase models to include the opposing WCF nucleobase, with hydrogen bonding occupying the WCF face. Nevertheless, WCF base pairing has a modest impact on the structure and binding strength of hydrated Li^+–nucleic acid interactions.

Incorporating both implicit and explicit solvent is important to provide a more complete picture of solute–solvent interactions and offers a balance for computational efficiency and accuracy. Given that solvent modulates Li^+–nucleic acid interactions, both explicit and implicit solvent models should be considered in future studies. Additionally, the impact of base pair models on hydrated Li^+–nucleic acid interactions highlight the importance of using larger models that account for other nucleic acid contacts. To reduce computational costs, we recommend considering the most common Li^+ coordination number (4) for future studies exploring Li^+–nucleic acid interactions to ensure capture of all possible binding modes. Overall, this work provides fundamental structural and energetic insights into how the presence of solvent impacts Li^+–nucleic acid interactions, which is information that can be utilized towards understanding hydrated metal ion behavior in biosystems, developing biosensors for metal detection, and designing therapeutics to treat diseases.

6.5. References

(1) Kim, Y.; Nguyen, T.-T. T.; Churchill, D. G. Bioinorganic Chemistry of the Alkali Metal Ions. In *The Alkali Metal Ions: Their Role for Life*, Sigel, A., Sigel, H., Sigel, R. K. O. Eds.; Springer International Publishing, 2016; pp 1-10.

(2) Jakobsson, E.; Argüello-Miranda, O.; Chiu, S.-W.; Fazal, Z.; Kruczek, J.; Nunez-Corrales, S.; Pandit, S.; Pritchet, L. Towards a Unified Understanding of Lithium Action in Basic Biology and Its Significance for Applied Biology. *J. Membr. Biol.* **2017**, *250* (6), 587-604.

(3) Villegas-Vázquez, E. Y.; Quintas-Granados, L. I.; Cortés, H.; González-Del Carmen, M.; Leyva-Gómez, G.; Rodríguez-Morales, M.; Bustamante-Montes, L. P.; Silva-Adaya, D.; Pérez-Plasencia, C.; Jacobo-Herrera, N.; et al. Lithium: A Promising Anticancer Agent. *Life* **2023**, *13* (2), 537.

(4) He, T.; Cao, H.; Chen, P. The Roles of Alkali/Alkaline Earth Metals in the Materials Design and Development for Hydrogen Storage. *Acc. Mater. Res.* **2021**, *2*, 726-738.

(5) Armand, M.; Axmann, P.; Bresser, D.; Copley, M.; Edström, K.; Ekberg, C.; Guyomard, D.; Lestriez, B.; Novák, P.; Petranikova, M.; et al. Lithium-Ion Batteries – Current State of the Art and Anticipated Developments. *J. Power Sources* **2020**, *479*, 228708.

(6) Machado-Vieira, R.; Manji, H. K.; Zarate Jr, C. A. The Role of Lithium in the Treatment of Bipolar Disorder: Convergent Evidence for Neurotrophic Effects as a Unifying Hypothesis. *Bipolar Disord.* **2009**, *11* (s2), 92-109.

(7) Chellan, P.; Sadler, P. J. The Elements of Life and Medicines. *Philos. Trans. Royal Soc. A.* **2015**, *373*, 20140182.

(8) Snitow, M. E.; Bhansali, R. S.; Klein, P. S. Lithium and Therapeutic Targeting of GSK-3. *Cells* **2021**, *10* (2).

(9) Gomes, H. I.; Mayes, W. M.; Rogerson, M.; Stewart, D. I.; Burke, I. T. Alkaline Residues and the Environment: A Review of Impacts, Management Practices and Opportunities. *J. Clean. Prod.* **2016**, *112*, 3571-3582.

(10) Bolan, N.; Hoang, S. A.; Tanveer, M.; Wang, L.; Bolan, S.; Sooriyakumar, P.; Robinson, B.; Wijesekara, H.; Wijesooriya, M.; Keerthanan, S.; et al. From Mine to Mind and Mobiles – Lithium Contamination and Its Risk Management. *Environ. Pollut.* **2021**, *290*, 118067.

(11) Murphy, O.; Haji, M. N. A Review of Technologies for Direct Lithium Extraction from Low Li^+ Concentration Aqueous Solutions. *Front. Chem. Eng.* **2022**, *4*.

(12) Vera, M. L.; Torres, W. R.; Galli, C. I.; Chagnes, A.; Flexer, V. Environmental Impact of Direct Lithium Extraction from Brines. *Nat. Rev. Earth Environ.* **2023**, *4* (3), 149-165.

(13) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(14) Torabi, S. F.; Wu, P.; McGhee, C. E.; Chen, L.; Hwang, K.; Zheng, N.; Cheng, J.; Lu, Y. In Vitro Selection of a Sodium-Specific DNAzyme and Its Application in Intracellular Sensing. *Proc Natl Acad Sci U S A* **2015**, *112* (19), 5903-5908.

(15) Reiger, S. *New Tech Aims to Extract Lithium for Electric Car Batteries from Oilfield Waste*. 2020. https://www.cbc.ca/news/canada/calgary/lithium-alberta-oilsands-1.5424527 (accessed 2020 03-13).

(16) Cerda, B. A.; Wesdemiotis, C. Li^+, Na^+, and K^+ Binding to the DNA and RNA Nucleobases. Bond Energies and Attachment Sites from the Dissociation of Metal Ion-Bound Heterodimers. *J. Am. Chem. Soc.* **1996**, *118* (47), 11884-11892.

(17) Rodgers, M. T.; Armentrout, P. B. Cationic Noncovalent Interactions: Energetics and Periodic Trends. *Chem. Rev.* **2016**, *116* (9), 5642-5687.

(18) Rodgers, M. T.; Armentrout, P. B. Noncovalent Interactions of Nucleic Acid Bases (Uracil, Thymine, and Adenine) with Alkali Metal Ions. Threshold Collision-Induced Dissociation and Theoretical Studies. *J. Am. Chem. Soc.* **2000**, *122* (35), 8548-8558.

(19) Yang, Z.; Rodgers, M. T. Tautomerization in the Formation and Collision-Induced Dissociation of Alkali Metal Cation-Cytosine Complexes. *Phys. Chem. Chem. Phys.* **2012**, *14* (13), 4517-4526.

(20) Pakiari, A. H.; Farrokhnia, M. Nature of Lithium Interactions with DNA Nucleobases: Theoretical Study. *Phys. Chem. Res.* **2014**, *2* (2), 229-243.

(21) Russo, N.; Toscano, M.; Grand, A. Lithium Affinity for DNA and RNA Nucleobases. The Role of Theoretical Information in the Elucidation of the Mass Spectrometry Data. *J. Phys. Chem. B.* **2001**, *105* (20), 4735-4741.

(22) Gillis, E. A. L.; Rajabi, K.; Fridgen, T. D. Structures of Hydrated Li^+−Thymine and Li^+−Uracil Complexes by IRMPD Spectroscopy in the N−H/O−H Stretching Region. *J. Phys. Chem. A* **2009**, *113* (5), 824-832.

(23) Rajabi, K.; Gillis, E. A. L.; Fridgen, T. D. Structures of Alkali Metal Ion−Adenine Complexes and Hydrated Complexes by IRMPD Spectroscopy and Electronic Structure Calculations. *J. Phys. Chem. A.* **2010**, *114* (10), 3449-3456.

(24) Burda, J. V.; Šponer, J.; Hobza, P. Ab Initio Study of the Interaction of Guanine and Adenine with Various Mono- and Bivalent Metal Cations (Li^+, Na^+, K^+, Rb^+, Cs^+; Cu^+, Ag^+, Au^+; Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}; Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem. A.* **1996**, *100* (17), 7250-7255.

(25) Boychuk, B. T. A.; Meyer, S. P.; Wetmore, S. D. Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* **2023**, *11*.

(26) Stasyuk, O. A.; Solà, M.; Swart, M.; Fonseca Guerra, C.; Krygowski, T. M.; Szatylowicz, H. Effect of Alkali Metal Cations on Length and Strength of Hydrogen Bonds in DNA Base Pairs. *ChemPhysChem.* **2020**, *21* (18), 2112-2126.

(27) Ruan, C.; Huang, H.; Rodgers, M. T. Modeling Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase Via Alkali Metal Cation-Triethyl Phosphate Complexes. *J. Phys. Chem. A* **2007**, *111* (51), 13521-13527.

(28) Ruan, C.; Huang, H.; Rodgers, M. T. A Simple Model for Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase: Alkali Metal Cations and Trimethyl Phosphate. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 305-314.

(29) Boychuk, B. T. A.; Jeong, R. Y. E.; Wetmore, S. D. Assessment of the Accuracy of DFT-Predicted Li^+–Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408.

(30) Zhu, W.; Luo, X.; Puah, C. M.; Tan, X.; Shen, J.; Gu, J.; Chen, K.; Jiang, H. The Multiplicity, Strength, and Nature of the Interaction of Nucleobases with Alkaline and Alkaline Earth Metal Cations: A Density Functional Theory Investigation. *J. Phys. Chem. A* **2004**, *108* (18), 4008-4018.

(31) Monajjemi, M.; Ghiasi, R.; Sadjadi, M. A. S. Metal-Stabilized Rare Tautomers: N4 Metalated Cytosine (M = Li^+, Na^+, K^+, Rb^+ and Cs^+), Theoretical Views. *Appl. Organomet. Chem.* **2003**, *17* (8), 635-640.

(32) Hashemianzadeh, S. M.; Faraji, S.; Amin, A. H.; Ketabi, S. Theoretical Study of the Interactions between Isolated DNA Bases and Various Groups IA and IIA Metal Ions by Ab Initio Calculations. *Monatsh. Chem.* **2008**, *139* (2), 89-100.

(33) Jay, R. M.; Vaz da Cruz, V.; Eckert, S.; Fondell, M.; Mitzner, R.; Föhlisch, A. Probing Solute–Solvent Interactions of Transition Metal Complexes Using L-Edge Absorption Spectroscopy. *J. Phys. Chem. B* **2020**, *124* (27), 5636-5645.

(34) Hush, N. S.; Reimers, J. R. Solvent Effects on the Electronic Spectra of Transition Metal Complexes. *Chem. Rev.* **2000**, *100* (2), 775-786.

(35) Mondal, S.; Bagchi, B. From Structure and Dynamics to Biomolecular Functions: The Ubiquitous Role of Solvent in Biology. *Curr. Opin. Struct. Biol.* **2022**, *77*, 102462.

(36) Mähler, J.; Persson, I. A Study of the Hydration of the Alkali Metal Ions in Aqueous Solution. *Inorg. Chem.* **2012**, *51* (1), 425-438.

(37) Persson, I. Hydrated Metal Ions in Aqueous Solution: How Regular Are Their Structures? *Pure and Applied Chemistry* **2010**, *82* (10), 1901-1917.

(38) Zeng, Z.; Liu, C.-W.; Hou, G.-L.; Feng, G.; Xu, H.-G.; Gao, Y. Q.; Zheng, W.-J. Photoelectron Spectroscopy and Ab Initio Calculations of $Li(H_2O)_N^-$ and $Cs(H_2O)_N^-$ (N = 1–6) Clusters. *J. Phys. Chem. A* **2015**, *119* (12), 2845-2856.

(39) Nagasaka, M.; Yuzawa, H.; Kosugi, N. Interaction between Water and Alkali Metal Ions and Its Temperature Dependence Revealed by Oxygen K-Edge X-Ray Absorption Spectroscopy. *J. Phys. Chem. B* **2017**, *121* (48), 10957-10964.

(40) Egan, C. K.; Bizzarro, B. B.; Riera, M.; Paesani, F. Nature of Alkali Ion–Water Interactions: Insights from Many-Body Representations and Density Functional Theory. II. *J. Chem. Theory Comput.* **2020**, *16* (5), 3055-3072.

(41) Rao, J. S.; Dinadayalane, T. C.; Leszczynski, J.; Sastry, G. N. Comprehensive Study on the Solvation of Mono- and Divalent Metal Cations: Li^+, Na^+, K^+, Be^{2+}, Mg^{2+} and Ca^{2+}. *J. Phys. Chem. A.* **2008**, *112* (50), 12944-12953.

(42) Gillis, E. A. L.; Fridgen, T. D. The Hydrated Li^+–Adenine–Thymine Complex by IRMPD Spectroscopy in the N–H/O–H Stretching Region. *Int. J. Mass Spectrom.* **2010**, *297* (1), 2-8.

(43) Dhaouadi, Z.; Nsangou, M.; Hernández, B.; Pflüger, F.; Liquier, J.; Ghomi, M. Geometrical and Vibrational Features of Phosphate, Phosphorothioate and Phosphorodithioate Linkages Interacting with Hydrated Cations: A DFT Study. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2009**, *73* (5), 805-814.

(44) Farrokhpour, H.; Mokhtari, N. Intermolecular Hydrogen Bonding in DNA Base Pairs Interacting with Different Numbers of Bare and Hydrated Li^+: NBO, QTAIM, and Computational Spectroscopic Studies. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2024**, *310*, 123896.

(45) Cruz-León, S.; Vanderlinden, W.; Müller, P.; Forster, T.; Staudt, G.; Lin, Y.-Y.; Lipfert, J.; Schwierz, N. Twisting DNA by Salt. *Nucleic Acids Res.* **2022**, *50* (10), 5726-5738.

(46) Gebala, M.; Bonilla, S.; Bisaria, N.; Herschlag, D. Does Cation Size Affect Occupancy and Electrostatic Screening of the Nucleic Acid Ion Atmosphere? *J. Am. Chem. Soc.* **2016**, *138* (34), 10925-10934.

(47) Savelyev, A.; MacKerell, A. D. Competition among Li^+, Na^+, K^+, and Rb^+ Monovalent Ions for DNA in Molecular Dynamics Simulations Using the Additive Charmm36 and Drude Polarizable Force Fields. *J. Phys. Chem. B.* **2015**, *119* (12), 4428-4440.

(48) Savelyev, A.; MacKerell, A. D. Differential Deformability of the DNA Minor Groove and Altered Bi/Bii Backbone Conformational Equilibrium by the Monovalent Ions Li^+, Na^+, K^+, and Rb^+ Via Water-Mediated Hydrogen Bonding. *J. Chem. Theory Comput.* **2015**, *11* (9), 4473-4485.

(49) Lohikoski, R. A.; Timonen, J.; Lyubartsev, A. P.; Laaksonen, A. Internal Structure and Dynamics of the Decamer $D(ATGCAGTCAG)_2$ in Li^+-H 2 O Solution: A Molecular Dynamics Simulation Study. *Molecular Simulation* **2003**, *29* (1), 47-62.

(50) Lyubartsev, A. P.; Laaksonen, A. Molecular Dynamics Simulations of DNA in Solutions with Different Counter-Ions. *J. Biomol. Struct. Dyn.* **1998**, *16* (3), 579-592.

(51) van Dam, L.; Lyubartsev, A. P.; Laaksonen, A.; Nordenskiöld, L. Self-Diffusion and Association of Li^+, Cs^+, and H_2O in Oriented DNA Fibers. An NMR and MD Simulation Study. *J. Phys. Chem. B* **1998**, *102* (51), 10636-10642.

(52) Ruan, C.; Huang, H.; Rodgers, M. T. Modeling Metal Cation−Phosphate Interactions in Nucleic Acids in the Gas Phase Via Alkali Metal Cation−Triethyl Phosphate Complexes. *J. Phys. Chem. A* **2007**, *111* (51), 13521-13527.

(53) Cuervo, A.; Dans, P. D.; Carrascosa, J. L.; Orozco, M.; Gomila, G.; Fumagalli, L. Direct Measurement of the Dielectric Polarization Properties of DNA. *Proc. Natl. Acad. Sci.* **2014**, *111* (35), E3624.

(54) Olsher, U.; Izatt, R. M.; Bradshaw, J. S.; Dalley, N. K. Coordination Chemistry of Lithium Ion: A Crystal and Molecular Structure Review. *Chem. Rev.* **1991**, *91* (2), 137-164.

(55) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.;; Robb M. A. Cheeseman J. R. Scalmani G.; Barone, V. P., G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision B.01. *Gaussian, Inc. Wallingford, CT* **2016**.

(56) Marynick, D. S.; Schaefer, H. F., 3rd. Theoretical Studies of Metal-Phosphate Interactions: Interaction of Li^{+}, Na^{+}, K^{+}, Be^{++}, Mg^{++}, and Ca^{++} with $H_2PO_4^{-}$ and $(CH_3O)_2PO_2^{-}$: Implications for Nucleic Acid Solvation. *Proc. Natl. Acad. Sci.* **1975**, *72* (10), 3794-3798.

Chapter 7: Manipulating Metal Binding to Nucleic Acids Through Chemical Modification: A Molecular Dynamics Simulation Study of PNA

7.1. Introduction

Metal–nucleic acid interactions are vital to numerous biological functions and enable the development of materials, therapeutics, and biotechnologies.[1-3] Specifically, the diverse metal binding sites in nucleic acid components (i.e., phosphate, sugar, or nucleobases; Figure 7.1) facilitate interactions, while the structural flexibility of nucleic acids affords the formation of new structural motifs for metal binding.[4] For example, the Mg^{2+} ions on a mica surface stabilize the structure of DNA-origami (i.e., single-stranded DNA that folds to form different shapes), with DNA-origami scaffolds being used to develop enhanced DNA nanowires to improve solar cell design.[5] Alternatively, cisplatin interacts with nucleic acids to form distinctive 1,2-intrastrand crosslinks and thereby provides chemotherapeutic effects.[6] Aside from material and therapeutic applications, nucleic acid-based biotechnologies have emerged such as nucleic acid sensing platforms to detect metals in the environment and cellular conditions.[3,7]

The functionality of nucleic acids for metal binding can be further enhanced through chemical modifications. Numerous modifications to nucleic acid subcomponents have been synthesized.[8] The type and location of modifications affect how metals interact with nucleic acids. Chemical modifications can induce nucleic acid selectivity for a particular metal. For example, introducing a phosphorothioate modification at the cleavage site in the Cd16 and PSCu10 DNAzymes leads to selective detection of Cd^{2+} and Cu^{2+}, respectively, which is due to the sulfur modification attracting the soft Lewis acids.[9,10] Introducing chemical alterations to nucleic acids can induce the formation of specific folded structures. As an

example, the 8–17 DNAzyme was chemically modified to contain 4-carboxyimidazole base pairs, which cause the metal binding pocket to fold into a compact structure upon specific recognition of Cu^{2+}.[11,12] Additionally, chemical modifications can facilitate the formation of structural motifs with distinct functions. For example, a nucleic acid sensor was designed in which the carbonyl and amino group positions in guanine were swapped to create a G-pentaplex that selectively binds Cs^{+}.[3] While there have been advancements in the development of chemically-modified nucleic acids to bind metals, the full potential of modifications to impart new functions to nucleic acids largely remains untapped, at least in part due to gaps in our understanding of how modified nucleic acids interact with metals.[13]

Among different modifications, chemical alterations to the sugar–phosphate backbone are of particular interest due to their ability to influence the charge, structure, and stability of nucleic acids.[14,15] Since positively charged metals are naturally drawn towards the negatively charged phosphate backbone, these metal–nucleic acid interactions lack selectivity and therefore modifying the sugar–backbone is useful for manipulating metal behavior in metal–nucleic acid-based technologies. Specifically, modifying the DNA sugar–phosphate backbone to a peptide backbone leads to peptide nucleic acids (PNA; Figure 7.1). PNA has uses in biotechnologies such as biosensing and gene therapies.[16] In particular, PNA probes have been used to detect bacterial pathogens in food,[17] and microRNAs (miRNAs) related to breast and liver cancer,[18,19] and have been interfaced with smartphone technology to sense miRNA linked to prostate cancer.[20] Aside from biosensing, antisense PNA agents have proven successful for inhibiting gene expression in bacteria, which results in greater antibacterial activity against *E. coli*.[21] Additionally, gene editing technologies have been designed by coupling PNA with T7 endonuclease to carry out site-

specific double-strand breaks in DNA sequences.[22] Overall, PNA has proven to be successful across various applications; however, further work is required to understand how these modified nucleic acids interact with metals to further advance these applications.

Since PNA has proven to be a valuable tool in medicine and biotechnologies, there have been many experimental and computational studies carried out to understand the structure and binding properties of PNA duplexes and PNA hybrids with DNA or RNA.[23-31] Crystallographic data revealed that the neutral peptide backbone of PNA significantly alters the nucleic acid shape,[23] with PNA adopting a P-type helix that is more unwound (smaller twist), exhibits less stacking between nucleobases, is wider in diameter, and is more flexible than canonical B-DNA. Additionally, gel electrophoresis has shown that PNA is more resistant to enzyme degradation compared to DNA or RNA.[14] When different DNA, RNA, and PNA (full and hybrid) duplexes were studied using UV-monitored thermal denaturation experiments in combination with molecular dynamics (MD) simulations, PNA duplexes were identified to be the most stable followed by hybrid PNA-DNA/RNA duplexes, while DNA/RNA duplexes were the least stable.[30] Molecular dynamics (MD) studies were initially performed to uncover the structural dynamics of PNA using preliminary parameters, which identified a more flexible backbone in PNA than DNA.[24,27,28,31] Improved parameters have been more recently developed for PNA[29] that more accurately describe the backbone and reliably reproduce the most-accurately resolved and most-widely studied crystallographic structure (PDB ID: 3MBS).[23] While the predicted backbone torsion parameters better match experimental data, the conclusion regarding the highly dynamic structure of PNA remains unchanged.[29] Although experimental and

computational works have uncovered the structural dynamics of PNA, no studies to date have explored metal–PNA interactions, which could unlock new applications.

In the present work, we uncover the binding distributions of select metals to PNA duplexes using molecular dynamics (MD) simulations (Figure 7.1). An initial emphasis was placed on understanding Na^+, which is prevalent in biological systems, and Li^+, which requires solutions for environmental remediation and material recovery for battery technologies.[32,33] Investigating these two metals will provide insights regarding how metal identity affects metal–PNA interactions. Subsequently, the effects of salt concentration on metal binding positions and PNA structural features are explored by varying the salt concentration for each metal from 100 mM to 750 mM. Additionally, two nucleobase sequences are investigated to uncover the possible sequence dependence of metal–PNA interactions. All conditions were also explored with DNA in comparison to PNA. Overall, this study provides the first insight into how metals bind to PNA and demonstrates how modifications can shift metal binding towards specific nucleic acid regions. The information gained from this study will help guide the future development of PNA-based therapeutics, nanomaterials, and biotechnologies that can be applied in different metal ion and salt environments.

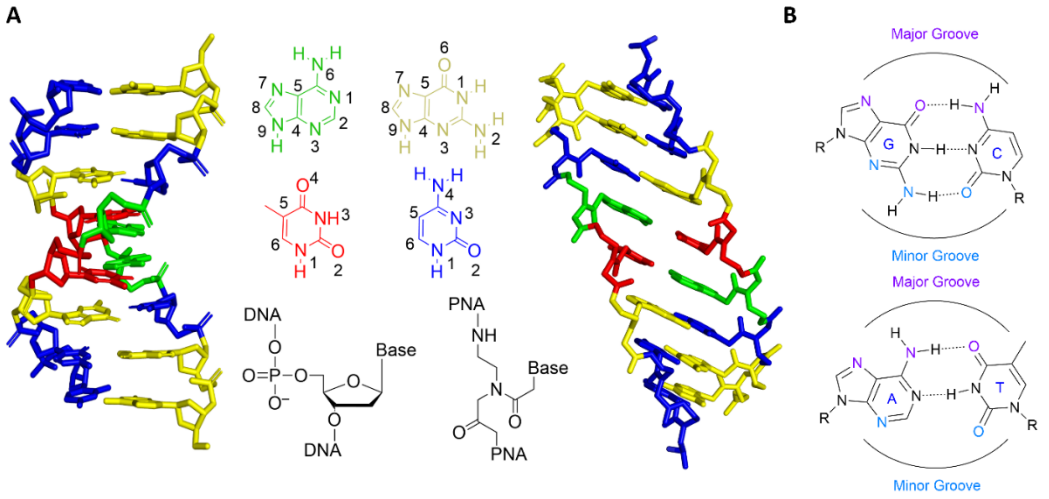


Figure 7.1. (**A**) 3D structure as well as nucleobase and backbone subcomponents of DNA (left) and PNA (right), and (**B**) the defined major and minor grooves for GC-rich and AT-rich Watson-Crick-Franklin base pairs.

7.2. Methodology

7.2.1. Model Preparation: The starting model for PNA was extracted from an X-ray crystal structure (PDB ID: 3MBS; 5′-GGCATCGG-3′), while a DNA model with the same sequence was constructed using PyMOL 2.5.[34] Additional models were generated for both PNA and DNA wherein the sequence was modified from GC to AT-rich (5′-AATGCTAA-3′). PNA strands were capped with an acetyl ($COCH_3$) group at the N-terminus and an amide group at the C-terminus, while terminal ends of DNA were capped with a hydroxy group. DNA models were neutralized with either Li^+ or Na^+ using tLEAP, and the same number of metal cations as the DNA models in addition to Cl^- anions were added to PNA models to maintain neutrality. Each model was solvated in a TIP4P-EW water box,[35] with the distance between the edge of the box and the solute being a minimum

of 24 Å for PNA (more stretched length) and 32 Å for DNA (more compressed in length) to ensure the volume (~77 000 $Å^3$) and salt concentration are comparable between systems. Subsequently, different LiCl and NaCl concentrations were considered (100, 250, 500, and 750 mM), with an additional NaCl concentration of 150 mM added to model a physiological environment. DNA was described using the AMBER OL15 forcefield,[36] while the Amber-type force field with improved backbone torsion parameters was used for PNA.[29] All ions (Li^+, Na^+, and Cl^-) were described using Joung and Cheatham parameters.[37]

7.2.2. MD Simulation Protocol: Each system was minimized in four stages: 1) solvent and ions were minimized while restraining the solute; 2) hydrogens and solute were minimized, with heavy atoms, ions, and water restrained; 3) all atoms of the solute were minimized, while the ions and solvent were restrained; and 4) the entire system was minimized. 2 000 cycles of the steepest descent algorithm followed by 2000 steps of conjugate gradient minimization were used in each step, and a force constant of 100 $kcal \cdot mol^{-1} \cdot Å^{-2}$ was applied as the restraint. Each system was heated gradually from 10 K to 310 K in 50 K increments, with a 25 $kcal \cdot mol^{-1} \cdot Å^{-2}$ force constant used to restrain the system during heating. Each system was equilibrated in 5 stages of 10 000 steps each by gradually decreasing the restraint on the solute from 20, 15, 10, 5, and 1.5 $kcal \cdot mol^{-1} \cdot Å^{-2}$. A 2 fs time step was used during production simulations with the SHAKE algorithm.[38] All production simulations were performed in 5 replicas of 500 ns each at 1 bar of pressure (Berendsen barostat)[39] and 310 K (Langevin thermostat) to mimic

human body temperature.[40] All simulations were performed using the AMBER 20 modeling package.[41]

7.2.3. MD Simulation Analysis: All frames from all production simulations were saved every 1 ns and included in our analysis. Averages and standard deviations were calculated over all data. The total ion density around PNA and DNA was mapped using the bounds and grid analysis with 0.1 Å spacing and visualized with ChimeraX using a distance cutoff of 5 Å from the solute. Radial distribution functions (RDF) were used to define the interaction cutoffs for inner- and outer-sphere metal binding events. Inner-sphere coordination was determined based on a coordination distance cutoff of 3 Å and total interaction cutoff of 5 Å. A metal binding event is defined as an instance where a metal interacts with a given binding site within PNA or DNA with these cutoffs. The binding sites considered include N3, N7, and N1 of the purines, as well as O6 of G and N6 of A, as well as the N3, N4, and O2 binding sites of C, and the O2 and O4 positions of T and U (Figure 7.1). These sites were selected according to established gas-phase data highlighting these sites as dominant for metal binding in nucleic acids.[3,42] Binding sites in the backbone of DNA included all phosphate oxygens (O1P, O2P, O3′, O5′), while PNA included nitrogen (N2′) and oxygen (O1′, O3′) sites. Subsequently, metal binding events were normalized according to the number of each nucleic acid component present in the duplex. The terminal base pairs were excluded from all metal binding analysis. Averages in the nucleic acid structural parameters for PNA and DNA were calculated for each base step. To minimize terminal end effects, two terminal base pairs in PNA and DNA were excluded from our structural analysis, with terminal end effects leading to observed perturbations in

an adjacent base pair in previous work on PNA.[29] Subsequently, the neutral PNA and DNA systems were used as a reference to calculate structural deviations in the presence of the different salt concentrations. The minor groove width was calculated from the interstrand P′ to P′ in DNA, and C3′ to N4′ in PNA (Figure F.1, Appendix F), as defined in previous work.[26] All MD analysis was carried out using CPPTRAJ of AMBERTools20 and X3DNA for the nucleic acid structural parameters of PNA and DNA.

7.3. Results

We have utilized MD simulations to uncover how metals (Na^+ and Li^+) differentially interact with PNA and DNA. The results and discussion are organized according to comparisons of the metal binding distributions between DNA and PNA for each metal at the lowest salt concentration (100 mM). Subsequently, the structural deviations in the nucleic acid parameters of DNA and PNA in 100 mM are discussed for each metal. Finally, the impact of changing salt concentration (100, 250, 500, and 750 mM) and sequence (GC vs. AT-rich) on the metal binding distributions and structural deviations is uncovered.

7.3.1. While the Major and Minor Grooves of DNA Exhibit Nearly Equivalent Na^+ Binding, the Peptide Backbone of PNA Pushes Na^+ Towards the Major Groove

Overall, the presence of Na^+ does not change the general B-form of DNA or P-form of PNA (Figure 8.2). Na^+ is densely populated throughout the DNA duplex, with an emphasis on the backbone (Figure 8.2), as previously reported in the literature.[43,44] In

contrast, there is overall less Na^+ density around PNA, with less Na^+ distributed around the backbone and more in the major groove. As a result, the magnitude of metal binding events significantly decreases from DNA to PNA, with DNA demonstrating ~10 times more total Na^+ binding events than PNA. Indeed, the average total normalized binding distribution throughout DNA and PNA reveals a noticeable shift of Na^+ binding from the backbone (22% DNA; 11% PNA) towards the nucleobases from DNA to PNA (Figure 8.3A, Table F.1, Appendix F). RDFs highlight that inner-sphere interactions to the backbone of DNA and PNA are dominant (Figure 8.4). In contrast, the inner-sphere nucleobase RDFs are prominent in DNA, while inner- and outer-sphere interactions are nearly equivalent in PNA, which signifies that altering the nucleic acid backbone affects Na^+–nucleobase interactions.

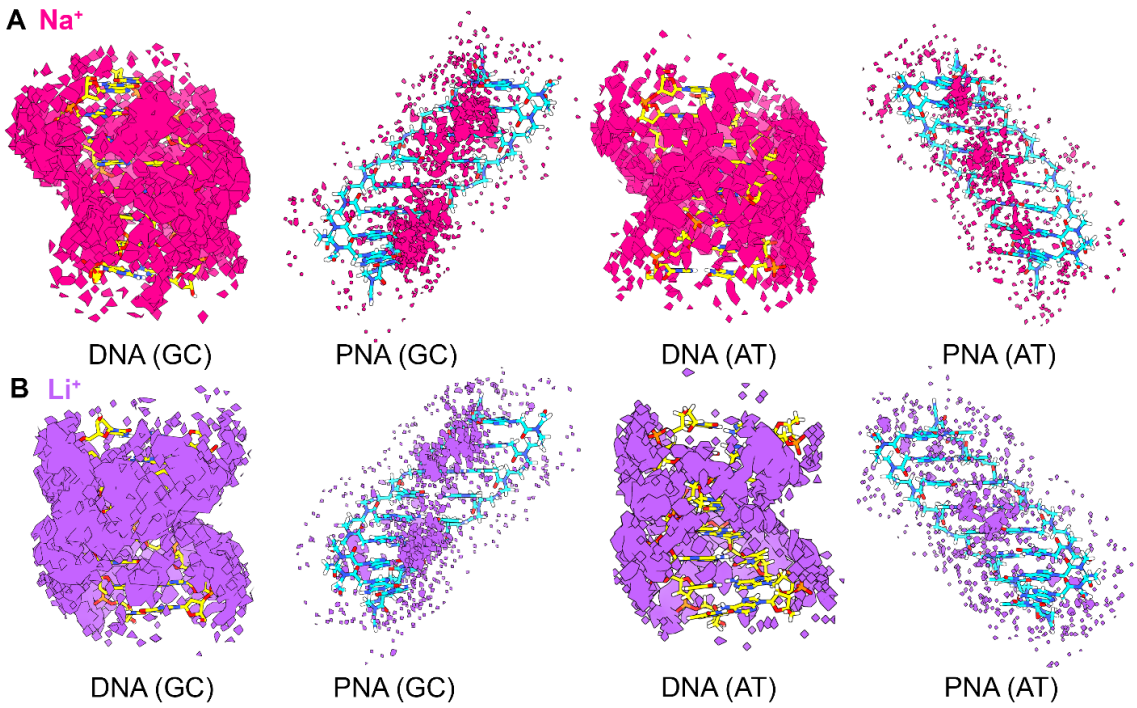


Figure 7.2. 3D density plots of the total metal binding distributions of (**A**) Na^+ and (**B**) Li^+ to DNA (left) and PNA (right) in 100 mM for the GC-rich (5′-GGCATCGG-3′) and AT-rich (5′- AATGCTAA-3′) sequences.

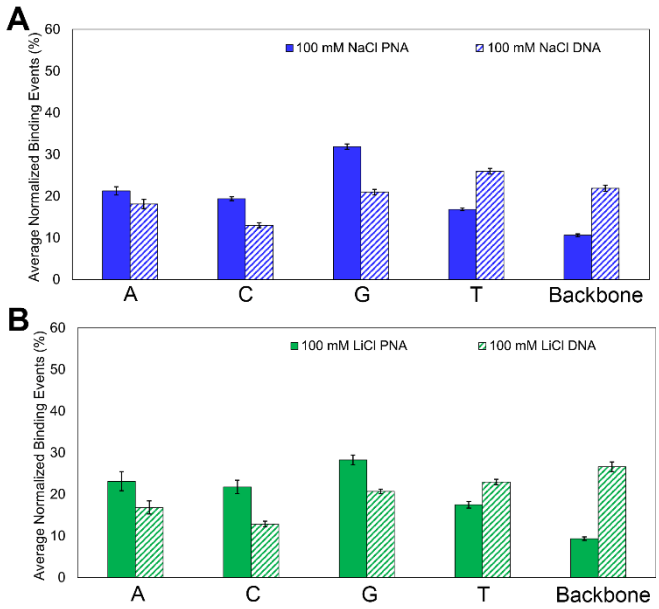


Figure 7.3. The average normalized (**A**) Na^+ and (**B**) Li^+ binding events (%) within 5 Å (total) of the backbone or each nucleobase in DNA (striped) or PNA (solid) in 100 mM of NaCl or LiCl.

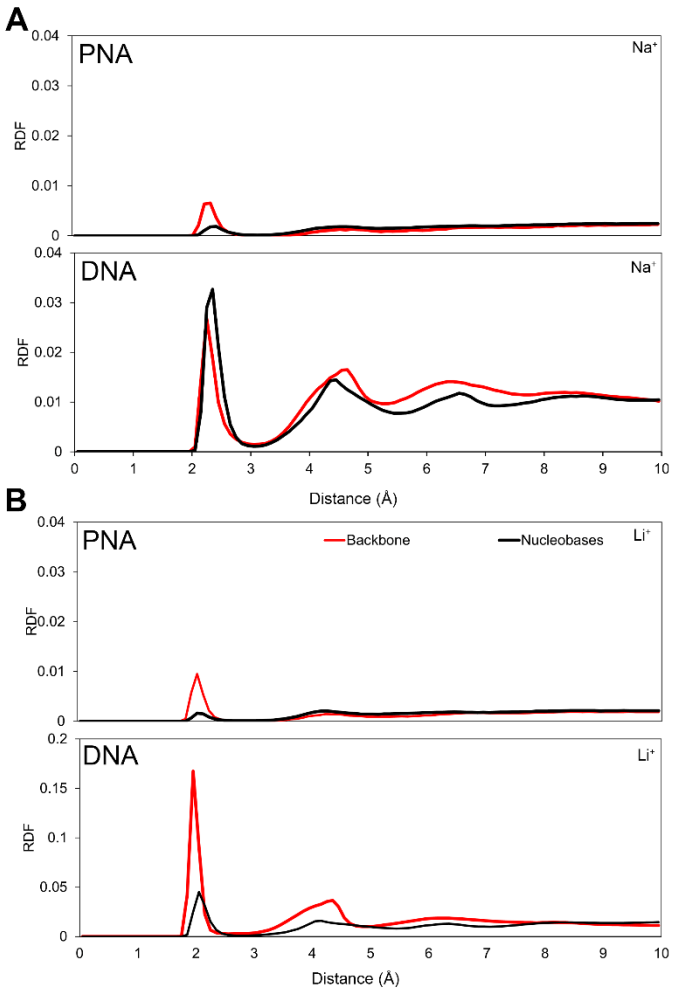


Figure 7.4. Radial distribution function (RDF) for (**A**) Na^+ or (**B**) Li^+ distance from all possible backbone (red) or nucleobase (black) binding sites in DNA or PNA.

To further investigate the differences in Na^+ binding to DNA and PNA, we next focus on the inner-sphere Na^+ interactions with each nucleobase binding site in the major and minor groove. The average normalized binding events demonstrate comparable binding in both grooves of DNA, with ~45% in the major groove and ~55% in the minor groove (Figure 8.5A, Table F.2, Appendix F), which agrees with previous computational and experimental results.[45,46] The preferred binding sites in the DNA major and minor grooves are G(O6) and T(O2), respectively, which is consistent with previous gas-phase

experimental and computational works that highlight these are the strongest metal binding sites in the major and minor groove.[42,47] In contrast, a drastic shift of Na^+ binding occurs from the minor (2%) to the major groove (98%) in PNA. In PNA, G(O6), G(N7), and T(O4) are the dominant sites for Na^+ binding, with average normalized binding events of 42%, 27%, and 25% respectively. This shift in Na^+ binding from the minor to major groove from DNA to PNA may be attributed to the narrower minor groove width of PNA (9.7 Å) compared to DNA (12.2 Å; Table F.3, Appendix F), with the narrower minor groove in PNA identified in previous work.[24,26] Overall, PNA demonstrates the ability to manipulate the position of Na^+ binding in a nucleic acid duplex by shifting metal binding away from the backbone and minor groove towards the major groove.

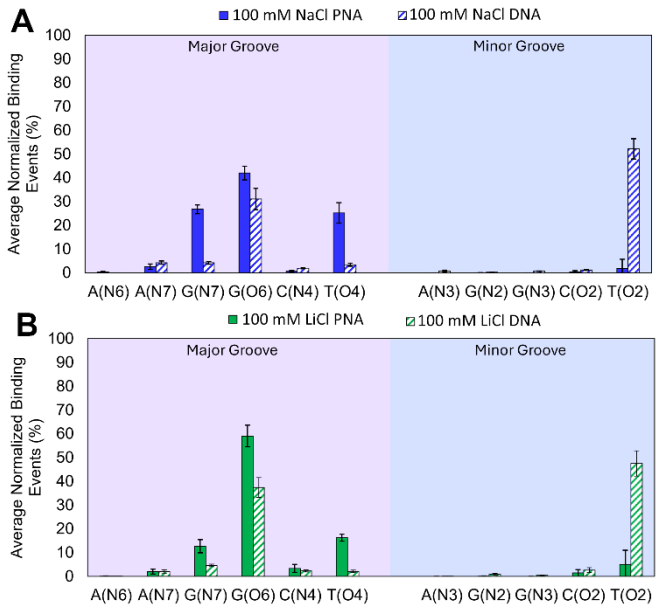


Figure 7.5. The average normalized (**A**) Na^+ or (**B**) Li^+ binding events (%) within 3 Å to each nucleobase site in the major and minor grooves of DNA (striped) or PNA (solid) in 100 mM of NaCl or LiCl.

7.3.2. Li^+ is pushed to the O6 site of G in PNA

Having explored Na^+ binding to PNA compared to DNA, the next step is to evaluate how metal binding differs from Na^+ to Li^+. Li^+ density around the backbone of DNA is smoother than Na^+ (Figure 8.2), indicating more interactions between Li^+ and the phosphate backbone. Like Na^+, the metal binding distributions show that Li^+ is pushed away from the backbone towards the major groove for PNA compared to DNA, with a greater localization at the G:C base pairs (Figure 8.3B). Overall, the number of metal binding events substantially decreases (by ~11 times) from DNA to PNA. Furthermore, there is ~1.4 and 1.2 times more binding between Li^+ and DNA or PNA than Na^+. These results also highlight that Li^+ binds more to DNA than Na^+, which agrees with previous experimental and computational literature.[48,49] The total normalized binding events demonstrate that Li^+ is more localized at the DNA phosphate backbone (27%) than Na^+, followed by T (23%), and G (21%; Figure 8.3B, Table F.1, Appendix F). In PNA, Li^+ binding to the peptide backbone is significantly reduced (9%) compared to the DNA phosphate backbone (27%). Li^+ is also more evenly distributed across the nucleobases in PNA (~17–28%). Comparing Li^+ binding interactions within the inner- and outer-coordination sphere of the nucleobases and backbone, Li^+ predominantly forms inner-sphere interactions with the backbone in DNA and PNA (Figure 8.4). Like Na^+, Li^+ shows equivalent binding between the inner- and outer-coordination sphere in PNA. Within the major and minor grooves of DNA and PNA, Li^+ binds in a similar fashion to Na^+ such that metal binding is distributed evenly between the grooves of DNA (48% major and 52% minor groove) but is dominant in the major groove in PNA (93% in major and 7% in minor groove; Figure 8.5B, Table F.2,

Appendix F). However, unlike Na^+, Li^+ is pushed to a specific binding site in the major groove of PNA, particularly the O6 site of G (59%). Indeed, the next most populated Li^+ binding site in PNA is the O4 site of T (16%). Altogether, these results highlight that modifying the backbone from DNA to PNA influences the Li^+ binding distribution, pushing the smaller metal to a particular nucleobase binding site. This suggests that PNA may be a step in the right direction for selectively binding Li^+.

7.3.3. At the Lowest Salt Concentration, PNA Structure is Less Sensitive to Metal Identity than DNA

The impact of Na^+ and Li^+ on the structure of DNA and PNA was investigated by calculating the difference in structural parameters for nucleic acids in the lowest salt concentration (100 mM) and no salt (Figure 8.6). To minimize artifacts in our structural analysis introduced by terminal end effects, the two terminal base pairs were omitted from the analysis. In general, there are minimal deviations in the nucleic acid parameters of DNA in the presence of Na^+ as shown by the light shades of red and blue. In contrast, there are deeper shades of red and blue throughout the helix in the presence of Li^+, signifying greater structural deviations in DNA, which agrees with previous literature.[50] In PNA, there are minimal structural deviations in the presence of either Na^+ or Li^+, except for one base pair above the two terminal ends within the duplex that demonstrated localized structural deviations in the X- and Y-displacement in the presence of Li^+, highlighting minor perturbation in the base pairs due to terminal end effects. Most importantly, the structure of PNA is less susceptible to deviations than DNA in the presence of metals. Overall, this highlights that PNA is likely an improved candidate for targeting Li^+ due to minimal

structural deviations upon metal binding, with structural stability being an important property in designing nucleic acid sensing and extraction platforms.[51]

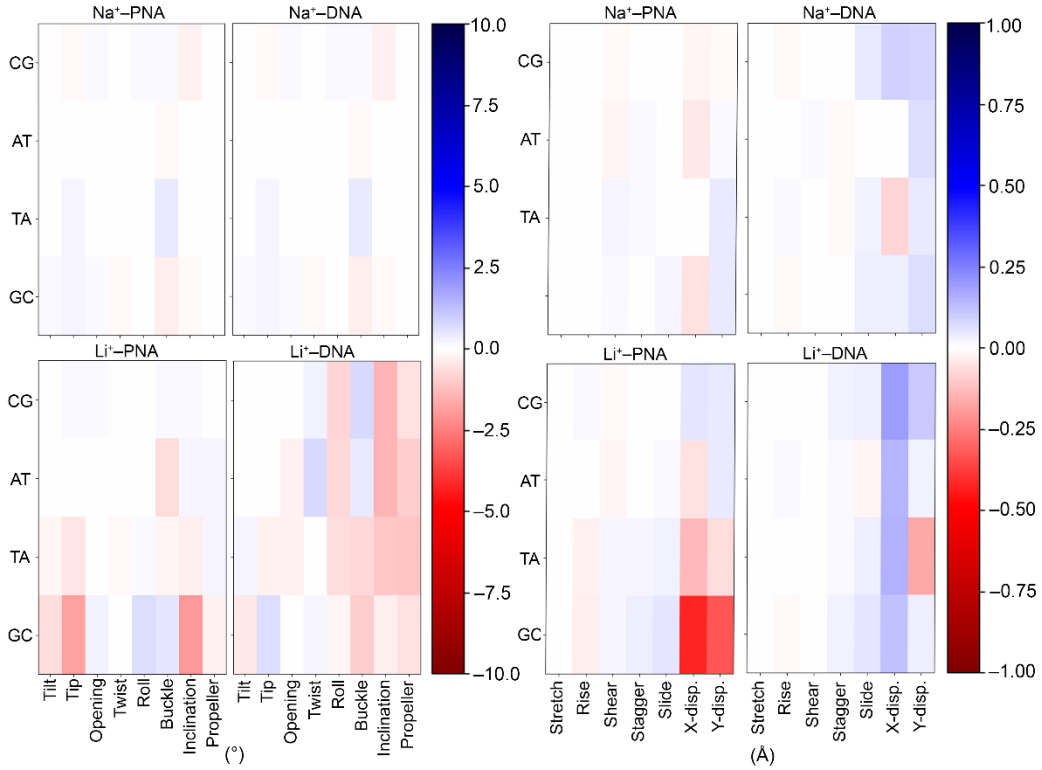


Figure 7.6. Deviations (°, left, and Å, right) in the nucleic acid structural parameters when Na^+ or Li^+ interacts with DNA or PNA in 100 mM of NaCl or LiCl relative to the corresponding duplex in the absence of salt.

7.3.4. Although the Preferred Nucleic Acid Binding Sites Remain Constant Regardless of Salt Concentration, the Structure of PNA Deviates Less than DNA in the Presence of Li^+ Compared to Na^+ Under Different Salt Concentrations

The effects of salt concentration on the structure of PNA and its ability to bind metals was studied, which is important information for designing therapeutics, materials, and nucleic acid sensing and extraction platforms. Increasing the salt concentration of NaCl from 100, 250, 500, to 750 mM had a minimal impact on metal binding distributions in

DNA and PNA (Figures 8.7A, F.2 and F.4, and Tables F.4–F.5, Appendix F). An additional salt concentration of 150 mM of NaCl was included to explore in physiological conditions, with this information being important in biological applications. Nevertheless, there was no change in the Na^+ binding events at this additional concentration compared to 100 mM (Figures 8.2 and 8.4). RDFs demonstrate that Na^+ binding interactions in the inner-sphere coordination remain the same regardless of salt concentration, while the outer-sphere interactions become saturated with Na^+ binding as the increasing salt levels (Figure F.3, Appendix F).

The total normalized binding events for Na^+ indicate there are minimal fluctuations in Na^+ binding to DNA and PNA (Figure F.2 and Table F.4, Appendix F). However, Na^+ binding at each nucleobase site is more susceptible to fluctuations caused by increasing the salt concentration. The Na^+ binding trends for the preferred sites remain the same relative to 100 mM, with both the DNA major and minor grooves being occupied by Na^+ (G(O6) and T(O2)), and G(O6) in PNA being dominant. The greatest fluctuation occurs at the T(O2) site in PNA, with ~22% more Na^+ binding from 100 to 750 mM. Increasing the NaCl concentration has a minimal effect on DNA structure (Figures F.6–F.8, Appendix F) but induces more structural deviations to PNA, especially highlighted by the inclination and roll (Figures F.6–F.8, Appendix F).

Like Na^+, increasing the LiCl concentration from 100, 250, 500, to 750 mM does not significantly affect the Li^+ binding distributions at each nucleobase site or DNA and PNA structure (Figure 8.7B and Figure F.4 and Tables F.4 and F.6, Appendix F). The RDFs show that inner-sphere Li^+ binding interactions remain consistent in varying LiCl concentrations, while outer-sphere interactions reach saturation as LiCl levels rise (Figure

F.3, Appendix F). The total normalized Li^+ binding events to DNA are more susceptible to fluctuations in changing LiCl concentrations. Indeed, as the salt levels rise, the Li^+ binding to the nucleobases and backbone increases (Figure F.2, Appendix F). The dominant nucleobase binding sites for Li^+ (G(O6) and T(O2) in DNA and G(O6) in PNA; Figure 8.7B and Figure F.5, Appendix F) are consistent regardless of LiCl concentration, with the difference being less than 10% at each binding site. Elevated LiCl levels cause slightly more deviations in DNA structure, while PNA is less susceptible than DNA to structural fluctuations in different LiCl concentrations, which is highlighted by the inclination, propeller, and buckle (Figures F.6 and F.9–F.10, Appendix F).

Overall, Na^+ and Li^+ binding distributions in PNA and DNA are generally consistent in changing salt concentrations, with the metal binding trends to each nucleobase site remaining consistent. Na^+ causes more structural deviations to PNA than Li^+. Alternatively, DNA is more distorted in the presence of Li^+ than Na^+. Since, Na^+ causes more structural deviations to PNA than Li^+, PNA may offer greater resistance to structural distortions caused by Li^+, suggesting that PNA is a more suitable material for Li^+-related applications in the development of therapeutics and materials.

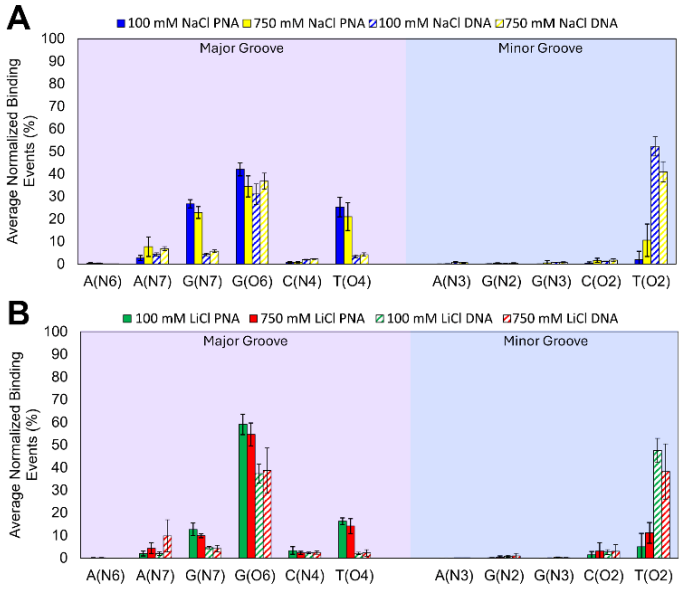


Figure 7.7. The average normalized (**A**) Na^+ or (**B**) Li^+ binding events (%) within 3 Å to each nucleobase site in the major and minor grooves of DNA (striped) or PNA (solid) in 100 or 750 mM of NaCl or LiCl.

7.3.5. Metal Binding to PNA is Sensitive to Sequence, with GC Pairs Being Favored

To ensure that metal binding distributions in PNA and DNA are the same regardless of sequence, the GC-rich sequence of PNA and DNA was altered to consider an AT-rich sequence. Changing the sequence context from GC- (5′-GGCATCGG-3′) to AT-rich (5′-AATGCTAA-3′) in DNA and PNA results in a visible change in the total Na^+ binding distributions throughout the PNA duplexes in 100 mM (Figure 8.2), with the Na^+ density exhibiting more localization at the G:C pairs. In contrast, sequence-dependent variations in Na^+ binding to DNA are minimal (Figure 8.2). The inner- and outer-sphere Na^+ binding interactions from the GC- to AT-rich sequence generally remain the same as shown by RDFs (Figures F.3 and F.12, Appendix F), with the magnitude of outer-sphere metal

binding to the nucleobases or backbone increasing with the salt concentration. The total normalized Na^+ binding distribution shows that substitution of G:C with A:T pairs in PNA leads to a reduction in Na^+ binding at G (by 13%) and increased total Na^+ binding at A (by 18%), while minimal Na^+ binding changes occur between G and A in DNA (Figure F.11A and Table F.7, Appendix F). The Na^+ binding distribution across nucleobase binding sites exhibit no change from the GC- to AT-rich sequence (Figure 8.8 and Figures F.4 and F.14, Appendix F).

While the metal binding distributions generally showed little variation between GC and AT-rich sequences in DNA and PNA, the largest sequence effect is the impact of metals on the structure of the duplex (Figures F.6–F.10 and F.15–F.18, Appendix F). Specifically, there are more deviations in the AT-rich than GC-rich sequence of DNA in the presence of Na^+, which is supported by previous findings that metal binding to different sequence contexts affects the conformational substates of DNA.[52] In contrast, Na^+ generally causes minimal deviations in PNA structure, except for a visible change in X-displacement (Figure F.16, Appendix F). This may be the result of ~1.6 times less Na^+ binding interactions from the GC- to AT-rich sequence in PNA.

Changing the sequence context from GC- to AT-rich in DNA and PNA alters Li^+ binding, with Li^+ density demonstrating even greater localization at the G:C pairs in PNA, while DNA displays minimal Li^+ binding differences between the GC- and AT-rich sequences (Figure 8.2). The RDFs show the inner- and outer-sphere Li^+ binding interactions are consistent in PNA and DNA, regardless of the sequence (Figures F.3 and F.12, Appendix F). The total normalized Li^+ binding distribution shows greater differences between GC- and AT-rich sequences for Li^+ binding at the nucleobases and backbone in

PNA, while minimal changes occur in DNA (Figure F.4 and Figure F.13 and Tables F.4 and F.7, Appendix F). Specifically, Li^+ binding is increased at A from the GC- to AT-rich sequence of PNA from 23% to 36% while only a subtle decrease of Li^+ binding occurs at G from 21% to 19%. Nevertheless, the Li^+ binding distribution at each nucleobase site remains the same regardless of the PNA or DNA sequence (Figure 8.8).

In general, there are similar levels of structural deviations caused by Li^+ binding in the AT-rich sequence of DNA compared to the GC sequence (Figures F.8–F.10 and F.17–F.18, Appendix F). Interestingly, the AT-rich sequence of PNA is more susceptible to structural deviations in the presence of Li^+ compared to the GC sequence, with the tip, inclination, twist, X- and Y-displacement demonstrating the largest deviations in the AT duplex (Figures F.8–F.10 and F.17–F.18, Appendix F). This result is consistent with prior crystallographic work that identified PNA as being generally susceptible to local disorder, while maintaining its global structure.[53]

Overall, Li^+ density is more localized at G:C regions of PNA compared to Na^+ regardless of a GC- or AT-rich sequence considered. Although the general metal binding behavior remains consistent regardless of sequence context of PNA or DNA, there are sequence-dependent differences in metal binding on the structure of PNA. The AT-rich sequence of PNA being more susceptible to structural deviations in the presence of Li^+ highlights the importance of selecting a GC-rich sequence to target Li^+.

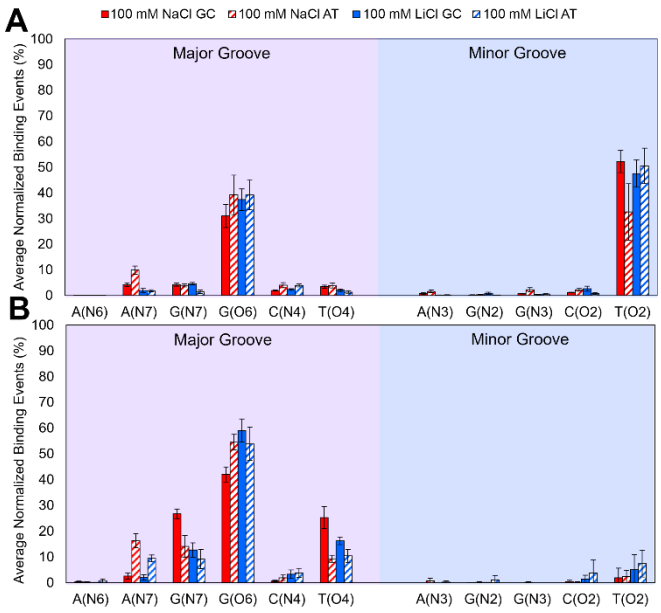


Figure 7.8. The average normalized Na^+ or Li^+ binding events (%) within 3 Å of (**A**) DNA or (**B**) PNA to each binding site of the nucleobases within the major and minor grooves of a GC- (solid) or AT-rich (striped) sequence in 100 mM of NaCl or LiCl.

7.4. Discussion

Although PNA is a powerful tool in gene therapies, pathogen detection, and biotechnologies,[14,18,22,54] how metals interact with PNA was poorly understood prior to this study. While studies utilizing experimental (NMR, X-ray crystallography) and computational techniques (MD) have uncovered the structural details and dynamics of PNA,[23,24,26,27,29] with the greater backbone flexibility of PNA over DNA or RNA revealed, in addition to PNA adopting a P-conformation. Furthermore, the helical twist of PNA in MD simulations deviated from X-ray crystal structures but aligned more closely with NMR data.[27] MD simulations also revealed new conformational features of PNA, which showed increased flexibility in the linker region between the nucleobases and backbone that was not detected in previous experimental studies.[27] Our MD simulations complement existing data by providing insights into how metals interact with different PNA components

(i.e., backbone, nucleobases) and influence the structure of PNA compared to DNA. Specifically, we have identified that PNA shifts metal binding from the backbone towards the nucleobases in the major groove, with emphasis on the O6 site in G for attracting Li^+. Additionally, the structure of PNA undergoes less deviations in the presence of metals compared to DNA, and the preferred nucleic acid binding sites for metals remain constant in different salt concentrations. Altogether, gaining these insights into metal–PNA interactions is necessary for understanding the behavior of PNA in the body and environment and expanding PNA functionalization to more applications in medicine, materials, and biotechnology.

7.4.1. Altering the charge of nucleic acids by modifying the backbone significantly affects metal distribution, paving the way for tuning nucleic acid function

Chemically modifying the negatively charged phosphate backbone of DNA to a neutral peptide backbone (PNA) significantly alters metal binding. Indeed, Na^+ and Li^+ binding is shifted from the backbone to localize at nucleobase binding sites. In DNA, Na^+ and Li^+ are distributed almost equally between the major groove and minor grooves, with both metals preferentially bound to O6 of G in the major and O2 of T in the minor groove. In contrast, metal binding in the minor groove of PNA is nearly absent, with both Na^+ and Li^+ preferentially binding in the major groove. Manipulation of metal binding to specific regions of nucleic acids can unlock enhanced functionality in different applications such as nucleic acid sensing platforms for metal detection in the environment and remediation for materials, and therapeutics that permit improved targeted delivery and gene regulation.[55,56] Chemically altering nucleic acids to enable the tuning of metal–nucleic acid interactions to

perform specific functions (e.g., metal capture, catalysis) has been done in the past.[3,57] For example, the metal binding preference of the Ce13d DNAzyme was altered from selectively targeting Ce^{3+} to Cu^{2+} by introducing a phosphorothioate modification.[58] As another example, a Hg^{2+}-specific nucleic acid sensor was designed by using noncanonical base pairs where the opposing A in an AT base pair is replaced with T, which enables the formation of T–T base mispairs that demonstrate strong selectivity for Hg^{2+}.[59,60] In material applications, DNA has been modified to form metal-mediated base pairs consisting of opposing hydroxypyridones to selectively bind Cu^{2+}, with this Cu^{2+}–DNA complex used in nanowire development.[61] Alternatively, DNA architectures have been created by introducing 2,2′-bipyridine modifications that form a three-way junction once Ni^{2+} is selectively bound.[62] A similar DNA architecture was constructed by introducing different metallohelices to form a Fe^{2+}-specific three-way junction, which was found to induce DNA damage in human colon cancer cells, thereby demonstrating therapeutic benefits.[63] Our results indicate that PNA demonstrates similar promise for region-specific metal binding, which may be useful in therapeutics, materials, or biosensing applications.

7.4.2. PNA demonstrates metal-specific behavior, with Li^+ being localized at the O6 site of G, while Na^+ is distributed across multiple binding sites

In addition to manipulating metal binding to favor the major groove side of PNA, we have identified metal binding site preferences for PNA. Specifically, Na^+ and Li^+ bind differently within the major groove of PNA. Na^+ is distributed across three major groove binding sites, O6 and N7 of G and O4 of T. In contrast, the dominant binding site for Li^+ is O6 of G, with significantly less binding to N7 of G and O4 of T. These results highlight

that chemically modifying the nucleic acid backbone can push certain metals (e.g., Li^+) to a specific binding site. This opens the possibility of PNA being a step in the right direction for developing a biosensor and extraction tool that targets Li^+.[3] There is a growing need for Li^+ sensing and extraction technologies since there are rising levels of Li^+ contamination in the environment, which can in turn result in excess levels of Li^+ in ecosystems and the human body, primarily due to mining activities and the production and disposal of lithium-ion batteries.[33] While biosensors have been designed to detect metals,[3] there is a growing need for environmentally-friendly solutions that also aid remediation and recovery efforts.[64] Given that nucleic acids are powerful environmentally-friendly biosensors, chemically modified nucleic acids can expand the functionality of biosensors to enhance sensing and extraction of metals, especially Li^+. Altering the charged backbone of canonical DNA to a neutral backbone (PNA) shifts Li^+ binding to a localized binding site (O6 in G), which demonstrates the promise of using PNA for Li^+ recognition. Nevertheless, additional modifications may be required to achieve selective capture of Li^+ since the O6 binding site in G is the preferred nucleobase binding site for many metals.[3,47] Our work suggests future studies should explore the introduction of chemical modifications to the major groove of PNA. Aside from sensing and extraction technologies, PNA could be used to transport Li^+ to specific therapeutic targets or to regulate Li^+ and Na^+ exchange in the body.[65] Indeed, Li^+ has known beneficial therapeutic effects for the treatment of Alzheimer's, Parkinson's, and Huntington's disease, bipolar disorder, and different types of cancers.[66-68] As a result, methods to selectively deliver Li^+ to cancer cells are required, with nucleic acid delivery systems having already been considered for canonical DNA.[69] PNA may be a solution to Li^+ delivery, with PNA having been shown to enter cells more easily than DNA.[70] Our results highlight that PNA demonstrates the potential to be used in platforms to selectively

administer Li^+ to target cells, in addition to being used towards the development of biosensors to detect and capture Li^+.

7.4.3. PNA demonstrates less structural distortion than DNA in the presence of metals at varying salt concentration

The impact of metal binding on the structure of DNA and PNA varies with metal identity. For DNA, Na^+ induces minimal structural deviations regardless of concentration, while Li^+ causes greater deviations. In contrast, PNA demonstrates minimal structural changes in varying Li^+ concentrations, while Na^+ leads to greater fluctuations in PNA structure. Nevertheless, the global structure of PNA is largely unaffected by varying salt concentrations compared to DNA, regardless of the metal. Additionally, the metal binding distributions to PNA are unaffected by different salt concentrations. Indeed, gradually increasing the salt concentration from 100 to 750 mM, including salt conditions that mimic a real-life environment such as sea water (~560–600 mM),[71] brine water for lithium (~200–600 mM),[72] and brackish water (~100–500 mM),[71] emphasizes the ability of PNA to resist structural distortions in diverse environments. The ability of PNA to resist structural deviations imposed by metals may be attributed to the neutral charged backbone, which reduces metal binding to the PNA backbone, with high levels of metal binding to DNA (particularly at the phosphate backbone) known to destabilize DNA duplexes.[73] The structural resilience of PNA in extreme salt concentrations as well as unwavering metal binding distributions highlights the ability of PNA to be used in different environmental conditions (sea and brine water, the human body), and therefore could be a robust platform for applications ranging from materials to therapeutics.

The minimal impact of metal binding on PNA structure may be attributed to the enhanced structural stability of PNA over DNA or RNA,[30] which has likely contributed to the success of PNA in therapeutic and biotechnological applications. For example, a PNA-based drug (OLP-1002) is used as a painkiller to treat osteoarthritis and works by PNA binding with pre-mRNA to block splicing enzymes and thereby preventing protein production.[70] Unlike anesthetics such as lidocaine that affect various sodium ion channels, PNA is a safer alternative, with greater selectivity and does not shut down sodium channels that control heart function.[70] PNA has also been used to create scaffolds for nanomaterial design.[74] For example, PNA has been incorporated into hydrogels to enhance hydrogel rigidity, with improved mechanical stiffness of hydrogels important for tissue engineering.[74] PNA has also proven to be effective in gene editing technologies such as CRISPR-Cas9, with PNA more easily entering cells and binding to DNA or RNA more strongly to more effectively silence genes.[75] Additionally, PNA does not cut the genomic DNA target,[70] which decreases the risk of off-target effects. Understanding how metals interact with PNA can contribute towards further advancing CRISPR-Cas9 gene-editing therapeutics. For example, previous work found that Mn^{2+} increases the catalytic efficiency of the Cas9 protein, thereby improving its gene editing capabilities.[76] Furthermore, the conformational state of Cas9 is dependent on metal and nucleic acid identity bound to Cas9.[77] Aside from the Cas9 protein, the structural stability of modified hybrid PNA-DNA/RNA duplexes for antisense applications was found to be enhanced under certain salt conditions,[78] which underscores how metal–PNA interactions could be used to strengthen gene-editing technology. Improving CRISPR performance can be useful towards advancing therapeutics as well as generating sustainable food sources. In therapeutics, reducing off-target effects can prevent harmful side-effects from CRISPR such as undesired

DNA damage and cytotoxicity in the cell.[79] Aside from therapeutics, sustainable food sources that can adapt to different environmental conditions due to climate change is a top global priority, and PNA coupled with gene editing technology may help create stress-tolerant plants.[80] Overall, our work provides insights as to why the structure of PNA is resistant to deviations in the presence of metals, information that can be used towards improving current PNA applications such as therapeutics and materials that function in varying salt conditions. Additionally, this highlights that PNA could be used under extreme salt conditions for biosensing and extraction applications such as sea and brine water.[75]

7.4.4. PNA demonstrates sequence-specific behavior in the presence of metals, with G:C base pairs being preferred metal binding sites regardless of sequence

Our work highlights that PNA demonstrates greater sequence-specific effects on metal binding than DNA. GC-rich sequences of PNA bind more metals than AT-rich sequences of PNA. Additionally, GC-rich sequences are less distorted in the presence of metals than AT-rich sequences. This highlights that choosing the sequence will be important for applications, with a GC-rich PNA sequence being more beneficial for detecting (biosensing), extracting, and delivering metals. This underscores that GC-rich regions are prime targets for introducing further modifications to specifically capture metals. It is also noteworthy that Li^+ binds more than Na^+ to GC-rich PNA, suggesting that such sequences provide a more optimal platform for developing tools to specifically detect Li^+. Understanding the impact of PNA sequence is also highly relevant for applications such as diagnostics and therapeutics, which rely on complementary nucleobase binding. For example, a PNA biosensor has been designed to detect the DNA sequence of the

hepatitis C virus.[81] PNA has been used to interfere with the splicing process by binding to specific pre-mRNA sequences, which prevents protein production that support tumor development in breast cancer.[82] PNA-based technologies have also shown promise in environmental applications, with a PNA probe coupled with fluorescence in situ hybridization (PNA-FISH) to detect *Legionella* pathogens in drinking water,[83] with these pathogens able to survive more saline environments.[84] Our data suggests such applications may be most successful when the PNA sequence is GC-rich due to enhanced metal binding and stability in the presence of varying metal concentrations.

7.5. Conclusions

In summary, we have carried out MD simulations to uncover how select metals applicable to biology and material applications interact with and affect the structure of PNA in comparison to DNA. Our simulations revealed that while Na^+ and Li^+ nearly equivalently bind to the DNA minor and major grooves, both Na^+ and Li^+ preferentially bind in the major groove due to the narrower minor groove width of PNA compared to DNA. Unlike Na^+, Li^+ binding in PNA is more localized at the O6 site of G than in DNA, which remains consistent in different salt concentrations. At the lowest salt concentration (100 mM), PNA undergoes less structural deviations in the presence of Li^+ than Na^+. Nevertheless, the global structure of PNA is maintained regardless of metal identity or salt concentration. Flipping the GC-rich to an AT-rich sequence does not impact the preferred binding sites of Li^+ and Na^+ in DNA or PNA. However, like DNA, the AT-rich duplex of PNA is susceptible to more structural deviations in the presence of Li^+ than Na^+, which highlights that a GC-rich sequence of PNA is preferable in applications requiring Li^+ binding. Our

results have provided a fundamental understanding of how metals bind to PNA, which can be used towards developing gene therapies and biotechnologies that can deliver metals and biosensors for metal detection and extraction. Furthermore, our results show that by modifying the nucleic acid backbone, we can manipulate metal binding towards specific sites in the duplex, which opens the door for future research into applications of PNA with Li^+ for developing Li^+-sensing and remediation technologies.

7.6. References

(1) Xu, W.; He, W.; Du, Z.; Zhu, L.; Huang, K.; Lu, Y.; Luo, Y. Functional Nucleic Acid Nanomaterials: Development, Properties, and Applications. *Angew. Chem. Int. Ed.* **2021**, *60* (13), 6890-6918.

(2) Luther, D. C.; Huang, R.; Jeon, T.; Zhang, X.; Lee, Y.-W.; Nagaraj, H.; Rotello, V. M. Delivery of Drugs, Proteins, and Nucleic Acids Using Inorganic Nanoparticles. *Adv. Drug Deliv. Rev.* **2020**, *156*, 188-213.

(3) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(4) Müller, J. Functional Metal Ions in Nucleic Acids. *Metallomics* **2010**, *2* (5), 318-327.

(5) Chen, Z.; Liu, C.; Cao, F.; Ren, J.; Qu, X. DNA Metallization: Principles, Methods, Structures, and Applications. *Chem. Soc. Rev.* **2018**, *47* (11), 4017-4072.

(6) Fantoni, N. Z.; Brown, T.; Kellett, A. DNA-Targeted Metallodrugs: An Untapped Source of Artificial Gene Editing Technology. *ChemBioChem* **2021**, *22* (13), 2184-2205.

(7) Yang, M.; Xie, Y.; Zhu, L.; Li, X.; Xu, W. Functional Nucleic Acid Enzymes: Nucleic Acid-Based Catalytic Factories. *ACS Catal.* **2024**, *14* (21), 16392-16422.

(8) McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. Recent Progress in Non-Native Nucleic Acid Modifications. *Chem. Soc. Rev.* **2021**, *50* (8), 5126-5164.

(9) Huang, P.-J. J.; Liu, J. An Ultrasensitive Light-up Cu^{2+} Biosensor Using a New DNAzyme Cleaving a Phosphorothioate-Modified Substrate. *Anal. Chem.* **2016**, *88* (6), 3341-3347.

(10) Huang, P.-J. J.; Liu, J. Rational Evolution of Cd^{2+}-Specific DNAzymes with Phosphorothioate Modified Cleavage Junction and Cd^{2+} Sensing. *Nucleic Acids Res.* **2015**, *43* (12), 6125-6133.

(11) Takezawa, Y.; Hu, L.; Nakama, T.; Shionoya, M. Metal-Dependent Activity Control of a Compact-Sized 8–17 DNAzyme Based on Metal-Mediated Unnatural Base Pairing. *Chem. Commun.* **2024**, *60* (3), 288-291.

(12) Rajasree, S. C.; Takezawa, Y.; Shionoya, M. Cu^{ii}-Mediated Stabilisation of DNA Duplexes Bearing Consecutive Ethenoadenine Lesions and Its Application to a Metal-Responsive DNAzyme. *Chem. Commun.* **2023**, *59* (8), 1006-1009.

(13) Chan, K. Y.; Kinghorn, A. B.; Hollenstein, M.; Tanner, J. A. Chemical Modifications for a Next Generation of Nucleic Acid Aptamers. *ChemBioChem* **2022**, *23* (15), e202200006.

(14) Dean, D. A. Peptide Nucleic Acids: Versatile Tools for Gene Therapy Strategies. *Adv. Drug Deliv. Rev.* **2000**, *44* (2), 81-95.

(15) Kauppinen, S.; Vester, B.; Wengel, J. Locked Nucleic Acid (LNA): High Affinity Targeting of RNA for Diagnostics and Therapeutics. *Drug Discov. Today Technol.* **2005**, *2* (3), 287-290.

(16) Saarbach, J.; Sabale, P. M.; Winssinger, N. Peptide Nucleic Acid (PNA) and Its Applications in Chemical Biology, Diagnostics, and Therapeutics. *Curr. Opin. Chem. Biol.* **2019**, *52*, 112-124.

(17) Oliveira, R.; Barbosa, A.; Sousa, M.; Azevedo, N. F.; Cerqueira, L.; Almeida, C. Using Peptide Nucleic Acid Fluorescence in Situ Hybridization (PNA-Fish) to Detect Campylobacter Spp. In Food Samples. *LWT* **2024**, *198*, 115922.

(18) Gao, A.; Wang, F.; Fan, Z.; Chen, S.; Wu, K. Reliable Breast Cancer Mirnas Detection with Enhanced Silicon Nanowire Biosensor by PNA Probe and Optical Calibration. *Sens. Actuat. B-Chem.* **2024**, *401*, 135011.

(19) Lavecchia di Tocco, F.; Botti, V.; Cannistraro, S.; Bizzarri, A. R. Detection of miR-155 Using Peptide Nucleic Acid at Physiological-Like Conditions by Surface Plasmon Resonance and Bio-Field Effect Transistor. *Biosens.* **2024**, *14* (2), 79.

(20) Barman, S. C.; Ali, M.; Hasan, E. A.; Wehbe, N.; Alshareef, H. N.; Alsulaiman, D. Smartphone-Interfaced Electrochemical Biosensor for Microrna Detection Based on Laser-Induced Graphene with π–π Stacked Peptide Nucleic Acid Probes. *ACS mater. lett.* **2024**, *6* (3), 837-846.

(21) Saini, S.; Goel, K.; Ghosh, S.; Das, A.; Saraogi, I. Effects of PNA Sequence and Target Site Selection on Function of a 4.5s Non-Coding RNA. *ChemBioChem* **2024**, *25* (11), e202400029.

(22) Aman, R.; Syed, M. M.; Saleh, A.; Melliti, F.; Gundra, Sivakrishna R.; Wang, Q.; Marsic, T.; Mahas, A.; Mahfouz, Magdy M. Peptide Nucleic Acid-Assisted Generation of Targeted Double-Stranded DNA Breaks with T7 Endonuclease I. *Nucleic Acids Res.* **2024**, *52* (6), 3469-3482.

(23) Yeh, J. I.; Pohl, E.; Truan, D.; He, W.; Sheldrick, G. M.; Du, S.; Achim, C. The Crystal Structure of Non-Modified and Bipyridine-Modified PNA Duplexes. *Chem. Eur. J.* **2010**, *16* (39), 11867-11875.

(24) Sen, S.; Nilsson, L. Molecular Dynamics of Duplex Systems Involving PNA: Structural and Dynamical Consequences of the Nucleic Acid Backbone. *J. Am. Chem. Soc.* **1998**, *120* (4), 619-631.

(25) Hatcher, E.; Balaeff, A.; Keinan, S.; Venkatramani, R.; Beratan, D. N. PNA Versus DNA: Effects of Structural Fluctuations on Electronic Structure and Hole-Transport Mechanisms. *J. Am. Chem. Soc.* **2008**, *130* (35), 11752-11761.

(26) He, W.; Hatcher, E.; Balaeff, A.; Beratan, D. N.; Gil, R. R.; Madrid, M.; Achim, C. Solution Structure of a Peptide Nucleic Acid Duplex from NMR Data: Features and Limitations. *J. Am. Chem. Soc.* **2008**, *130* (40), 13264-13273.

(27) Autiero, I.; Saviano, M.; Langella, E. Molecular Dynamics Simulations of PNA–PNA and PNA–DNA Duplexes by the Use of New Parameters Implemented in the Gromacs Package: A Conformational and Dynamics Study. *Phys. Chem. Chem. Phys.* **2014**, *16* (5), 1868-1874.

(28) Verona, M. D.; Verdolino, V.; Palazzesi, F.; Corradini, R. Focus on PNA Flexibility and RNA Binding Using Molecular Dynamics and Metadynamics. *Sci. Rep.* **2017**, *7* (1), 42799.

(29) Jasiński, M.; Feig, M.; Trylska, J. Improved Force Fields for Peptide Nucleic Acids with Optimized Backbone Torsion Parameters. *J. Chem. Theory Comput.* **2018**, *14* (7), 3603-3620.

(30) Jasiński, M.; Miszkiewicz, J.; Feig, M.; Trylska, J. Thermal Stability of Peptide Nucleic Acid Complexes. *J. Phys. Chem. B* **2019**, *123* (39), 8168-8177.

(31) Weroński, P.; Jiang, Y.; Rasmussen, S. Molecular Dynamics Study of Small PNA Molecules in Lipid-Water System. *Biophys J* **2007**, *92* (9), 3081-3091.

(32) Robinson, B. H.; Yalamanchali, R.; Reiser, R.; Dickinson, N. M. Lithium as an Emerging Environmental Contaminant: Mobility in the Soil-Plant System. *Chemosphere* **2018**, *197*, 1-6.

(33) Bolan, N.; Hoang, S. A.; Tanveer, M.; Wang, L.; Bolan, S.; Sooriyakumar, P.; Robinson, B.; Wijesekara, H.; Wijesooriya, M.; Keerthanan, S.; et al. From Mine to Mind and Mobiles – Lithium Contamination and Its Risk Management. *Environ. Pollut.* **2021**, *290*, 118067.

(34) DeLano, W. L. Pymol: An Open-Source Molecular Graphics Tool. *CCP4 Newslet. Pro. Crystallogr.* **2002**, (40), 82-92.

(35) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-EW. *J. Chem. Phys.* **2004**, *120* (20), 9665-9678.

(36) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12* (8), 4114-4127.

(37) Peng, T.; Chang, T.-M.; Sun, X.; Nguyen, A. V.; Dang, L. X. Development of Ions-TIP4P-EW Force Fields for Molecular Processes in Bulk and at the Aqueous Interface Using Molecular Simulations. *J. Mol. Liq.* **2012**, *173*, 47-54.

(38) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast Shake Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* **2001**, *22* (5), 501-508.

(39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684-3690.

(40) Davidchack, R. L.; Ouldridge, T. E.; Tretyakov, M. V. New Langevin and Gradient Thermostats for Rigid Body Dynamics. *J. Chem. Phys.* **2015**, *142* (14), 144114/144111.

(41) D.A. Case, H. M. A., K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, N. Forouzesh, M. Ghazimirsaeed, G. Giambaşu, T. Giese, M.K. Gilson, H. Gohlke, A.W. Goetz, J. Harris, Z. Huang, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, A. Kovalenko, T. Kurtzman, T.S. Lee, P. Li, Z. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, X. Wu, Y. Wu, Y. Xiong, Y. Xue, D.M. York, C. Zhao, Q. Zhu, and P.A. Kollman (2020), Amber 2020, University of California, San Francisco.

(42) Sigel, R. K. O.; Sigel, H. A Stability Concept for Metal Ion Coordination to Single-Stranded Nucleic Acids and Affinities of Individual Sites. *Acc. Chem. Res.* **2010**, *43* (7), 974-984.

(43) Várnai, P.; Zakrzewska, K. DNA and Its Counterions: A Molecular Dynamics Study. *Nucleic Acids Res.* **2004**, *32* (14), 4269-4280.

(44) Xi, K.; Wang, F.-H.; Xiong, G.; Zhang, Z.-L.; Tan, Z.-J. Competitive Binding of Mg^{2+} and Na^{+} Ions to Nucleic Acids: From Helices to Tertiary Structures. *Biophys J* **2018**, *114* (8), 1776-1790.

(45) Smith, J. A.; Keene, F. R.; Li, F.; Collins, J. G. 3.24 - Noncovalent DNA Binding of Metal Complexes. In *Comprehensive Inorganic Chemistry II (Second Edition)*, Reedijk, J., Poeppelmeier, K. Eds.; Elsevier, 2013; pp 709-750.

(46) Giambaşu, George M.; Luchko, T.; Herschlag, D.; York, Darrin M.; Case, David A. Ion Counting from Explicit-Solvent Simulations and 3D-RISM. *Biophys J* **2014**, *106* (4), 883-894.

(47) Boychuk, B. T. A.; Meyer, S. P.; Wetmore, S. D. Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* **2023**, *11*.

(48) Gebala, M.; Bonilla, S.; Bisaria, N.; Herschlag, D. Does Cation Size Affect Occupancy and Electrostatic Screening of the Nucleic Acid Ion Atmosphere? *J. Am. Chem. Soc.* **2016**, *138* (34), 10925-10934.

(49) Yoo, J.; Aksimentiev, A. Competitive Binding of Cations to Duplex DNA Revealed through Molecular Dynamics Simulations. *J. Phys. Chem. B.* **2012**, *116* (43), 12946-12954.

(50) Cruz-León, S.; Vanderlinden, W.; Müller, P.; Forster, T.; Staudt, G.; Lin, Y.-Y.; Lipfert, J.; Schwierz, N. Twisting DNA by Salt. *Nucleic Acids Res.* **2022**, *50* (10), 5726-5738.

(51) Yu, X.; Zhang, S.; Guo, W.; Li, B.; Yang, Y.; Xie, B.; Li, K.; Zhang, L. Recent Advances on Functional Nucleic-Acid Biosensors. *Sens.* **2021**, *21* (21), 7109.

(52) Heddi, B.; Foloppe, N.; Hantz, E.; Hartmann, B. The DNA Structure Responds Differently to Physiological Concentrations of K(+) or Na(+). *J Mol Biol* **2007**, *368* (5), 1403-1411.

(53) Kiliszek, A.; Banaszak, K.; Dauter, Z.; Rypniewski, W. The First Crystal Structures of RNA–PNA Duplexes and a PNA-PNA Duplex Containing Mismatches—toward Anti-Sense Therapy against Treds. *Nucleic Acids Res.* **2015**, *44* (4), 1937-1943.

(54) El-Fateh, M.; Chatterjee, A.; Zhao, X. A Systematic Review of Peptide Nucleic Acids (Pnas) with Antibacterial Activities: Efficacy, Potential and Challenges. *Int. J. Antimicrob. Agents* **2024**, *63* (3), 107083.

(55) Ochoa, S.; Milam, V. T. Modified Nucleic Acids: Expanding the Capabilities of Functional Oligonucleotides. *Mol.* **2020**, *25* (20), 4659.

(56) Egli, M.; Manoharan, M. Chemistry, Structure and Function of Approved Oligonucleotide Therapeutics. *Nucleic Acids Res.* **2023**, *51* (6), 2529-2573.

(57) Ma, L.; Liu, J. Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology. *iScience* **2020**, *23* (1).

(58) Saran, R.; Huang, Z.; Liu, J. Phosphorothioate Nucleic Acids for Probing Metal Binding, Biosensing and Nanotechnology. *Coord. Chem. Rev.* **2021**, *428*, 213624.

(59) Torigoe, H.; Ono, A.; Kozasa, T. Hgii Ion Specifically Binds with T:T Mismatched Base Pair in Duplex DNA. *Chem. Eur. J.* **2010**, *16* (44), 13218-13225.

(60) Wang, X.; Xu, C.; Wang, Y.; Li, W.; Chen, Z. Electrochemical DNA Sensor Based on T-Hg-T Pairs and Exonuclease III for Sensitive Detection of Hg^{2+}. *Sens. Actuat. B-Chem.* **2021**, *343*, 130151.

(61) Yang, H.; Metera, K. L.; Sleiman, H. F. DNA Modified with Metal Complexes: Applications in the Construction of Higher Order Metal–DNA Nanostructures. *Coord. Chem. Rev.* **2010**, *254* (19), 2403-2415.

(62) Takezawa, Y.; Sakakibara, S.; Shionoya, M. Bipyridine-Modified DNA Three-Way Junctions with Amide Linkers: Metal-Dependent Structure Induction and Self-Sorting. *Chem. Eur. J.* **2021**, *27* (67), 16626-16633.

(63) Malina, J.; Kostrhunova, H.; Scott, P.; Brabec, V. Metallohelices Stabilize DNA Three-Way Junctions and Induce DNA Damage in Cancer Cells. *Nucleic Acids Res.* **2023**, *51* (14), 7174-7183.

(64) Dhaliwal, S. S.; Singh, J.; Taneja, P. K.; Mandal, A. Remediation Techniques for Removal of Heavy Metals from the Soil Contaminated through Different Sources: A Review. *Environ. Sci. Pollut. Res. Int.* **2020**, *27* (2), 1319-1333.

(65) Dudev, T.; Mazmanian, K.; Lim, C. Competition between Li^+ and Na^+ in Sodium Transporters and Receptors: Which Na^+-Binding Sites Are "Therapeutic" Li^+ Targets? *Chem. Sci.* **2018**, *9* (17), 4093-4103.

(66) Singulani, M. P.; Ferreira, A. F. F.; Figueroa, P. S.; Cuyul-Vásquez, I.; Talib, L. L.; Britto, L. R.; Forlenza, O. V. Lithium and Disease Modification: A Systematic Review and Meta-Analysis in Alzheimer's and Parkinson's Disease. *Ageing Res. Rev.* **2024**, *95*, 102231.

(67) Villegas-Vázquez, E. Y.; Quintas-Granados, L. I.; Cortés, H.; González-Del Carmen, M.; Leyva-Gómez, G.; Rodríguez-Morales, M.; Bustamante-Montes, L. P.; Silva-Adaya, D.; Pérez-Plasencia, C.; Jacobo-Herrera, N.; et al. Lithium: A Promising Anticancer Agent. *Life* **2023**, *13* (2), 537.

(68) Shim, S. S. B., K.; Yu, S. P. Lithium: An Old Drug for New Therapeutic Strategy for Alzheimer's Disease and Related Dementia. *J. Neurodegener. Dis.* **2023**, *23*, 1-12.

(69) Gupta, A.; Andresen, J. L.; Manan, R. S.; Langer, R. Nucleic Acid Delivery for Therapeutic Applications. *Adv. Drug Deliv. Rev.* **2021**, *178*, 113834.

(70) Brazil, R. Peptide Nucleic Acids Promise New Therapeutics and Gene Editing Tools. *ACS Cent. Sci.* **2023**, *9* (1), 3-6.

(71) Du, J. R.; Zhang, X.; Feng, X.; Wu, Y.; Cheng, F.; Ali, M. E. A. Desalination of High Salinity Brackish Water by an Nf-Ro Hybrid System. *Desal.* **2020**, *491*, 114445.

(72) Bradley, D.; Munk, L.; Jochens, H.; Hynek, S.; Labay, K. A. *A Preliminary Deposit Model for Lithium Brines*; Reston, VA, 2013. https://pubs.usgs.gov/publication/ofr20131006

(73) Anastassopoulou, J. Metal–DNA Interactions. *J. Mol. Struct.* **2003**, *651-653*, 19-26.

(74) Sarkar, S. Recent Advancements in Bionanomaterial Applications of Peptide Nucleic Acid Assemblies. *Biopolymers* **2024**, *115* (2), e23567.

(75) Economos, N. G.; Quijano, E.; Carufe, K. E. W.; Perera, J Dinithi R.; Glazer, Peter M. Antispacer Peptide Nucleic Acids for Sequence-Specific CRISPR-Cas9 Modulation. *Nucleic Acids Res.* **2022**, *50* (10), e59-e59.

(76) Zhang, C.; Wang, X.; Liu, G.; Ren, H.; Li, J.; Jiang, Z.; Liu, J.; Lovell, J. F.; Zhang, Y. Metal Coordination Micelles for Anti-Cancer Treatment by Gene-Editing and Phototherapy. *J. Contr. Release* **2023**, *357*, 210-221.

(77) Bhattacharya, S.; Satpati, P. Insights into the Mechanism of Crispr/Cas9-Based Genome Editing from Molecular Dynamics Simulations. *ACS Omega* **2023**, *8* (2), 1817-1837.

(78) De Costa, N. T. S.; Heemstra, J. M. Evaluating the Effect of Ionic Strength on Duplex Stability for PNA Having Negatively or Positively Charged Side Chains. *PLOS ONE* **2013**, *8* (3), e58670.

(79) Asmamaw Mengstie, M.; Teshome Azezew, M.; Asmamaw Dejenie, T.; Teshome, A. A.; Tadele Admasu, F.; Behaile Teklemariam, A.; Tilahun Mulu, A.; Mekonnen Agidew, M.; Adugna, D. G.; Geremew, H.; et al. Recent Advancements in Reducing the Off-Target Effect of CRISPR-Cas9 Genome Editing. *Biologics* **2024**, *18*, 21-28.

(80) Angon, P. B.; Mondal, S.; Akter, S.; Sakil, M. A.; Jalil, M. A. Roles of Crispr to Mitigate Drought and Salinity Stresses on Plants. *Plant Stress* **2023**, *8*, 100169.

(81) Singh, G.; Monga, V. Peptide Nucleic Acids: Recent Developments in the Synthesis and Backbone Modifications. *Bioorg. Chem.* **2023**, *141*, 106860.

(82) Pradeep, S. P.; Malik, S.; Slack, F. J.; Bahal, R. Unlocking the Potential of Chemically Modified Peptide Nucleic Acids for RNA-Based Therapeutics. *RNA* **2023**, *29* (4), 434-445.

(83) Nácher-Vázquez, M.; Barbosa, A.; Armelim, I.; Azevedo, A. S.; Almeida, G. N.; Pizarro, C.; Azevedo, N. F.; Almeida, C.; Cerqueira, L. Development of a Novel Peptide Nucleic Acid Probe for the Detection of *Legionella* Spp. In Water Samples. *Microorganisms* **2022**, *10* (7), 1409.

(84) Heller; Höller; Süßmuth; Gundermann. Effect of Salt Concentration and Temperature on Survival of Legionella Pneumophila. *Lett. Appl. Microbiol.* **1998**, *26* (1), 64-68.

Chapter 8: Designing a Crown Ether–Nucleic Acid Adduct for Targeting Lithium: A Density Functional Theory and Ab Initio Molecular Dynamics Approach

8.1. Introduction

Lithium is an emerging metal contaminant in the environment, which is largely due to the production and disposal of lithium-ion batteries, in addition to mining activities for lithium extraction.[1,2] Lithium is a highly mobile element in soil and plants, causing the element to readily travel through ecosystems.[3,4] As a result, increasing levels of lithium are found in plants and water,[4] leading to adverse effects on ecosystems and human health.[3] Excess levels of lithium in the environment have demonstrated toxic effects such as oxidative stress in plants, animals, and aquatic life.[5-7] Although lithium is used as a drug to treat bipolar disorder[8] and was more recently found to be effective for slowing cancer tumor and ALS progression,[8-10] lithium is only beneficial to human health in a narrow concentration range of 0.5–1.0 mM.[11] Accumulation of lithium in the body above the recommended dosage can cause hypothyroidism, hyperparathyroidism, permanent brain damage, and kidney disease.[3,11] Since lithium contamination is a growing global concern, it is important to have technologies that can detect and capture lithium in the environment and the human body.

Nucleic acids are a promising onsite green-tool to aid metal capture and extraction.[12] Indeed, nucleic acids contain a negatively charged phosphate moiety, nitrogenous bases (nucleobases), and a sugar, with each component capable of attracting metal ions with varying binding affinity.[13] Additionally, nucleic acids are structurally versatile, demonstrating the ability to fold into diverse 3D structures with distinct functions. For example, G-quadruplexes, DNAzymes, and aptamers can interact with and use metals

in different ways (e.g., stability, catalysis, and metal capture).[14,15] As a result, nucleic acids have been designed to detect various metals (e.g., Pb^{2+}, Cd^{2+}, Ni^{2+}, Cu^{2+}, and Cs^{+}).[12,16] Although nucleic acid platforms exist for capture of several metals, there are no nucleic acid solutions to date that for the detection and capture of lithium.

Designing a nucleic acid platform that is specific for lithium requires fundamental structural and energetic information about how lithium binds with nucleic acid components. Previous experimental and computational gas-phase studies uncovered that group I metals preferentially interact with the phosphate moiety,[17-19] however, this preference does not provide selectivity. Among the nucleobase sites, guanine (G) is the strongest binding position for lithium.[19-21] Within G, the strongest binding sites for lithium are O6 and N7.[21] Additional work that explored the role of water on lithium–nucleic acid interactions concluded the preferred binding sites remain consistent in different solvents.[22] Nevertheless, other metals are known to compete with lithium for nucleic acid binding sites,[19,23,24] which interferes with selective lithium binding for targeted applications.

To enhance the selectivity of nucleic acids for particular metals, incorporation of chemical modifications has been shown to be an effective approach. Nucleobase modifications have shown great potential for developing new materials and biosensing technologies, especially in applications involving metals.[12,25] For example, replacing canonical nucleobases with hydroxypyridone nucleobases enabled the formation of Cu^{2+}–mediated base pairs, which have proven useful for developing DNA nanowires and magnets.[26] Alternatively, a Cs^{+}-specific detection strategy was designed by swapping carbonyls with amino groups in a G-quadruplex.[27] A Cd^{2+}-specific nucleic acid sensing

platform was developed by introducing a phosphorothioate modification at a DNAzyme cleavage site.[25]

To achieve the goal of lithium capture using nucleic acid technology, a nucleobase modification that selectively binds lithium is required. In this light, crown ethers (CE) have proven to afford designs that can selectively bind alkali metals,[28,29] with the ring size adjustable to capturing smaller-radius metals. Indeed, several computational studies have employed density functional theory (DFT) and ab initio molecular dynamics (AIMD) to uncover how metals, including group I metals, interact with crown ethers and the dynamics of these interactions.[28,30] These studies have provided insights into how ring size and metal ion radius impact metal binding selectivity and affinity, in addition to how metal interactions within the crown ether fluctuate over time, emphasizing the importance of using DFT and AIMD techniques for investigating metal–crown ether interactions.

Among crown ether candidates, experimental and computational studies have highlighted that 14-crown-4 is particularly selective towards lithium.[31-33] In fact, 14-crown-4 has been successfully used in a nanofiber for lithium recovery from discarded lithium-ion batteries.[34] The demonstrated success of the 14-crown-4 for capturing lithium emphasizes its potential for incorporating into nucleic acids and developing a nucleic acid platform for sensing and extraction lithium. This platform would enable monitoring of lithium in the environment and in the human body. Building upon the success of 14-crown-4, a previous combined experimental and computational study identified improved crown ether variants, with dibenzo-14-crown-4 and its derivatives emerging as strong candidates due to ideal cavity dimensions, while maintaining the preferred coordination numbers of lithium (4 or 6). The CE 3d was identified as the most selective

variant, making it a strong candidate for incorporation into a nucleic acid. Attachment of crown ethers to nucleic acids has been done in prior work to monitor DNA damage,[35] with a synthetic route to conjugate crown ethers to nucleobases established, highlighting the viability of modifying nucleic acids with crown ethers.

To work towards developing a nucleic acid sensing and extraction platform for lithium, we explore the ability of CE incorporated into nucleic acids to capture lithium. Specifically, the most selective CE 3d variant for lithium is attached to C5 site in C to form a crown-ether-cytosine (CEC) adduct (Figure 8.1). C was selected as the modification site since G is a strong binder for lithium. The CEC adduct is assessed to whether it can effectively retain lithium compared to the strongest nucleic acid binding sites (i.e., G(O6–N7) and the phosphate moiety). Initially, DFT calculations are performed on nucleobase models of lithium interacting with the CEC center or the O6–N7 site in G. Subsequently, the nucleobase models are extended to incorporate the complementary Watson-Crick-Franklin (WCF) hydrogen-bonded partner (G), forming a crown ether base pair (CEBP). To consider the extended nucleic acid component, part of the nucleic acid backbone is considered for canonical (DNA) and a select modified (peptide nucleic acid; PNA) nucleic acid. PNA, which is a modification that replaces the negatively charged phosphate backbone in DNA to a neutral peptide backbone. In fact, PNA has been found in previous work to push lithium from the minor groove to the major groove in PNA, particularly the O6 site in G. DFT calculations are performed on these models to gauge Li^+–CE and Li^+–G(O6,N7) binding strengths in a larger nucleic acid framework, and AIMD simulations are subsequently carried out to understand the dynamics of Li^+ interactions with CECNBP (Figure 8.1). Finally, Our work builds upon previous DFT and ab initio

studies that have strived to understand lithium–solvent interactions and to improve our understanding of electrolyte solutions to enhance battery performance[36-38] by providing insights into lithium interactions with nucleic acids in explicit solvent. Our findings highlight a novel modified nucleic acid capable of capturing lithium as a proof of principle for the development of a lithium sensing and extraction method.

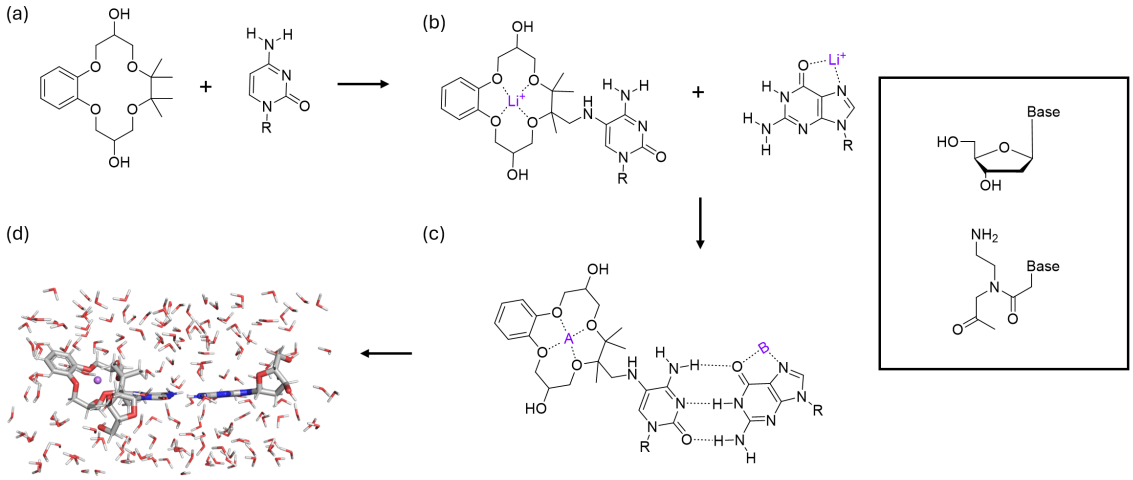


Figure 8.1. Schematic outlining the general computational workflow in the present work, with calculations performed on models of (a) CE, (b) CEC and G, (c) CECBP (A shows Li^+ positioned at the crown ether center and B shows Li^+ bound to G(O6–N7) and (d) CECBP, CECNBP, or CECPBP. DFT calculations were performed on parts (a) to (c) to elucidate structural and binding information, while AIMD simulations were carried out on part (d) for the CECNBP model to provide insights into the dynamics of these complexes. See Figure G.1 in Appendix G for a diagram of the nucleoside and peptide connection in the CECNBP and CECPBP models.

8.2. Computational Methodology

8.2.1. Model Building

The 14-crown-ether-4 model was built using the BEGVAL crystal structure (Figure G.2, Appendix G) from the Cambridge Structural Database,[39] with the additional OH group

added to yield the structure shown in Figure 8.1 to create the CE 3d variant. Models were constructed where Li^+ was placed at the center of the 14-crown-ether, which is attached through an alkylamine linkage to cytosine (CEC) to mirror the synthesis of the 2′-deoxycytidine system.[40] For comparison, a model was also constructed with Li^+ bound at the O6–N7 site of G, the most favorable nucleobase binding site.[18] Subsequently, the nucleobase models were expanded to include the opposing Watson-Crick-Franklin nucleobase (G) to form a base pair (CECBP). Methyl caps at N1 of C and N9 of G were used in DFT calculations on nucleobase and base pair models. To mimic a broader nucleic acid environment, the base pair models were expanded to incorporate the deoxyribose to form nucleoside models (CECNBP) for canonical and peptide models (CECPBP) for modified nucleic acids, respectively. All nucleobase, base pair, and nucleoside and peptide models were built using GaussView 5. Finally, the CE modification is incorporated into an 8-mer duplex of DNA and PNA to ensure the feasibility of including this modification in a nucleic acid. For comparison, additional models were constructed of unmodified DNA and PNA duplexes, and CE-modified DNA and PNA were built with Li^+ in the CE center. The CE modification was incorporated into an 8-mer of a DNA and PNA duplex model using PyMOL (Version 3.0.3), with the PNA structure used from a crystal structure (PDB ID: 3MBS).[41] To neutralize the DNA duplex, 14 Na^+ ions were added near the phosphate backbone to provide charge stabilization.

8.2.2. DFT Calculations

DFT calculations were initially performed in the gas phase on the 14-crown-ether-4 model to compare to previous literature.[42] Geometry optimizations were

carried out using MN15[43] and PBE-D3(BJ)[44] with def2-SVP[45] and frequency calculations at the same level of theory confirmed all stationary points are minima (i.e., no negative/imaginary frequencies). Both MN15 and PBE-D3(BJ) have been identified to be reliable functionals for describing Li^+–nucleic acid interactions.[18] Single-point calculations were performed using MN15 and PBE-D3(BJ) with def2-TZVPP. All subsequent DFT calculations were carried out in implicit solvent (i.e., water) using the Integral Equation Formalism Polarizable Continuum Model (IEF-PCM).[46] The binding energy between Li^+ and each component was calculated as follows

$$\Delta E_{BE} = E_{Li-NA} - E_{Li} - E_{NA} \quad (8.1)$$

where E_{Li-NA} is the energy of Li^+ complexed with CE, CEC, or CECBP, while E_{Li} represents bare Li^+, and E_{NA} is CE, CEC, CECBP, or CECNBP. All DFT calculations were carried out using Gaussian 16.[47] The DNA and PNA duplex models were optimized using the GFN2-xTB[48] semi-empirical method in implicit water using the conductor polarizable continuum (CPCM) model in ORCA 6.0.1.[49] GFN2-xTB has been shown to provide accurate structural descriptions of larger molecules (~1000 atoms).[48]

8.2.3. AIMD simulations

The CECNBP model with Li^+ was placed in a periodic cell of dimension 25.0 Å x 21.0 Å x 14.0 Å, which contained 171 waters. Waters were randomly placed using tLEAP from the AMBER suite.[50] To reduce computational cost, PBE-D3(BJ) was employed. The molecular optimized double-zeta basis set was used with core-electron Goedecker-Teter-Hutter (GTH) pseudopotentials (DZVP-MOLOPT-GTH). The plane-wave grid cutoff was

set to 400 Ry and the self-consistent-field (SCF) convergence threshold was set to 1.0 x 10^{-7} au with 30 SCF cycles per step. AIMD simulations in the NVT ensemble were performed at 310 K using the GLE thermostat. A 1.0 fs time step was employed to integrate the equations of motion. All simulations were carried out for 30 ps. The coordination number of Li^+ over the course of each AIMD simulation was calculated based on a 3.0 Å cutoff for Li^+–CEC(O) distances and Li^+–G(O6–N7).[51] Additionally, all distances between Li^+ and CE or G are calculated using all data points. All AIMD simulations were carried out using the Quickstep engine of CP2K (Version 8.2). The visualization software used was PyMOL (Version 3.0.3).

8.3. Results and Discussion

8.3.1. CE Strongly Binds to Li^+ in Implicit Water

The calculated gas-phase structure of CE bound to Li^+ shows 5 coordination distances between Li^+ and the 4 central oxygens of CE and one hydroxyl oxygen (Figure 8.2), with an average coordination distance of 2.034 Å (Table G.1, Appendix G). In contrast, a prior study on the same Li^+–CE system identified that Li^+ interacts with the four oxygens within the CE, while a hydroxyl proton is pointed towards Li^+.[42] This discrepancy is most likely due to different computational methodologies used, with the prior study[42] utilizing plane-wave DFT with the PBE functional and a relatively low energy cutoff, which have been attributed to incorrect description of electron density and bonding interactions,[52] whereas the calculations in present work used DFT with atomic orbital basis sets. In the absence of Li^+, the CE monomer resembles the previous study structure where

a hydroxyl proton is facing the 4 center oxygens (Figure G.3a, Appendix G). The calculated gas-phase binding strength of Li^+–CE is –101.0 kcal/mol, which is 28.5 kcal/mol stronger than Li^+ bound at the O6–N7 site in G, the strongest Li^+–nucleobase interaction.[18]

Since the goal in this study is to design a biosensor for detecting and capturing Li^+ in the human body and ecosystems, a gas-phase environment is not realistic, therefore, the structure and binding strength of Li^+–CE in implicit water is calculated. Overlay comparisons of gas-phase and implicit water optimized Li^+–CE structures show that Li^+–CE complexes remain largely unchanged in either environment (Figures G.3b and G.3c, Appendix G). The structure of the unbound CE in implicit water closely resembles the Li^+-CE complex (Figure G.3d, Appendix G), while the gas-phase structure of unbound CE shows greater distortion compared to the bound Li^+–CE complex (Figure G.3a, Appendix G), highlighting that implicit water stabilizes the CE structure. This stabilization is further supported by the overlay of gas-phase and implicit water structures of the unbound CE (Figure G.3d, Appendix G). Like the gas-phase structures, a Li^+–CEC complex forms a 5-coordinate structure in implicit water, with Li^+ interacting with the 4 center oxygens and a hydroxyl oxygen at an average coordination distance of 2.100 Å (Table G.1, Appendix G). The binding strength of the Li^+–CE complex is reduced significantly by ~6 times from gas-phase (–101.0 kcal/mol) to implicit water results (–18.2 kcal/mol). This reduced binding strength is supported by increased coordination distances between Li^+ and CE (Table G.1, Appendix G), which is due to the dielectric water medium screening electrostatic interactions that weakens binding. Given that implicit solvent impacts Li^+–CE interactions, all further calculations are conducted in implicit water.

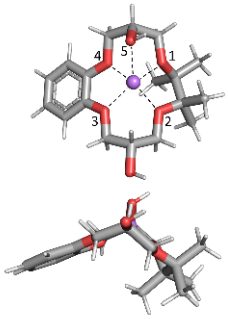


Figure 8.2. DFT-optimized Li^+–CE complex in implicit water. The numbering of the oxygen atoms shown here is consistently used throughout our analysis.

8.3.2. The CE-modified Nucleobase Binds to Li^+ Stronger than the most Prominent Canonical Nucleobase Site (G(O6–N7))

Using the synthetic approach outlined in previous work,[40] the CE is subsequently incorporated into the nucleic acid through an alkylamino linkage between C5 of C and CE to form CEC. Like the isolated CE, Li^+ forms a 5-coordinate structure at the center of the CE, with an average Li^+–CEC distance of 2.095 Å (Figure 8.3a; Tables G.2–G.3, Appendix G). The binding strength of the Li^+–CEC complex is slightly weaker than in the isolated CE by 2.1 kcal/mol. G binds to Li^+ through a bidentate interaction at the O6–N7 site, where the Li^+–O6 and Li^+–N7 distances are 2.080 Å and 2.318 Å, respectively. Notably, the Li^+–CEC binding strength is ~3 times greater than Li^+–G(O6–N7). These results highlight the potential of this novel crown ether adduct to create a Li^+-selective site in a nucleic acid.

8.3.3. The Opposing Nucleobase in a WCF Base Pair does not Alter the Stronger Binding Affinity of Li^+ for CEC Compared to G

Since there are stronger interactions between Li^+ and CEC than G, we expanded our CEC model to include the opposing nucleobase, forming a modified base pair (CECBP).

CEC is well accommodated within the base pair, with no disruption in hydrogen bonding when CEC is paired with G compared to a canonical G:C base pair (Figure 8.3b; Tables G.2–G.3, Appendix G). The Li^+–CEC interaction does not change from the isolated CEC to the CEC:G base pair, with an average Li^+–CEC distance of 2.098 Å and a modest increase in the Li^+–CEC binding strength by 0.3 kcal/mol. Like the isolated Li^+–G complex, the bidentate interaction between Li^+ and the CEC:G base pair at the O6–N7 site is maintained with a slightly elongated Li^+–O6 distance by 0.011 Å and a shortened Li^+–N7 distance by 0.046 Å compared to isolated G. These distances are enhanced in a canonical G:C pair where Li^+–O6 elongates by 0.034 Å and Li^+–N7 shortens by 0.033 Å compared to the CEC:G base pair. Comparing Li^+ bound at either the center of CEC or O6–N7 of G, we found that the binding strengths are –16.4 and –6.0 kcal/mol, respectively. Given that Li^+–CEC interactions are greater than the strongest nucleobase site being G:C by ~3 times, this suggests the CE will likely be the nucleic acid site to capture Li^+.

The base pair model is further expanded to include a part of the nucleic acid backbone. Specifically, the CEC model is expanded to include the deoxyribose to form the nucleoside model (CECNBP; Figure 8.3), enabling the investigation of Li^+–CEC interactions within a larger nucleic acid framework, which is important information for designing a lithium-specific sensing and extraction platform. The Li^+–CE and Li^+–G interactions remain unchanged in the presence of the sugar components (Figure 8.3), with slightly weaker binding strengths than the base pair models of –13.6 and –5.9 kcal/mol, respectively. Even when the nucleoside component is replaced with a peptide linkage to mimic PNA, the Li^+ interactions with the CE or G remains unchanged, with binding strengths of –14.0 and –5.8 kcal/mol for Li^+–CECNBP and Li^+–CECPBP (Figure G.5,

Appendix G), respectively. As a result, the CE remains the strongest site to attract Li^+, even within a larger nucleic acid framework.

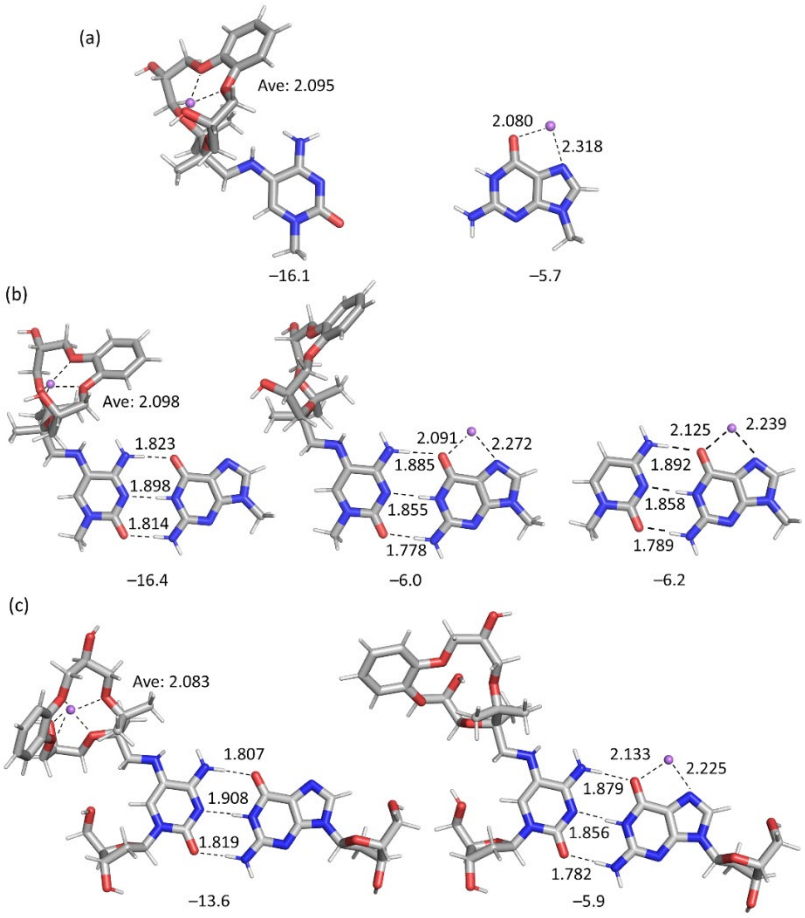


Figure 8.3. DFT-optimized structures in implicit water of nucleobase models with (a) Li^+ interacting at the CE center (left) or G(O6–N7) (right), and base pair models of (b) Li^+–CEC (left) and Li^+–G(O6–N7) (right). Distances and binding energies are reported in angstrom (Å) and kcal/mol.

8.3.4. AIMD Simulations Show that Li^+ Remains Centered in the CE in Solvated Nucleoside Models

The next step is to identify whether Li^+ remains bound at the center of CECNBP through the course of the natural dynamics of the Li^+–CECNBP complex in an explicitly

solvated environment, as well as compare this coordination to the complex with Li^+ bound at the O6–N7 site in G. Since there are no differences in Li^+ binding to the CE or G(O6–N7) in the extended nucleoside and peptide models, CECNBP was selected as a representative system. AIMD simulations were initially conducted with Li^+ positioned at the CE center in the CECNBP. After extensive 30 ps simulations, Li^+ remains bound to the CE center (Figure 8.4), with consistent coordination distances (average of 2.010 ± 0.131 Å between Li^+ and the center CE oxygens) relative to the implicitly solvated Li^+–CECNBP (Figure 8.5). However, the OH group of the CE bound to the metal seen in implicit solvent DFT calculations dissociated from Li^+, being displaced from the opposite side by a nearby water molecule to form a square pyramidal geometry (Figure 8.4). For most of the simulation, Li^+ adopts a coordination number (CN) of 5, with a square pyramidal geometry (Figure 8.5), while a coordination number of 6 is occasionally adopted, which featured simultaneous coordination of a water and an OH group attached to the CE to the Li^+ center (Figure 8.5). The WCF base pair is maintained over the course of the simulation as demonstrated by consistent average hydrogen bonding distances (<2 Å; Table G.4, Appendix G). Overall, the structural changes between implicit and explicit solvent mediums for Li^+ interacting with the CE center highlight the importance of simulating metal complexes in an explicitly solvated environment.

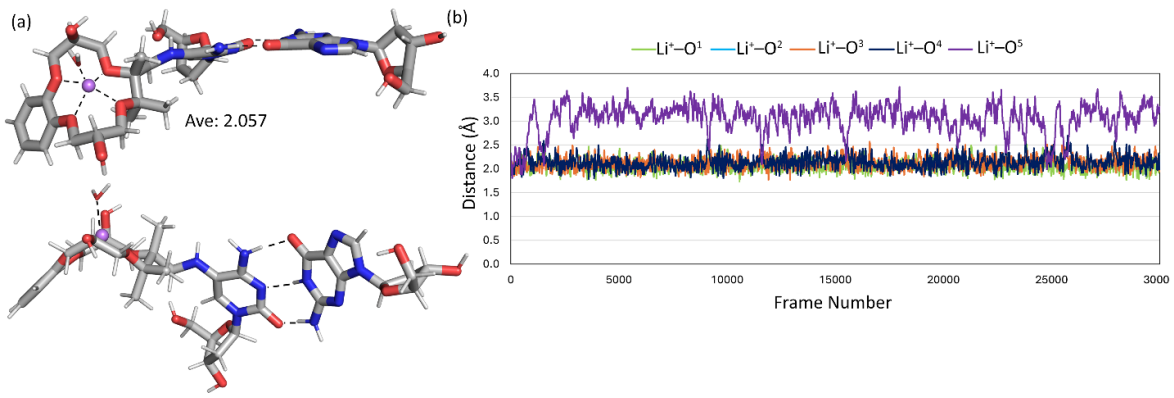


Figure 8.4. AIMD-simulated Li^+–CECNBP complex showing (a) the structure of the dominant 5-coordinate complex and (b) the distance (Å) between each oxygen within the CE and Li^+ as a function of frame.

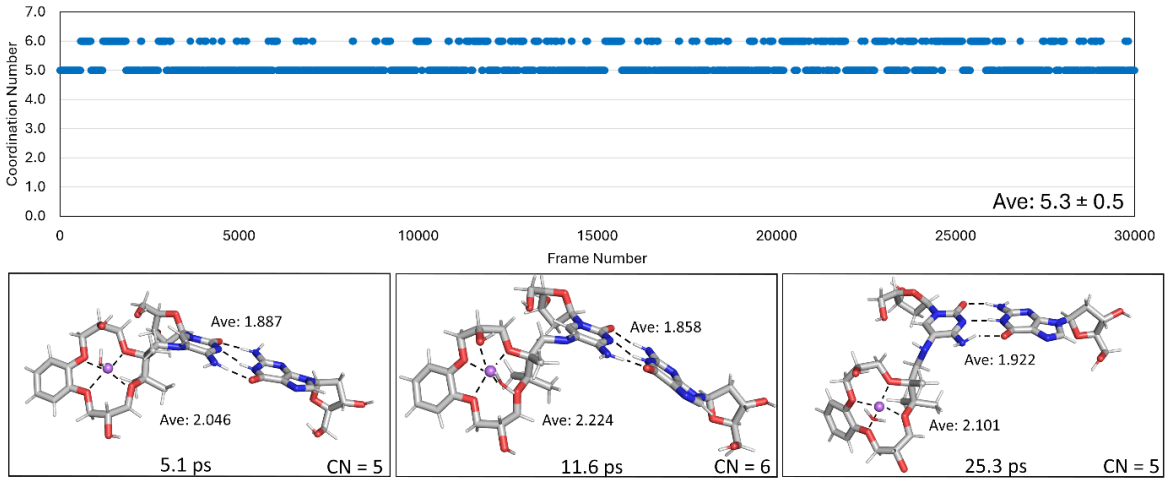


Figure 8.5. Li^+ coordination number over each AIMD-simulated frame for Li^+–CECNBP. The calculated average coordination number over each simulation is shown in the bottom right-hand corner. Example snapshots of Li^+–CECNBP interactions from different frames are provided below, in addition to average coordination distances (Å) between Li^+ and CE and average WCF hydrogen bond distances over the selected frame.

Next, AIMD simulations were performed with Li^+ placed at the O6–N7 site of G. After ~2 ps, Li^+ dissociates from the O6 site of G (Figures 8.6–8.7). After 14 ps, Li^+ also detaches from N7 (Figures 8.6–8.7). While the Li^+ coordination numbers fluctuate from 3 to 6 during the initial phase of the simulation (< 5 ps), Li^+ generally maintains the average

coordination number of 4 with waters after Li^+ dissociates from G (Figure 8.7). These results align with prior experimental and computational studies on Li^+–nucleic acid interactions, which found that Li^+ can bind to and dissociate from nucleic acids before diffusing into surrounding solvent,[24,53,54] emphasizing the importance of performing computational simulations to capture the dynamics of these interactions. The retention of Li^+ bound to the CE center couple Li^+ dissociation from G(O6–N7) emphasizes the ability of crown ether adducts in nucleic acids to target Li^+.

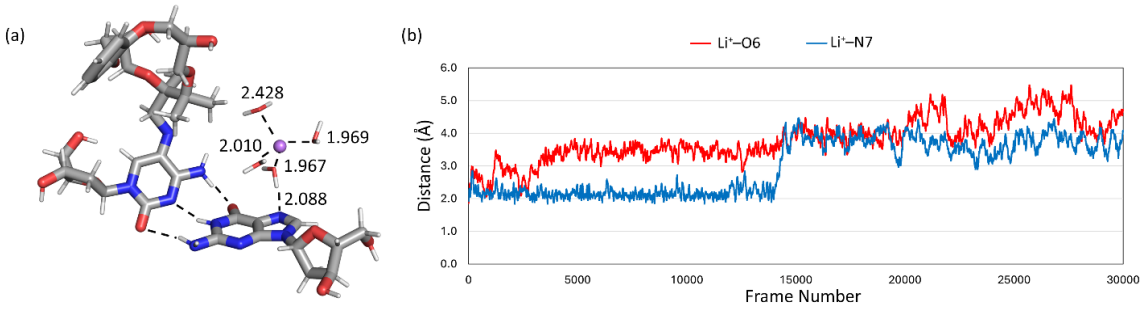


Figure 8.6. AIMD-simulated Li^+–CECNBP complex showing (a) the structure of the dominant 4-coordinate complex and (b) the distance (Å) between O6 or N7 in G and Li^+ as a function of frame.

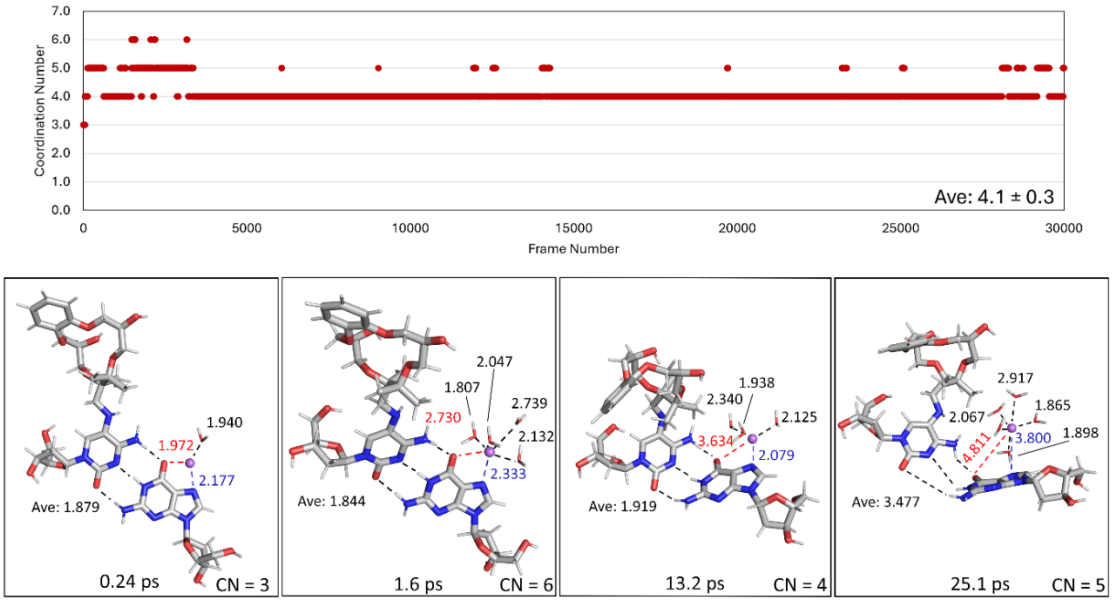


Figure 8.7. Li^+ coordination number over each AIMD-simulated frame for Li^+–G(O6–N7) in the CECNBP model. The calculated average coordination number over each simulation is shown in the bottom right-hand corner. Example snapshots of Li^+–G(O6–N7) interactions in the CECNBP model from different frames are provided below, in addition to coordination distances (Å) between Li^+ and O6–N7 and surrounding waters, and the average WCF hydrogen bond distances over the selected frame are provided.

8.3.5. The CE-modified Nucleobase is Accommodated within a Duplex of Canonical and Modified Nucleic Acids

To confirm whether this CE modification can be incorporated into a nucleic acid, this modification was introduced into an 8-mer DNA duplex and a modified nucleic acid (PNA), and subsequently optimized using GFN2-xTB in implicit water. The CE-modified DNA and PNA duplex remain intact (Figures 8.8 and G.6, Appendix G). Comparing the unmodified to CE-modified DNA structures yielded an RMSD of 1.101 Å, while the CE-modified DNA structures with and without Li^+ showed a smaller RMSD of 0.303 Å

(Figure G.7, Appendix G). In contrast, the RMSDs are consistently small (~0.001 Å) regardless of CE-modification or Li^+ binding in PNA (Figure G.8, Appendix G). This discrepancy in RMSDs between DNA (negatively charged) and PNA (neutral) may be due to their differences in charge, with GFN2-xTB found to have limitations in accurately modeling highly charge-sensitive systems.[55] Nevertheless, the DNA and PNA duplexes remain structurally intact after optimization, demonstrating that the CE modification can be incorporated into either canonical or modified nucleic acids.

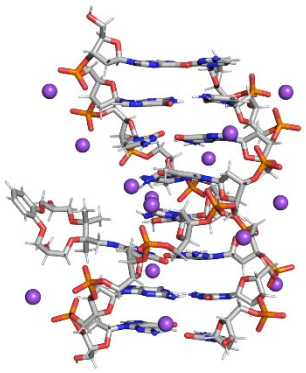


Figure 8.8. GFN2-xTB-optimized structure of a CE-modified DNA duplex in implicit water.

Overall, the results from this study identified that CE-modified nucleic acids are useful for selectively binding Li^+. The information gained from this study can direct the design of a Li^+-specific nucleic acid platform for environmental monitoring, including sensing and remediation of lithium in our drinking water and industrial waste. Aside from biosensing, CE-nucleic acid scaffolds can be useful towards developing novel therapeutics to treat diseases.[56] For example, CEs have shown significant promise in biomedical applications as anticancer agents.[56] Indeed, chemotherapeutics that utilize CE adducts can aid disruption of ion homeostasis in cancer cells or serve as a controlled-release drug

delivery system.[56] Since CE adducts demonstrate high selectivity and stability, as found in our work, future studies should expand CE adduct research to explore a broader range of metals, diverse crown ether structures and nucleic acid environments to reach the full potential of CE adducts in biosensing, remediation, and pharmaceuticals. Altogether, the present work provides fundamental insight for engineering CE adducts to selectively bind Li^+, with our DFT and AIMD results demonstrating how a targeted chemical modification can not only preserve nucleic acid structures but also introduce binding sites that capture Li^+.

8.4. Conclusions

Overall, our results highlight the potential of incorporating CE modifications into nucleic acids to enhance Li^+ capture capabilities. We propose a targeted modification (CEC) designed to selectively bind Li^+, which offers a foundation for developing Li^+-specific nucleic acid sensors. This modification is structurally stable within a base pair context, with comparable binding energy to a canonical GC base pair. Additionally, Li^+ exhibits a stronger affinity for CEC than the preferred nucleobase site, G(O6–N7). Furthermore, in a dynamic solvated environment, Li^+ remains within the CE center, but dissociates from G. These results underscore the potential of this modification for sensing and extracting Li^+. Such platforms are crucial for detecting Li^+ contamination in both environmental and biological systems. Beyond biosensing applications, a system capable of detecting and capturing Li^+ could be used for environmental remediation efforts.

8.5. References

(1) Mrozik, W.; Rajaeifar, M. A.; Heidrich, O.; Christensen, P. Environmental Impacts, Pollution Sources and Pathways of Spent Lithium-Ion Batteries. *Energy Environ Sci.* **2021**, *14* (12), 6099-6121.

(2) Bolan, N.; Hoang, S. A.; Tanveer, M.; Wang, L.; Bolan, S.; Sooriyakumar, P.; Robinson, B.; Wijesekara, H.; Wijesooriya, M.; Keerthanan, S.; et al. From Mine to Mind and Mobiles – Lithium Contamination and Its Risk Management. *Environ. Pollut.* **2021**, *290*, 118067.

(3) Shakoor, N.; Adeel, M.; Ahmad, M. A.; Zain, M.; Waheed, U.; Javaid, R. A.; Haider, F. U.; Azeem, I.; Zhou, P.; Li, Y.; et al. Reimagining Safe Lithium Applications in the Living Environment and Its Impacts on Human, Animal, and Plant System. *Environ. Sci. Ecotechnology.* **2023**, *15*, 100252.

(4) Robinson, B. H.; Yalamanchali, R.; Reiser, R.; Dickinson, N. M. Lithium as an Emerging Environmental Contaminant: Mobility in the Soil-Plant System. *Chemosphere* **2018**, *197*, 1-6.

(5) Barbosa, H.; Leite, C.; Pinto, J.; Soares, A. M. V. M.; Pereira, E.; Freitas, R. Are Lithium Batteries So Eco-Friendly? Ecotoxicological Impacts of Lithium in Estuarine Bivalves. *Environ. Toxicol. Pharmacol.* **2023**, *101*, 104197.

(6) Ruocco, N.; Costantini, M.; Santella, L. New Insights into Negative Effects of Lithium on Sea Urchin Paracentrotus Lividus Embryos. *Sci. Rep.* **2016**, *6* (1), 32157.

(7) Shahzad, B.; Mughal, M. N.; Tanveer, M.; Gupta, D.; Abbas, G. Is Lithium Biologically an Important or Toxic Element to Living Organisms? An Overview. *Environ. Sci. Pollut. Res. Int.* **2017**, *24* (1), 103-115.

(8) Machado-Vieira, R.; Manji, H. K.; Zarate Jr, C. A. The Role of Lithium in the Treatment of Bipolar Disorder: Convergent Evidence for Neurotrophic Effects as a Unifying Hypothesis. *Bipolar Disord.* **2009**, *11* (s2), 92-109.

(9) Villegas-Vázquez, E. Y.; Quintas-Granados, L. I.; Cortés, H.; González-Del Carmen, M.; Leyva-Gómez, G.; Rodríguez-Morales, M.; Bustamante-Montes, L. P.; Silva-Adaya, D.; Pérez-Plasencia, C.; Jacobo-Herrera, N.; et al. Lithium: A Promising Anticancer Agent. *Life* **2023**, *13* (2), 537.

(10) Fornai, F.; Longone, P.; Cafaro, L.; Kastsiuchenka, O.; Ferrucci, M.; Manca, M. L.; Lazzeri, G.; Spalloni, A.; Bellio, N.; Lenzi, P.; et al. Lithium Delays Progression of Amyotrophic Lateral Sclerosis. *Proc. Natl. Acad. Sci.* **2008**, *105* (6), 2052-2057.

(11) Ramoju, S.; Andersen, M.; Poddalgoda, D.; Nong, A.; Karyakina, N.; Shilnikova, N.; Krishnan, K.; Krewski, D. Derivation of Whole Blood Biomonitoring Equivalents for Lithium for the Interpretation of Biomonitoring Data. *Regul. Toxicol. Pharmacol.* **2020**, *111*, 104581.

(12) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(13) Sigel, H.; Sigel, R. Metal-Ion Interactions with Nucleic Acids and Their Constituents. *Comprehensive Inorganic Chemistry II (Second Edition): From Elements to Applications* **2013**, *3*.

(14) Ward, W. L.; Plakos, K.; DeRose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chem. Rev.* **2014**, *114* (8), 4318-4342.

(15) Xu, W.; He, W.; Du, Z.; Zhu, L.; Huang, K.; Lu, Y.; Luo, Y. Functional Nucleic Acid Nanomaterials: Development, Properties, and Applications. *Angew. Chem. Int. Ed.* **2021**, *60* (13), 6890-6918.

(16) Du, Y.; Dong, S. Nucleic Acid Biosensors: Recent Advances and Perspectives. *Anal. Chem.* **2017**, *89* (1), 189-215.

(17) Ruan, C.; Huang, H.; Rodgers, M. T. A Simple Model for Metal Cation-Phosphate Interactions in Nucleic Acids in the Gas Phase: Alkali Metal Cations and Trimethyl Phosphate. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (2), 305-314.

(18) Boychuk, B. T. A.; Jeong, R. Y. E.; Wetmore, S. D. Assessment of the Accuracy of DFT-Predicted Li^+–Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408.

(19) Boychuk, B. T. A.; Meyer, S. P.; Wetmore, S. D. Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes. *Front. Chem.* **2023**, *11*.

(20) Rodgers, M. T.; Armentrout, P. B. Cationic Noncovalent Interactions: Energetics and Periodic Trends. *Chem. Rev.* **2016**, *116* (9), 5642-5687.

(21) Boychuk, B. T. A.; Jeong, R. Y. E.; Wetmore, S. D. Assessment of the Accuracy of DFT-Predicted Li(+)-Nucleic Acid Binding Energies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5392-5408.

(22) Boychuk, B. T. A.; Wagner, L. W.; Lea, M. J.; Wetmore, S. D. A DFT Investigation into the Structure and Strength of Hydrated Li^+–Nucleic Acid Interactions. *Unpublished Results*, (Chapter 3).

(23) Zhu, W.; Luo, X.; Puah, C. M.; Tan, X.; Shen, J.; Gu, J.; Chen, K.; Jiang, H. The Multiplicity, Strength, and Nature of the Interaction of Nucleobases with Alkaline and Alkaline Earth Metal Cations: A Density Functional Theory Investigation. *J. Phys. Chem. A* **2004**, *108* (18), 4008-4018.

(24) Korolev, N.; Lyubartsev, A. P.; Rupprecht, A.; Nordenskiöld, L. Experimental and Monte Carlo Simulation Studies on the Competitive Binding of Li^+, Na^+, and K^+ Ions to DNA in Oriented DNA Fibers. *J. Phys. Chem. B.* **1999**, *103* (42), 9008-9019.

(25) Ma, L.; Liu, J. Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology. *iScience* **2020**, *23* (1).

(26) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shionoya, M. A Discrete Self-Assembled Metal Array in Artificial DNA. *Science* **2003**, *299* (5610), 1212-1213.

(27) Lin, S.; Yang, C.; Mao, Z.; He, B.; Wang, Y.-T.; Leung, C.-H.; Ma, D.-L. A G-Pentaplex-Based Assay for Cs^+ Ions in Aqueous Solution Using a Luminescent Ir(III) Complex. *Biosens Bioelectron.* **2016**, *77*, 609-612.

(28) Jing, Z.; Zhou, Y.; Yamaguchi, T.; Ohara, K.; Pan, J.; Wang, G.; Zhu, F.; Liu, H. Alkali Metal Ion Recognition by 18-Crown-6 in Aqueous Solutions: Evidence from Local Structures. *J. Phys. Chem. B* **2023**, *127* (21), 4858-4869.

(29) Jing, Z.; Wang, G.; Zhou, Y.; Pang, D.; Zhu, F.; Liu, H. Selectivity of 18-Crown-6 Ether to Alkali Ions by Density Functional Theory and Molecular Dynamics Simulation. *Journal of Molecular Liquids* **2020**, *311*, 113305.

(30) Diao, K.-S.; Wang, H.-J.; Qiu, Z.-M. A DFT Study on the Selective Extraction of Metallic Ions by 12-Crown-4. *J. Solution Chem.* **2009**, *38* (6), 713-724.

(31) Xiong, Y.; Ge, T.; Xu, L.; Wang, L.; He, J.; Zhou, X.; Tian, Y.; Zhao, Z. A Fundamental Study on Selective Extraction of Li^+ with Dibenzo-14-Crown-4 Ether: Toward New Technology Development for Lithium Recovery from Brines. *J. Environ. Manage.* **2022**, *310*, 114705.

(32) Li, X.; Chen, G.; Li, X.; Song, Z.; Liu, W.; Liu, B. Research on the Extraction Mechanism of Dibenzo-14-Crown-4 Ether to Specific Recovery Li^+ from Spent Lithium-Ion Battery Raffinate. *Desal.* **2025**, *601*, 118581.

(33) Tian, Y.; Chen, W.; Zhao, Z.; Xu, L.; Tong, B. Interaction and Selectivity of 14-Crown-4 Derivatives with Li^+, Na^+, and Mg^{2+} Metal Ions. *J. Mol. Model.* **2020**, *26* (4), 67.

(34) Yang, F.; Li, L.; Hua, J.; He, J.; Ma, X.; Li, J. Electrospun Nanofiber Membrane of Dibenzo 14-Crown-4-Ether Polyimide for Efficient Selective Lithium Recovery from Discarded Lithium-Ion Batteries. *Sep. Purif. Technol.* **2024**, *334*, 126018.

(35) An, N.; Fleming, A. M.; White, H. S.; Burrows, C. J. Crown Ether–Electrolyte Interactions Permit Nanopore Detection of Individual DNA Abasic Sites in Single Molecules. *Proc. Natl. Acad. Sci.* **2012**, *109* (29), 11504-11509.

(36) Liu, N.; Liang, J.; Lu, Y.; Xu, Z.; Liu, H. Interfacial Structures and Dissociation Reactions of Protic Ionic Liquids on Li Metal Surface: A Combined First-Principles and Ab Initio Molecular Dynamics Investigation. *Chem. Phys. Lett.* **2024**, *836*, 141026.

(37) Liu, P.; Liu, F.; Ren, J. Lithium Functionalized Heteroborospherene Si4b32 as a Promising Hydrogen Storage Material:A Density Functional Theory Study. *Chem. Phys. Lett.* **2024**, *836*, 141021.

(38) Chaban, V. Solvation of Lithium Ion in Dimethoxyethane and Propylene Carbonate. *Chem. Phys. Lett.* **2015**, *631-632*, 1-5.

(39) Olsher, U.; Dalley, N. K.; Frolow, F.; Shoham, G.; Bartsch, R. A. Evolution of Organized Lariat Ether Alcohol and Diol Hydrate Macrostructures in the Solid State. *J. Incl. Phenom. Macrocycl. Chem.* **2003**, *45* (3), 251-255.

(40) An, N.; Fleming, A. M.; Rosecrans, N. C.; Liao, Y.; Burrows, C. J. Synthesis of Site-Specific Crown Ether Adducts to DNA Abasic Sites: 8-Oxo-7,8-Dihydro-2'-Deoxyguanosine and 2'-Deoxycytidine. *Methods Mol Biol* **2019**, *1973*, 15-25.

(41) Yeh, J. I.; Pohl, E.; Truan, D.; He, W.; Sheldrick, G. M.; Du, S.; Achim, C. The Crystal Structure of Non-Modified and Bipyridine-Modified PNA Duplexes. *Chem. Eur. J.* **2010**, *16* (39), 11867-11875.

(42) Torrejos, R. E. C.; Nisola, G. M.; Song, H. S.; Limjuco, L. A.; Lawagon, C. P.; Parohinog, K. J.; Koo, S.; Han, J. W.; Chung, W.-J. Design of Lithium Selective Crown Ethers: Synthesis, Extraction and Theoretical Binding Studies. *Chem. Eng. J.* **2017**, *326*, 921-933.

(43) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032-5051.

(44) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.

(45) Rappoport, D.; Furche, F. Property-Optimized Gaussian Basis Sets for Molecular Response Calculations. *J. Chem. Phys.* **2010**, *133* (13), 134105.

(46) Cammi, R. The Pcm Model. In *Molecular Response Functions for the Polarizable Continuum Model: Physical Basis and Quantum Mechanical Formalism*, Springer International Publishing, 2013; pp 1-11.

(47) *Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16 Rev. B.01*; Wallingford, CT, 2016.

(48) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—an Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652-1671.

(49) Neese, F. The Orca Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73-78.

(50) D.A. Case, I. Y. B.-S., S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. Amber. **2018**.

(51) Boychuk, B. T. A.; Deak, T. K.; Wetmore, S. D. Manipulating Metal Binding to Nucleic Acids through Chemical Modification: A Molecular Dynamics Simulation Study of PNA. *Unpublished Results* **2024**.

(52) Raynaud, C.; Maron, L.; Jolibois, F.; Daudey, J.-P.; Esteves, P. M.; Ramírez-Solís, A. Ab Initio Molecular Dynamics: Plane Waves Vs. Local Basis: The Role of Energy Cutoff on the Convergence of Molecular Properties. *Chem. Phys. Lett.* **2005**, *414* (1), 161-165.

(53) Yu, B.; Bien, K. G.; Wang, T.; Iwahara, J. Diffusion NMR-Based Comparison of Electrostatic Influences of DNA on Various Monovalent Cations. *Biophys J* **2022**, *121* (18), 3562-3570.

(54) Cruz-León, S.; Schwierz, N. RNA Captures More Cations Than DNA: Insights from Molecular Dynamics Simulations. *J. Phys. Chem. B.* **2022**, *126* (43), 8646-8654.

(55) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11* (2), e1493.

(56) Kralj, M.; Tušek-Božić, L.; Frkanec, L. Biomedical Potentials of Crown Ethers: Prospective Antitumor Agents. *ChemMedChem* **2008**, *3* (10), 1478-1492.

Chapter 9: Conclusions and Future Directions

9.1 Summary

Nucleic acid sensors have proven to be an effective green-material solution for detecting metal contaminants. However, there are challenges developing nucleic acid sensors for metals. A specific nucleic acid sensor for lithium detection and extraction is in growing demand because of rising levels of lithium contamination in the environment. To design nucleic acid sensors with enhanced metal selectivity, an understanding of how metals interact with canonical and modified nucleic acids is required. This thesis uses computational techniques to uncover how to accurately model metal–nucleic acid interactions, while also gaining fundamental structural and binding information about how metals interact with canonical nucleic acids that has otherwise been challenging to obtain experimentally. Finally, this thesis contributed information on how select modified nucleic acids interact with Li^+, with the goal of designing a lithium-specific nucleic acid sensing and extraction platform. Overall, this chapter provides a summary of the key conclusions from each chapter and suggests future research directions.

9.2 Thesis Contributions

Chapter 2 provided highly accurate energetic data on Li^+–nucleic acid interactions using the 'gold-standard' CCSD(T)/CBS computational technique to uncover the binding energies of Li^+ interacting at each binding site in the nucleobases and phosphate moiety. The generated data set offers fundamental insights into how Li^+ interacts with each nucleic acid component, which were unclear prior to my work. Specifically, the

CCSD(T)/CBS data showed that Li^+ prefers to bind to each nucleic acid component in the order of phosphate ≫ G > C ≫ A ~ T = U. Among the two strongest Li^+ binding components (i.e., the phosphate moiety and G), the 2 non-bridging oxygens in the phosphate moiety and the O6–N7 position of G result in the strongest interaction energies. The CCSD(T)/CBS data set was subsequently used as a reference to test the performance of 54 DFT methods in conjunction with 3 basis sets for reproducing the binding energies of Li^+–nucleic acid interactions. The results showed that Li^+–nucleic acid binding energies have a greater dependence on functional than basis set. Nevertheless, several functionals (B2-PLYP, PBE-QIDH, ωB97, ωB97X-D, MN15, B3PW91, B97-2, TPSS, BP86-D3(BJ), and PBE) were found to accurately describe Li^+–nucleic acid interactions. Additionally, Chapter 2 underscores the importance of selecting accurate computational methods to study Li^+–nucleic acid interactions, with many functionals demonstrating significant fluctuations in performance. Overall, the DFT methods identified to be reliable for describing Li^+–nucleic acid interactions were used in subsequent chapters and can be applied to future projects to investigate Li^+ behavior in nucleic acid systems.

Using similar methodologies as applied in Chapter 2, an accurate database of group I metals (Li^+, Na^+, K^+, Rb^+, and Cs^+) interacting with each nucleic acid component (A, C, G, T, U, and the phosphate moiety) were generated using gas-phase CCSD(T)/CBS binding strengths in Chapter 3. These results provided key insights into how each group I metal binds to nucleic acids, with the binding strengths decreasing as $Li^+ >> Na^+ >> K^+ > Rb^+ > Cs^+$, which follows the well-established trend of decreasing metal binding affinity with increasing ionic radius. When considering the strongest binding sites in

each nucleic acid component, the trend in group I metal–nucleic acid binding strengths are phosphate moiety >> G > C >> A > T = U. Upon assessment of 61 DFT methods for reproducing the CCSD(T)/CBS binding strengths of group I metal–nucleic acid interactions, the top performing functionals are mPW2-PLYP and ωB97M-V. The results from this chapter highlight that functional performance is dependent on metal identity, emphasizing the importance of testing computational methods for reliably describing metal–nucleic acid interactions. Overall, information on metal interactions with nucleic acid binding sites as well as identifying reliable methodologies for investigating group I metal–nucleic acid interactions can be used towards developing biosensors with enhanced target-specificity, in addition to designing materials and therapeutics.

Using the same methodology from Chapter 3, the systems considered were expanded to include group II metals (Be^{2+}, Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}) interacting with each nucleic acid component (A, C, G, T, U, and the phosphate moiety) in Chapter 4. This study provides for the first time a comprehensive database of gas-phase CCSD(T)/CBS group II metal–nucleic acid binding strengths, which have otherwise been elusive to gain by experimental techniques. The binding strength for group II metals interacting with nucleic acids follows size-to-charge ratio, decreasing as Be^{2+} >> Mg^{2+} >> Ca^{2+} > Sr^{2+} > Ba^{2+}, and the strongest binding decrease as a function of the nucleobase as G >> C >> A >> T ~ U. The performance of 61 DFT methods was assessed over group I and II metal–nucleic acid interactions to identify a method that can robustly study metal competition effects. The results from this chapter uncovered the functional dependence of group II metal–nucleic acid interactions. Overall,

ωB97M-V is the top performing functional for exploring metal competition effects over group I and II metal–nucleic acid interactions, while mPW91 is a computationally cheaper alternative. This information on how group II metals interact with nucleic acids in addition to the accurate computational methods identified in this chapter can be used to explore how these metals compete for modified nucleic acids for future development of a nucleic acid sensing and extraction platform to target different metals.

In Chapter 5, DFT models of transition and post-transition metal–nucleic acid complexes were used to identify reliable methods for investigating these interactions. Unlike main group metals, transition and post-transition metal complexes are more challenging to describe using computational techniques. Using crystal structure geometries from the Protein Data Bank (PDB) and Cambridge Structural Database (CSD), a dataset featuring 53 transition and post-transition (Cr^{3+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, $Cu^{+/2+}$, Zn^{2+}, Cd^{2+}, Pd^{2+}, $Ag^{+/2+}$, Pt^{2+}, Hg^{2+}, Pb^{2+}, $Au^{+/3+}$) metal–nucleic acid complexes was constructed. The results from this study highlighted the importance of exploring the ability of DFT to describe the structures of transition and post-transition metal–nucleic acid complexes, revealing that 12 of the 53 crystallographic structures are challenging for DFT to reproduce. The ability of DFT to describe structures of transition and post-transition metal–nucleic acid complexes vary with metal identity. Among the metals studied, complexes containing Cu were the most challenging to accurately describe across many functionals. Nevertheless, the top methods for describing transition and post-transition metal–nucleic acid complexes are ωB97M-V, ωB97X-D3(BJ), and MN15. Overall, ωB97M-V emerges as the most robust method for accurately describing a broad range of metals (i.e., groups I and II, transition and post-

transition metals) across Chapters 2–5. This result highlights the reliability of ωB97M-V to be applied as a unified computational approach in future studies on these different metals that may necessitate a single reliable functional, especially for studies involving the development of a nucleic acid sensing and extraction platform. Although ωB97M-V offers the highest accuracy for describing these interactions, it is computationally more expensive for more demanding calculations like AIMD simulations. Therefore, a practical alternative functional that offers a compromise between computational cost and accuracy for exploring a broad range of metals (i.e., groups I and II, transition and post-transition metals) interacting with nucleic acids in more expensive calculations is MN15-L.

While thesis chapters 2–5 have focused on identifying reliable computational methods to investigate metal–nucleic acid interactions. Chapter 6 shifts to applying these methods to further explore Li^+ binding to nucleic acids for developing sensing and extraction platforms. In Chapter 6, an accurate DFT method and basis set combination (MN15/aug-cc-pVTZ) identified in Chapter 2 was used to investigate the impact of solvent on Li^+–nucleic acid interactions using implicitly and explicitly solvated models. This chapter revealed that increasing the solvent polarity from gas phase to water alters Li^+ coordination with the nucleic acid component from bidentate to monodentate and decreases the magnitude of Li^+–nucleic acid binding strengths. The inclusion of explicit solvent further alters Li^+ binding to nucleic acids in the gas phase, with Li^+ coordination to the nucleic acid component changing from bidentate to monodentate upon addition of 3 Li^+-bound waters.

The displacement of an Li^+-coordinated water by a nucleic acid component is feasible in the gas phase regardless of the number of inner-shell waters. In contrast, such displacement is only feasible when 2 or more inner-shell waters are present, suggesting that at lower hydration numbers, the nucleic acid can bind to Li^+ without displacing waters. Additionally, implicit and explicit solvent significantly influence the trends in Li^+ binding energies to the preferred nucleic acid sites. Specifically, Li^+ binds most strongly to C due to additional hydrogen bonding between Li^+-bound waters and C, with the new binding strengths decreasing as a function of nucleobase as C > G > T ~ U > A. In contrast, the gas-phase binding strengths decrease as G > C >> A > T = U, underscoring how solvent can influence Li^+–nucleic acid interactions. However, when explicitly hydrated Li^+ coordination to the nucleobase can become restricted in the presence of the opposing WCF base pair. Since the binding sites in the WCF are blocked by the complementary opposing nucleobase, there is a loss of preference for Li^+ binding to C over G. Nevertheless, within the context of WCF base pairs, the binding strengths as a function of nucleobase are decreasing as G > C > A > T ~ U. Overall, the information gained in this study demonstrates the importance of understanding how solvent can modulate Li^+–nucleic acid interactions, with the binding preferences of Li^+ for nucleobase sites altered in a solvated environment. These details assist in our understanding of Li^+ behavior in biological systems and has practical implications for developing biosensing and extraction technologies for Li^+.

Chapter 7 used MD simulations to understand how a chemically modified nucleic acid (PNA) interacts with Li^+ and Na^+ compared to DNA. This chapter showed that modifying the sugar-backbone from DNA to PNA manipulates the position of where

Li^+ and Na^+ binding occurs in the nucleic acid duplex. Although metal binding is nearly equivalent in the major and minor grooves of DNA, metal binding is shifted from the minor groove to favor the major groove in PNA. Among binding sites in the major groove of PNA, Na^+ binding is distributed across G(O6), G(N7), and T(O4), while Li^+ binding is pushed towards the O6 site in G. The behavior of Li^+ and Na^+ binding to PNA is unaffected by changing the salt concentration (up to 750 mM). Additionally, the PNA structure is less susceptible to deviations in the presence of Li^+ compared to Na^+, regardless of salt concentration. Flipping the sequence from GC- to AT-rich demonstrated that metal binding to PNA is sequence-dependent, with GC-rich sequences of PNA showing greater preference for binding Li^+. Overall, these results highlight how PNA is a step in the right direction towards creating a Li^+-specific sensing and extraction platform.

While Chapter 7 explored how chemically modifying the backbone of nucleic acids influence metal binding, Chapter 8 is focused on understanding how introducing a nucleobase modification alters metal binding further. In particular, crown ethers (CE) are well-established for selectively binding alkali metals, and a Li^+-specific crown ether variant has been designed[1] that can be incorporated into nucleic acids. Specifically, a previously published synthetic protocol successfully conjugated a crown ether to N4 of C,[2] with crown ethers as well-established Li^+ binding motifs.[1,3] Using the strongest Li^+ binding crown-ether candidate identified in previous work,[1] DFT and AIMD simulations were performed to uncover the structures, binding strengths, and dynamics of Li^+ binding to crown-ether-modified C (CEC). The results from this study demonstrate that the structure of Li^+ bound to CE is stable in a nucleobase and a base-

pair model. Additionally, Li^+ has a stronger binding affinity for CEC than the strongest nucleic acid site in the G:C pair, namely G(O6). In a solvated and dynamic environment, Li^+ remains bound to CEC, while Li^+ dissociates from G(O6), suggesting that CEC is a promising modification for capturing Li^+. Overall, this chapter presents a nucleobase modification that can target Li^+, and when incorporated into a nucleic acid, this modified nucleic acid can be used to develop a sensing and extraction platform that could be extended to cellular environments.

9.3 Future Directions

The computational methodologies identified in this thesis to accurately model metal–nucleic acid interactions can be applied to investigate the roles of different metals in biology, and support the development of new materials, biotechnologies, and therapeutics. Despite numerous studies carried out on exploring how metals behave in biological systems, the interactions of metals in diverse biological networks are not fully understood.[4] For example, alkali and alkaline earth metals are utilized by ribozymes to facilitate the catalytic cleavage of introns during gene splicing.[5] While there have been many studies aiming to uncover the chemistry of ribozymes, the precise role of metals in ribozyme catalysis remains poorly understood. As a specific example, Mg^{2+} is a known co-factor in the GlmS ribozyme; however, how Mg^{2+} facilitates catalysis is not well understood.[5] The computational methods outlined in my thesis can be used towards studying the chemistry and roles of metals and metal competition in ribozyme-catalyzed intron splicing. Additionally, understanding the chemistry of ribozymes can aid development of future therapeutics. In material applications, these

methods can be used towards improving metal binding specificity in DNA nanowires.[6-8] Improving metal binding specificity is crucial for increasing the structural stability of DNA nanowires, which is key to DNA nanowire performance in electronics, biomedical technologies such as drug delivery, and energy storage.[6-8] For example, Ag^+ can bind to C or G in DNA duplexes, which can form C-Ag-C or G-Ag-G pairs that cause different structural variations in DNA, effectively tuning the functionality of DNA nanowires.[8] In therapeutics, these methods can be used to explore different metal binding modes to nucleic acids and this information can be applied to develop more stable and potent anticancer drugs.[9] For example, cisplatin is a well-known therapeutic for treating cancers, which binds irreversibly to DNA through 1,2-intrastrand crosslinks between cisplatin and DNA.[10] However, cisplatin can be harmful to the body due to targeting healthy cells. Ruthenium complexes have been found to be alternative therapeutics, which have excellent selectivity and reduced toxicity compared to cisplatin.[11] This improvement in therapeutic performance is due to reversible binding between ruthenium complexes and DNA by intercalation, emphasizing the value of computational studies of these systems. Overall, the studies carried out in this thesis provide important guidelines for exploring metal–nucleic acid interactions using computational modeling.

This thesis also provides a foundation for future work on rationally designing nucleic acid sensing and extraction platforms to capture metal contaminants using computational modeling. A proposal for constructing a Li^+-specific nucleic acid platform involves combining the sugar–backbone modification from Chapter 7 (PNA) and the nucleobase modification from Chapter 8 (crown ether). This approach will

enable the generation of a nucleic acid platform that can manipulate Li^+ binding to the major groove, and subsequently, the crown ether modification can capture Li^+. Individually, crown ethers are toxic to cells, causing paralysis, tremors, and neurobehavioral symptoms (e.g., aggression) in mammals.[12] Incorporating crown ethers into a nucleic acid scaffold can improve the functionality of crown ethers and may also reduce their toxicity in biosystems.[12] In fact, there has been research demonstrating that linking crown ethers to DNA can enhance antitumor activity.[12] Since nucleic acids are biocompatible and stable in biofluids,[13] with PNA offering enhanced resistance to enzymatic degradation,[14] these biomolecules are excellent candidates for developing a reliable and safe sensing platform. Incorporating a Li^+-specific crown ether into a nucleic acid could also be beneficial for increasing the stability of a sensing and extraction platform for the environment, while also ensuring biosafety. Extending this strategy into PNA could provide an enhanced sensing and capturing approach for Li^+ in the environment and in cells. Chapter 8 demonstrates that crown ethers can be successfully incorporated into duplex structures of canonical (DNA) and modified (PNA) nucleic acids (Figure 9.1). Ideally, a nucleic acid sensing and extraction platform could be developed to provide a rapid response in real-time, with a successful example shown for a miniature device that detects Cd^{2+} using an aptamer combined with an exonuclease.[13]

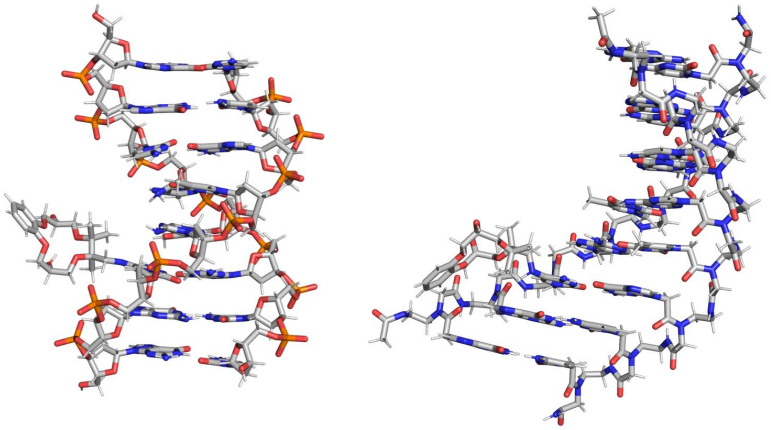


Figure 9.1. Optimized duplex structures of DNA (left) and PNA (right) with the Li^+-specific crown ether incorporated into C.

The next step in developing a Li^+-specific nucleic acid platform is to design a sensing component that emits a response upon capturing Li^+. An attractive approach for sensing metals in a crown ether is using fluorescence-based detection.[15,16] In fact, crown ethers can be chemically modified to emit a fluorescence signal upon binding a metal.[15] For example, crown ether derivatives containing benzoxazolyl-alanines have been shown to capture Cu^{2+}, Pd^{2+}, Hg^{2+}, and Fe^{3+}, while emitting a fluorescent signal.[15] Fluorescent probes have proven to be highly effective in a medical setting for detecting conditions such as epilepsy, Parkinson's disease, and glioma.[17] As a specific example, a DNAzyme-based sensor emits a fluorescent response in the presence of *Vibrio cholerae*, which is the bacterium that causes cholera.[18] Altogether, developing a chemically modified crown ether that maintains Li^+ specificity while also providing a fluorescent response upon Li^+ capture would provide a functional nucleic acid sensing platform for targeting Li^+. This approach could also be utilized to create a Li^+ extraction platform by incorporating multiple crown ether modifications into nucleic acid strands.

Beyond Li^+, the information gained in this thesis has provided structural and binding details of other metals interacting with nucleic acids, which can be used to explore the impact of introducing chemical modifications aimed to selectively bind other metals. There are different chemical modifications that can be incorporated into nucleic acids to tune metal binding specificity. For example, phosphorothioate modifications have proven successful for targeting soft metals such as Cd^{2+},[19] suggesting that soft donor groups (e.g., S^{2-} and Se^{2-}) could be introduced into nucleobases as a next step toward enhancing Cd^{2+} specificity or targeting other soft metals such as Cu^+ and Ag^+. For capturing group I metals beyond Li^+, the cavity size of crown ethers can be adjusted to accommodate larger group I metals. For example, the 18-crown-6 is known to prefer binding to K^+ over other group I metals.[20] Further modifications can be done to crown ethers to tune metal selectivity. Notably, chemically modifying the 18-crown-6 into dibenzo-18-crown-6 creates a crown ether variant that specifically binds to Sr^{2+}.[21] There are numerous other chemical modifications that can be added to nucleic acids that can be explored for binding metals such as ammonium, guanidinium, and imidazole groups.[22] Additional studies can investigate other ligand-attachments to DNA such as phenanthroline or terpyridine to chelate metals.[23] Alternatively, metal-mediated base pairing with modified nucleobases to capture metals can be explored.[24] Overall, expanding the range of chemical modifications will permit tuning metal-specificity for developing nucleic acid sensing and extraction platforms to target other metals.

9.4 Final Remarks

In summary, this thesis used multiple computational approaches (Ab Initio, DFT, MD, and AIMD) to investigate how metals interact with nucleic acids. The early chapters addressed gaps in the literature on how different metals interact with nucleic acids as well as identified reliable computational methods for describing these interactions. Additionally, the role of solvent on Li^+–nucleic acid interactions was uncovered, which was unclear prior to my thesis work. The later chapters of this thesis explored modifications that can enhance Li^+–nucleic acid binding, contributing fundamental information for the rational design of a nucleic acid sensor with high Li^+ specificity. Aside from biosensing applications, this thesis provides valuable insights into metal–nucleic acid interactions that can aid advancements in developing enhanced pharmaceuticals, gene therapies, and materials.

9.5 References

(1) Torrejos, R. E. C.; Nisola, G. M.; Song, H. S.; Limjuco, L. A.; Lawagon, C. P.; Parohinog, K. J.; Koo, S.; Han, J. W.; Chung, W.-J. Design of Lithium Selective Crown Ethers: Synthesis, Extraction and Theoretical Binding Studies. *Chem. Eng. J.* **2017**, *326*, 921-933.

(2) An, N.; Fleming, A. M.; Rosecrans, N. C.; Liao, Y.; Burrows, C. J. Synthesis of Site-Specific Crown Ether Adducts to DNA Abasic Sites: 8-Oxo-7,8-Dihydro-2'-Deoxyguanosine and 2'-Deoxycytidine. *Methods Mol Biol* **2019**, *1973*, 15-25.

(3) Pei, H.; Yan, F.; Liu, H.; He, B.; Li, J. The Selective Complexation of Crown Ethers for Lithium Isotope Separation: A Critical Review. *Sep. Purif. Technol.* **2024**, *341*, 126857.

(4) Buccella, D.; Lim, M. H.; Morrow, J. R. Metals in Biology: From Metallomics to Trafficking. *Inorg. Chem.* **2019**, *58* (20), 13505-13508.

(5) Ward, W. L.; Plakos, K.; DeRose, V. J. Nucleic Acid Catalysis: Metals, Nucleobases, and Other Cofactors. *Chem. Rev.* **2014**, *114* (8), 4318-4342.
(6) Ijiro, K.; Mitomo, H. Metal Nanoarchitecture Fabrication Using DNA as a Biotemplate. *Polym. J.* **2017**, *49* (12), 815-824.

(7) Bayrak, T.; Jagtap, N. S.; Erbe, A. Review of the Electrical Characterization of Metallic Nanowires on DNA Templates. *Int J Mol Sci* **2018**, *19* (10), 3019.

(8) Sun, Q.; Xie, X.; Song, Y.; Sun, L. A Review on Silver-Mediated DNA Base Pairs: Methodology and Application. *Biomater. Res.* **2022**, *26* (1), 9.

(9) Hangan, A. C.; Oprean, L. S.; Dican, L.; Procopciuc, L. M.; Sevastre, B.; Lucaciu, R. L. Metal-Based Drug–DNA Interactions and Analytical Determination Methods. *Mol.* **2024**, *29* (18), 4361.

(10) Mantri, Y.; Lippard, S. J.; Baik, M.-H. Bifunctional Binding of Cisplatin to DNA: Why Does Cisplatin Form 1,2-Intrastrand Cross-Links with Ag but Not with Ga? *J. Am. Chem. Soc.* **2007**, *129* (16), 5023-5030.

(11) Adhikari, S.; Nath, P.; Das, A.; Datta, A.; Baildya, N.; Duttaroy, A. K.; Pathak, S. A Review on Metal Complexes and Its Anti-Cancer Activities: Recent Updates from in Vivo Studies. *Biomed. Pharmacother.* **2024**, *171*, 116211.

(12) Kralj, M.; Tušek-Božić, L.; Frkanec, L. Biomedical Potentials of Crown Ethers: Prospective Antitumor Agents. *ChemMedChem* **2008**, *3* (10), 1478-1492.

(13) Liu, S.; Xu, B.; Li, C.; Ren, Y.; Gan, H.; Kuang, S.; Lei, C.; Nie, Z. Advancements and Prospects in DNA-Based Bioanalytical Technology for Environmental Toxicant Detection. *JACS Au* **2025**, *5* (6), 2443-2462.

(14) Koppelhus, U.; Nielsen, P. E. Cellular Delivery of Peptide Nucleic Acid (PNA). *Adv. Drug Deliv. Rev.* **2003**, *55* (2), 267-280.
(15) Batista, P. M. R.; Martins, C. D. F.; Raposo, M. M. M.; Costa, S. P. G. Novel Crown Ether Amino Acids as Fluorescent Reporters for Metal Ions. *Mol.* **2023**, *28* (8), 3326.

(16) Li, J.; Yim, D.; Jang, W.-D.; Yoon, J. Recent Progress in the Design and Applications of Fluorescence Probes Containing Crown Ethers. *Chem. Soc. Rev.* **2017**, *46* (9), 2437-2458.

(17) Wang, X.; Ding, Q.; Groleau, R. R.; Wu, L.; Mao, Y.; Che, F.; Kotova, O.; Scanlan, E. M.; Lewis, S. E.; Li, P.; et al. Fluorescent Probes for Disease Diagnosis. *Chem. Rev.* **2024**, *124* (11), 7106-7164.

(18) Miao, Q.; Ding, W.; Bao, X.; Wang, S.; Lin, Q.; Xu, Y.; Lu, J.; Lyu, M.; Wang, S. An Efficient DNAzyme for the Fluorescence Detection of *Vibrio Cholerae*. *Food Sci. Nutr.* **2023**, *11* (6), 3235-3245.

(19) Huang, P.-J. J.; Liu, J. Rational Evolution of Cd^{2+}-Specific DNAzymes with Phosphorothioate Modified Cleavage Junction and Cd^{2+} Sensing. *Nucleic Acids Res.* **2015**, *43* (12), 6125-6133.

(20) Jing, Z.; Zhou, Y.; Yamaguchi, T.; Ohara, K.; Pan, J.; Wang, G.; Zhu, F.; Liu, H. Alkali Metal Ion Recognition by 18-Crown-6 in Aqueous Solutions: Evidence from Local Structures. *J. Phys. Chem. B* **2023**, *127* (21), 4858-4869.

(21) Wang, L.; Xiong, Y.; Zhou, X.; Xu, L.; Tian, Y.; Zhao, Z. Complexation Mechanism of Crown Ethers with Rubidium and Cesium Ions Using Density Functional Theory. *Comput. Theor. Chem.* **2023**, *1225*, 114139.

(22) Zhou, W.; Saran, R.; Liu, J. Metal Sensing by DNA. *Chem. Rev.* **2017**, *117* (12), 8272-8325.

(23) Yang, H.; Metera, K. L.; Sleiman, H. F. DNA Modified with Metal Complexes: Applications in the Construction of Higher Order Metal–DNA Nanostructures. *Coord. Chem. Rev.* **2010**, *254* (19), 2403-2415.

(24) Hu, L.; Takezawa, Y.; Shionoya, M. Metal-Mediated DNA Base Pairing of Easily Prepared 2-Oxo-Imidazole-4-Carboxylate Nucleotides. *Chem. Sci.* **2022**, *13* (14), 3977-3983.

Appendix A
Supplementary Information for Chapter 2: An Assessment of the Accuracy of DFT Predicted Li^+–Nucleic Acid Binding Energies
Contains Tables A.1–A.13 and Figures A.1–A.22

Table A.1. The percent error (PE), mean percent error (MPE) and average mean percent error (AMPE) across nucleic acid components and/or basis sets (%) for Li^+–nucleobases and Li^+–phosphate DFT binding energies without counterpoise corrections.[a,b]

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
Double-Hybrid	B2-PLYP	2.49	0.92	2.22	2.80	0.59	2.43	2.94	1.74	2.74	2.74	1.08	2.46
	B2-PLYP-D3	4.15	1.77	3.75	4.46	1.44	3.96	4.60	2.60	4.27	4.40	1.94	3.99
	mPW2-PLYP	4.35	1.78	3.92	4.67	1.45	4.13	4.87	2.63	4.50	4.63	1.95	4.18
	PBE-QIDH	1.92	1.43	1.84	2.06	1.06	1.89	1.82	1.88	1.83	1.93	1.46	1.85
	PBE0-DH	2.87	1.86	2.71	3.23	1.57	2.95	3.07	2.35	2.95	3.06	1.93	2.87
	revDSD-PBEP86-D3	–	–	–	–	–	–	2.11	1.86	2.07	–	–	–
RSH	M11	4.58	2.28	4.20	4.60	2.08	4.18	4.92	2.37	4.49	4.70	2.25	4.29
	MN12-SX	2.69	0.57	2.34	2.96	0.56	2.56	2.49	1.10	2.26	2.72	0.74	2.39
	MN12-SX-D3	3.09	0.98	2.74	3.41	0.97	3.00	2.89	1.52	2.66	3.13	1.16	2.80
RSH meta-GGA	ωB97M-V	2.35	0.93	2.12	2.71	0.66	2.37	3.08	1.94	2.89	2.72	1.18	2.46
RSH GGA	ωB97	1.14	0.27	1.00	1.54	0.32	1.34	1.20	0.41	1.07	1.29	0.33	1.13
	ωB97X	2.05	1.17	1.91	1.46	1.29	1.43	1.69	0.47	1.49	1.73	0.98	1.61
	ωB97X-D	1.04	1.08	1.04	0.83	1.20	0.89	0.83	0.44	0.76	0.90	0.91	0.90
	ωB97X-V	2.07	0.94	1.88	2.51	0.65	2.20	2.52	1.72	2.39	2.37	1.10	2.16
	HSE06	3.34	1.50	3.04	3.79	1.28	3.37	3.65	2.12	3.39	3.59	1.64	3.27
	HSE06-D3	4.92	2.25	4.48	5.40	2.03	4.84	5.26	2.87	4.86	5.20	2.38	4.73
	LC-PBE	5.61	3.34	5.23	6.03	3.12	5.55	5.29	3.55	5.00	5.65	3.34	5.26
	LC-ωPBE	2.86	1.77	2.68	3.35	1.56	3.05	2.89	2.25	2.78	3.03	1.86	2.84
	LC-ωPBE-XDM	–	–	–	4.51	3.14	4.28	3.45	2.51	3.29	2.65	1.88	2.52
	LC-ωPBE-D3	4.84	2.74	4.49	5.34	2.52	4.87	4.88	3.21	4.60	5.02	2.82	4.65
	CAM-B3LYP	6.58	2.66	5.93	7.13	2.48	6.35	6.97	3.42	6.38	6.89	2.85	6.22
	CAM-B3LYP-D3	8.41	3.52	7.59	8.95	3.35	8.02	8.79	4.29	8.04	8.72	3.72	7.88
GH meta-GGA	M06	1.82	1.00	1.68	1.83	0.81	1.66	1.63	0.01	1.36	1.76	0.61	1.57
	M06-2X	3.30	2.40	3.15	4.03	2.01	3.70	3.95	2.72	3.74	3.76	2.38	3.53

	M06-HF	8.07	5.84	7.70	8.44	5.33	7.92	8.44	5.49	7.95	8.31	5.56	7.85
	MN15	1.76	0.55	1.55	2.13	0.55	1.86	1.42	0.89	1.33	1.77	0.66	1.58
	BMK	2.09	0.31	1.79	2.03	0.47	1.77	1.75	0.15	1.48	1.96	0.31	1.68
GH GGA	BH&HLYP	7.15	3.45	6.53	7.67	3.27	6.94	7.50	3.97	6.91	7.44	3.56	6.79
	SOGGA11-X	3.58	1.48	3.23	4.32	1.51	3.85	3.93	2.18	3.64	3.94	1.73	3.57
	B3PW91	1.72	0.20	1.47	2.05	0.01	1.71	1.82	0.78	1.65	1.86	0.33	1.61
	B3PW91-D3	3.57	1.84	3.28	4.14	1.65	3.73	3.89	2.41	3.65	3.87	1.97	3.55
	PBE0	4.16	1.57	3.73	4.60	1.35	4.05	4.61	2.23	4.22	4.46	1.72	4.00
	PBE0-D3	5.41	2.36	4.90	5.89	2.15	5.26	5.90	3.03	5.42	5.73	2.51	5.19
	B3LYP	4.04	1.24	3.57	4.62	1.07	4.03	4.56	2.08	4.14	4.41	1.46	3.92
	B3LYP-D3	7.17	2.83	6.45	7.75	2.66	6.90	7.69	3.66	7.02	7.54	3.05	6.79
	X3LYP	5.36	1.81	4.77	5.93	1.63	5.22	5.92	2.69	5.38	5.74	2.04	5.12
	X3LYP-D3	8.94	3.79	8.08	9.50	3.61	8.52	9.49	4.67	8.69	9.31	4.02	8.43
	O3LYP	2.19	0.48	1.90	2.49	0.22	2.12	2.60	1.77	2.46	2.43	0.83	2.16
	O3LYP-D3	5.79	3.11	5.34	6.36	2.85	5.77	6.96	4.40	6.53	6.37	3.45	5.88
	B97-2	1.51	0.04	1.27	1.80	0.14	1.53	1.67	0.78	1.52	1.66	0.32	1.44
Local meta-GGA	TPSS	2.12	0.50	1.85	2.48	0.25	2.11	2.65	1.36	2.43	2.42	0.70	2.13
	TPSS-D3	4.31	1.61	3.86	4.77	1.36	4.20	4.93	2.47	4.52	4.67	1.81	4.19
	M06-L	2.53	1.31	2.33	2.07	1.45	1.97	1.91	0.69	1.70	2.17	1.15	2.00
	M11-L	1.94	0.12	1.64	1.71	0.35	1.48	2.23	0.27	1.91	1.96	0.25	1.68
	MN12-L	3.00	0.67	2.61	3.33	0.78	2.90	2.72	1.16	2.46	3.01	0.87	2.66
	MN15-L	2.21	0.96	2.00	1.30	0.98	1.25	0.98	0.33	0.87	1.50	0.75	1.37
Local GGA	mPW91	2.74	1.17	2.48	3.06	0.97	2.71	2.88	1.81	2.70	2.89	1.31	2.63
	BLYP	2.45	0.39	2.10	2.78	0.55	2.41	2.85	0.80	2.51	2.69	0.58	2.34
	BLYP-D3	5.84	1.50	5.12	6.50	1.33	5.64	6.59	2.69	5.94	6.31	1.84	5.56
	BP86	2.35	1.67	2.23	1.83	1.85	1.84	1.90	0.85	1.72	2.03	1.46	1.93
	BP86-D3	1.62	0.00	1.35	1.85	0.19	1.57	1.63	0.81	1.49	1.70	0.33	1.47
	PBE	2.33	0.04	1.95	2.42	0.18	2.05	2.33	1.11	2.13	2.36	0.45	2.04
	PBE-D3	3.30	0.85	2.89	3.89	0.63	3.35	3.95	1.92	3.61	3.71	1.13	3.28
Local LDA	SVWN5	6.92	2.52	6.19	7.27	2.23	6.43	7.14	3.28	6.50	7.11	2.68	6.37
AMPE		3.70	1.56	3.35	4.00	1.43	3.57	3.91	2.05	3.59	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).
[b)] revDSD-PBEP86-D3 is excluded from the reported average.

Table A.2. The mean unsigned errors (MUE), root-mean squared errors (RMSE), average mean unsigned errors (AMUE) and average root-mean squared errors (ARMSE) across nucleic acid components and/or basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections.[a,b]

		6-311+G(3df,2p)			aug-cc-pVTZ			def2-TZVPP					
		Nucleobases		PO_4^-	Nucleobases		PO_4^-	Nucleobases		PO_4^-	AMUE		ARMSE
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Nucleobases**	**PO_4^-**	**Base**
Double-Hybrid	B2-PLYP	1.41	1.54	1.50	1.58	1.70	0.96	1.68	1.73	2.85	1.55	1.77	1.66
	B2-PLYP-D3	2.35	2.40	2.90	2.52	2.57	2.36	2.62	2.64	4.25	2.50	3.17	2.54
	mPW2-PLYP	2.48	2.57	2.90	2.66	2.73	2.37	2.80	2.84	4.30	2.65	3.19	2.71
	PBE-QIDH	1.13	1.21	2.34	1.21	1.31	1.73	1.10	1.23	3.08	1.15	2.38	1.25
	PBE0-DH	1.70	1.90	3.04	1.90	2.10	2.58	1.83	2.01	3.85	1.81	3.16	2.00
	rev-DSD-PBEP86-D3	–	–	–	–	–	–	1.23	1.55	3.04	–	–	–
RSH	M11	2.72	2.96	3.74	2.72	2.94	3.41	2.92	3.15	3.87	2.79	3.67	3.02
	MN12-SX	1.48	1.56	0.93	1.65	1.79	0.91	1.40	1.47	1.80	1.51	1.21	1.61
	MN12-SX-D3	1.74	1.91	1.61	1.93	2.19	1.59	1.66	1.85	2.49	1.78	1.90	1.98
RSH meta-GGA	ωB97M-V	1.36	1.43	1.53	1.56	1.63	1.07	1.79	1.84	3.17	1.57	1.92	1.63
RSH GGA	ωB97	0.63	0.70	0.44	0.88	0.88	0.52	0.69	0.72	0.67	0.73	0.54	0.77
	ωB97X	1.15	1.52	1.92	0.82	1.29	2.11	0.93	1.30	0.78	0.97	1.60	1.37
	ωB97X-D	0.57	0.82	1.77	0.47	0.62	1.97	0.47	0.63	0.71	0.50	1.48	0.69
	ωB97X-V	1.23	1.33	1.54	1.48	1.57	1.07	1.49	1.58	2.81	1.40	1.81	1.49
	HSE06	1.96	2.19	2.46	2.21	2.43	2.10	2.14	2.35	3.48	2.10	2.68	2.32
	HSE06-D3	2.85	2.99	3.68	3.12	3.25	3.32	3.06	3.18	4.70	3.01	3.90	3.14
	LC-PBE	3.39	3.71	5.46	3.62	3.92	5.11	3.21	3.56	5.81	3.40	5.46	3.73
	LC-ωPBE	1.74	1.97	2.90	2.02	2.22	2.55	1.77	2.01	3.69	1.84	3.05	2.07
	LC-ωPBE-XDM	–	–	–	2.64	2.92	5.13	2.10	2.30	4.11	2.37	4.62	2.61

	LC-ωPBE-D3	2.87	2.98	4.47	3.14	3.24	4.12	2.89	3.01	5.26	2.97	4.62	3.08
	CAM-B3LYP	3.80	3.91	4.34	4.11	4.21	4.06	4.03	4.12	5.60	3.98	4.67	4.08
	CAM-B3LYP-D3	4.83	4.90	5.76	5.14	5.20	5.48	5.06	5.12	7.02	5.01	6.09	5.07
GH meta-GGA	M06	0.97	1.46	1.63	0.97	1.22	1.32	0.87	1.24	0.02	0.94	0.99	1.31
	M06-2X	1.91	1.97	3.93	2.33	2.40	3.29	2.29	2.34	4.45	2.18	3.89	2.24
	M06-HF	4.71	4.78	9.55	4.91	4.96	8.72	4.91	4.98	8.99	4.84	9.09	4.91
	MN15	0.99	1.08	0.90	1.21	1.22	0.89	0.80	0.91	1.46	1.00	1.08	1.07
	BMK	1.18	1.70	0.51	1.08	1.44	0.77	0.97	1.46	0.24	1.08	0.51	1.53
GH GGA	BH&HLYP	4.13	4.26	5.64	4.43	4.56	5.34	4.34	4.46	6.49	4.30	5.82	4.43
	SOGGA11-X	2.05	2.18	2.43	2.47	2.56	2.47	2.27	2.35	3.57	2.26	2.82	2.36
	B3PW91	0.93	1.07	0.33	1.12	1.17	0.02	1.00	1.06	1.28	1.02	0.54	1.10
	B3PW91-D3	2.05	2.19	3.00	2.37	2.49	2.69	2.24	2.35	3.95	2.22	3.22	2.34
	PBE0	2.57	3.25	2.56	2.81	3.39	2.21	2.86	3.55	3.66	2.75	2.81	3.40
	PBE0-D3	3.21	3.51	3.87	3.47	3.71	3.51	3.51	3.82	4.96	3.40	4.11	3.68
	B3LYP	2.30	2.52	2.03	2.63	2.82	1.75	2.61	2.76	3.40	2.51	2.39	2.70
	B3LYP-D3	4.07	4.14	4.63	4.40	4.46	4.35	4.38	4.42	5.99	4.28	4.99	4.34
	X3LYP	3.06	3.22	2.96	3.39	3.53	2.67	3.39	3.51	4.40	3.28	3.34	3.42
	X3LYP-D3	5.10	5.14	6.19	5.42	5.46	5.90	5.43	5.46	7.63	5.31	6.58	5.35
	O3LYP	1.19	1.27	0.79	1.38	1.49	0.36	1.49	1.68	2.90	1.35	1.35	1.48
	O3LYP-D3	3.30	3.35	5.08	3.62	3.66	4.66	3.99	4.01	7.20	3.64	5.65	3.67
	B97-2	0.79	1.03	0.07	0.97	1.07	0.23	0.91	0.98	1.28	0.89	0.53	1.03
Local meta-GGA	TPSS	1.20	1.36	0.82	1.41	1.58	0.40	1.51	1.64	2.23	1.38	1.15	1.53
	TPSS-D3	2.44	2.51	2.64	2.70	2.75	2.22	2.80	2.84	4.05	2.65	2.97	2.70
	M06-L	1.45	2.01	2.15	1.14	1.59	2.38	1.06	1.70	1.13	1.22	1.88	1.77
	M11-L	1.04	1.32	0.20	0.87	1.11	0.57	1.27	1.71	0.44	1.06	0.40	1.38
	MN12-L	1.71	1.77	1.09	1.94	2.12	1.28	1.58	1.64	1.89	1.74	1.42	1.84
	MN15-L	1.25	1.49	1.57	0.74	1.10	1.60	0.53	0.75	0.54	0.84	1.23	1.11

Local GGA	mPW91	1.58	1.65	1.91	1.77	1.91	1.58	1.68	1.83	2.96	1.68	2.15	1.80
	BLYP	1.32	1.59	0.63	1.53	1.87	0.90	1.58	1.84	1.32	1.47	0.95	1.77
	BLYP-D3	3.25	3.37	2.45	3.63	3.73	2.18	3.69	3.76	4.40	3.52	3.01	3.62
	BP86	1.36	1.73	2.73	1.06	1.46	3.03	1.09	1.46	1.39	1.17	2.38	1.55
	BP86-D3	0.88	0.99	0.01	1.00	1.24	0.30	0.90	1.09	1.33	0.93	0.55	1.11
	PBE	1.31	1.43	0.07	1.34	1.49	0.29	1.32	1.48	1.82	1.33	0.73	1.47
	PBE-D3	1.88	2.06	1.39	2.21	2.36	1.02	2.26	2.38	3.14	2.12	1.85	2.27
Local LDA	SVWN5	4.08	4.47	4.12	4.27	4.63	3.64	4.21	4.57	5.37	4.19	4.38	4.56
AMUE		2.13	–	2.56	2.30	–	2.40	2.25	–	3.36	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).
[b)] revDSD-PBEP86-D3 is excluded from the reported average.

Table A.3. The percent error (PE), mean percent error (MPE) and average mean percent error (AMPE) across nucleic acid components and/or Dunning basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies.[a]

		aug-cc-pVDZ			**aug-cc-pVTZ**			**aug-cc-pVQZ**					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
Double-Hybrid	B2-PLYP	0.76	1.20	0.84	2.32	0.33	1.99	2.82	0.81	2.48	1.97	0.78	1.77
	B2-PLYP-D3	2.36	0.34	2.03	3.99	1.19	3.52	4.48	1.67	4.02	3.61	1.07	3.19
	mPW2-PLYP	2.55	0.31	2.18	4.19	1.2	3.69	4.67	1.67	4.17	3.81	1.06	3.35
	PBE-QIDH	0.49	0.68	0.53	1.85	0.83	1.68	2.27	1.32	2.11	1.54	0.94	1.44
	PBE0-DH	1.54	0.00	1.29	2.96	1.41	2.70	3.41	1.86	3.15	2.64	1.09	2.38
RSH	M11	3.97	0.69	3.42	4.02	1.80	3.65	4.89	2.11	4.43	4.29	1.53	3.83
	MN12-SX	1.33	0.98	1.27	2.65	0.32	2.26	2.71	0.71	2.38	2.23	0.67	1.97
	MN12-SX-D3	1.71	0.56	1.52	3.05	0.74	2.66	2.86	1.13	2.57	2.54	0.81	2.25
RSH meta-GGA	ωB97M-V	6.11	0.52	5.18	2.60	0.76	2.29	2.49	0.62	2.18	3.73	0.63	3.22
RSH GGA	ωB97	1.28	1.72	1.35	1.29	0.51	1.16	1.53	0.11	1.30	1.37	0.78	1.27
	ωB97X	3.25	2.65	3.15	1.70	1.39	1.65	1.42	0.94	1.34	2.12	1.66	2.05

	ωB97X-D	2.15	2.59	2.22	0.85	1.30	0.93	0.94	0.83	0.92	1.31	1.57	1.36
	ωB97X-V	0.82	0.53	0.77	2.31	0.69	2.04	2.46	0.71	2.17	1.86	0.64	1.66
	HSE06	2.12	0.24	1.81	3.44	1.11	3.05	3.91	1.52	3.52	3.16	0.96	2.79
	HSE06-D3	3.52	0.51	3.02	5.05	1.86	4.52	5.53	2.27	4.98	4.70	1.54	4.17
	LC-PBE	3.85	1.54	3.46	5.49	2.88	5.05	6.04	3.36	5.60	5.13	2.59	4.70
	LC-ωPBE	1.39	0.01	1.16	2.73	1.29	2.49	3.25	1.73	2.99	2.45	1.01	2.21
	LC-ωPBE-XDM	2.06	0.01	1.72	3.86	3.14	3.50	–	–	–	2.96	1.58	2.61
	LC-ωPBE-D3	3.17	0.95	2.80	4.72	2.25	4.31	5.23	2.69	4.81	4.38	1.97	3.97
	CAM-B3LYP	4.97	0.88	4.28	6.58	2.24	5.86	7.05	2.67	6.32	6.20	1.93	5.49
	CAM-B3LYP-D3	6.79	1.74	5.95	8.41	3.11	7.52	8.88	3.54	7.99	8.03	2.80	7.15
GH meta-GGA	M06	2.43	2.22	2.40	1.75	0.94	1.62	1.78	0.20	1.52	1.99	1.12	1.84
	M06-2X	2.04	0.88	1.85	3.65	1.84	3.35	4.50	2.46	4.16	3.40	1.73	3.12
	M06-HF	6.61	3.74	6.13	7.32	4.56	6.86	8.48	5.16	7.92	7.47	4.49	6.97
	MN15	1.22	1.10	1.20	1.8	0.17	1.53	2.03	0.58	1.79	2.39	0.62	2.09
	BMK	3.43	1.74	3.15	2.05	0.57	1.80	1.93	0.01	1.61	1.80	0.77	1.60
GH GGA	BH&HLYP	5.77	1.81	5.11	7.36	3.11	6.65	7.78	3.51	7.07	6.97	2.81	6.28
	SOGGA11-X	1.88	0.02	1.57	3.94	1.38	3.51	4.04	1.72	3.65	3.29	1.04	2.91
	B3PW91	1.11	1.52	1.18	1.87	0.15	1.58	2.24	0.29	1.92	1.74	0.65	1.56
	B3PW91-D3	2.19	0.12	1.84	3.83	1.49	3.44	4.32	1.92	3.92	3.45	1.18	3.07
	PBE0	1.95	0.18	1.66	4.25	1.19	3.74	3.82	1.61	3.45	3.34	0.99	2.95
	PBE0-D3	3.34	0.61	2.88	5.49	1.98	4.91	5.47	2.40	4.96	4.77	1.66	4.25
	B3LYP	2.52	0.54	2.19	4.14	0.85	3.59	4.60	1.27	4.04	3.75	0.89	3.27
	B3LYP-D3	5.65	1.05	4.88	7.27	2.43	6.46	7.73	2.86	6.91	6.88	2.11	6.08
	X3LYP	3.80	0.01	3.17	5.42	1.40	4.75	5.87	1.82	5.20	5.03	1.08	4.37
	X3LYP-D3	7.37	1.99	6.48	8.99	3.38	8.06	8.75	3.56	7.88	8.37	2.98	7.47
	O3LYP	0.99	1.46	1.07	2.17	0.06	1.81	2.49	0.34	2.13	1.88	0.62	1.67
	O3LYP-D3	4.06	1.17	3.57	5.74	2.57	5.21	6.16	2.96	5.62	5.32	2.23	4.80
	B97-2	1.44	1.64	1.47	1.63	0.29	1.41	1.97	0.12	1.66	1.68	0.68	1.51
Local meta-GGA	TPSS	0.97	1.34	1.03	2.19	0.08	1.84	2.69	0.51	2.33	1.95	0.64	1.73
	TPSS-D3	2.31	0.45	2.00	4.47	1.19	3.93	4.98	1.62	4.42	3.92	1.09	3.45
	M06-L	3.80	2.43	3.57	2.09	1.60	2.01	1.86	0.87	1.63	2.56	1.63	2.40
	M11-L	3.60	1.87	3.31	2.05	0.60	1.81	3.98	0.10	3.33	3.23	0.86	2.84

	MN12-L	1.48	0.89	1.38	2.95	0.45	2.53	2.95	0.81	2.60	2.46	0.72	2.17
	MN15-L	3.58	2.23	3.36	1.64	1.22	1.57	1.49	0.88	1.39	2.24	1.44	2.11
Local GGA	mPW91	1.53	0.56	1.37	2.86	0.8	2.52	3.23	1.23	2.90	2.54	0.86	2.26
	BLYP	1.49	2.32	1.63	2.4	0.88	2.14	2.74	9.32	3.84	2.21	4.18	2.54
	BLYP-D3	4.10	0.44	3.49	5.75	1.00	4.96	6.23	1.44	5.43	5.36	0.96	4.63
	BP86	3.83	3.48	3.77	2.11	2.05	2.10	1.87	1.59	1.83	2.60	2.37	2.56
	BP86-D3	0.92	1.81	1.07	1.67	0.38	1.46	2.00	0.07	1.68	1.53	0.75	1.40
	PBE	1.11	1.90	1.24	2.31	0.44	2.00	2.63	0.00	2.19	2.02	0.78	1.81
	PBE-D3	1.60	1.09	1.51	3.31	0.37	2.82	3.80	0.80	3.30	2.90	0.75	2.54
Local LDA	SVWN5	5.17	0.52	4.39	6.94	2.04	6.13	7.53	2.54	6.69	6.54	1.70	5.74
AMPE		2.82	1.21	2.54	3.68	1.32	3.29	4.02	1.67	3.62	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.4. The percent error (PE), mean percent error (MPE) and average mean percent error (AMPE) across nucleic acid components and/or Karlsruhe basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies.[a]

		def2-SVP			**def2-TZVPP**			**def2-QZVPP**					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
Double-Hybrid	B2-PLYP	4.50	0.91	3.90	1.97	0.72	1.76	2.61	0.82	2.31	5.50	2.75	5.04
	B2-PLYP-D3	6.17	1.77	5.43	3.63	1.58	3.29	4.57	1.68	4.00	7.26	3.61	6.62
	mPW2-PLYP	6.47	1.76	5.68	3.90	1.61	3.51	4.98	1.69	4.32	7.60	3.62	6.90
	PBE-QIDH	2.69	0.49	2.32	1.40	0.99	1.33	2.09	1.29	1.96	4.06	2.54	3.80
	PBE0-DH	4.12	1.24	3.64	2.56	1.62	2.41	4.13	1.84	3.56	5.47	3.10	5.01
RSH	M11	6.68	2.22	5.93	4.14	1.75	3.74	4.73	2.24	4.31	7.32	3.77	6.73
	MN12-SX	3.61	0.21	3.04	2.10	0.48	1.83	2.60	0.64	2.27	4.51	1.95	4.08
	MN12-SX-D3	4.28	0.63	3.67	2.50	0.90	2.23	3.00	1.06	2.67	5.07	2.37	4.62
RSH meta-GGA	ωB97M-V	10.14	1.53	8.71	2.50	1.30	2.30	4.20	1.17	3.59	6.26	3.16	5.71
RSH GGA	ωB97	1.91	0.07	1.61	0.91	0.24	0.80	1.42	0.09	1.20	3.58	1.89	3.30
	ωB97X	2.64	0.90	2.35	2.25	1.12	2.06	1.44	0.93	1.36	3.69	2.06	3.42
	ωB97X-D	2.39	0.97	2.15	1.21	1.07	1.18	0.85	0.81	0.85	3.25	1.87	3.01

	ωB97X-V	4.12	1.22	3.63	2.03	1.12	1.88	2.51	1.03	2.14	5.06	1.22	2.55
	HSE06	5.02	1.53	4.44	3.14	1.44	2.86	3.70	1.54	3.34	5.94	3.12	5.47
	HSE06-D3	6.63	2.27	5.91	4.64	2.19	4.23	5.31	2.29	4.81	7.51	3.87	6.90
	LC-PBE	6.42	2.87	5.83	4.71	2.93	4.41	5.78	3.33	5.37	7.78	4.72	7.27
	LC-ωPBE	4.14	1.59	3.71	2.37	1.52	2.23	3.01	1.74	2.80	5.36	3.33	5.02
	LC-ωPBE-XDM	–	–	–	2.77	1.75	2.26	–	–	–	–	–	–
	LC-ωPBE-D3	6.13	2.55	5.53	4.20	2.48	3.91	5.00	2.70	4.62	7.30	4.29	6.79
	CAM-B3LYP	8.88	3.15	7.92	6.24	2.63	5.63	6.85	2.69	6.15	3.30	4.62	5.00
	CAM-B3LYP-D3	10.70	4.01	9.59	8.06	3.50	7.30	8.67	3.56	7.82	4.16	6.15	6.85
GH meta-GGA	M06	2.46	0.63	2.15	1.55	0.57	1.38	1.75	0.36	1.51	5.43	7.82	8.67
	M06-2X	4.95	2.03	4.46	3.42	2.15	3.21	4.22	2.45	3.92	0.87	1.51	1.75
	M06-HF	17.28	4.61	15.17	7.63	4.76	7.15	7.75	4.88	7.27	2.83	3.92	4.22
	MN15	2.22	0.61	1.96	1.17	0.25	1.02	1.87	0.56	1.65	6.01	7.27	7.75
	BMK	1.69	0.22	1.44	1.89	0.32	1.63	1.89	0.04	1.58	0.79	1.67	1.89
GH GGA	BH&HLYP	9.17	3.28	8.19	6.95	3.35	6.35	7.57	3.51	6.90	0.84	1.56	1.87
	SOGGA11-X	5.16	1.39	4.53	3.38	1.65	3.09	3.93	1.75	3.57	4.93	6.90	7.57
	B3PW91	2.66	0.21	2.25	1.53	0.12	1.30	2.07	0.31	1.77	2.42	3.57	3.93
	B3PW91-D3	5.16	1.84	4.61	3.29	1.76	3.04	4.09	1.94	3.73	0.80	1.77	2.07
	PBE0	4.79	1.50	4.24	4.06	1.54	3.64	3.59	1.62	3.27	2.49	3.73	4.09
	PBE0-D3	6.44	2.30	5.75	5.22	2.33	4.74	5.24	2.42	4.77	2.63	3.27	3.59
	B3LYP	6.46	1.85	5.69	3.83	1.29	3.41	4.39	1.30	3.87	3.58	4.77	5.24
	B3LYP-D3	9.59	3.44	8.56	6.96	2.88	6.28	7.52	2.89	6.75	2.36	3.87	4.39
	X3LYP	7.85	2.43	6.94	5.17	1.88	4.62	5.67	1.86	5.03	4.58	6.75	7.52
	X3LYP-D3	11.42	4.41	10.25	8.74	3.86	7.93	9.24	3.83	8.34	3.24	5.03	5.67
	O3LYP	3.71	0.83	3.23	2.10	0.74	1.87	2.37	0.39	2.04	5.88	8.34	9.24
	O3LYP-D3	8.12	3.46	7.34	5.91	3.37	5.49	6.02	3.02	5.52	1.13	2.04	2.37
	B97-2	2.52	0.01	2.10	1.36	0.08	1.15	1.82	0.15	1.54	4.26	5.52	6.02
Local meta-GGA	TPSS	3.90	0.89	3.40	1.97	0.54	1.73	2.52	0.55	2.19	0.65	1.54	1.82
	TPSS-D3	6.18	2.00	5.49	4.14	1.66	3.73	4.80	1.66	4.28	1.14	2.19	2.52
	M06-L	1.64	1.23	1.57	2.35	1.08	2.14	1.87	0.88	1.71	2.69	4.28	4.80
	M11-L	2.15	1.68	2.08	3.03	0.84	2.67	2.44	0.54	2.12	1.51	1.71	1.87
	MN12-L	3.48	0.12	2.92	2.36	0.59	2.06	2.91	0.73	2.54	1.60	2.12	2.44

	MN15-L	1.13	1.35	1.16	1.63	0.40	1.43	1.70	0.87	1.56	1.39	2.54	2.91
Local GGA	mPW91	3.91	1.10	3.44	2.57	1.12	2.33	3.07	1.26	2.76	1.15	1.56	1.70
	BLYP	5.09	0.94	4.40	2.23	0.18	1.89	2.59	0.38	2.22	1.79	2.76	3.07
	BLYP-D3	8.83	2.82	7.83	5.65	1.71	4.99	6.03	1.51	5.28	1.13	2.22	2.59
	BP86	1.31	1.01	1.26	2.61	1.62	2.44	1.89	1.54	1.83	3.25	5.28	6.03
	BP86-D3	3.00	0.65	2.61	1.45	0.05	1.22	1.83	0.12	1.55	1.99	1.83	1.89
	PBE	3.80	0.77	3.29	2.23	0.20	1.89	2.45	0.05	2.05	0.67	1.55	1.83
	PBE-D3	5.44	1.58	4.79	3.09	1.01	2.74	3.81	1.35	3.40	0.97	2.05	2.45
Local LDA	SVWN5	8.48	3.07	7.58	6.38	2.48	5.73	7.26	2.57	6.48	2.05	3.40	3.81
AMPE		5.34	1.66	4.72	3.48	1.51	3.15	3.92	1.55	3.51	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.5. The percent error (PE), mean percent error (MPE) and average mean percent error (AMPE) across nucleic acid components and/or Dunning basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections.[a]

		aug-cc-pVDZ			**aug-cc-pVTZ**			**aug-cc-pVQZ**					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
Double-Hybrid	B2-PLYP	1.47	0.60	1.33	2.8	0.59	2.43	3.16	0.99	2.79	2.48	0.73	2.18
	B2-PLYP-D3	3.13	0.25	2.65	4.46	1.44	3.96	4.82	1.85	4.32	4.14	1.18	3.65
	mPW2-PLYP	3.31	0.28	2.81	4.67	1.45	4.13	5.02	1.85	4.49	4.33	1.19	3.81
	PBE-QIDH	0.69	0.04	0.59	2.06	1.06	1.89	2.38	1.45	2.23	1.71	0.85	1.57
	PBE0-DH	1.86	0.50	1.63	3.23	1.57	2.95	3.60	1.97	3.33	2.90	1.34	2.64
RSH	M11	4.59	1.03	3.99	4.60	2.08	4.18	5.38	2.41	4.88	4.85	1.84	4.35
	MN12-SX	1.56	0.54	1.39	2.96	0.56	2.56	3.16	1.09	2.81	2.56	0.73	2.26
	MN12-SX-D3	1.96	0.13	1.66	3.41	0.97	3.00	3.15	1.51	2.87	2.84	0.87	2.51
RSH meta-GGA	ωB97M-V	5.54	0.22	4.66	2.71	0.66	2.37	2.79	0.92	2.48	3.68	0.60	3.17
RSH GGA	ωB97	0.79	1.31	0.87	1.54	0.32	1.34	1.76	0.07	1.48	1.36	0.57	1.23
	ωB97X	2.78	2.26	2.69	1.46	1.29	1.43	1.41	0.86	1.32	1.88	1.47	1.81
	ωB97X-D	1.68	2.22	1.77	0.83	1.2	0.89	0.96	0.75	0.93	1.16	1.39	1.20
	ωB97X-V	1.21	0.19	1.04	2.51	0.65	2.2	2.69	0.93	2.40	2.14	0.59	1.88

	HSE06	2.38	0.15	2.01	3.79	1.28	3.37	4.25	1.67	3.82	3.47	1.04	3.07
	HSE06-D3	4.00	0.90	3.48	5.4	2.03	4.84	5.86	2.42	5.29	5.09	1.78	4.54
	LC-PBE	4.67	2.04	4.24	6.03	3.12	5.55	6.59	3.59	6.09	5.76	2.92	5.29
	LC-ωPBE	1.79	0.41	1.56	3.35	1.56	3.05	3.89	1.99	3.57	3.01	1.32	2.73
	LC-ωPBE-XDM	2.29	0.41	2.29	4.51	3.14	4.28	–	–	–	3.40	1.78	3.28
	LC-ωPBE-D3	3.78	1.37	3.38	5.34	2.52	4.87	5.88	2.95	5.39	5.00	2.28	4.55
	CAM-B3LYP	5.55	1.29	4.84	7.13	2.48	6.35	7.61	2.91	6.83	6.76	2.23	6.01
	CAM-B3LYP-D3	7.38	2.16	6.51	8.95	3.35	8.02	9.43	3.78	8.49	8.59	3.09	7.67
GH meta-GGA	M06	1.95	1.83	1.93	1.83	0.81	1.66	1.86	0.32	1.86	1.88	1.32	1.82
	M06-2X	2.53	1.25	2.32	4.03	2.01	3.70	4.88	2.64	4.51	3.81	1.97	3.51
	M06-HF	7.51	4.17	6.95	8.44	5.33	7.92	9.90	5.91	9.24	8.62	5.14	8.04
	MN15	0.91	0.63	0.87	2.13	0.55	1.86	2.53	1.03	2.28	2.36	0.74	2.09
	BMK	3.01	1.36	2.74	2.03	0.47	1.77	1.95	0.15	1.65	1.83	0.66	1.61
GH GGA	BH&HLYP	6.23	2.22	5.57	7.67	3.27	6.94	8.04	3.64	7.31	7.32	3.04	6.61
	SOGGA11-X	2.31	0.38	1.99	4.32	1.51	3.85	4.34	1.87	3.93	3.66	1.25	3.26
	B3PW91	0.82	1.14	0.88	2.05	0.01	1.71	2.40	0.43	2.07	1.76	0.53	1.55
	B3PW91-D3	2.65	0.49	2.29	4.14	1.65	3.73	4.60	2.06	4.18	3.80	1.40	3.40
	PBE0	2.18	0.21	1.85	4.60	1.35	4.05	4.14	1.75	3.74	3.64	1.11	3.21
	PBE0-D3	3.81	1.01	3.34	5.89	2.15	5.26	5.79	2.55	5.25	5.16	1.90	4.62
	B3LYP	3.07	0.13	2.58	4.62	1.07	4.03	5.09	1.48	4.49	4.26	0.90	3.70
	B3LYP-D3	6.20	1.45	5.41	7.75	2.66	6.9	8.22	3.07	7.36	7.39	2.40	6.56
	X3LYP	4.37	0.43	3.71	5.93	1.63	5.22	6.39	2.05	5.67	5.56	1.37	4.87
	X3LYP-D3	7.94	2.40	7.02	9.50	3.61	8.52	9.27	3.78	8.36	8.90	3.27	7.96
	O3LYP	1.05	0.99	1.04	2.49	0.22	2.12	2.81	0.57	2.43	2.11	0.59	1.86
	O3LYP-D3	4.67	1.64	4.17	6.36	2.85	5.77	6.74	3.20	6.15	5.92	2.56	5.36
	B97-2	0.96	1.24	1.01	1.8	0.14	1.53	2.10	0.23	1.79	1.62	0.54	1.44
Local meta-GGA	TPSS	1.15	0.98	1.12	2.48	0.25	2.11	2.87	0.61	2.49	2.17	0.61	1.91
	TPSS-D3	2.80	0.09	2.35	4.77	1.36	4.2	5.16	1.72	4.58	4.24	1.06	3.71
	M06-L	3.39	2.05	3.17	2.07	1.45	1.97	1.81	0.66	1.62	2.42	1.39	2.25
	M11-L	2.23	0.99	2.02	1.71	0.35	1.48	2.76	1.39	2.53	2.23	0.91	1.97
	MN12-L	1.91	0.31	1.64	3.33	0.78	2.9	3.59	1.37	3.22	2.94	0.82	2.59
	MN15-L	3.10	1.69	2.86	1.3	0.98	1.25	1.20	0.67	1.11	1.87	1.11	1.74

Local GGA	mPW91	1.69	0.18	1.43	3.06	0.97	2.71	3.40	1.37	3.06	2.71	0.84	2.40
	BLYP	1.47	1.88	1.54	2.78	0.55	2.41	3.31	8.98	4.26	2.52	3.81	2.74
	BLYP-D3	4.79	0.00	3.99	6.5	1.33	5.64	7.05	1.78	6.17	6.11	1.04	5.27
	BP86	3.32	3.10	3.28	1.83	1.85	1.84	1.83	1.41	1.76	2.32	2.12	2.29
	BP86-D3	0.83	1.44	0.93	1.85	0.19	1.57	2.22	0.26	1.90	1.64	0.63	1.47
	PBE	1.13	1.48	1.18	2.42	0.18	2.05	2.97	0.26	2.52	2.17	0.64	1.92
	PBE-D3	2.18	0.67	1.93	3.89	0.63	3.35	4.43	1.06	3.87	3.50	0.79	3.05
Local LDA	SVWN5	6.66	1.23	5.75	7.27	2.23	6.43	7.84	2.70	6.98	7.26	2.05	6.39
AMPE		3.04	1.08	2.72	4.00	1.43	3.57	4.33	1.86	3.93	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.6. The percent error (PE), mean percent error (MPE) and average mean percent error (AMPE) across nucleic acid components and/or Karlsruhe basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections.[a]

		def2-SVP			**def2-TZVPP**			**def2-QZVPP**					
Family	**Functional**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base MPE**	**PO_4^- PE**	**All MPE**	**Base AMPE**	**PO_4^- AMPE**	**All AMPE**
Double-Hybrid	B2-PLYP	11.92	6.70	11.05	2.94	1.74	2.74	2.95	1.27	2.67	5.94	3.24	5.49
	B2-PLYP-D3	13.59	7.56	12.58	4.60	2.60	4.27	4.61	2.12	4.20	7.60	4.09	7.02
	mPW2-PLYP	13.92	7.56	12.86	4.87	2.63	4.50	4.81	2.13	4.36	7.87	4.11	7.24
	PBE-QIDH	8.68	5.33	8.13	1.82	1.88	1.83	2.26	1.67	2.16	4.25	2.96	4.04
	PBE0-DH	9.71	5.82	9.06	3.07	2.35	2.95	3.39	2.15	3.18	5.39	3.44	5.06
RSH	M11	13.09	7.30	12.13	4.92	2.37	4.49	5.00	2.58	4.60	7.67	4.09	7.07
	MN12-SX	8.82	4.74	8.14	2.49	1.10	2.26	2.80	1.09	2.51	4.70	2.31	4.31
	MN12-SX-D3	9.72	5.16	8.96	2.89	1.52	2.66	3.20	1.51	2.92	5.27	2.73	4.85
RSH meta-GGA	ωB97M-V	12.09	1.53	11.24	3.08	1.94	2.89	2.68	1.22	2.43	5.95	1.56	5.52
RSH GGA	ωB97	8.41	5.33	7.90	1.20	0.41	1.07	1.49	0.11	1.26	3.70	1.95	3.41
	ωB97X	7.39	4.12	6.84	1.69	0.47	1.49	1.35	0.71	1.25	3.48	1.77	3.19
	ωB97X-D	7.68	3.72	7.02	0.83	0.44	0.76	0.88	0.54	0.82	3.13	1.57	2.87
	ωB97X-V	10.66	1.22	3.63	2.52	1.72	2.39	2.31	1.10	2.11	5.16	1.35	2.71
	HSE06	10.98	6.37	10.21	3.65	2.12	3.39	3.84	1.83	3.51	6.16	3.44	5.70

	HSE06-D3	12.59	7.12	11.68	5.26	2.87	4.86	5.45	2.58	4.98	7.77	4.19	7.17
	LC-PBE	12.85	7.90	12.02	5.29	3.55	5.00	5.91	3.56	5.52	8.01	5.00	7.51
	LC-ωPBE	10.70	6.75	10.04	2.89	2.25	2.78	3.16	2.02	2.97	5.58	3.67	5.26
	LC-ωPBE-XDM	–	–	–	3.45	2.51	3.29	–	–	–	3.45	2.51	3.29
	LC-ωPBE-D3	12.69	7.71	11.86	4.88	3.21	4.60	5.15	2.98	4.79	7.57	4.63	7.08
	CAM-B3LYP	16.32	8.88	15.08	6.97	3.42	6.38	7.01	3.00	6.34	10.10	5.10	9.27
	CAM-B3LYP-D3	18.14	9.75	16.74	8.79	4.29	8.04	8.83	3.87	8.00	11.92	5.97	10.93
GH meta-GGA	M06	6.62	4.02	6.19	1.63	0.01	1.36	1.76	0.03	1.47	3.34	1.35	3.01
	M06-2X	10.65	6.67	9.99	3.95	2.72	3.74	4.36	2.70	4.08	6.32	4.03	5.94
	M06-HF	24.03	10.14	21.71	8.44	5.49	7.95	8.21	5.40	7.74	13.56	7.01	12.47
	MN15	8.05	5.79	7.67	1.42	0.89	1.33	1.93	0.77	1.74	2.94	2.48	2.86
	BMK	5.56	4.29	5.35	1.75	0.15	1.48	1.83	0.21	1.56	3.91	1.55	3.52
GH GGA	BH&HLYP	14.96	7.99	13.80	7.50	3.97	6.91	7.70	3.74	7.04	10.05	5.23	9.25
	SOGGA11-X	10.59	5.90	9.81	3.93	2.18	3.64	4.11	1.98	3.76	6.21	3.35	5.73
	B3PW91	7.66	4.92	7.21	1.82	0.78	1.65	2.13	0.60	1.88	3.87	2.10	3.58
	B3PW91-D3	10.96	6.55	10.22	3.89	2.41	3.65	4.24	2.23	3.91	6.36	3.73	5.93
	PBE0	10.70	6.30	9.97	4.61	2.23	4.22	3.74	1.92	3.44	6.35	3.48	5.88
	PBE0-D3	12.35	7.10	11.48	5.90	3.03	5.42	5.39	2.71	4.95	7.88	4.28	7.28
	B3LYP	13.59	7.41	12.56	4.56	2.08	4.14	4.55	1.63	4.07	7.57	3.71	6.92
	B3LYP-D3	16.72	9.00	15.43	7.69	3.66	7.02	7.68	3.22	6.94	10.70	5.29	9.80
	X3LYP	15.13	8.10	13.96	5.92	2.69	5.38	5.84	2.20	5.23	8.96	4.33	8.19
	X3LYP-D3	18.70	10.08	17.26	9.49	4.67	8.69	9.41	4.17	8.54	12.53	6.31	11.50
	O3LYP	10.73	6.29	9.99	2.60	1.77	2.46	2.47	0.89	2.21	5.27	2.98	4.89
	O3LYP-D3	15.13	8.92	14.10	6.96	4.40	6.53	6.27	3.52	5.81	9.45	5.61	8.81
	B97-2	7.25	4.63	6.81	1.67	0.78	1.52	1.89	0.46	1.66	3.60	1.96	3.33
Local meta-GGA	TPSS	10.35	6.08	9.63	2.65	1.36	2.43	2.73	0.98	2.44	5.24	2.80	4.84
	TPSS-D3	12.63	7.19	11.72	4.93	2.47	4.52	5.02	2.09	4.53	7.53	3.92	6.92
	M06-L	3.58	2.35	3.37	1.91	0.69	1.70	1.77	0.59	1.58	2.42	1.21	2.22
	M11-L	3.26	1.57	2.98	2.23	0.27	1.91	1.81	0.05	1.51	2.43	0.63	2.13
	MN12-L	7.27	3.96	6.71	2.72	1.16	2.46	3.12	1.16	2.79	4.37	2.09	3.99
	MN15-L	3.97	2.92	3.79	0.98	0.33	0.87	1.42	0.57	1.28	2.12	1.27	1.98
Local GGA	mPW91	9.77	5.84	9.12	2.88	1.81	2.70	3.14	1.57	2.88	5.26	3.07	4.90

	BLYP	13.70	7.40	12.65	2.85	0.80	2.51	2.68	0.08	2.25	6.41	2.76	5.80
	BLYP-D3	17.43	9.29	16.08	6.59	2.69	5.94	6.26	1.97	5.55	10.09	4.65	9.19
	BP86	6.44	4.37	6.10	1.90	0.85	1.72	1.77	1.19	1.67	3.37	2.14	3.16
	BP86-D3	9.81	6.04	9.18	1.63	0.81	1.49	1.91	0.48	1.67	4.45	2.44	4.11
	PBE	11.23	6.55	10.45	2.33	1.11	2.13	2.54	0.49	2.20	5.37	2.72	4.93
	PBE-D3	12.87	7.35	11.95	3.95	1.92	3.61	3.60	0.91	3.15	6.81	3.40	6.24
Local LDA	SVWN5	8.48	3.07	7.58	7.14	3.28	6.50	7.45	2.94	6.70	7.69	3.10	6.93
AMPE		11.16	6.13	10.23	3.91	2.05	3.59	3.96	1.77	3.59	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.7. The mean unsigned errors (MUE), root-mean squared errors (RMSE), average mean unsigned errors (AMUE) and average root-mean squared errors (ARMSE) across nucleic acid components and/or Dunning basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies.[a]

		aug-cc-pVDZ			**aug-cc-pVTZ**			**aug-cc-pVQZ**					
		Nucleobases		**PO_4^-**	**Nucleobases**		**PO_4^-**	**Nucleobases**		**PO_4^-**	**AMUE**		**ARMSE**
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Nucleobases**	**PO_4^-**	**Base**
Double-Hybrid	B2-PLYP	0.41	0.49	1.96	1.31	1.46	0.54	1.59	1.74	1.32	1.11	1.27	1.23
	B2-PLYP-D3	1.32	1.34	0.56	2.25	2.30	1.95	2.54	2.60	2.73	2.04	1.74	2.08
	mPW2-PLYP	1.45	1.49	0.51	2.38	2.47	1.96	2.67	2.76	2.73	2.17	1.73	2.24
	PBE-QIDH	0.28	0.37	1.12	1.08	1.12	1.35	1.33	1.42	2.16	0.90	1.54	0.97
	PBE0-DH	0.93	1.01	0.01	1.75	1.95	2.30	2.01	2.23	3.03	1.56	1.78	1.73
RSH	M11	2.37	2.53	1.13	2.39	2.63	2.94	2.90	3.13	3.44	2.55	2.50	2.76
	MN12-SX	0.68	0.84	1.61	1.46	1.54	0.52	1.50	1.57	1.17	1.21	1.10	1.32
	MN12-SX-D3	0.93	1.02	0.92	1.72	1.90	1.20	1.59	1.82	1.85	1.41	1.32	1.58
RSH meta-GGA	ωB97M-V	3.42	3.57	0.85	1.49	1.56	1.24	1.45	1.55	1.02	2.12	1.02	2.23
RSH GGA	ωB97	0.69	0.79	2.82	0.72	0.75	0.83	0.87	0.87	0.18	0.76	1.28	0.80
	ωB97X	1.84	1.95	4.33	0.95	1.38	2.27	0.78	1.23	1.54	1.19	2.71	1.52

	ωB97X-D	1.21	1.25	4.24	0.48	0.67	2.12	0.52	0.62	1.35	0.74	2.57	0.85
	ωB97X-V	0.51	0.61	0.87	1.36	1.44	1.13	1.44	1.55	1.15	1.10	1.05	1.20
	HSE06	1.26	1.36	0.39	2.01	2.25	1.82	2.28	2.53	2.48	1.85	1.57	2.05
	HSE06-D3	2.06	2.15	0.83	2.92	3.06	3.04	3.19	3.34	3.71	2.72	2.53	2.85
	LC-PBE	2.38	2.73	2.52	3.31	3.64	4.71	3.63	3.95	5.50	3.10	4.24	3.44
	LC-ωPBE	0.90	1.13	0.01	1.67	1.91	2.11	1.96	2.19	2.83	1.51	1.65	1.74
	LC-ωPBE-XDM	1.30	1.55	0.01	2.32	2.50	2.79	–	–	–	1.30	0.01	1.55
	LC-ωPBE-D3	1.92	2.05	1.56	2.79	2.90	3.68	3.08	3.19	4.40	2.60	3.22	2.72
	CAM-B3LYP	2.88	2.96	1.43	3.80	3.91	3.66	4.06	4.19	4.37	3.58	3.15	3.69
	CAM-B3LYP-D3	3.91	3.95	2.85	4.83	4.90	5.08	5.10	5.17	5.79	4.61	4.57	4.67
GH meta-GGA	M06	1.38	1.76	3.64	0.93	1.24	1.54	0.94	1.25	0.32	1.08	1.83	1.42
	M06-2X	1.20	1.24	1.43	2.11	2.19	3.00	2.60	2.67	4.03	1.97	2.82	2.03
	M06-HF	3.89	4.05	6.12	4.27	4.32	7.46	4.93	4.99	8.44	4.37	7.34	4.45
	MN15	0.64	0.96	1.80	1.01	1.09	0.28	1.14	1.16	0.96	0.93	1.01	1.07
	BMK	1.97	2.24	2.84	1.11	1.49	0.92	1.01	1.38	0.01	1.36	1.26	1.70
GH GGA	BH&HLYP	3.35	3.44	2.96	4.25	4.38	5.09	4.49	4.63	5.75	4.03	4.60	4.15
	SOGGA11-X	1.08	1.16	0.03	2.25	2.35	2.25	2.31	2.42	2.81	1.88	1.70	1.98
	B3PW91	0.60	0.97	2.48	1.01	1.11	0.24	1.23	1.28	0.47	0.95	1.07	1.12
	B3PW91-D3	1.26	1.35	0.19	2.19	2.32	2.44	2.47	2.61	3.15	1.98	1.93	2.09
	PBE0	1.16	1.24	0.30	2.61	3.23	1.94	2.23	2.46	2.63	2.00	1.62	2.31
	PBE0-D3	1.95	2.04	1.00	3.25	3.50	3.24	3.16	3.28	3.93	2.79	2.72	2.94
	B3LYP	1.44	1.61	0.88	2.35	2.57	1.38	2.61	2.83	2.08	2.13	1.45	2.34
	B3LYP-D3	3.21	3.24	1.71	4.12	4.19	3.98	4.38	4.46	4.67	3.90	3.45	3.96
	X3LYP	2.18	2.29	0.02	3.09	3.25	2.29	3.35	3.52	2.98	2.87	1.76	3.02
	X3LYP-D3	4.21	4.23	3.26	5.13	5.17	5.52	4.99	5.04	5.82	4.78	4.87	4.81
	O3LYP	0.52	0.73	2.39	1.18	1.26	0.10	1.37	1.46	0.55	1.02	1.01	1.15
	O3LYP-D3	2.33	2.35	1.91	3.27	3.31	4.20	3.51	3.56	4.85	3.03	3.65	3.07
	B97-2	0.80	1.07	2.68	0.86	1.04	0.48	1.06	1.15	0.19	0.91	1.12	1.09

Local meta-GGA	TPSS	0.53	0.58	2.20	1.25	1.43	0.13	1.53	1.72	0.83	1.10	1.05	1.24
	TPSS-D3	1.30	1.34	0.74	2.53	2.59	1.95	2.81	2.88	2.65	2.21	1.78	2.27
	M06-L	2.18	2.40	3.97	1.17	1.69	2.61	1.00	1.48	1.42	1.45	2.66	1.86
	M11-L	2.09	2.22	3.06	1.18	1.66	0.99	2.53	4.01	0.17	1.93	1.40	2.63
	MN12-L	0.81	0.88	1.46	1.70	1.78	0.74	1.70	1.76	1.33	1.40	1.17	1.47
	MN15-L	2.03	2.06	3.65	0.94	1.26	2.00	0.85	1.17	1.43	1.27	2.36	1.50
Local GGA	mPW91	0.88	0.88	0.92	1.65	1.75	1.31	1.87	2.01	2.01	1.46	1.42	1.55
	BLYP	0.79	0.87	3.80	1.28	1.55	1.44	1.49	1.79	15.25	1.19	6.83	1.40
	BLYP-D3	2.27	2.36	0.72	3.20	3.32	1.64	3.47	3.59	2.35	2.98	1.57	3.09
	BP86	2.20	2.34	5.69	1.23	1.64	3.35	1.07	1.47	2.61	1.50	3.88	1.82
	BP86-D3	0.52	0.66	2.97	0.90	1.08	0.62	1.09	1.32	0.12	0.84	1.23	1.02
	PBE	0.61	0.70	3.10	1.29	1.40	0.72	1.47	1.61	0.00	1.12	1.28	1.24
	PBE-D3	0.91	1.08	1.78	1.88	2.06	0.60	2.16	2.34	1.31	1.65	1.23	1.83
Local LDA	SVWN5	3.07	3.45	0.84	4.08	4.46	3.34	4.41	4.79	4.15	3.86	2.78	4.23
AMUE		1.62	–	1.91	2.11	–	2.16	2.30	–	2.73	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.8. The mean unsigned errors (MUE), root-mean squared errors (RMSE), average mean unsigned errors (AMUE) and average root-mean squared errors (ARMSE) across nucleic acid components and/or Karlsruhe basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies.[a]

		def2-SVP			**def2-TZVPP**			**def2-QZVPP**					
		Nucleobases		**PO_4^-**	**Nucleobases**		**PO_4^-**	**Nucleobases**		**PO_4^-**	**AMUE**		**ARMSE**
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Nucleobases**	**PO_4^-**	**Base**
Double-Hybrid	B2-PLYP	2.61	2.63	1.49	1.11	1.22	1.17	1.48	1.62	1.34	1.73	1.33	1.82
	B2-PLYP-D3	3.55	3.56	2.89	2.06	2.09	2.58	2.42	2.47	2.75	2.68	2.74	2.71
	mPW2-PLYP	3.74	3.77	2.88	2.23	2.29	2.63	2.56	2.65	2.76	2.84	2.76	2.90
	PBE-QIDH	1.62	1.74	0.80	0.82	0.84	1.63	1.22	1.29	2.11	1.22	1.51	1.29

	PBE0-DH	2.46	2.65	2.03	1.52	1.67	2.65	1.88	2.10	3.01	1.95	2.56	2.14
RSH	M11	3.96	4.11	3.63	2.47	2.73	2.87	2.81	3.05	3.67	3.08	3.39	3.30
	MN12-SX	2.09	2.29	0.34	1.15	1.19	0.78	1.44	1.49	1.05	1.56	0.72	1.66
	MN12-SX-D3	2.49	2.72	1.02	1.41	1.50	1.47	1.70	1.85	1.73	1.86	1.41	2.02
RSH meta-GGA	ωB97M-V	6.50	7.97	2.50	1.45	1.51	2.13	1.98	2.33	1.91	3.31	2.18	3.94
RSH GGA	ωB97	1.21	1.49	0.12	0.51	0.68	0.39	0.80	0.81	0.15	0.84	0.22	1.00
	ωB97X	1.69	2.72	1.47	1.25	1.56	1.83	0.80	1.29	1.52	1.25	1.61	1.86
	ωB97X-D	1.57	2.40	1.59	0.66	0.86	1.74	0.48	0.63	1.33	0.90	1.55	1.30
	ωB97X-V	2.45	2.56	1.99	1.21	1.30	1.83	1.28	1.43	1.68	1.65	1.83	1.76
	HSE06	2.98	3.20	2.49	1.85	2.04	2.36	2.16	2.41	2.52	2.33	2.46	2.55
	HSE06-D3	3.90	4.04	3.72	2.70	2.84	3.58	3.07	3.22	3.74	3.22	3.68	3.36
	LC-PBE	3.89	4.23	4.69	2.87	3.25	4.79	3.48	3.81	5.45	3.41	4.98	3.76
	LC-ωPBE	2.52	2.78	2.60	1.46	1.67	2.48	1.83	2.06	2.84	1.94	2.64	2.17
	LC-ωPBE-XDM	–	–	–	1.71	1.95	2.69	–	–	–	1.71	2.69	1.95
	LC-ωPBE-D3	3.65	3.80	4.17	2.50	2.64	4.05	2.95	3.07	4.41	3.03	4.21	3.17
	CAM-B3LYP	5.17	5.27	5.15	3.61	3.72	4.31	3.95	4.07	4.40	4.24	4.62	4.35
	CAM-B3LYP-D3	6.21	6.27	6.57	4.64	4.71	5.73	4.98	5.05	5.82	5.28	6.04	5.34
GH meta-GGA	M06	1.38	1.43	1.03	0.83	1.38	0.94	0.93	1.22	0.59	1.05	0.85	1.34
	M06-2X	2.90	2.96	3.32	1.99	2.05	3.52	2.44	2.51	4.02	2.44	3.62	2.51
	M06-HF	9.42	11.88	7.54	4.45	4.51	7.78	4.51	4.56	7.98	6.13	7.77	6.98
	MN15	1.36	1.51	1.00	0.64	0.92	0.40	1.05	1.10	0.92	2.17	0.77	1.18
	BMK	0.92	1.14	0.35	1.08	1.63	0.53	1.00	1.39	0.07	1.00	0.32	1.39
GH GGA	BH&HLYP	5.33	5.44	5.37	4.02	4.15	5.48	4.37	4.51	5.74	4.58	5.53	4.70
	SOGGA11-X	3.01	3.10	2.28	1.95	2.05	2.69	2.26	2.36	2.86	2.41	2.61	2.50
	B3PW91	1.56	1.61	0.34	0.82	1.02	0.20	1.13	1.19	0.51	1.17	0.35	1.27

	B3PW91-D3	3.02	3.14	3.01	1.90	2.03	2.88	2.34	2.47	3.18	2.42	3.02	2.55
	PBE0	2.85	3.07	2.46	2.53	3.26	2.51	2.10	2.33	2.65	2.50	2.54	2.89
	PBE0-D3	3.78	3.92	3.76	3.13	3.47	3.81	3.03	3.16	3.95	3.32	3.84	3.51
	B3LYP	3.75	3.87	3.03	2.19	2.39	2.12	2.50	2.71	2.13	2.81	2.43	2.99
	B3LYP-D3	5.52	5.57	5.62	3.96	4.02	4.71	4.27	4.34	4.72	4.58	5.02	4.64
	X3LYP	4.55	4.65	3.98	2.96	3.10	3.08	3.24	3.40	3.04	3.58	3.37	3.72
	X3LYP-D3	6.59	6.63	7.22	5.00	5.03	6.31	5.27	5.31	6.27	5.62	6.60	5.66
	O3LYP	2.19	2.46	1.35	1.17	1.23	1.21	1.30	1.38	0.65	1.55	1.07	1.69
	O3LYP-D3	4.71	4.78	5.65	3.39	3.42	5.51	3.43	3.48	4.94	3.84	5.37	3.89
	B97-2	1.47	1.49	0.02	0.71	0.97	0.13	0.98	1.09	0.24	1.05	0.13	1.18
Local meta-GGA	TPSS	2.29	2.43	1.45	1.12	1.26	0.89	1.43	1.62	0.89	1.61	1.08	1.77
	TPSS-D3	3.58	3.64	3.27	2.35	2.41	2.71	2.72	2.78	2.71	2.88	2.90	2.94
	M06-L	0.86	1.58	2.01	1.32	1.88	1.76	1.02	1.55	1.45	1.07	1.74	1.67
	M11-L	1.17	1.90	2.74	1.72	2.08	1.37	1.42	1.79	0.88	1.43	1.66	1.92
	MN12-L	2.02	2.09	0.20	1.35	1.37	0.97	1.68	1.73	1.19	1.68	0.79	1.73
	MN15-L	0.58	0.99	2.20	0.91	1.06	0.66	0.97	1.24	1.42	0.82	1.43	1.10
Local GGA	mPW91	2.33	2.58	1.79	1.49	1.54	1.83	1.77	1.90	2.06	1.86	1.89	2.01
	BLYP	2.94	3.07	1.53	1.21	1.44	0.29	1.40	1.68	0.62	1.85	0.81	2.06
	BLYP-D3	5.05	5.08	4.62	3.16	3.25	2.80	3.36	3.47	2.47	3.86	3.30	3.94
	BP86	0.72	1.02	1.65	1.50	1.81	2.65	1.09	1.52	2.52	1.10	2.27	1.45
	BP86-D3	1.76	1.92	1.07	0.79	0.87	0.08	0.99	1.20	0.20	1.18	0.45	1.33
	PBE	2.26	2.51	1.26	1.28	1.41	0.33	1.37	1.49	0.08	1.64	0.56	1.80
	PBE-D3	3.18	3.33	2.58	1.77	1.93	1.65	2.09	2.30	2.21	2.35	2.15	2.52
Local LDA	SVWN5	5.06	5.48	5.03	3.78	4.18	4.06	4.27	4.64	4.20	4.37	4.43	4.77
AMUE		3.12	–	2.71	1.99	–	2.44	2.21	–	2.53	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.9. The mean unsigned errors (MUE), root-mean squared errors (RMSE), average mean unsigned errors (AMUE) and average root-mean squared errors (ARMSE) across nucleic acid components and/or Dunning basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections.[a]

		aug-cc-pVDZ			aug-cc-pVTZ			aug-cc-pVQZ					
		Nucleobases		PO_4^-	Nucleobases		PO_4^-	Nucleobases		PO_4^-	AMUE		ARMSE
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Nucleobases**	**PO_4^-**	**Base**
Double-Hybrid	B2-PLYP	0.82	0.85	0.99	1.58	1.70	0.96	1.79	1.92	1.62	1.40	1.19	1.49
	B2-PLYP-D3	1.77	1.77	0.42	2.52	2.57	2.36	2.73	2.78	3.02	2.34	1.93	2.37
	mPW2-PLYP	1.89	1.90	0.45	2.66	2.73	2.37	2.86	2.95	3.03	2.47	1.95	2.53
	PBE-QIDH	0.44	0.52	0.07	1.21	1.31	1.73	1.40	1.52	2.38	1.02	1.39	1.12
	PBE0-DH	1.13	1.26	0.81	1.90	2.10	2.58	2.12	2.33	3.22	1.72	2.20	1.90
RSH	M11	2.73	2.85	1.68	2.72	2.94	3.41	3.18	3.40	3.93	2.88	3.01	3.06
	MN12-SX	0.83	0.92	0.89	1.65	1.79	0.91	1.78	1.93	1.78	1.42	1.19	1.55
	MN12-SX-D3	1.09	1.23	0.21	1.93	2.19	1.59	1.73	2.15	2.46	1.58	1.42	1.86
RSH meta-GGA	ωB97M-V	3.17	3.18	0.35	1.56	1.63	1.07	1.62	1.70	1.50	2.12	0.97	2.17
RSH GGA	ωB97	0.41	0.51	2.15	0.88	0.88	0.52	1.01	1.02	0.11	0.77	0.93	0.80
	ωB97X	1.57	1.68	3.70	0.82	1.29	2.11	0.76	1.19	1.40	1.05	2.40	1.39
	ωB97X-D	0.94	0.98	3.63	0.47	0.62	1.97	0.54	0.62	1.23	0.65	2.28	0.74
	ωB97X-V	0.74	0.82	0.31	1.48	1.57	1.07	1.58	1.68	1.51	1.27	0.96	1.36
	HSE06	1.42	1.57	0.25	2.21	2.43	2.10	2.47	2.70	2.73	2.03	1.69	2.23
	HSE06-D3	2.33	2.40	1.47	3.12	3.25	3.32	3.38	3.52	3.96	2.94	2.92	3.06
	LC-PBE	2.85	3.15	3.33	3.62	3.92	5.11	3.93	4.23	5.87	3.47	4.77	3.77
	LC-ωPBE	1.14	1.39	0.68	2.02	2.22	2.55	2.33	2.52	3.26	1.83	2.16	2.04
	LC-ωPBE-XDM	1.65	1.85	0.68	2.64	2.92	5.13	–	–	–	2.15	2.91	2.39
	LC-ωPBE-D3	2.27	2.38	2.25	3.14	3.24	4.12	3.45	3.55	4.83	2.95	3.73	3.06
	CAM-B3LYP	3.22	3.27	2.11	4.11	4.21	4.06	4.38	4.49	4.76	3.90	3.64	3.99

	CAM-B3LYP-D3	4.25	4.28	3.53	5.14	5.20	5.48	5.41	5.48	6.18	4.93	5.06	4.99
GH meta-GGA	M06	1.11	1.51	3.00	0.97	1.22	1.32	0.98	1.23	0.32	1.02	1.55	1.32
	M06-2X	1.48	1.51	2.05	2.33	2.40	3.29	2.82	2.89	4.31	2.21	3.22	2.27
	M06-HF	4.41	4.58	6.83	4.91	4.96	8.72	5.75	5.80	9.66	5.02	8.40	5.11
	MN15	0.51	0.75	1.04	1.21	1.22	0.89	1.45	1.47	0.25	1.06	0.73	1.15
	BMK	1.73	1.99	2.23	1.08	1.44	0.77	1.01	1.34	1.69	1.27	1.56	1.59
GH GGA	BH&HLYP	3.61	3.69	3.63	4.43	4.56	5.34	4.64	4.78	5.95	4.23	4.97	4.34
	SOGGA11-X	1.32	1.37	0.62	2.47	2.56	2.47	2.49	2.59	3.06	2.09	2.05	2.17
	B3PW91	0.44	0.78	1.87	1.12	1.17	0.02	1.33	1.36	0.70	0.96	0.86	1.10
	B3PW91-D3	1.53	1.59	0.80	2.37	2.49	2.69	2.63	2.76	3.37	2.18	2.29	2.28
	PBE0	1.30	1.45	0.35	2.81	3.39	2.21	2.41	2.62	2.87	2.17	1.81	2.49
	PBE0-D3	2.22	2.29	1.65	3.47	3.71	3.51	3.34	3.46	4.17	3.01	3.11	3.15
	B3LYP	1.75	1.86	0.22	2.63	2.82	1.75	2.90	3.09	2.43	2.43	1.47	2.59
	B3LYP-D3	3.52	3.54	2.38	4.40	4.46	4.35	4.67	4.73	5.03	4.20	3.92	4.24
	X3LYP	2.50	2.58	0.70	3.39	3.53	2.67	3.65	3.80	3.35	3.18	2.24	3.30
	X3LYP-D3	4.54	4.55	3.93	5.42	5.46	5.90	5.29	5.33	6.19	5.08	5.34	5.11
	O3LYP	0.56	0.64	1.61	1.38	1.49	0.36	1.56	1.70	0.93	1.17	0.97	1.28
	O3LYP-D3	2.68	2.70	2.68	3.62	3.66	4.66	3.84	3.88	5.23	3.38	4.19	3.41
	B97-2	0.53	0.84	2.03	0.97	1.07	0.23	1.14	1.20	0.38	0.88	0.88	1.04
Local meta-GGA	TPSS	0.65	0.76	1.60	1.41	1.58	0.40	1.63	1.81	1.00	1.23	1.00	1.38
	TPSS-D3	1.58	1.61	0.15	2.70	2.75	2.22	2.92	2.98	2.82	2.40	1.73	2.45
	M06-L	1.95	2.18	3.35	1.14	1.59	2.38	0.95	1.39	1.08	1.35	2.27	1.72
	M11-L	1.32	1.50	1.61	0.87	1.11	0.57	1.55	1.70	2.28	1.25	1.49	1.44
	MN12-L	1.07	1.10	0.50	1.94	2.12	1.28	2.10	2.33	2.24	1.70	1.34	1.85
	MN15-L	1.76	1.78	2.76	0.74	1.10	1.60	0.68	1.01	1.09	1.06	1.82	1.30
Local GGA	mPW91	0.98	1.02	0.29	1.77	1.91	1.58	1.97	2.16	2.25	1.57	1.37	1.70
	BLYP	0.79	0.93	3.08	1.53	1.87	0.90	1.82	2.16	14.69	1.38	6.22	1.65
	BLYP-D3	2.66	2.73	0.00	3.63	3.73	2.18	3.94	4.04	2.91	3.41	1.70	3.50

	BP86	1.91	2.04	5.08	1.06	1.46	3.03	1.02	1.33	2.30	1.33	3.47	1.61
	BP86-D3	0.47	0.53	2.35	1.00	1.24	0.30	1.23	1.50	0.42	0.90	1.02	1.09
	PBE	0.62	0.70	2.42	1.34	1.49	0.29	1.68	1.89	0.42	1.21	1.04	1.36
	PBE-D3	1.25	1.35	1.10	2.21	2.36	1.02	2.51	2.67	1.74	1.99	1.29	2.13
Local LDA	SVWN5	3.92	4.20	2.00	4.27	4.63	3.64	4.59	4.95	4.41	4.26	3.35	4.59
AMUE		1.76	–	1.77	2.30	–	2.40	2.48	–	3.04	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.10. The mean unsigned errors (MUE), root-mean squared errors (RMSE), average mean unsigned errors (AMUE) and average root-mean squared errors (ARMSE) across nucleic acid components and/or Karlsruhe basis sets (kcal/mol) for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections.[a]

		def2-SVP			**def2-TZVPP**			**def2-QZVPP**					
		Nucleobases		**PO_4^-**	**Nucleobases**		**PO_4^-**	**Nucleobases**		**PO_4^-**	**AMUE**		**ARMSE**
Family	**Functional**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**MUE**	**RMSE**	**UE**	**Nucleobases**	**PO_4^-**	**Base**
Double-Hybrid	B2-PLYP	6.86	6.89	10.96	1.68	1.73	2.85	1.67	1.79	2.07	3.40	5.29	3.47
	B2-PLYP-D3	7.80	7.83	12.36	2.62	2.64	4.25	2.62	2.66	3.48	4.35	6.70	4.38
	mPW2-PLYP	8.02	8.04	12.37	2.80	2.84	4.30	2.75	2.83	3.49	4.52	6.72	4.57
	PBE-QIDH	5.06	5.12	8.72	1.10	1.23	3.08	1.33	1.45	2.73	2.50	4.85	2.60
	PBE0-DH	5.66	5.75	9.53	1.83	2.01	3.85	2.00	2.21	3.52	3.16	5.63	3.32
RSH	M11	7.62	7.71	11.95	2.92	3.15	3.87	2.97	3.20	4.23	4.50	6.68	4.69
	MN12-SX	5.09	5.20	7.75	1.40	1.47	1.80	1.57	1.66	1.79	2.69	3.78	2.78
	MN12-SX-D3	5.60	5.69	8.44	1.66	1.85	2.49	1.83	2.05	2.48	3.03	4.47	3.20
RSH meta-GGA	ωB97M-V	6.99	7.03	11.49	1.79	1.84	3.17	1.56	1.63	1.99	3.45	5.55	3.50
RSH GGA	ωB97	4.92	5.01	8.72	0.69	0.72	0.67	0.85	0.85	0.18	2.15	3.19	2.19
	ωB97X	4.48	5.12	6.74	0.93	1.30	0.78	0.75	1.24	1.16	2.05	2.89	2.55
	ωB97X-D	4.62	5.02	6.09	0.47	0.63	0.71	0.49	0.61	0.89	1.86	2.56	2.09
	ωB97X-V	6.13	6.20	10.42	1.49	1.58	2.81	1.37	1.48	1.80	3.00	5.01	3.09

	HSE06	6.39	6.49	10.42	2.14	2.35	3.48	2.24	2.48	3.00	3.59	5.63	3.77
	HSE06-D3	7.30	7.38	11.64	3.06	3.18	4.70	3.16	3.29	4.22	4.51	6.85	4.62
	LC-PBE	7.56	7.74	12.92	3.21	3.56	5.81	3.55	3.88	5.82	4.77	8.19	5.06
	LC-ωPBE	6.27	6.39	11.03	1.77	2.01	3.69	1.92	2.14	3.30	3.32	6.01	3.52
	LC-ωPBE-XDM	–	–	–	2.10	2.30	4.11	–	–	–	2.10	4.11	2.30
	LC-ωPBE-D3	7.40	7.49	12.61	2.89	3.01	5.26	3.04	3.15	4.88	4.44	7.58	4.55
	CAM-B3LYP	9.43	9.48	14.52	4.03	4.12	5.60	4.04	4.16	4.91	5.83	8.35	5.92
	CAM-B3LYP-D3	10.46	10.50	15.94	5.06	5.12	7.02	5.07	5.14	6.33	6.86	9.77	6.92
GH meta-GGA	M06	3.85	4.04	6.57	0.87	1.24	0.02	0.94	1.18	0.06	1.89	2.22	2.15
	M06-2X	6.16	6.20	10.91	2.29	2.34	4.45	2.52	2.59	4.41	3.66	6.59	3.71
	M06-HF	13.27	15.10	16.59	4.91	4.98	8.99	4.77	4.82	8.83	7.65	11.47	8.30
	MN15	4.73	4.85	9.47	0.80	0.91	1.46	1.09	1.13	1.26	2.21	4.06	2.29
	BMK	3.23	3.37	7.01	0.97	1.46	0.24	0.95	1.34	0.35	1.72	2.53	2.06
GH GGA	BH&HLYP	8.65	8.71	13.06	4.34	4.46	6.49	4.44	4.58	6.11	5.81	8.55	5.92
	SOGGA11-X	6.11	6.16	9.65	2.27	2.35	3.57	2.36	2.46	3.24	3.58	5.49	3.65
	B3PW91	4.47	4.60	8.04	1.00	1.06	1.28	1.18	1.22	0.98	2.22	3.43	2.29
	B3PW91-D3	6.34	6.39	10.71	2.24	2.35	3.95	2.43	2.55	3.65	3.67	6.10	3.77
	PBE0	6.23	6.33	10.31	2.86	3.55	3.66	2.19	2.41	3.14	3.76	5.70	4.10
	PBE0-D3	7.17	7.24	11.62	3.51	3.82	4.96	3.12	3.24	4.44	4.60	7.00	4.76
	B3LYP	7.83	7.88	12.13	2.61	2.76	3.40	2.60	2.79	2.67	4.35	6.07	4.48
	B3LYP-D3	9.60	9.63	14.73	4.38	4.42	5.99	4.37	4.43	5.27	6.12	8.66	6.16
	X3LYP	8.72	8.77	13.25	3.39	3.51	4.40	3.34	3.49	3.59	5.15	7.08	5.26
	X3LYP-D3	10.75	10.79	16.48	5.43	5.46	7.63	5.37	5.41	6.83	7.18	10.31	7.22
	O3LYP	6.20	6.29	10.29	1.49	1.68	2.90	1.37	1.46	1.45	3.02	4.88	3.14
	O3LYP-D3	8.72	8.77	14.59	3.99	4.01	7.20	3.57	3.62	5.75	5.43	9.18	5.46
	B97-2	4.23	4.35	7.58	0.91	0.98	1.28	1.02	1.11	0.75	2.05	3.20	2.15

Local meta-GGA	TPSS	5.98	6.03	9.94	1.51	1.64	2.23	1.55	1.72	1.60	3.01	4.59	3.13
	TPSS-D3	7.27	7.30	11.76	2.80	2.84	4.05	2.84	2.90	3.42	4.30	6.41	4.35
	M06-L	2.07	2.18	3.85	1.06	1.70	1.13	0.94	1.44	0.96	1.36	1.98	1.77
	M11-L	1.84	1.89	2.56	1.27	1.71	0.44	1.05	1.49	0.08	1.39	1.03	1.70
	MN12-L	4.24	4.46	6.47	1.58	1.64	1.89	1.81	1.91	1.89	2.55	3.42	2.67
	MN15-L	2.35	2.46	4.77	0.53	0.75	0.54	0.81	1.10	0.93	1.23	2.08	1.44
Local GGA	mPW91	5.69	5.79	9.55	1.68	1.83	2.96	1.82	1.98	2.57	3.06	5.03	3.20
	BLYP	7.86	7.90	12.11	1.58	1.84	1.32	1.46	1.76	0.14	3.63	4.52	3.83
	BLYP-D3	9.97	9.99	15.19	3.69	3.76	4.40	3.49	3.60	3.22	5.72	7.60	5.78
	BP86	3.75	3.86	7.16	1.09	1.46	1.39	1.02	1.43	1.94	1.95	3.49	2.25
	BP86-D3	5.65	5.70	9.88	0.90	1.09	1.33	1.05	1.27	0.78	2.53	4.00	2.69
	PBE	6.51	6.60	10.71	1.32	1.48	1.82	1.43	1.57	0.80	3.09	4.44	3.22
	PBE-D3	7.44	7.50	12.03	2.26	2.38	3.14	1.97	2.19	1.49	3.89	5.55	4.02
Local LDA	SVWN5	9.57	9.79	14.80	4.21	4.57	5.37	4.38	4.74	4.82	6.05	8.33	6.37
AMUE		6.54	–	10.55	2.25	–	3.36	2.27	–	2.88	–	–	–

[a)] Calculated with respect to CCSD(T)/CBS values (Figure 2.2).

Table A.11. The percent error (PE) and average mean percent error (AMPE) across nucleic acid components (%) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies evaluated with aug-cc-pVTZ.

		Adenine			Guanine			Cytosine		Thymine			Uracil			Phosphate[a]		Favored[b] Coor. Sites	Other[c] Coor. Sites	All[d] Coor. Sites	Cation–π[e] Coor. Sites	ALL (w/ Cation–π)[f] Coor. Sites
Family	**Functional**	N6,N7	N3	Li^+–π	N6,N1	O6,N7	N2,N3	O2,N3	N3,N4	O4	O2	Li^+–π	O4	O2	2NB	1B1NB	2B	**AMPE**	**AMPE**	**AMPE**	**AMPE**	**AMPE**
Double-Hybrid	B2PLYP	0.42	3.10	0.72	1.00	1.95	1.82	1.53	1.22	3.98	1.96	4.31	3.75	1.47	0.33	0.25	0.72	1.99	1.44	1.68	2.52	1.78
	PBE-QIDH	0.89	1.79	3.25	0.15	1.78	0.87	1.87	0.03	2.40	0.27	1.63	2.31	0.18	0.83	0.22	1.63	1.68	0.64	1.09	1.99	1.20
RSH GGA	ωB97	2.26	0.90	2.58	2.37	0.81	4.07	0.86	1.78	1.25	0.61	13.02	1.27	0.86	0.51	0.67	2.04	1.16	1.66	1.45	7.80	2.24
	ωB97X-D	2.64	1.31	1.44	3.08	0.32	4.20	0.85	2.07	0.25	2.70	5.08	0.21	3.22	1.30	1.51	2.83	0.93	2.62	1.89	3.26	2.06
GH meta-GGA	M06	4.96	1.52	3.81	4.54	0.57	4.53	0.50	3.86	1.39	1.58	12.36	1.35	1.89	0.94	0.98	2.56	1.62	2.68	2.23	8.09	2.96
	MN15	3.51	0.14	3.29	2.75	1.47	3.27	0.82	2.78	1.65	0.37	6.12	1.56	0.66	0.17	0.17	2.16	1.53	1.54	1.53	4.71	1.93
GH GGA	B3PW91	3.35	2.12	0.81	2.32	1.17	1.98	0.49	1.87	2.27	0.52	9.37	2.07	1.26	0.15	0.62	1.93	1.58	1.58	1.58	5.09	2.02
	B97-2	3.60	1.72	0.48	2.49	0.71	2.71	0.16	2.12	1.95	0.86	9.85	1.76	1.57	0.29	0.86	2.24	1.41	1.82	1.65	5.17	2.09
Local meta-GGA	TPSS	0.02	5.03	5.71	1.21	2.18	4.13	1.37	2.12	3.93	0.82	5.31	3.47	0.14	0.08	0.32	0.31	1.84	1.76	1.80	5.51	2.26
	MN15-L	4.90	2.59	3.84	3.05	1.26	4.77	1.33	3.04	0.28	3.72	1.70	0.41	4.17	1.22	1.88	4.18	1.57	3.43	2.63	2.77	2.65
Local GGA	BP86-D3	0.92	4.86	5.76	0.54	1.28	2.97	0.12	0.86	3.60	0.26	5.82	2.45	1.16	0.38	0.06	0.09	1.46	1.35	1.40	5.79	1.95
	PBE	1.49	4.92	7.10	0.42	2.36	4.82	0.77	0.71	3.72	0.52	6.93	3.22	0.41	0.44	0.32	0.88	2.00	1.63	1.79	7.02	2.44

[a] For Li^+-phosphate interactions, 2NB denotes interactions involving 2 non-bridging oxygen atoms, 1B1NB denotes interactions involving 1 bridging and 1 non-bridging oxygen atom and 2B denotes interactions involving 2 bridging oxygen atoms.
[b] AMPEs calculated over the lowest energy Li^+ coordination positions for Li^+–nucleobase and Li^+–phosphate interactions.
[c] AMPEs calculated over all Li^+ coordinated positions excluding the preferred binding locations.
[d] AMPEs calculated over all coordinated positions.
[e] AMPEs calculated over cation–π interactions.
[f] AMPEs calculated over all Li^+ complexes, including cation–π interactions.

Table A.12. The percent error (PE) and average mean percent error (AMPE) across nucleic acid components (%) for Li^+–nucleobase and Li^+–phosphate DFT binding energies evaluated with the aug-cc-pVTZ basis set without counterpoise corrections.

		Adenine			Guanine			Cytosine		Thymine			Uracil			Phosphate[a]		Favored[b] Coor. Sites	Other[c] Coor. Sites	All[d] Coor. Sites	Cation–π[e] Coor. Sites	ALL (w/ Cation–π)[f] Coor. Sites
Family	**Functional**	N6,N7	N3	Li^+–π	N6,N1	O6,N7	N2,N3	O2,N3	N3,N4	O4	O2	Li^+–π	O4	O2	2NB	1B1NB	2B	**AMPE**	**AMPE**	**AMPE**	**AMPE**	**AMPE**
Double-Hybrid	B2PLYP	0.99	3.64	1.92	1.54	2.36	2.53	1.93	1.81	4.47	2.44	2.14	4.23	1.97	0.59	0.51	0.42	2.43	1.86	2.10	2.03	2.09
	PBE-QIDH	0.42	2.21	4.36	0.28	2.13	0.30	2.21	0.51	2.81	0.66	0.93	2.71	0.24	1.06	0.45	1.36	1.89	0.75	1.24	2.04	1.34
RSH GGA	ωB97	1.68	1.52	1.51	1.79	1.21	3.31	1.27	1.16	1.78	0.10	3.78	1.78	0.34	0.32	0.47	1.80	1.34	1.31	1.32	6.18	1.93
	ωB97X-D	2.43	1.52	1.88	2.88	0.47	3.93	0.69	1.84	0.50	2.45	4.06	0.03	2.96	1.20	1.40	2.69	0.89	2.46	1.79	2.97	1.93
GH meta-GGA	M06	4.58	1.91	2.86	4.16	0.82	4.02	0.22	3.41	1.78	1.19	10.42	1.75	1.47	0.81	0.83	2.36	1.66	2.42	2.09	6.64	2.66
	MN15	2.70	0.98	4.69	1.94	2.02	2.17	1.38	1.95	2.32	0.27	3.28	2.21	0.01	0.55	0.19	1.75	1.86	1.16	1.46	3.98	1.78
GH GGA	B3PW91	3.00	2.50	1.49	1.98	1.43	1.51	0.74	1.49	2.63	0.17	8.03	2.42	0.90	0.01	0.46	1.74	1.71	1.34	1.50	4.76	1.91
	B97-2	3.25	2.08	1.12	2.16	0.96	2.27	0.41	1.75	2.30	0.52	8.56	2.10	1.21	0.14	0.70	2.06	1.53	1.59	1.57	4.84	1.97
Local meta-GGA	TPSS	0.3	5.36	6.32	1.52	2.41	4.55	1.62	2.46	4.27	1.15	6.53	3.8	0.48	0.25	0.48	0.11	2.11	2.01	2.05	6.42	2.60
	MN15-L	4.42	2.13	4.69	2.60	0.94	4.10	0.99	2.55	0.11	3.36	0.42	0.03	3.81	0.98	1.62	3.89	1.25	3.00	2.25	2.56	2.29
Local GGA	BP86-D3	0.44	5.38	6.64	0.06	1.62	3.61	0.23	1.37	4.06	0.19	7.52	2.89	0.70	0.19	0.25	0.14	1.57	1.46	1.51	7.08	2.21
	PBE	0.79	5.66	8.3	0.27	1.88	3.92	1.27	1.43	4.35	1.14	9.2	3.83	0.23	0.18	0.07	0.58	2.05	1.66	1.83	8.75	2.69

[a] For Li^+-phosphate interactions, 2NB denotes interactions involving 2 non-bridging oxygen atoms, 1B1NB denotes interactions involving 1 bridging and 1 non-bridging oxygen atom and 2B denotes interactions involving 2 bridging oxygen atoms.
[b] AMPEs calculated over the lowest energy Li^+ coordination positions for Li^+–nucleobase and Li^+–phosphate interactions.
[c] AMPEs calculated over all Li^+ coordinated positions excluding the preferred binding locations.
[d] AMPEs calculated over all coordinated positions.
[e] AMPEs calculated over cation–π interactions.
[f] AMPEs calculated over all Li^+ complexes, including cation–π interactions.

Table A.13. The percent error (PE) and absolute difference (Abs. Differ.) across MP2/aug-cc-pVTZ and DFT/aug-cc-pVTZ optimized nucleic acid components (%) for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies evaluated with the aug-cc-pVTZ basis set.

		Adenine			Guanine			Cytosine			Thymine			Uracil			Phosphate			AMPE		
Family	**Functional**	MP2	DFT	Abs. Differ.	MP2	DFT	Abs. Differ.	MP2	DFT	Abs. Differ.	MP2	DFT	Abs. Differ.	MP2	DFT	Abs. Differ.	MP2	DFT	Abs. Differ.	MP2	DFT	Abs. Differ.
Double-Hybrid	B2PLYP	0.42	0.36	0.07	1.95	1.99	0.04	1.53	1.58	0.05	3.98	4.05	0.08	3.75	3.83	0.08	0.33	0.31	0.02	1.99	2.02	0.03
	PBE-QIDH	0.89	1.00	0.10	1.78	1.73	0.05	1.87	1.84	0.03	2.40	2.35	0.05	2.31	2.25	0.06	0.83	0.78	0.05	1.68	1.66	0.02
RSH GGA	ωB97	2.26	2.42	0.15	0.81	0.88	0.07	0.86	0.92	0.06	1.25	1.27	0.01	1.27	1.28	0.01	0.51	0.50	0.01	1.16	1.21	0.05
	ωB97X-D	2.64	2.66	0.03	0.32	0.48	0.17	0.85	0.69	0.16	0.25	0.30	0.04	0.21	0.22	0.00	1.30	1.32	0.03	0.93	0.95	0.02
GH meta-GGA	M06	4.96	5.13	0.17	0.57	0.49	0.08	0.50	0.61	0.11	1.39	1.33	0.06	1.35	1.27	0.08	0.94	1.01	0.07	1.62	1.64	0.02
	MN15	3.51	3.79	0.28	1.47	1.47	0.00	0.82	0.74	0.09	1.65	1.72	0.08	1.56	1.64	0.08	0.17	0.14	0.03	1.53	1.58	0.05
GH GGA	B3PW91	3.35	3.56	0.20	1.17	1.17	0.00	0.49	0.46	0.02	2.27	2.32	0.04	2.07	2.11	0.04	0.15	0.22	0.08	1.58	1.64	0.06
	B97-2	3.60	3.73	0.13	0.71	0.73	0.02	0.16	0.17	0.01	1.95	1.99	0.05	1.76	1.79	0.04	0.29	0.34	0.05	1.41	1.46	0.05
Local meta-GGA	TPSS	0.02	0.11	0.09	2.18	2.08	0.10	1.37	1.42	0.04	3.93	4.12	0.18	3.47	3.65	0.18	0.08	0.15	0.08	1.84	1.92	0.08
	MN15-L	4.90	4.78	0.12	1.26	1.02	0.24	1.33	1.20	0.12	0.28	0.11	0.17	0.41	0.02	0.40	1.22	1.23	0.02	1.57	1.39	0.17
Local GGA	BP86-D3	0.92	1.05	0.12	1.28	1.31	0.03	0.12	0.09	0.03	3.60	3.83	0.23	2.45	2.67	0.22	0.38	0.52	0.14	1.46	1.58	0.12
	PBE	1.49	1.61	0.12	2.36	2.27	0.09	0.77	0.83	0.05	3.72	3.93	0.21	3.22	3.43	0.21	0.44	0.58	0.14	2.00	2.11	0.11

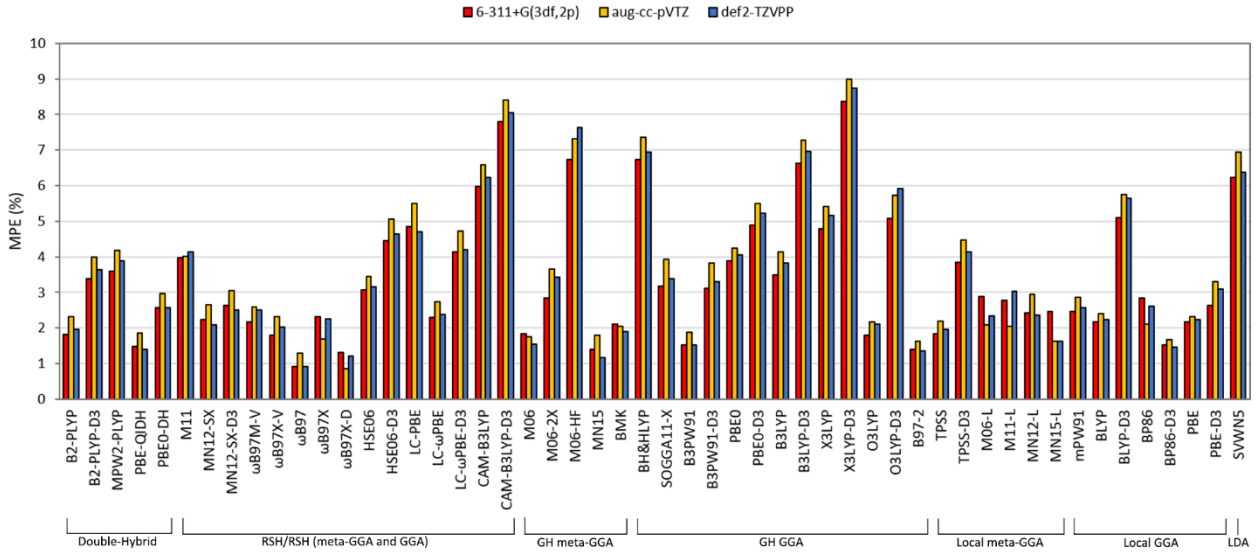


Figure A.1. MPEs for Li^+–nucleobase counterpoise-corrected DFT binding energies evaluated with the 6-311+G(3df,2p) (red), aug-cc-pVTZ (yellow) and def2- TZVPP (blue) basis sets relative to CCSD(T)/CBS reference values.

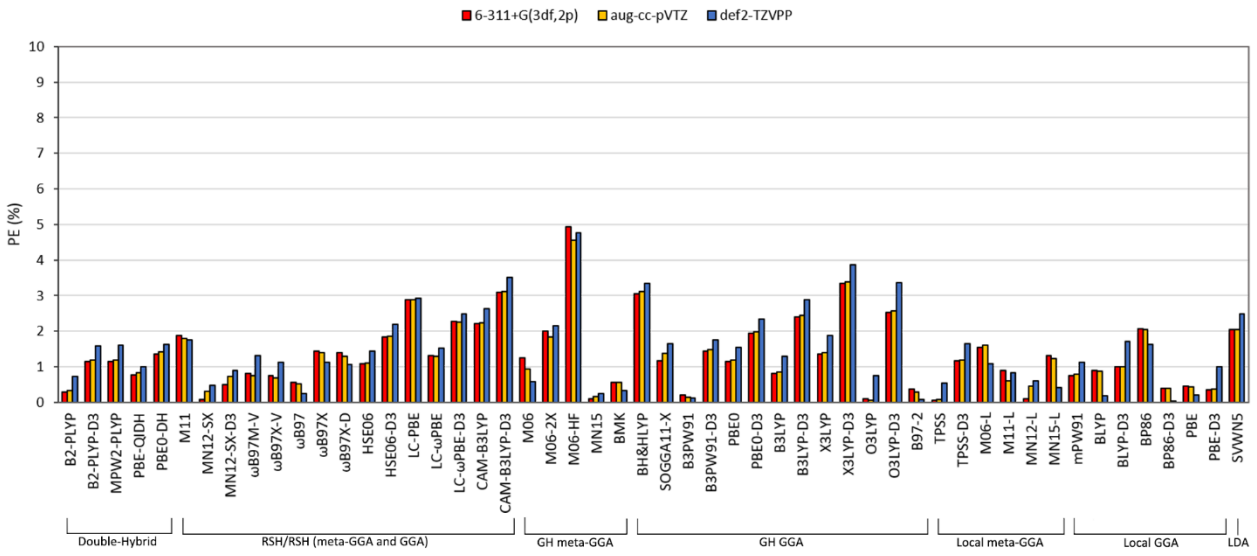


Figure A.2. PEs for Li^+–phosphate counterpoise-corrected DFT binding energies evaluated with the 6-311+G(3df,2p) (red), aug-cc-pVTZ (yellow) and def2-TZVPP (blue) basis sets relative to CCSD(T)/CBS reference values.

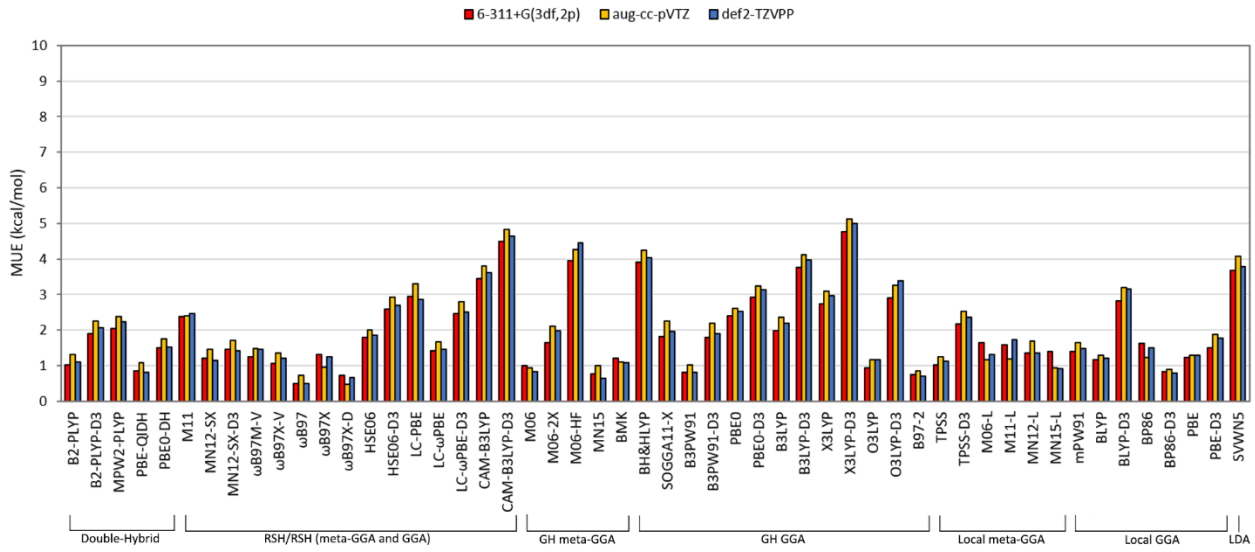


Figure A.3. MUEs for Li^+–nucleobase counterpoise-corrected DFT binding energies evaluated with the 6-311+G(3df,2p) (red), aug-cc-pVTZ (yellow) and def2-TZVPP (blue) basis sets relative to CCSD(T)/CBS reference values.

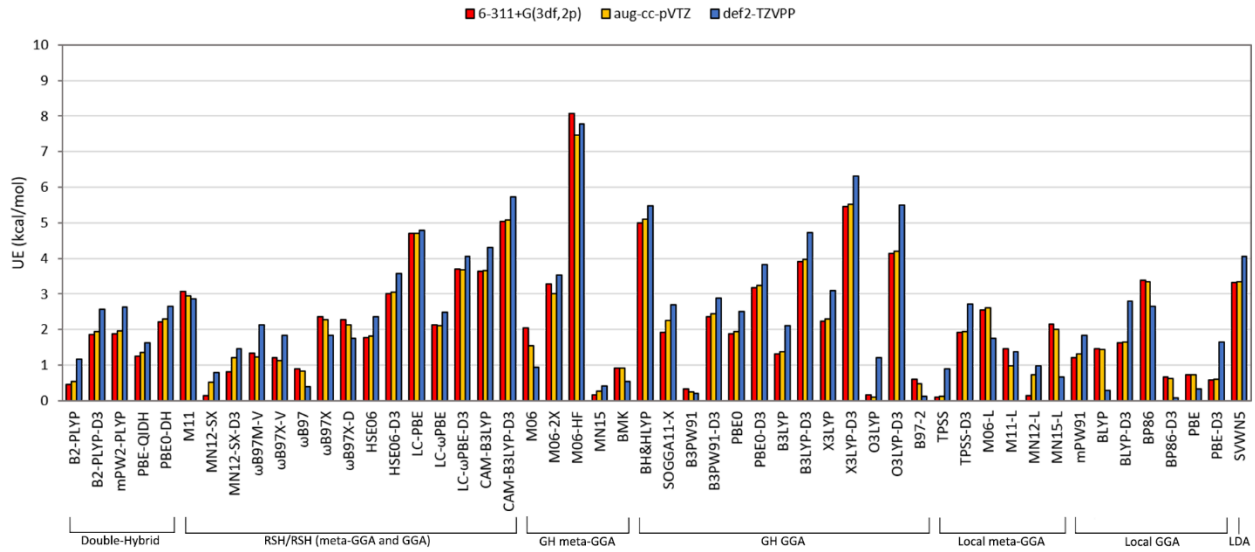


Figure A.4. UEs for Li^+–phosphate counterpoise-corrected DFT binding energies evaluated with the 6-311+G(3df,2p) (red), aug-cc-pVTZ (yellow) and def2-TZVPP (blue) basis sets relative to CCSD(T)/CBS reference values.

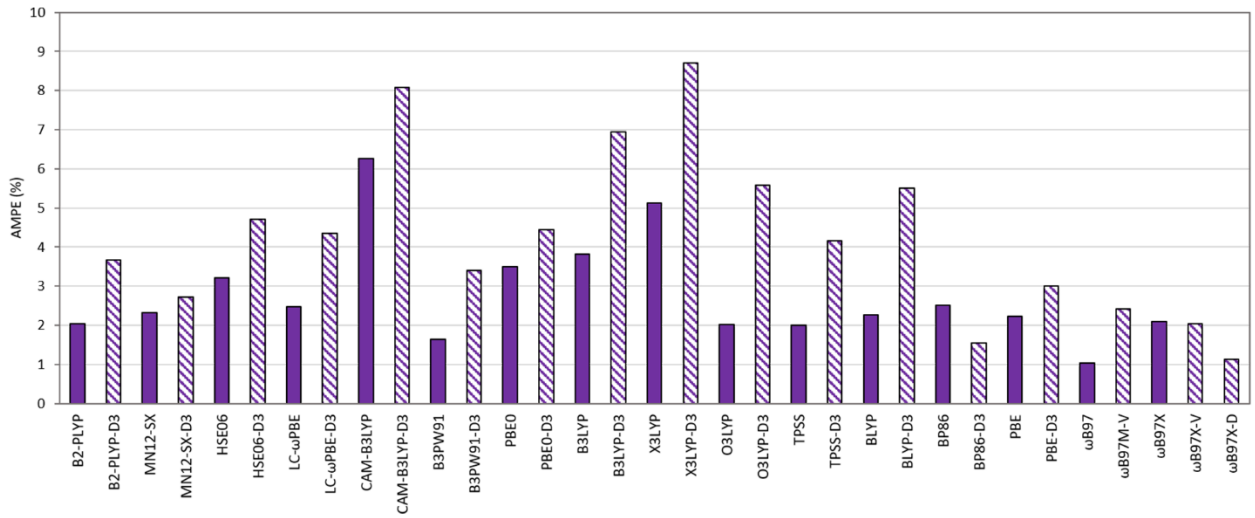


Figure A.5. AMPEs overall basis sets for Li^+–nucleobase counterpoise and dispersion-corrected DFT binding energies relative to CCSD(T)/CBS reference values.

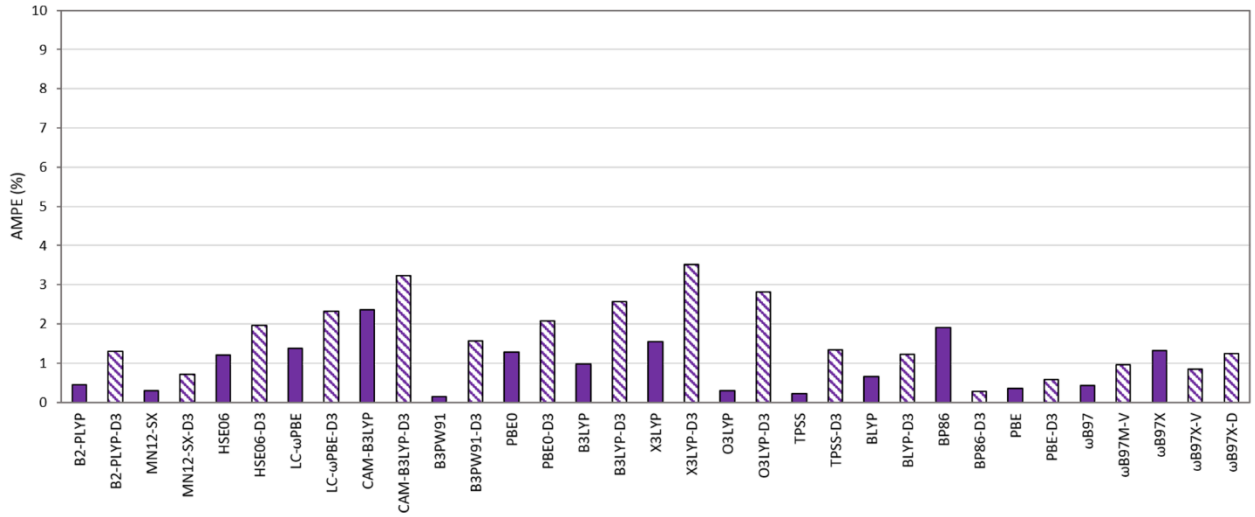


Figure A.6. AMPEs overall basis sets for Li^+–phosphate counterpoise and dispersion-corrected DFT binding energies relative to CCSD(T)/CBS reference values.

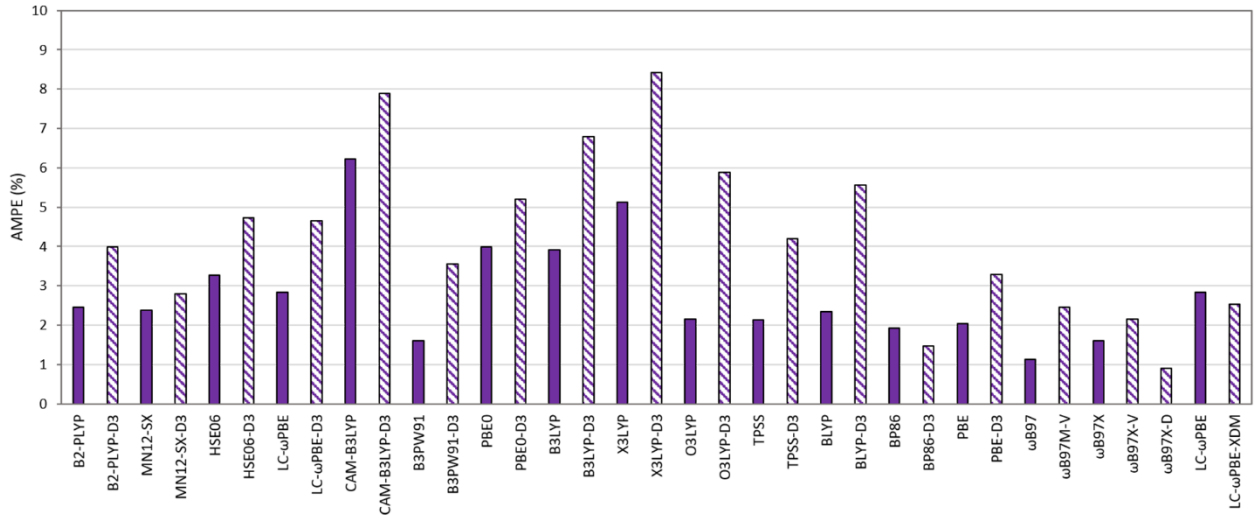


Figure A.7. AMPEs over all basis sets for Li^+–nucleobase and Li^+–phosphate DFT dispersion-corrected binding energies relative to CCSD(T)/CBS reference values without counterpoise corrections.

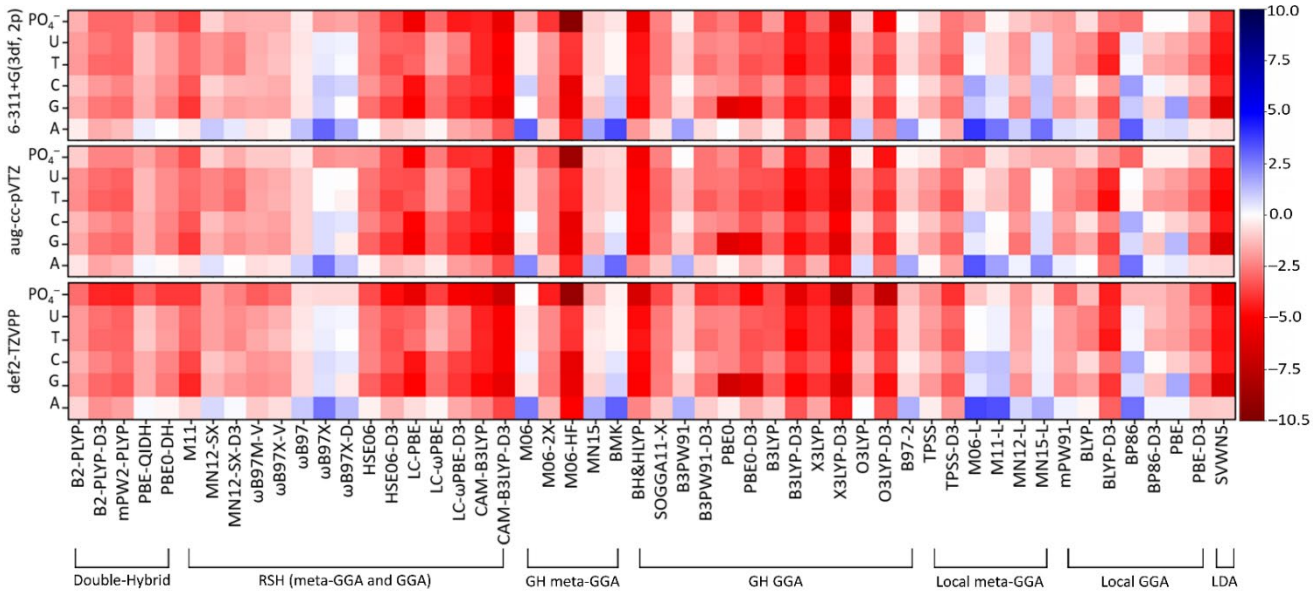


Figure A.8. Signed errors for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections, where negative values (red shades) represent over binding relative to CCSD(T)/CBS reference values.

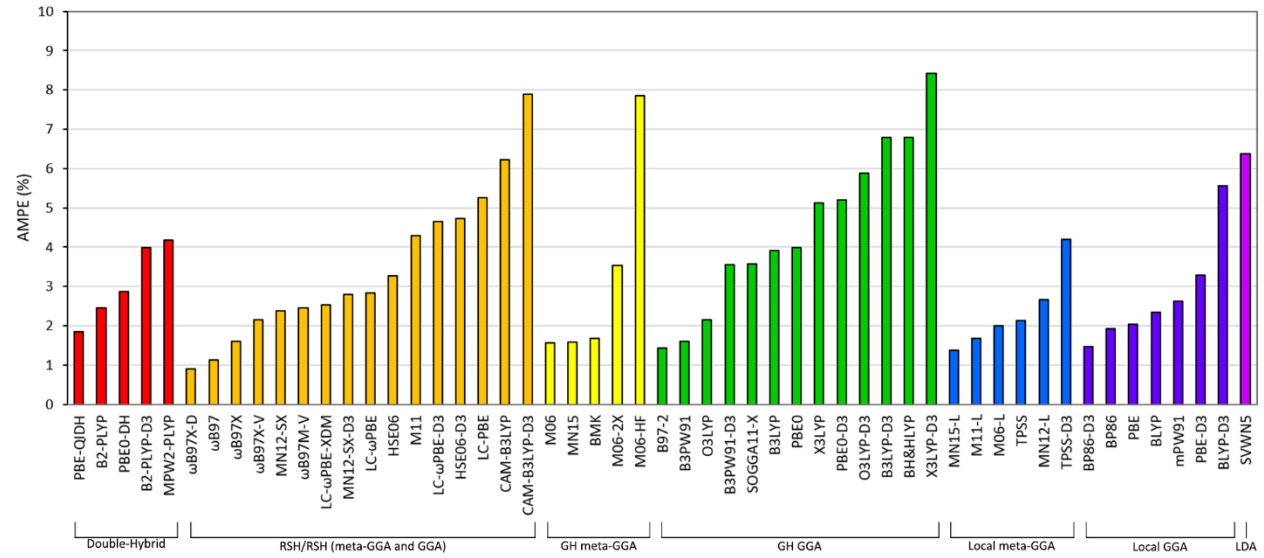


Figure A.9. AMPEs over all basis sets for Li^+–nucleobase and Li^+–phosphate DFT binding energies without counterpoise corrections relative to the CCSD(T)/CBS reference values arranged in ascending order of accuracy within families.

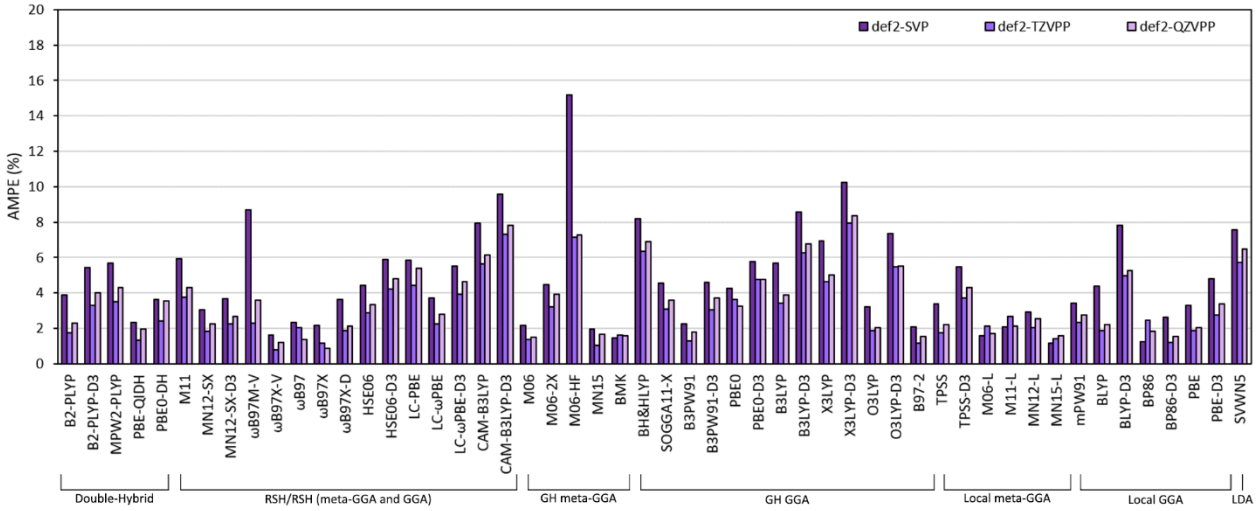


Figure A.10. AMPEs over Karlsruhe basis sets for Li^+–nucleobase and Li^+–phosphate counterpoise-corrected DFT binding energies relative to CCSD(T)/CBS reference values.

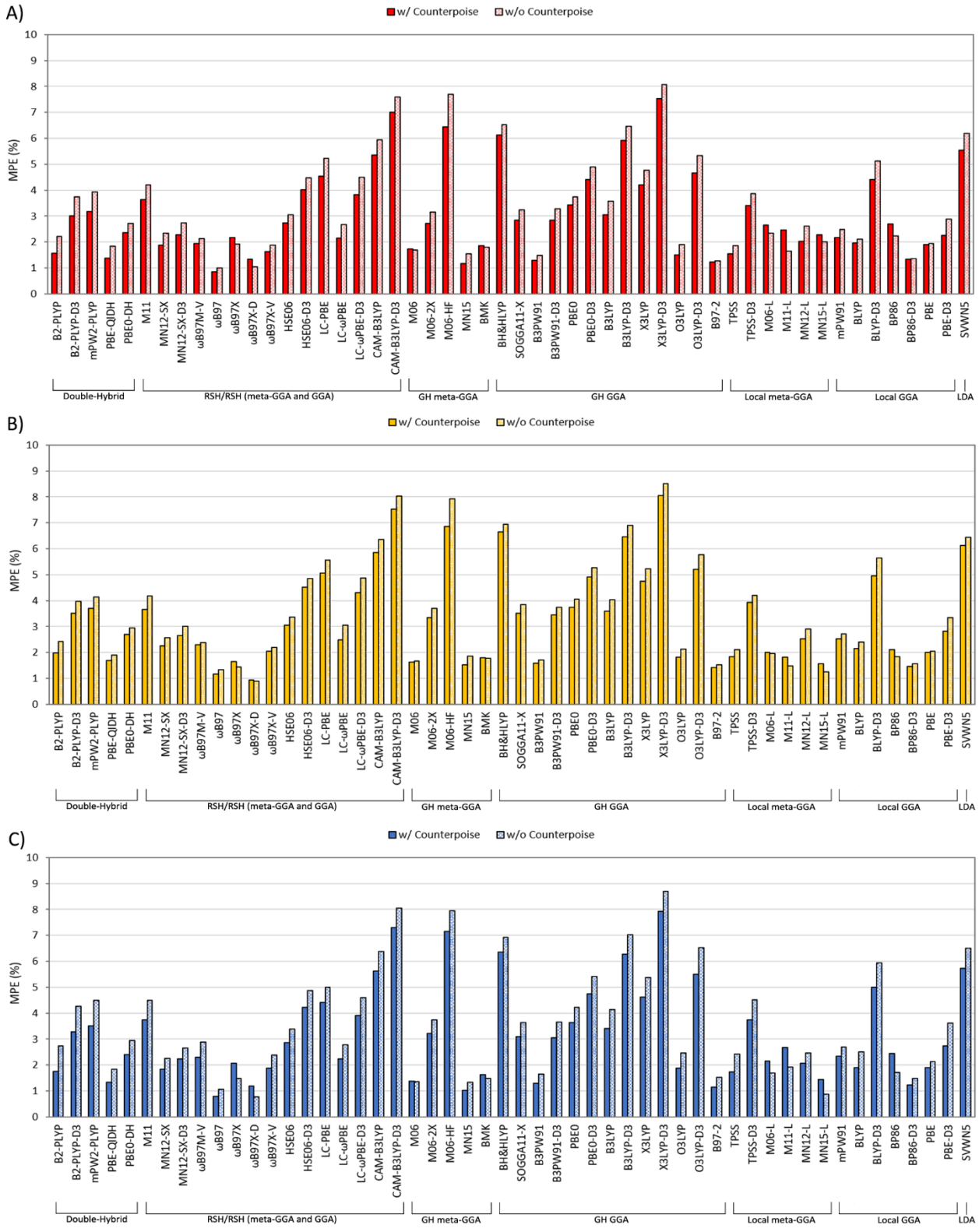


Figure A.11. MPEs of Li^+–nucleobase and Li^+–phosphate DFT binding energies with (dark) and without (light) counterpoise corrections relative to CCSD(T)/CBS reference values for A) 6-311+G(3df,2p), B) aug-cc-pVTZ, and C) def2-TZVPP.

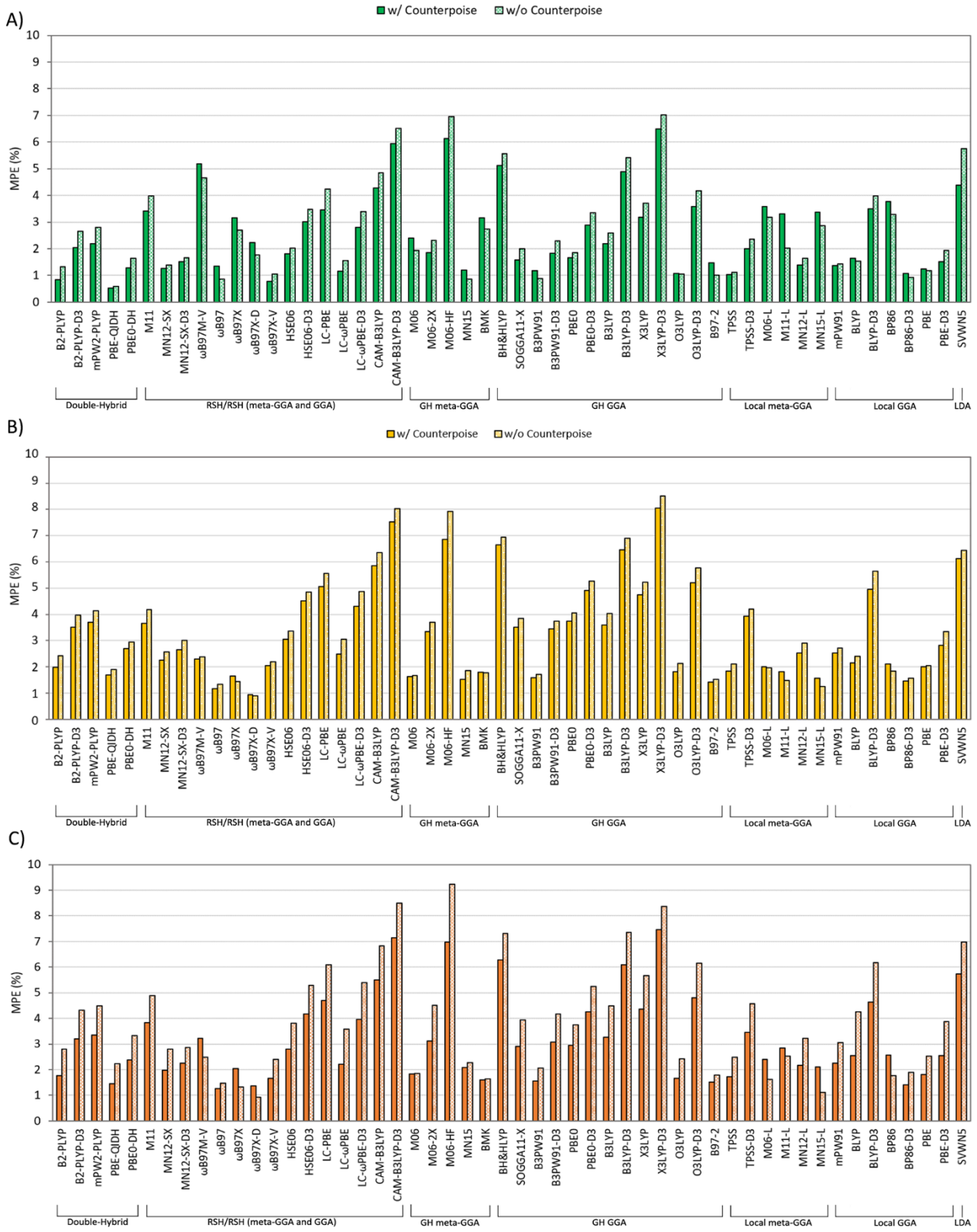


Figure A.12. MPEs of Li^+–nucleobase and Li^+–phosphate DFT binding energies with (dark) and without (light) counterpoise corrections relative to CCSD(T)/CBS reference values for A) aug-cc-pVDZ, B) aug-cc-pVTZ, and C) aug-cc-pVQZ.

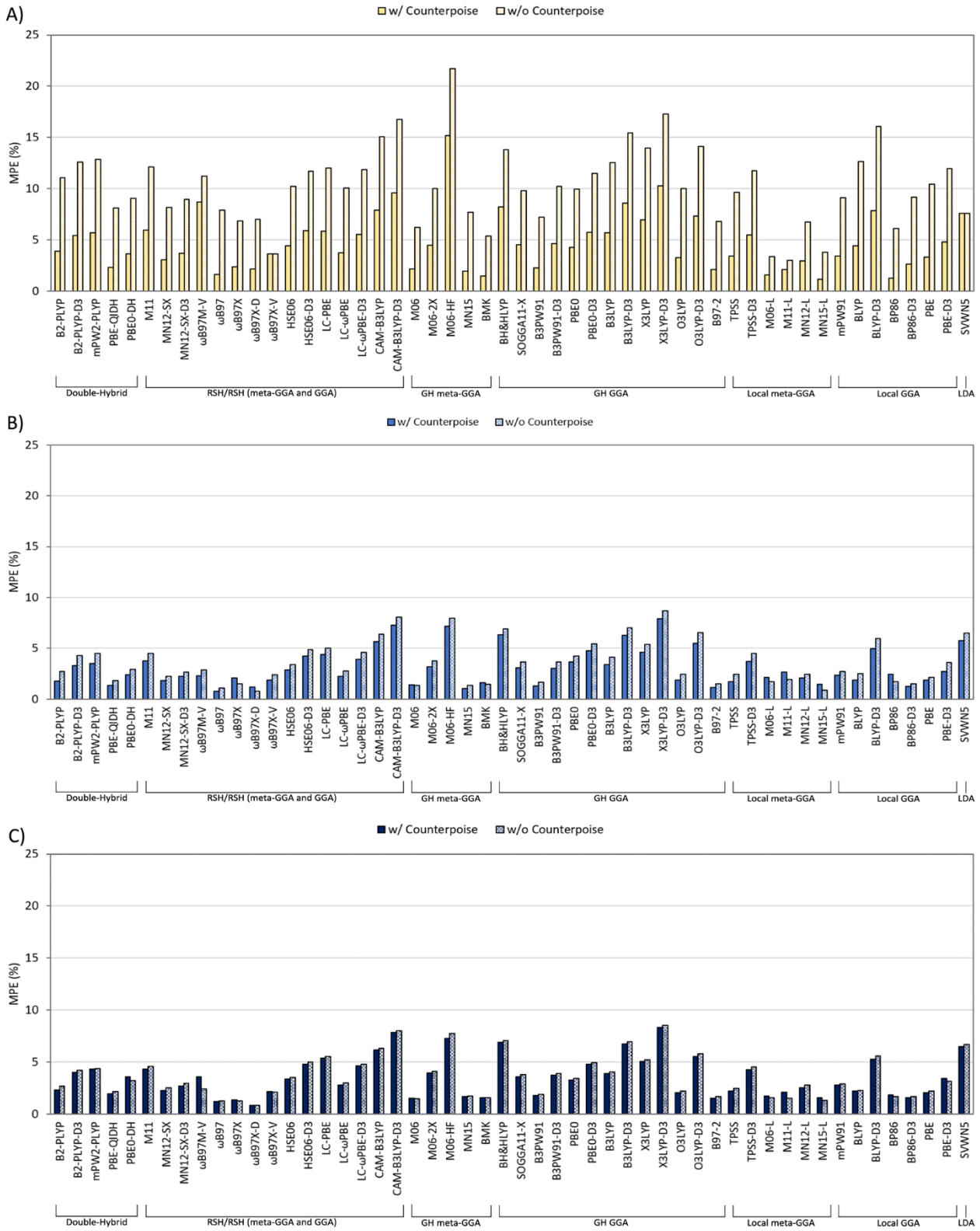


Figure A.13. MPEs of Li^+–nucleobase and Li^+–phosphate DFT binding energies with (dark) and without (light) counterpoise corrections relative to CCSD(T)/CBS reference values for A) def2-SVP, B) def2-TZVPP, and C) def2-QZVPP.

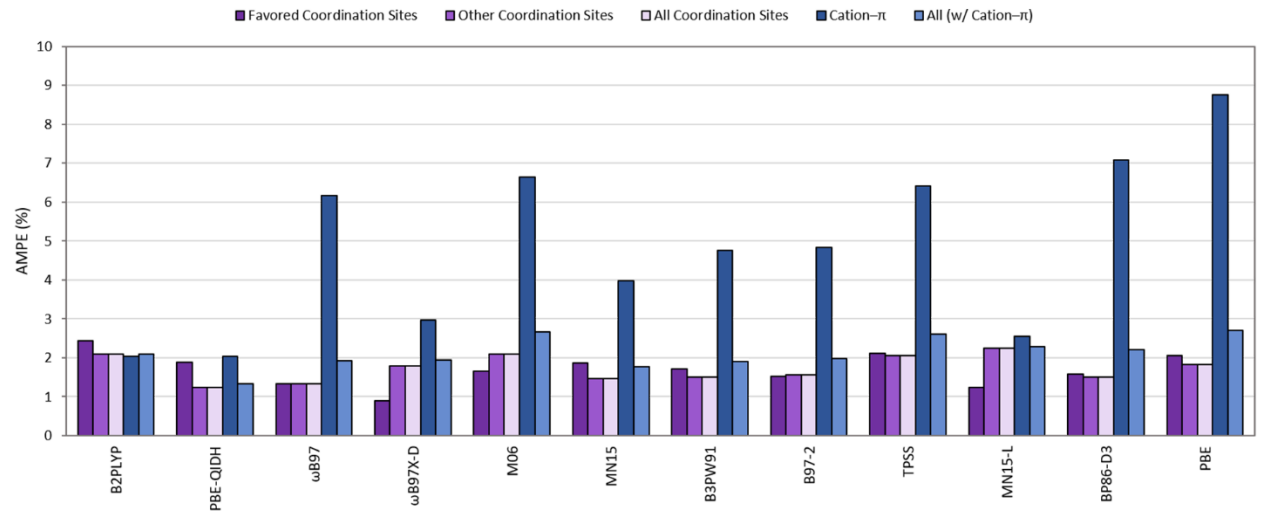


Figure A.14. AMPEs across the favored nucleic acid coordination sites (dark purple), other (nonpreferred) coordination sites (medium purple), all coordination sites (light purple), only cation–π interactions (dark blue) and all binding locations (including cation–π, medium blue), for Li^+–nucleobase and Li^+–phosphate DFT binding energies evaluated with the aug-cc-pVTZ basis set relative to CCSD(T)/CBS reference values without counterpoise corrections.

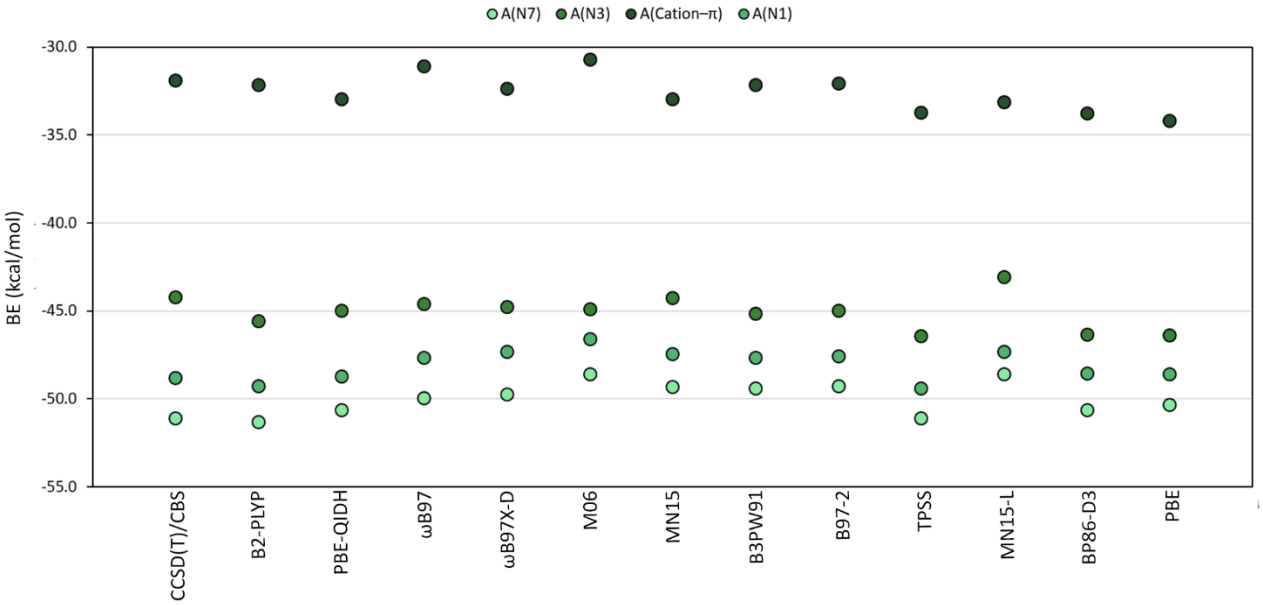


Figure A.15. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for all A binding sites (including cation–π interactions).

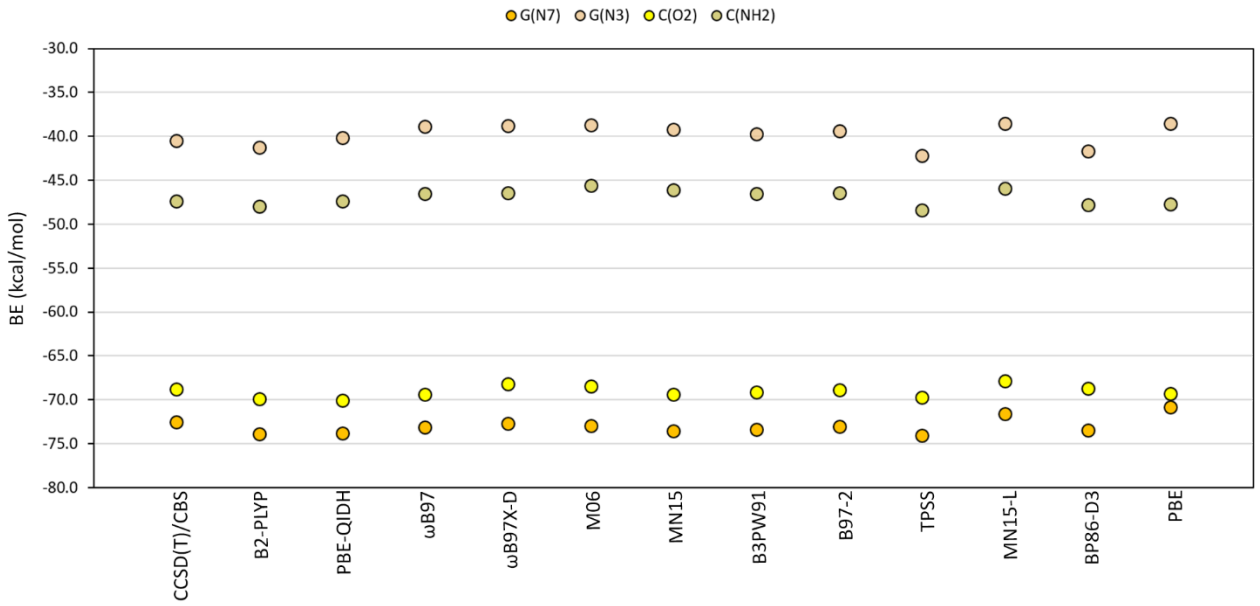


Figure A.16. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for all G and C coordination sites.

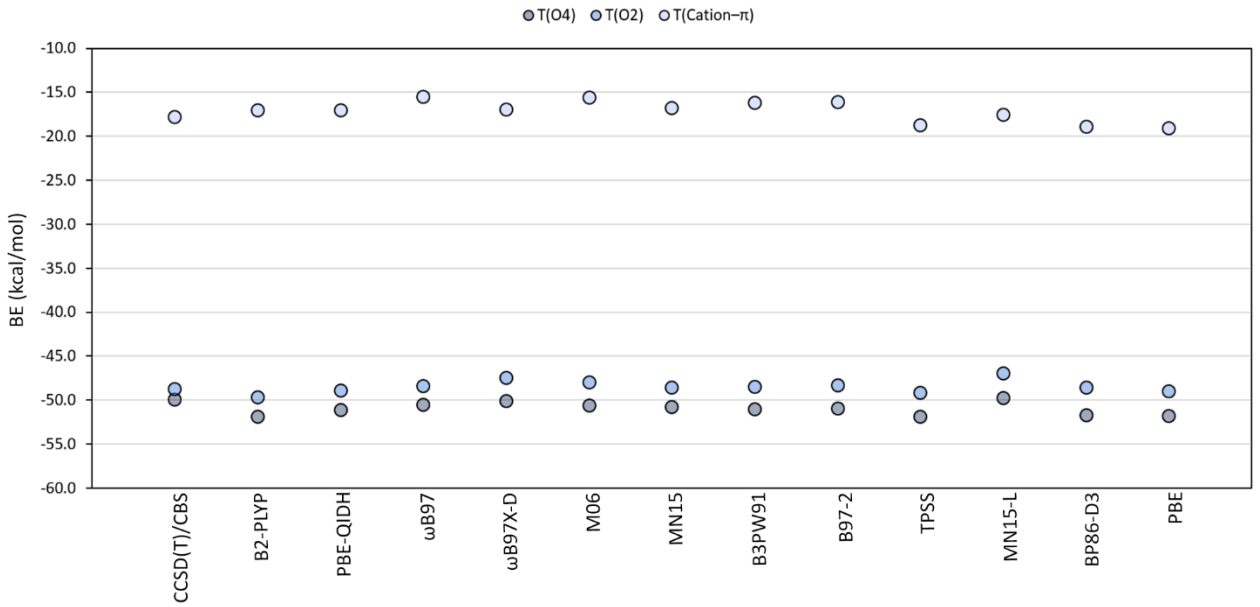


Figure A.17. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for all T binding sites (including cation–π interactions).

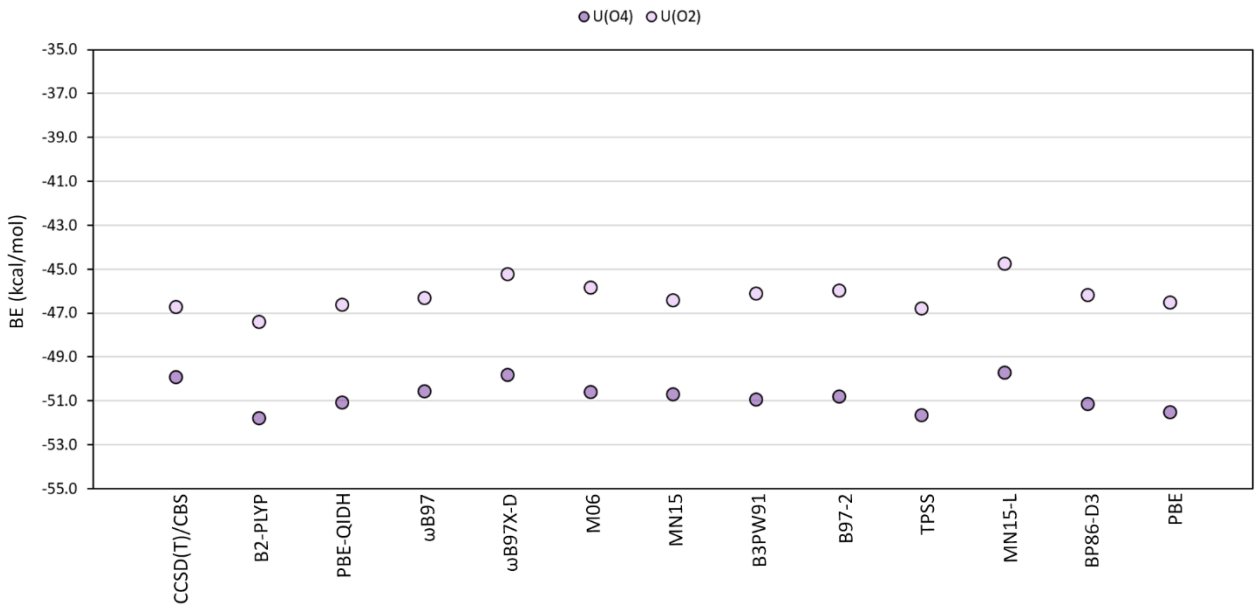


Figure A.18. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for all U coordination sites.

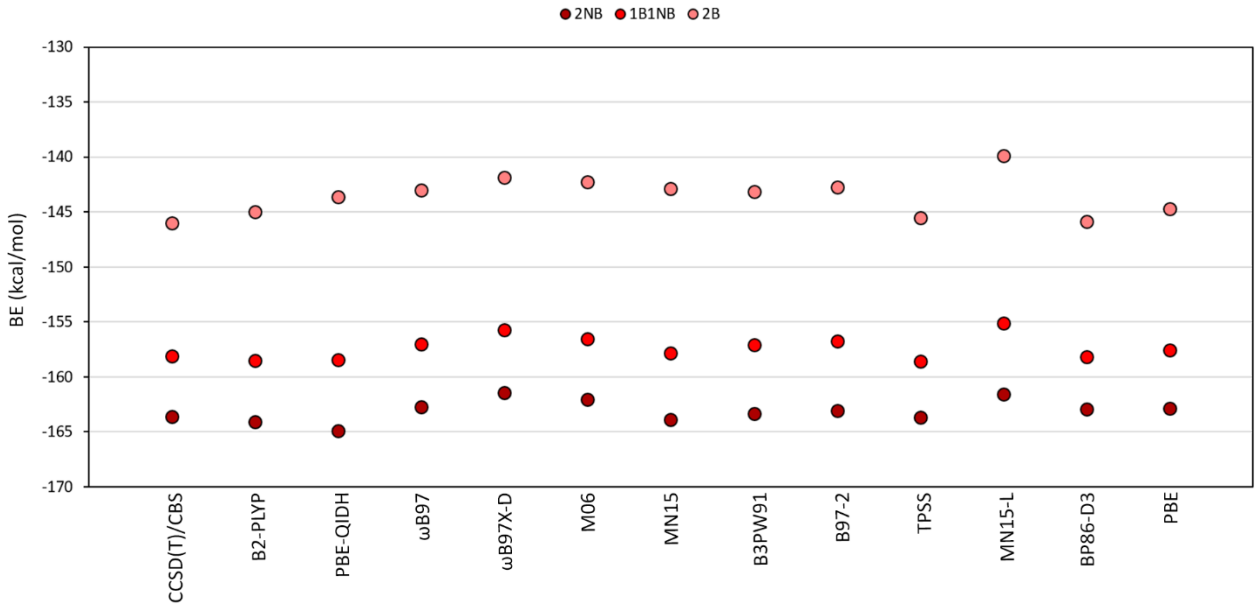


Figure A.19. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for all phosphate coordination sites (2NB denotes the interaction involving 2 non-bridging oxygen atoms, 1B1NB denotes 1 bridging and 1 non-bridging oxygen, and 2B denotes 2 bridging oxygens).

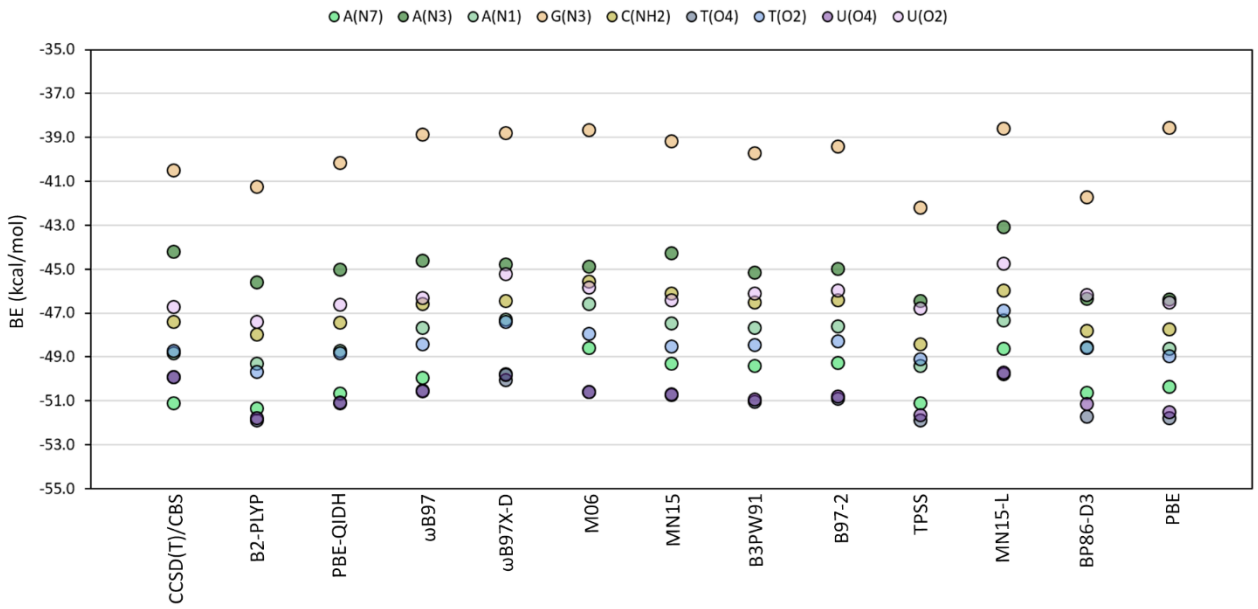


Figure A.20. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for A (N7, N3, N1), G (N3), C (NH2), T (O4, O2) and U (O4, O2) coordination sites.

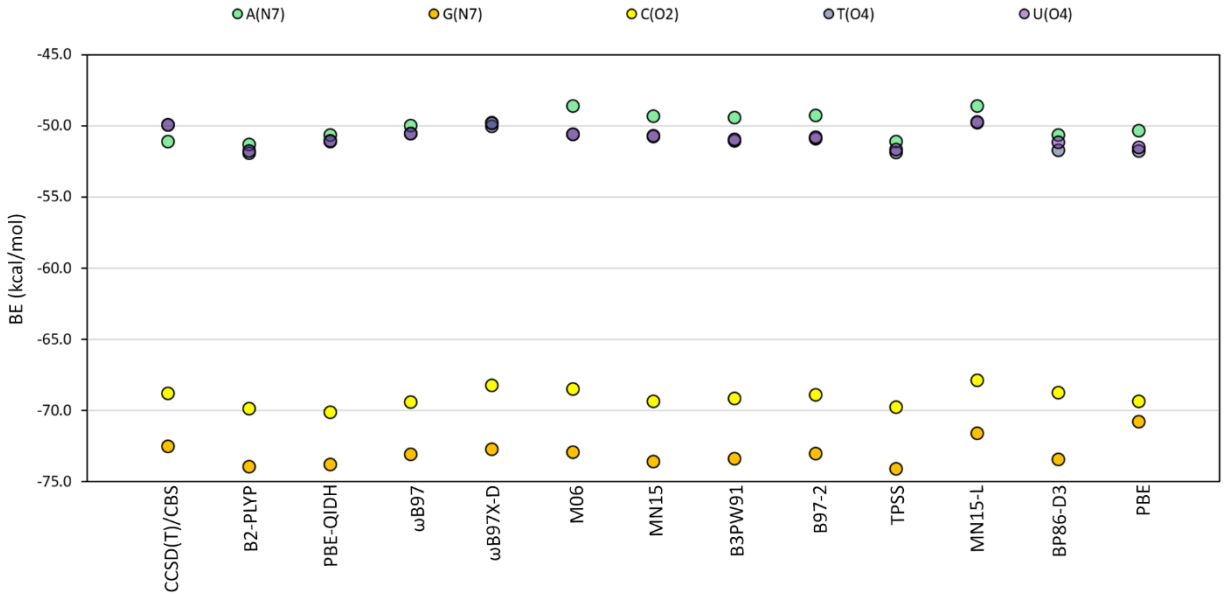


Figure A.21. Comparison of CCSD(T)/CBS and counterpoise-corrected DFT/aug-cc-pVTZ binding energies (BE) for preferred nucleobase coordination sites.

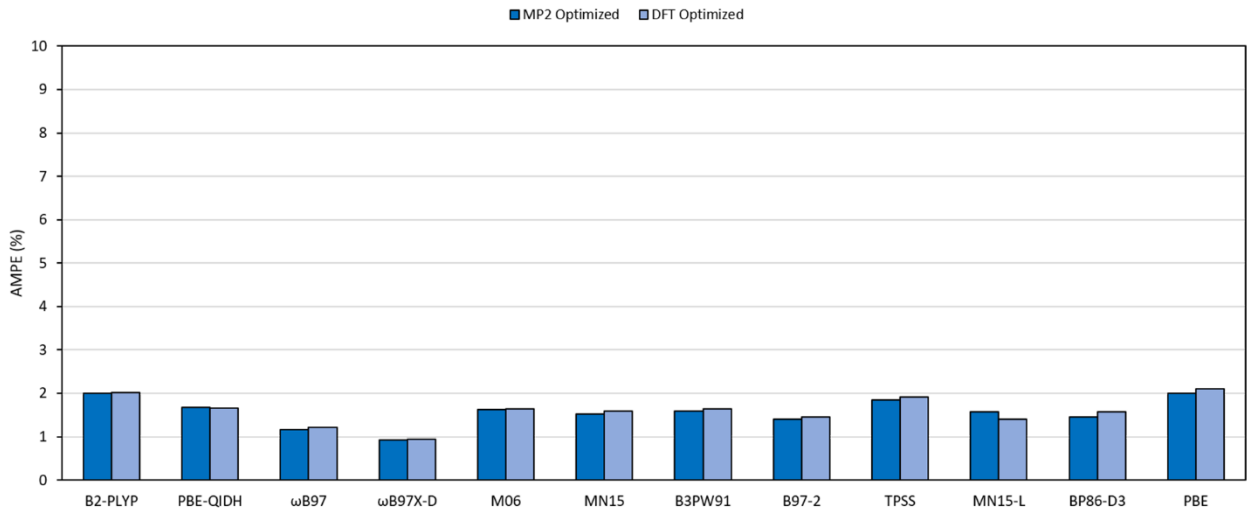


Figure A.22. Comparison of AMPE relative to the CCSD(T)/CBS reference values for the counterpoise-corrected DFT/aug-cc-pVTZ binding energies calculated with MP2/aug-cc-pVTZ and DFT/aug-cc-pVTZ optimized geometries of the preferred nucleobase and phosphate coordination complexes.

Appendix B

Supplementary Information for Chapter 3: Generation of an Accurate CCSD(T)/CBS Data Set and Assessment of DFT Methods for the Binding Strengths of Group I Metal–Nucleic Acid Complexes

Contains Tables B.1–B.9 and Figures B.1–B.6

Table B.1. Metal–nucleic acid coordination distances (Å) between each group I metal (Li^+, Na^+, K^+, Rb^+, or Cs^+) and nucleic acid component (A, C, T, G, U, or P).[a]

Binding Site		Li^+	Na^+	K^+	Rb^+	Cs^+
A(N6–N7)	X···N6	2.081	2.467	2.884	3.032	3.195
	X···N7	1.961	2.349	2.777	2.947	3.138
A(N1–N6)	X···N1	1.965	2.360	2.765	2.923	3.074
	X···N6	2.042	2.460	2.896	3.057	3.236
A(N3)	X···N3	1.887	2.283	2.696	2.862	3.013
G(O6–N7)	X···O6	1.927	2.278	2.635	2.782	2.923
	X···N7	2.030	2.387	2.790	2.951	3.121
G(N2–N3)	X···N2	2.050	2.437	2.921		
	X···N3	1.949	2.352	2.783		
C(O2–N3)	X···O2	1.859	2.209	2.560	2.707	2.836
	X···N3	2.058	2.471	2.911	3.082	3.268
C(N3–N4)	X···N3	1.938				
	X···N4	2.114				
T(O4)	X···O4	1.706	2.100	2.490	2.651	2.798
T(O2)	X···O2	1.709	2.103	2.497	2.659	2.813
U(O4)	X···O4	1.710	2.103	2.495	2.656	2.808
U(O2)	X···O2	1.713	2.109	2.506	2.669	2.826
P(2NB)	X···O(NB)	1.860	2.212	2.545	2.684	2.790
	X···O(NB)	1.862	2.212	2.545	2.684	2.790
P(1B1NB)	X···O(B)	1.875	2.253	2.625	2.772	2.926
	X···O(NB)	1.784	2.121	2.446	2.582	2.669
P(2B)	X···O(B)	1.800	2.400	2.802	2.915	3.103
	X···O(B)	1.800	2.606	2.801	2.946	3.103
	X···O(NB)[b]		2.191	2.510	2.648	2.726

[a]Empty spaces within the table indicate that specific metal–nucleic acid complex could not be isolated. See Figure 3.2 for definitions of complexes considered in the present work. [b]Coordination to a non-bridging oxygen in the dimethylphosphate model also occurs for all metals except Li^+ as shown in Figure B.1.

Table B.2. CCSD(T)/CBS binding strengths of group I metal–nucleic acid complexes in kcal/mol.[a]

Binding Site	Li^+	Na^+	K^+	Rb^+	Cs^+
A(N6–N7)	–51.1	–36.0	–19.4	–16.5	–14.2
A(N1–N6)	–48.9	–34.7	–19.6	–17.1	–15.1
A(N3)	–44.4	–31.5	–18.0	–15.7	–13.8
G(O6–N7)	–73.0	–58.0	–39.9	–36.4	–33.4
G(N2–N3)	–40.6	–26.1	–11.3		
C(O2–N3)	–69.3	–53.8	–36.6	–33.2	–30.5
C(N3–N4)	–47.5				
T(O4)	–50.4	–36.3	–23.8	–21.4	–19.5
T(O2)	–48.8	–34.9	–22.7	–20.4	–18.6
U(O4)	–50.3	–36.4	–24.1	–21.7	–19.8
U(O2)	–46.7	–33.2	–21.3	–19.1	–17.3
P(2NB)	–163.9	–140.3	–115.3	–110.0	–105.5
P(1B1NB)	–157.8	–135.4	–111.8	–106.8	–102.7
P(2B)	–143.3	–133.3	–110.6	–105.8	–102.0

[a]Empty spaces within the table indicate that specific metal–nucleic acid complex could not be isolated. See Figure 3.2 for definitions of complexes considered in the present work.

Table B.3. MPEs (%) and MUEs (kcal/mol) in DFT binding strengths along with corresponding standard (SD), minimum (Min), and maximum (Max) deviations evaluated across Li^+–nucleic acid complexes for each functional relative to CCSD(T)/CBS reference values, as well as outliers based on the boxplot analysis.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Double-Hybrid	B2-PLYP	1.0	0.6	0.2	2.2		0.7	0.5	0.1	1.9	
	B2-PLYP-D3	2.7	0.9	1.3	4.2		1.8	0.7	0.9	3.7	P(1B1NB,2B)
	mPW2-PLYP	2.8	0.9	1.3	4.3	P(2NB), T(O4), U(O4), A(N3)	1.8	0.5	1.2	2.7	
	PBE-QIDH	0.7	0.4	0.1	1.4		0.4	0.2	0.1	0.8	
	PBE0-DH	1.4	1.1	0.1	3.5		0.9	0.7	0.0	1.9	
	DSD-PBEP86	0.7	0.4	0.1	1.5	U(O2)	0.5	0.4	0.1	1.5	
RSH	M11	2.3	1.5	0.3	4.6		1.6	1.0	0.2	3.3	
	MN12-SX	1.1	0.8	0.1	2.3		0.7	0.5	0.0	2.1	P(2B)
	MN12-SX-D3	1.1	1.0	0.1	3.0		0.7	0.5	0.1	1.5	
RSH meta-GGA	ωB97M-V	1.3	0.6	0.4	2.3		0.9	0.4	0.2	1.6	
RSH GGA	ωB97	1.5	1.5	0.0	5.2		0.8	0.6	0.0	2.1	
	ωB97X	3.4	2.1	1.5	8.3		2.0	0.8	0.9	3.4	
	ωB97X-D	2.4	1.4	0.5	5.0		1.5	0.6	0.3	2.4	
	ωB97X-D3	4.9	1.4	3.0	7.0		3.3	1.5	1.5	5.8	P(2B,2NB)
	ωB97X-V	0.9	0.6	0.1	2.1		0.6	0.4	0.0	1.3	
	ωB97X-D4	2.2	0.8	1.3	4.2		1.5	0.7	0.5	2.8	
	HSE06	1.8	1.3	0.0	4.6		1.2	0.7	0.0	2.3	
	HSE06-D3	3.4	1.4	1.8	6.7		2.2	0.8	1.1	3.2	
	LC-PBE	3.6	2.1	0.9	6.9		2.3	1.3	0.6	4.8	
	LC-ωPBE	1.2	1.2	0.2	3.7		0.9	0.8	0.1	2.2	
	LC-ωPBE-D3	3.2	1.2	1.9	6.2	A(N3)	2.1	0.9	1.0	3.3	
	CAM-B3LYP	4.7	1.7	2.2	7.1		3.0	0.8	1.7	4.1	
	CAM-B3LYP-D3	6.6	2.0	3.1	9.4	P(1B1NB,2B,2NB), A(N3), T/U(O4)	4.1	0.9	2.8	5.5	
GH meta-GGA	M06	2.4	2.0	0.4	5.6		1.4	0.9	0.2	2.8	
	M06-2X	2.3	0.7	1.2	3.7	P(2B), A(N3)	1.5	0.5	0.9	2.8	
	M06-HF	6.7	2.1	2.7	9.5	P(1B1NB,2B)	4.2	1.1	3.1	7.0	
	MN15	1.6	1.4	0.1	4.1		0.8	0.6	0.1	1.9	
	BMK	3.0	2.3	0.7	7.5		1.7	0.9	0.5	3.3	

	PW6B95	2.4	1.1	0.9	4.4		1.6	0.7	0.4	2.6	
GH GGA	BH&HLYP	5.3	2.2	1.9	8.4		3.3	1.0	1.2	4.7	
	SOGGA11-X	1.9	1.0	0.1	3.6		1.2	0.6	0.0	2.0	
	B3PW91	1.4	1.3	0.2	3.7		0.7	0.6	0.1	1.9	
	B3PW91-D3	2.3	1.3	0.9	5.5	A(N3)	1.6	0.9	0.4	3.3	
	PBE0	1.8	1.3	0.1	4.5		1.2	0.7	0.0	2.2	
	PBE0-D3	3.4	1.4	2.0	6.5		2.2	0.8	1.2	3.2	
	B3LYP	2.4	1.4	0.7	4.9		1.6	0.8	0.4	2.7	
	B3LYP-D3	5.6	1.7	2.6	8.7		3.6	1.1	2.6	6.1	P(2B)
	B3LYP-D4	3.5	1.6	1.6	6.5		2.2	0.9	1.0	3.6	
	X3LYP	3.7	1.5	1.6	6.3		2.3	0.8	1.2	3.5	
	X3LYP-D3	7.3	2.0	3.6	10.6		4.7	1.2	3.4	7.5	P(2B)
	O3LYP	1.0	0.9	0.1	3.3	A(N3)	0.6	0.4	0.1	1.5	
	O3LYP-D3	5.1	1.8	3.0	8.7		3.4	1.4	1.4	6.4	P(2B)
	B97-2	1.4	1.3	0.0	3.7		0.8	0.5	0.0	1.9	
	TPSSh	1.4	1.1	0.1	3.9		1.0	0.7	0.0	2.4	
Local meta-GGA	revTPSS	1.1	0.8	0.0	2.7		0.8	0.6	0.0	2.4	P(2B)
	TPSS	1.3	1.0	0.0	3.6	A(N3)	0.9	0.8	0.0	2.9	P(2B)
	TPSS-D3	3.2	1.7	1.2	6.5		2.2	1.3	0.6	5.3	P(2B)
	TPSS-D4	1.8	1.5	0.0	5.1		1.3	1.1	0.0	3.7	
	M06-L	3.3	2.4	0.8	7.6		1.9	1.0	0.4	3.9	
	M11-L	3.5	1.9	1.0	6.4		2.3	1.4	0.4	5.5	
	MN12-L	1.3	0.7	0.1	2.1		0.9	0.6	0.2	2.5	P(2B)
	MN15-L	3.3	1.6	0.7	5.6		1.9	0.6	1.2	3.1	
Local GGA	mPW91	1.1	1.1	0.0	3.5		0.7	0.6	0.0	1.6	
	BLYP	1.2	1.0	0.0	3.2		0.8	0.9	0.0	3.3	
	BLYP-D3	4.4	1.9	1.8	7.8		2.9	1.6	1.2	7.3	P(2B)
	BP86	3.6	2.2	0.2	7.2		2.1	1.0	0.3	3.7	
	BP86-D3	1.4	1.0	0.1	3.0		1.0	0.8	0.1	3.2	P(2B)
	PBE	1.3	0.9	0.0	3.6	A(N3)	0.8	0.6	0.0	2.4	P(2B)
	PBE-D3	2.0	1.7	0.0	5.6		1.4	1.2	0.0	4.2	P(2B)
	PBE-D4	1.6	1.2	0.2	4.7		1.0	0.8	0.1	3.0	P(2B)

Local LDA	SVWN5	5.0	2.7	1.4	10.1		3.2	1.5	0.7	5.6	

[a]Complexes resulting in an outlier in the boxplot are listed according to the nucleic acid component (A, C, T, G, U, or P) and binding site (in parentheses). See Figure 3.2 for definitions of the complexes considered in present work and Figure 3.4 for the outliers.

Table B.4. MPEs (%) and MUEs (kcal/mol) in DFT binding strengths along with corresponding standard (SD), minimum (Min), and maximum (Max) deviations evaluated across Na^+–nucleic acid complexes for each functional relative to CCSD(T)/CBS reference values, as well as outliers based on the boxplot analysis.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outlier[a]
Double-Hybrid	B2-PLYP	3.5	2.1	1.5	7.9	A(N3), G(N3)	1.6	0.6	0.6	2.5	
	B2-PLYP-D3	1.1	0.9	0.0	2.9		0.4	0.3	0.0	0.9	
	mPW2-PLYP	1.6	1.6	0.5	5.3	A(N3,N7), G(N3)	0.7	0.5	0.2	1.5	
	PBE-QIDH	4.9	3.2	1.4	12.7	A(N3,N7)	2.1	0.7	1.1	3.3	
	PBE0-DH	3.4	3.2	0.6	11.2	A(N7), G(N3)	1.4	0.8	0.3	2.9	G(N3)
	DSD-PBEP86	3.6	1.7	0.9	6.5	P(1B1NB,2B,2NB)	1.6	0.4	1.1	2.2	
RSH	M11	2.2	3.0	0.1	9.4	A(N1,N7), G(N3)	0.8	0.9	0.0	2.5	A(N1,N7), G(N3)
	MN12-SX	4.3	3.4	1.3	12.5	A(N7), G(N3)	1.9	0.9	0.5	3.3	
	MN12-SX-D3	2.7	2.7	0.2	9.4	A(N1,N7), G(N3)	1.1	0.8	0.1	2.5	
RSH meta-GGA	ωB97M-V	1.2	1.5	0.0	5.2	G(N3)	0.4	0.4	0.0	1.3	A(N7), G(N3)
RSH GGA	ωB97	3.5	3.1	1.2	11.6	A(N1,N7), G(N3)	1.5	0.7	0.5	3.0	G(N3)
	ωB97X	5.5	4.1	2.0	15.7	A(N1,N7), G(N3)	2.4	1.0	1.1	4.1	
	ωB97X-D	5.3	3.2	2.4	13.2	G(N3)	2.5	0.8	1.1	3.4	
	ωB97X-D3	3.6	1.1	2.0	5.7	U(O2), A(N3,N7)	2.1	1.5	0.7	5.2	P(1B1NB,2B,2NB)
	ωB97X-V	1.8	2.0	0.3	7.3	G(N3)	0.7	0.5	0.1	1.9	A(N1,N7), G(N3)
	ωB97X-D4	1.5	0.7	0.8	2.8		0.9	0.6	0.3	1.9	P(2B)
	HSE06	2.3	2.7	0.1	8.2	A(N7), G(N3)	1.0	0.8	0.0	2.4	
	HSE06-D3	1.6	1.3	0.1	3.4		0.7	0.3	0.0	1.1	
	LC-PBE	2.7	3.0	0.3	10.8	G(N3)	1.1	0.8	0.1	2.8	
	LC-ωPBE	3.8	3.5	0.8	12.5	A(N7), G(N3)	1.5	0.9	0.7	3.3	A(N7), G(N3)
	LC-ωPBE-D3	1.6	1.8	0.3	6.7	A(N1,N7), G(N3)	0.7	0.4	0.3	1.7	G(N3)
	CAM-B3LYP	2.4	1.3	0.4	4.4	U(O4)	1.0	0.3	0.5	1.6	
	CAM-B3LYP-D3	4.1	2.1	1.3	7.2		2.0	0.8	0.3	2.9	G(N3)
GH meta-GGA	M06	7.9	5.0	2.9	18.9	A(N7), G(N3)	3.5	1.1	2.0	5.4	

	M06-2X	1.1	1.4	0.0	4.8	A(N1,N7),G(N3)	0.4	0.4	0.1	1.2	
	M06-HF	2.6	0.8	1.4	3.8		1.4	0.8	0.5	3.7	P(2B)
	MN15	5.2	3.7	1.5	13.7	G(N3)	2.2	0.8	1.2	3.8	A(N7), G(N3)
	BMK	6.7	4.9	1.9	17.8	G(N3)	2.9	1.1	1.3	4.8	
	PW6B95	2.0	2.5	0.1	7.3	A(N1,N7),G(N3)	0.7	0.7	0.1	2.2	A(N1,N7), G(N3)
GH GGA	BH&HLYP	2.8	2.2	0.1	8.5	G(N3)	1.2	0.6	0.1	2.2	
	SOGGA11-X	3.1	3.0	0.4	11.2	G(N3)	1.3	0.8	0.1	2.9	
	B3PW91	6.3	4.0	2.2	14.9	A(N7), G(N3)	2.8	1.0	1.4	4.4	
	B3PW91-D3	1.8	1.7	0.1	5.4	G(N3)	0.7	0.5	0.1	1.4	
	PBE0	2.4	2.7	0.0	8.4	A(N7), G(N3)	1.0	0.8	0.0	2.4	
	PBE0-D3	1.5	1.2	0.1	3.8		0.6	0.3	0.0	1.0	
	B3LYP	2.1	2.3	0.1	7.0	A(N1,N7), G(N3)	0.9	0.7	0.0	2.2	
	B3LYP-D3	3.0	1.8	1.3	6.3	A(N3)	1.5	0.6	0.5	2.3	P(2B)
	B3LYP-D4	1.9	1.5	0.3	4.4	A(N3), T(O4), U(O4)	0.8	0.4	0.4	1.6	T(O4)
	X3LYP	1.7	1.5	0.1	4.4		0.6	0.4	0.1	1.4	
	X3LYP-D3	6.0	2.1	2.9	9.9		2.9	0.9	1.9	4.6	
	O3LYP	5.6	3.2	2.3	11.6		2.6	1.0	1.0	3.9	
	O3LYP-D3	2.6	1.9	0.1	7.3	A(N3)	1.5	1.1	0.1	3.9	
	B97-2	6.5	4.0	2.4	15.4	A(N7), G(N3)	2.9	1.0	1.5	4.4	
	TPSSh	3.5	2.5	1.1	7.9		1.6	0.7	0.5	2.8	A(N3,N7), U(O4)
Local meta-GGA	revTPSS	4.0	2.2	1.1	7.8		1.8	0.6	0.6	2.8	
	TPSS	3.7	2.3	0.9	8.0		1.7	0.7	0.3	2.9	A(N3)
	TPSS-D3	1.4	1.0	0.0	3.6	A(N3)	0.6	0.3	0.0	1.1	
	TPSS-D4	1.8	1.4	0.0	4.7		0.8	0.5	0.0	1.7	
	M06-L	8.0	5.0	2.6	18.0	A(N7), G(N3)	3.5	1.1	1.9	5.6	
	M11-L	7.4	4.3	2.0	16.2		3.3	0.9	1.8	4.7	
	MN12-L	5.9	4.4	1.7	16.5	A(N7), G(N3)	2.5	1.0	1.0	4.3	
	MN15-L	5.7	3.1	1.8	12.5		2.5	0.6	1.5	3.3	A(N7), T(O4), U(O4)
Local GGA	mPW91	4.0	3.4	1.2	11.6	A(N1), G(N3)	1.8	0.9	0.6	3.3	
	BLYP	3.2	2.1	0.2	7.4		1.6	0.9	0.1	3.1	
	BLYP-D3	2.5	2.2	0.3	7.6	A(N3)	1.1	0.7	0.1	2.4	
	BP86	7.4	3.5	3.0	13.3		3.5	1.1	1.6	4.9	

	BP86-D3	2.6	1.7	0.1	4.7		1.1	0.6	0.1	1.9
	PBE	3.0	1.9	1.0	6.9		1.4	0.7	0.4	2.5
	PBE-D3	1.5	1.3	0.2	4.4		0.6	0.4	0.1	1.4
	PBE-D4	1.7	1.4	0.2	4.4		0.8	0.5	0.2	1.6
Local LDA	SVWN5	4.4	3.0	0.4	10.5		2.2	1.2	0.1	4.0

[a]Complexes resulting in an outlier in the boxplot are listed according to the nucleic acid component (A, C, T, G, U, or P) and binding site (in parentheses). See Figure 3.2 for definitions of the complexes considered in present work and Figure 3.4 for the outliers.

Table B.5. MPEs (%) and MUEs (kcal/mol) in DFT binding strengths along with corresponding standard (SD), minimum (Min), and maximum (Max) deviations evaluated across K^+–nucleic acid complexes for each functional relative to CCSD(T)/CBS reference values, as well as outliers based on the boxplot analysis.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Double-Hybrid	B2-PLYP	1.1	0.9	0.2	2.9		0.3	0.2	0.1	0.6	P(2NB)
	B2-PLYP-D3	4.4	3.1	1.2	11.7		1.2	0.5	0.5	2.3	P(2B)
	mPW2-PLYP	1.2	1.3	0.2	4.6	G(N3)	0.3	0.1	0.1	0.5	
	PBE-QIDH	1.1	0.8	0.1	3.1	G(N3)	0.3	0.2	0.1	0.7	G(N7)
	PBE0-DH	1.6	1.2	0.1	4.6	G(N3)	0.4	0.2	0.1	1.0	G(N7)
	DSD-PBEP86	3.0	2.8	0.5	9.7		0.9	0.6	0.1	2.1	P(2B)
RSH	M11	2.8	2.2	0.0	5.6		1.0	0.6	0.0	2.1	G(N7)
	MN12-SX	2.8	1.9	0.6	6.5		0.9	0.6	0.1	2.1	
	MN12-SX-D3	6.6	3.1	2.1	11.5		2.0	0.7	1.1	3.5	P(2NB)
RSH meta-GGA	ωB97M-V	1.0	1.2	0.2	4.5	G(N3)	0.3	0.2	0.0	0.5	
RSH GGA	ωB97	1.2	1.0	0.0	4.1	G(N3)	0.3	0.2	0.0	0.6	
	ωB97X	1.4	1.2	0.0	4.4	G(N3)	0.3	0.2	0.0	0.8	G(N7)
	ωB97X-D	5.7	3.8	1.1	13.3	G(N3)	1.5	0.5	0.6	2.6	
	ωB97X-D3	12.2	7.2	5.3	30.0	G(N3)	3.8	1.9	1.8	7.8	P(2B,2NB)
	ωB97X-V	1.1	1.0	0.1	4.0	G(N3)	0.3	0.1	0.1	0.5	
	ωB97X-D4	9.6	5.2	3.4	22.2	G(N3)	2.9	1.0	1.6	4.8	
	HSE06	1.9	1.5	0.1	5.1		0.4	0.2	0.1	0.8	
	HSE06-D3	4.9	2.9	1.4	10.4		1.4	0.5	0.6	2.2	
	LC-PBE	2.7	2.4	0.2	7.9		0.7	0.5	0.0	1.9	G(N7)

	LC-ωPBE	1.7	1.5	0.0	4.7		0.4	0.3	0.1	1.1	
	LC-ωPBE-D3	5.8	3.1	2.0	12.2	A(N3), G(N3)	1.8	0.6	0.9	2.8	
	CAM-B3LYP	2.1	2.2	0.1	7.6	G(N3)	0.5	0.3	0.0	0.9	
	CAM-B3LYP-D3	5.4	2.5	1.6	9.7	A(N3), P(2NB)	1.6	0.5	1.0	2.4	
GH meta-GGA	M06	2.0	1.5	0.1	5.7	A(N3)	0.5	0.3	0.2	1.0	
	M06-2X	0.9	0.7	0.0	2.2		0.3	0.3	0.0	0.8	
	M06-HF	2.0	1.3	0.3	4.0		0.7	0.6	0.1	2.0	
	MN15	1.7	1.8	0.1	6.3	A(N7), G(N3)	0.4	0.2	0.1	0.8	
	BMK	1.7	1.5	0.2	4.3		0.5	0.4	0.0	1.3	
	PW6B95	1.8	2.0	0.1	6.7	G(N3)	0.4	0.3	0.1	0.9	
GH GGA	BH&HLYP	3.0	2.4	0.2	9.3	G(N3)	0.8	0.4	0.2	1.6	
	SOGGA11-X	1.4	1.3	0.1	4.0		0.4	0.3	0.0	0.9	
	B3PW91	1.7	1.3	0.0	4.4		0.4	0.3	0.0	0.9	
	B3PW91-D3	9.5	5.9	3.0	23.6	G(N3)	2.7	0.8	1.4	4.5	
	PBE0	1.8	1.4	0.1	4.8		0.4	0.2	0.1	0.8	
	PBE0-D3	4.3	2.5	1.3	9.2		1.3	0.5	0.5	2.1	
	B3LYP	1.8	1.7	0.1	5.4		0.4	0.2	0.0	0.8	
	B3LYP-D3	8.4	4.8	2.5	19.4	G(N3)	2.4	0.7	1.4	4.0	P(2B)
	B3LYP-D4	5.5	3.0	1.3	10.5		1.5	0.4	1.0	2.2	
	X3LYP	1.9	1.9	0.2	6.3		0.5	0.3	0.0	0.9	
	X3LYP-D3	11.6	6.9	4.0	28.3		3.4	1.2	1.8	6.0	P(2B)
	O3LYP	2.1	1.8	0.1	4.9		0.6	0.3	0.0	0.9	T(O4)
	O3LYP-D3	15.5	11.0	6.2	43.6		4.6	2.1	1.8	8.7	P(2B)
	B97-2	1.6	1.3	0.0	3.8		0.4	0.2	0.1	0.7	
	TPSSh	1.4	1.1	0.0	4.2	A(N3)	0.4	0.2	0.0	0.8	A(N3), C(O2), G(N3), P(1B1NB)
Local meta-GGA	revTPSS	1.5	1.4	0.1	4.0		0.4	0.3	0.1	1.1	
	TPSS	1.6	1.4	0.1	4.5		0.4	0.3	0.0	1.0	
	TPSS-D3	5.8	4.8	1.1	17.8	A(N3), G(N3)	1.6	0.8	0.5	3.3	P(2B)
	TPSS-D4	4.3	3.7	0.6	13.1	G(N3)	1.1	0.6	0.4	2.3	
	M06-L	2.0	2.0	0.5	8.1	A(N3)	0.6	0.5	0.1	1.6	P(2B), A(N3)
	M11-L	2.9	2.2	0.4	9.4	A(N3)	0.9	0.6	0.1	2.4	P(2NB)
	MN12-L	3.2	2.4	0.5	8.4		1.0	0.7	0.1	2.2	

	MN15-L	1.3	0.8	0.2	2.5		0.4	0.3	0.1	1.2	P(1B1NB,2B)
Local GGA	mPW91	1.8	1.4	0.0	4.5		0.4	0.2	0.0	0.9	
	BLYP	2.3	1.7	0.3	5.5		0.7	0.5	0.1	1.9	
	BLYP-D3	8.7	6.9	1.7	26.3	G(N3)	2.3	0.9	1.1	4.2	
	BP86	1.8	1.7	0.1	4.8		0.6	0.4	0.0	1.4	
	BP86-D3	8.9	6.9	2.1	25.1		2.4	0.9	0.9	4.4	A(N3)
	PBE	2.7	1.9	0.6	6.0		0.8	0.5	0.2	1.9	
	PBE-D3	2.6	3.0	0.3	9.3	A(N3), G(N3)	0.6	0.5	0.2	1.6	
	PBE-D4	2.1	2.6	0.0	8.0	A(N3), G(N3)	0.5	0.4	0.0	1.4	
Local LDA	SVWN5	3.6	3.4	0.3	10.5		0.9	0.6	0.1	2.0	

[a]Complexes resulting in an outlier in the boxplot are listed according to the nucleic acid component (A, C, T, G, U, or P) and binding site (in parentheses). See Figure 3.2 for definitions of the complexes considered in present work and Figure 3.5 for the outliers.

Table B.6. MPEs (%) and MUEs (kcal/mol) in DFT binding strengths along with corresponding standard (SD), minimum (Min), and maximum (Max) deviations evaluated across Rb^+–nucleic acid complexes for each functional relative to CCSD(T)/CBS reference values, as well as outliers based on the boxplot analysis.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Double-Hybrid	B2-PLYP	0.9	0.7	0.0	2.2		0.2	0.1	0.0	0.5	
	B2-PLYP-D3	5.5	3.2	1.5	11.1		1.6	0.5	0.7	2.7	P(2B)
	mPW2-PLYP	1.0	0.7	0.0	1.9		0.2	0.1	0.0	0.4	
	PBE-QIDH	1.2	0.9	0.3	2.5		0.4	0.2	0.1	0.8	
	PBE0-DH	1.4	1.4	0.0	3.5		0.4	0.4	0.0	1.1	
	DSD-PBEP86	3.7	2.6	1.4	8.7	A(N7)	1.1	0.6	0.3	2.5	
RSH	M11	3.9	2.5	0.5	6.8		1.2	0.6	0.1	2.2	
	MN12-SX	4.1	2.2	1.2	8.1		1.2	0.5	0.6	2.3	
	MN12-SX-D3	8.5	4.0	2.9	15.0		2.5	0.8	1.3	4.0	
RSH meta-GGA	ωB97M-V	0.8	0.5	0.0	1.7		0.2	0.1	0.0	0.4	
RSH GGA	ωB97	1.1	1.1	0.1	2.8		0.3	0.3	0.0	0.9	
	ωB97X	1.3	1.3	0.0	3.0		0.3	0.4	0.0	1.0	
	ωB97X-D	7.3	4.6	1.6	14.8		1.9	0.6	0.9	3.0	P(2B)
	ωB97X-D3	13.7	6.9	6.2	28.1	A(N1,N7)	4.4	2.2	2.1	8.8	
	ωB97X-V	0.9	0.9	0.0	2.0		0.2	0.2	0.0	0.7	
	ωB97X-D4	11.7	5.6	4.5	22.2	A(N7)	3.6	1.4	2.0	6.2	
	HSE06	1.7	1.4	0.0	4.5		0.4	0.3	0.0	1.0	
	HSE06-D3	6.2	3.4	1.8	13.0		1.7	0.5	0.8	2.5	
	LC-PBE	2.7	2.0	0.4	5.5		0.8	0.6	0.2	2.0	G(N7), C(O2)
	LC-ωPBE	1.6	1.4	0.1	3.9		0.4	0.4	0.0	1.2	A(N3), C(O2)
	LC-ωPBE-D3	6.9	3.4	2.3	13.2		2.0	0.7	1.1	3.1	
	CAM-B3LYP	2.0	1.4	0.0	3.6		0.5	0.3	0.0	1.1	
	CAM-B3LYP-D3	6.9	3.3	2.0	12.4		2.0	0.5	1.3	2.7	
GH meta-GGA	M06	2.3	1.9	0.6	7.1	A(N3)	0.7	0.4	0.1	1.2	
	M06-2X	1.6	0.9	0.6	3.0		0.5	0.3	0.2	0.9	
	M06-HF	2.8	1.5	0.7	5.1		0.9	0.5	0.1	2.0	P(2NB)
	MN15	1.5	1.2	0.1	3.0		0.4	0.2	0.0	0.7	
	BMK	2.6	1.8	0.7	5.4		0.8	0.4	0.2	1.5	

	PW6B95	1.7	1.3	0.0	3.6		0.4	0.2	0.0	0.7	
GH GGA	BH&HLYP	2.9	2.0	0.3	5.8		0.9	0.5	0.2	1.8	
	SOGGA11-X	2.0	1.3	0.3	3.9		0.6	0.3	0.2	1.1	
	B3PW91	1.7	1.7	0.0	5.5		0.4	0.3	0.0	1.1	
	B3PW91-D3	10.5	5.6	3.4	19.6		3.0	0.9	1.6	4.8	
	PBE0	1.7	1.3	0.0	4.3		0.4	0.3	0.0	0.9	
	PBE0-D3	5.3	3.0	1.6	11.2		1.5	0.5	0.6	2.3	
	B3LYP	1.8	1.5	0.0	4.2		0.4	0.3	0.0	0.8	
	B3LYP-D3	9.5	4.9	2.9	17.3		2.7	0.8	1.6	4.4	P(2B)
	B3LYP-D4	6.9	3.5	1.7	13.0		1.9	0.5	1.2	2.7	
	X3LYP	1.8	1.5	0.1	3.7		0.4	0.3	0.0	0.8	
	X3LYP-D3	13.7	7.3	5.0	27.3		4.1	1.5	2.1	7.0	
	O3LYP	1.8	1.2	0.5	3.6		0.5	0.2	0.2	0.7	
	O3LYP-D3	18.6	11.4	8.1	42.1		5.7	2.7	2.3	10.7	
	B97-2	1.6	1.4	0.0	4.5		0.4	0.3	0.0	0.9	
	TPSSh	1.5	1.4	0.2	5.1	A(N3)	0.4	0.2	0.1	0.8	
Local meta-GGA	revTPSS	1.5	1.4	0.2	5.0	A(N3)	0.4	0.3	0.1	1.0	
	TPSS	1.6	1.4	0.0	5.4	A(N3)	0.4	0.2	0.0	0.9	A(N3), P(1B1NB)
	TPSS-D3	6.3	4.1	1.3	14.7		1.8	0.7	0.7	3.5	P(2B)
	TPSS-D4	5.1	3.4	0.8	12.6		1.4	0.6	0.6	2.6	
	M06-L	2.4	2.7	0.1	9.8	A(N3)	0.8	0.6	0.0	1.8	
	M11-L	3.8	2.6	1.3	10.9	A(N3)	1.2	0.6	0.4	2.5	
	MN12-L	4.3	2.7	1.3	10.0		1.3	0.7	0.5	2.3	
	MN15-L	1.8	1.3	0.0	3.6		0.4	0.2	0.0	0.7	
Local GGA	mPW91	1.7	1.5	0.1	4.6		0.4	0.3	0.0	1.0	
	BLYP	2.1	1.5	0.1	4.6		0.6	0.4	0.0	1.7	P(2NB)
	BLYP-D3	9.4	5.9	2.1	19.8		2.6	0.9	1.4	4.6	P(2B)
	BP86	1.8	1.6	0.2	6.0	A(N3)	0.5	0.3	0.2	1.2	A(N3), P(2NB)
	BP86-D3	10.2	6.8	2.6	22.1		2.8	1.0	1.3	4.9	
	PBE	2.3	1.8	0.1	4.7		0.7	0.5	0.0	1.7	P(2NB)
	PBE-D3	2.8	3.0	0.1	10.6	A(N3)	0.7	0.6	0.1	1.7	
	PBE-D4	2.5	2.8	0.2	10.0	A(N3)	0.6	0.5	0.0	1.6	
Local LDA	SVWN5	3.3	2.9	0.2	9.1		0.8	0.5	0.1	1.5	

[a]Complexes resulting in an outlier in the boxplot statistics are listed according to the nucleic acid component (A, C, T, G, U, or P) and binding site (in parentheses). See Figure 3.2 for definitions of the complexes considered in present work and Figure 3.6 for the outliers.

Table B.7. MPEs (%) and MUEs (kcal/mol) in DFT binding strengths along with corresponding standard (SD), minimum (Min), and maximum (Max) deviations evaluated across Cs^+–nucleic acid complexes for each functional relative to CCSD(T)/CBS reference values, as well as outliers based on the boxplot analysis.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Double-Hybrid	B2-PLYP	1.0	0.7	0.0	2.1		0.2	0.1	0.0	0.5	P(2NB)
	B2-PLYP-D3	6.1	4.0	1.8	13.9		1.6	0.6	0.7	2.8	
	mPW2-PLYP	1.1	0.8	0.0	2.4		0.2	0.1	0.0	0.4	
	PBE-QIDH	1.1	0.8	0.1	2.3		0.3	0.2	0.0	0.6	A(N1), G(N7), C(O2)
	PBE0-DH	1.5	1.1	0.3	3.5		0.4	0.2	0.1	0.9	
	DSD-PBEP86	3.6	2.9	1.0	9.5	A(N7)	1.0	0.6	0.2	2.3	
RSH	M11	4.0	2.5	0.7	7.7		1.2	0.6	0.1	2.1	
	MN12-SX	5.0	2.4	2.0	10.1		1.5	0.9	0.6	3.4	
	MN12-SX-D3	10.8	5.3	4.3	19.4		3.0	1.2	1.5	5.5	
RSH meta-GGA	ωB97M-V	0.9	0.6	0.1	1.9		0.2	0.1	0.1	0.3	
RSH GGA	ωB97	1.2	0.7	0.1	2.3		0.3	0.2	0.1	0.6	
	ωB97X	1.4	1.0	0.1	2.7		0.3	0.2	0.1	0.8	G(N7), C(O2)
	ωB97X-D	18.8	13.9	4.0	47.2		4.4	1.3	2.3	6.7	
	ωB97X-D3	15.1	8.2	6.9	33.2	A(N1,N7)	4.5	2.3	2.1	9.0	
	ωB97X-V	1.0	0.5	0.1	1.7		0.2	0.1	0.1	0.4	
	ωB97X-D4	14.4	7.5	5.6	29.1	A(N7)	4.0	1.7	2.1	7.1	
	HSE06	1.8	1.4	0.1	4.8		0.4	0.2	0.1	0.8	
	HSE06-D3	7.1	4.2	2.1	15.2		1.8	0.6	0.8	2.8	
	LC-PBE	2.4	2.1	0.1	5.0		0.6	0.5	0.0	1.7	G(N7), C(O2)
	LC-ωPBE	1.7	1.3	0.0	3.7		0.4	0.3	0.0	0.9	
	LC-ωPBE-D3	7.4	4.2	2.4	14.9		2.0	0.6	0.9	3.0	
	CAM-B3LYP	1.9	1.7	0.0	4.5		0.4	0.3	0.0	0.9	
	CAM-B3LYP-D3	7.6	3.9	2.2	14.5		2.0	0.5	1.2	2.9	
GH meta-GGA	M06	2.6	2.1	0.8	8.6	A(N3)	0.8	0.5	0.1	1.8	
	M06-2X	1.5	1.0	0.5	3.4		0.4	0.3	0.1	1.0	
	M06-HF	2.5	1.5	0.3	4.7		0.7	0.5	0.1	1.9	P(2NB)

	MN15	1.5	1.3	0.1	4.1		0.3	0.2	0.1	0.6	
	BMK	2.3	1.9	0.4	5.7		0.6	0.4	0.1	1.3	
	PW6B95	1.8	1.4	0.1	4.2		0.4	0.2	0.1	0.6	
GH GGA	BH&HLYP	2.8	1.7	0.4	5.4		0.8	0.4	0.3	1.5	
	SOGGA11-X	1.8	1.3	0.4	3.7		0.5	0.3	0.1	0.9	
	B3PW91	1.9	1.6	0.3	5.9		0.4	0.2	0.1	0.9	
	B3PW91-D3	11.8	6.9	3.9	24.1		3.1	1.0	1.5	5.1	P(2B)
	PBE0	1.7	1.3	0.0	4.5	A(N3)	0.4	0.2	0.0	0.8	G(N7)
	PBE0-D3	5.8	3.5	1.9	12.7		1.5	0.6	0.6	2.5	
	B3LYP	1.8	1.7	0.0	4.5		0.4	0.3	0.0	0.7	
	B3LYP-D3	10.6	5.9	3.4	20.7		2.8	0.9	1.5	4.7	P(2B)
	B3LYP-D4	8.2	4.2	2.4	16.0		2.2	0.6	1.3	3.2	
	X3LYP	1.8	1.7	0.1	4.4		0.4	0.3	0.0	0.7	
	X3LYP-D3	16.6	9.9	6.2	36.8	A(N7)	4.5	1.8	2.2	7.9	
	O3LYP	1.9	1.5	0.2	4.0		0.4	0.2	0.1	0.6	
	O3LYP-D3	23.9	16.1	10.2	59.0		6.7	3.4	2.7	12.7	
	B97-2	1.7	1.3	0.2	4.7	A(N3)	0.4	0.2	0.2	0.7	
	TPSSh	1.6	1.5	0.0	5.5	A(N3)	0.4	0.2	0.0	0.8	
Local meta-GGA	revTPSS	1.6	1.6	0.1	5.4	A(N3)	0.4	0.2	0.1	0.8	
	TPSS	1.8	1.5	0.4	5.9	A(N3)	0.4	0.2	0.2	0.8	A(N3)
	TPSS-D3	7.1	4.8	1.9	16.7		1.9	0.9	0.7	3.9	P(2B)
	TPSS-D4	6.3	4.1	1.6	15.1		1.7	0.7	0.7	3.3	P(2B)
	M06-L	2.9	3.0	0.3	11.3	A(N3)	1.0	0.8	0.0	2.6	
	M11-L	4.5	3.0	1.9	12.9	A(N3)	1.4	1.0	0.4	3.7	P(2NB)
	MN12-L	4.9	2.9	1.9	11.7		1.5	0.9	0.5	3.3	
	MN15-L	1.8	1.4	0.1	3.9		0.4	0.1	0.0	0.6	C(O2)
Local GGA	mPW91	1.8	1.3	0.2	4.8		0.4	0.2	0.2	0.9	
	BLYP	2.2	1.9	0.1	5.4		0.5	0.3	0.1	1.2	
	BLYP-D3	10.8	7.0	2.9	23.0		2.8	1.1	1.4	5.2	P(2B)
	BP86	1.9	1.8	0.1	6.7	A(N3)	0.4	0.2	0.1	0.9	A(N3)
	BP86-D3	12.1	8.5	3.4	27.5		3.1	1.2	1.4	5.5	
	PBE	2.4	2.1	0.1	5.0		0.6	0.4	0.1	1.3	
	PBE-D3	3.2	3.4	0.0	11.9	A(N3)	0.8	0.6	0.0	2.0	
	PBE-D4	3.3	3.4	0.0	11.9	A(N3)	0.8	0.6	0.0	1.8	

Local LDA	SVWN5	3.3	3.1	0.0	9.1		0.7	0.4	0.0	1.3

[a]Complexes resulting in an outlier in the boxplot are listed according to the nucleic acid component (A, C, T, G, U, or P) and binding site (in parentheses). See Figure 3.2 for definitions of the complexes considered in present work and Figure 3.6 for the outliers.

Table B.8. MPEs (%) and MUEs (kcal/mol) in DFT binding strengths along with corresponding standard (SD), minimum (Min), and maximum (Max) deviations evaluated across group I metal–nucleic acid complexes for each functional relative to CCSD(T)/CBS reference values, as well as outliers based on the boxplot analysis.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Double-Hybrid	B2-PLYP	1.5	1.5	0.0	7.9	Na(A(N1,N7), G(N3))	0.6	0.7	0.0	2.5	Na(A(N7), G(N3), P(2B,2NB))
	B2-PLYP-D3	3.9	3.2	0.0	13.9	KG(N3), RbA(N3,N7), CsA(N1,N3,N7)	1.3	0.7	0.0	3.7	LiP(2B)
	mPW2-PLYP	1.6	1.3	0.0	5.3	NaG(N3)	0.7	0.7	0.0	2.7	LiP(2B)
	PBE-QIDH	1.8	2.2	0.1	12.7	Na(A(N1,N3,N7),G(N3), U(O2))	0.7	0.8	0.0	3.3	Na((A(N1,N3,N7), C(O2), G(N3,N7), P(1B1NB,2B,2NB), T/U(O2))
	PBE0-DH	1.9	1.9	0.0	11.2	Na(A(N1,N7), G(N3)),	0.7	0.7	0.0	2.9	Na(A(N1,N7), G(N3))
	DSD-PBEP86	2.9	2.5	0.1	9.7	KG(N3), RbA(N7), CsA(N7)	1.0	0.6	0.1	2.5	
RSH	M11	3.0	2.4	0.0	9.4		1.2	0.8	0.0	3.3	
	MN12-SX	3.4	2.6	0.1	12.5	NaG(N3)	1.3	0.8	0.0	3.4	NaG(N3), CsP(2NB)
	MN12-SX-D3	5.8	4.9	0.1	19.4		1.8	1.2	0.1	5.5	CsP(2NB)
RSH meta-GGA	ωB97M-V	1.1	1.0	0.0	5.2	NaG(N3),KG(N3)	0.4	0.4	0.0	1.6	Li(G(N7), P(1B1NB,2B,2NB))
RSH GGA	ωB97	1.7	1.9	0.0	11.6	Na(A(N1,N7),G(N3)), LiG(N3)	0.7	0.7	0.0	3.0	LiG(N3), Na(A(N1,N7), G(N3))
	ωB97X	2.7	2.8	0.0	15.7	Na(A(N1,N7), G(N3)), LiG(N3)	1.1	1.1	0.0	4.1	
	ωB97X-D	7.6	8.6	0.5	47.2	Cs(A(N1,N3,N7),T(O4))	2.3	1.3	0.3	6.7	Cs(A(N1,N3), G(N7))
	ωB97X-D3	9.7	7.3	2.0	33.2	KG(N3), RbA(N1,N7), CsA(N1,N3,N7)	3.6	2.0	0.7	9.0	RbP(2B), CsP(2B,2NB)
	ωB97X-V	1.1	1.2	0.0	7.3	Na(A(N7),G(N3)), KG(N3)	0.4	0.4	0.0	1.9	Li(A(N1,N3,N7), P(2NB),C(O2))
	ωB97X-D4	7.6	6.9	0.8	29.1	CsA(N7)	2.5	1.6	0.3	7.1	CsP(2B,2NB)
	HSE06	1.9	1.7	0.0	8.2	Na(A(N7), G(N3))	0.7	0.6	0.0	2.4	Li(A(N1,N7), G(N3), P(2NB), U(O4)), Na(A(N1,N7), G(N3))
	HSE06-D3	4.5	3.3	0.1	15.2	CsA(N3)	1.6	0.8	0.0	3.2	
	LC-PBE	2.8	2.3	0.1	10.8	NaG(N3)	1.1	1.0	0.0	4.8	Li(C(O2), G(N7), P(2NB))
	LC-ωPBE	2.0	2.1	0.0	12.5	Na(A(N1,N7), G(N3)),	0.7	0.7	0.0	3.3	LiG(N7), Na(A(N1,N7), G(N3))

	LC-ωPBE-D3	4.9	3.6	0.3	14.9		1.7	0.9	0.3	3.3	
	CAM-B3LYP	2.7	2.0	0.0	7.6		1.1	1.1	0.0	4.1	Li(C(O2), G(N7), P(1B1NB,2B,2NB), T/U(O4))
	CAM-B3LYP-D3	6.1	3.0	1.3	14.5	CsA(N3)	2.4	1.2	0.3	5.5	Li(G(N7),P(1B1NB,2B,2NB))
GH meta-GGA	M06	3.4	3.5	0.1	18.9	Na(A(N1,N7),G(N3))	1.4	1.3	0.1	5.4	Na(A(N1,N3), G(N3))
	M06-2X	1.5	1.1	0.0	4.8		0.7	0.6	0.0	2.8	LiP(1B1NB,2NB)
	M06-HF	3.4	2.3	0.3	9.5	Li(A(N1,N7), G(N3))	1.7	1.6	0.1	7.0	Li(C(O2), P(1B1NB,2NB))
	MN15	2.3	2.5	0.1	13.7	Na(A(N1,N7), G(N3))	0.8	0.9	0.0	3.8	Na(A(N1,N7), G(N3))
	BMK	3.3	3.2	0.2	17.8	Na(A(N1,N7), G(N3))	1.3	1.1	0.0	4.8	Na(A(N1,N7), G(N3))
	PW6B95	2.0	1.7	0.0	7.3	NaG(N3)	0.7	0.7	0.0	2.6	Li(A(N3), C(O2), G(N7), P(1B1NB,2B,2NB), T(O4))
GH GGA	BH&HLYP	3.4	2.3	0.1	9.3		1.4	1.2	0.1	4.7	Li(C(O2), G(N7), P(1B1NB,2B), T/U(O4)
	SOGGA11-X	2.0	1.8	0.1	11.2	Na(A(N7), G(N3))	0.8	0.6	0.0	2.9	Na(A(N7),G(N3))
	B3PW91	2.6	2.9	0.0	14.9	Na(A(N1,N7), G(N3)),	1.0	1.1	0.0	4.4	Na(A(N1,N7), G(N3), P(1B1NB,2B,2NB))
	B3PW91-D3	7.0	6.3	0.1	24.1	CsA(N3,N7), KG(N3)	2.2	1.3	0.1	5.1	
	PBE0	1.9	1.7	0.0	8.4	Na(A(N1,N7), G(N3))	0.7	0.6	0.0	2.4	Li(A(N3),G(N7)), NaG(N3)
	PBE0-D3	4.0	2.8	0.1	12.7	CsA(N3,N7)	1.4	0.8	0.0	3.2	LiG(N7),
	B3LYP	2.0	1.7	0.0	7.0		0.8	0.7	0.0	2.7	LiP(2B)
	B3LYP-D3	7.3	4.8	1.3	20.7	KG(N3), CsA(N3,N7)	2.6	1.1	0.5	6.1	LiP(1B1NB,2B)
	B3LYP-D4	5.1	3.6	0.3	16.0	CsA(N3)	1.7	0.8	0.4	3.6	
	X3LYP	2.2	1.8	0.1	6.3		0.9	0.9	0.0	3.5	Li(A(N3), P(1B1NB,2B), T/U(O4))
	X3LYP-D3	10.9	7.3	2.9	36.8	KG(N3), RbA(N3,N7), CsA(N1,N3,N7)	3.9	1.5	1.8	7.9	LiP(2B), CsP(2B)
	O3LYP	2.5	2.5	0.1	11.6	Na(A(N1,N7), G(N3))	0.9	1.0	0.0	3.9	Na(A(N1,N7), C(O2), G(N3,N7),P(1B1NB,2B,2NB))
	O3LYP-D3	12.8	12.6	0.1	59.0	KG(N3), RbA(N7), CsA(N1,N3,N7)	4.3	2.8	0.1	12.7	CsP(2B)
	B97-2	2.6	2.9	0.0	15.4	Na(A(N1,N7), G(N3))	1.0	1.1	0.0	4.4	Na(A(N1,N7), G(N3), P(1B1NB,2B,2NB))

	TPSSh	1.9	1.8	0.0	7.9	Na(A(N1), G(N3), U(O2)), RbA(N3), CsA(N3)	0.7	0.7	0.0	2.8	NaA(N7)
Local meta-GGA	revTPSS	1.9	1.8	0.0	7.8	Na(A(N1,N7), G(N3), U(O2))	0.8	0.7	0.0	2.8	LiP(2B), Na(A(N7), C(O2), P(1B1NB,2B))
	TPSS	2.0	1.8	0.0	8.0	Na(A(N7), G(N3), U(O2)), RbA(N3), CsA(N3)	0.8	0.7	0.0	2.9	LiP(1B1NB,2B), Na(A(N1,N7), C(O2), G(N7), P(1B1NB,2NB), U(O2)
	TPSS-D3	4.7	4.1	0.0	17.8	KG(N3), RbA(N3), CsA(N3,N7)	1.6	1.0	0.0	5.3	LiP(2B)
	TPSS-D4	3.8	3.4	0.0	15.1	K(A(N3), G(N3), RbA(N3), CsA(N3,N7)	1.2	0.8	0.0	3.7	LiP(2B), CsP(2B)
	M06-L	3.8	3.8	0.1	18.0	Na(A(N1,N7), G(N3)), CsA(N3)	1.6	1.4	0.0	5.6	NaA(N7)
	M11-L	4.4	3.2	0.4	16.2	Na(A(N7), G(N3), U(O2)), RbA(N3), CsA(N3)	1.9	1.3	0.1	5.5	
	MN12-L	3.9	3.2	0.1	16.5	Na(A(N7), G(N3)), CsA(N3)	1.4	1.0	0.1	4.3	Na(A(N7), G(N3))
	MN15-L	2.8	2.4	0.0	12.5	Na(A(N7), G(N3))	1.2	1.0	0.0	3.3	
Local GGA	mPW91	2.1	2.1	0.0	11.6	Na(A(N1,N7), G(N3))	0.8	0.7	0.0	3.3	Na(A(N1,N7),G(N3), P(1B1NB)
	BLYP	2.2	1.7	0.0	7.4	NaA(N7)	0.9	0.8	0.0	3.3	LiP(2B),Na(A(N7), P(1B1NB,2NB))
	BLYP-D3	7.0	6.0	0.3	26.3	KG(N3), RbA(N3,N7), CsA(N3,N7)	2.3	1.2	0.1	7.3	LiP(2B)
	BP86	3.4	3.1	0.1	13.3	Na(A(N1,N7), G(N3))	1.4	1.4	0.0	4.9	
	BP86-D3	6.8	7.0	0.1	27.5	K(A(N3), G(N3)), Rb/CsA(N1,N3,N7)	2.0	1.2	0.1	5.5	CsP(2B)
	PBE	2.3	1.8	0.0	6.9		0.9	0.6	0.0	2.5	LiP(2B), Na(A(N7), P(1B1NB,2NB))
	PBE-D3	2.4	2.6	0.0	11.9	K(A(N3), G(N3)), RbA(N3), CsA(N3)	0.8	0.8	0.0	4.2	Li(A(N3), P(1B1NB,2B))
	PBE-D4	2.2	2.4	0.0	11.9	KA(N3), RbA(N3), CsA(N3)	0.8	0.6	0.0	3.0	LiP(2B)
Local LDA	SVWN5	4.0	3.0	0.0	10.5		1.6	1.4	0.0	5.6	LiG(N7)

[a]Complexes resulting in an outlier in the boxplot are listed according to the metal (Li^+, Na^+, K^+, Rb^+, or Cs^+) and nucleic acid component (A, C, T, G, U, or P), with the binding site indicated in parentheses. See Figure 3.2 for definitions of the complexes considered in present work and Figure 3.7 for the outliers.

Table B.9. MPEs and MUEs for each DFT functional (with and without counterpoise corrections) relative to CCSD(T)/CBS reference values for group I metal complexes.[a]

Metal	Functional	MPE		MUE		Absolute Diff.[b]	
		w/o CP	w/ CP	w/o CP	w/ CP	MPE	MUE
Li^+	PBE-QIDH	1.0	0.7	0.8	0.4	0.3	0.4
	ωB97X-V	1.5	0.9	1.1	0.6	0.6	0.5
	O3LYP	1.7	1.0	1.2	0.6	0.7	0.6
Na^+	B2-PLYP-D3	0.9	1.1	0.6	0.4	0.2	0.2
	HSE06-D3	2.0	1.6	1.1	0.7	0.4	0.4
	PBE0-D3	1.8	1.5	1.0	0.6	0.3	0.4
	TPSS-D3	1.4	1.4	0.8	0.6	0.0	0.2
	PBE-D3	1.6	1.5	0.8	0.6	0.1	0.2
K^+	B2-PLYP	1.3	1.1	0.5	0.3	0.2	0.2
	ωB97X-V	1.4	1.1	0.6	0.3	0.3	0.3
	M06-2X	2.0	0.9	0.8	0.3	1.1	0.5
	MN15-L	1.8	1.3	0.5	0.4	0.5	0.1
Rb^+	mPW2-PLYP	2.5	1.0	0.9	0.4	1.5	0.5
	ωB97M-V	2.3	0.8	0.8	0.2	1.5	0.6
	MN15-L	2.9	1.8	0.9	0.4	1.1	0.5
Cs^+	B2-PLYP	1.8	1.0	0.6	0.2	0.8	0.4
	ωB97M-V	1.7	0.9	0.6	0.2	0.8	0.4
	ωB97X-V	1.7	1.0	0.6	0.2	0.7	0.4
	MN15-L	2.9	1.8	0.8	0.4	1.1	0.4
	mPW91	2.3	1.8	0.8	0.4	0.5	0.4
Group I	mPW2-PLYP	2.1	1.6	1.1	0.7	0.5	0.4
	ωB97M-V	1.7	1.1	0.8	0.4	0.6	0.4
	revTPSS	2.1	1.9	0.9	0.8	0.2	0.1
	TPSS	2.1	2.0	0.9	0.8	0.1	0.1

[a]See Figure 3.2 for the complexes considered in present work. [b]Absolute differences between counterpoise uncorrected and corrected for each functional are calculated for each metric.

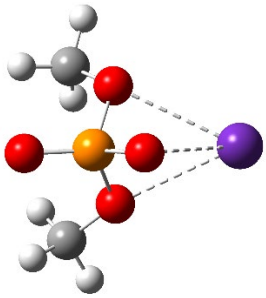


Figure B.1. CCSD(T)/CBS//MP2/def2-TZVPP characterized P(2B) complex between each metal and the dimethylphosphate model, which includes an additional contact with the nonbridging oxygen for all metals except Li^+.

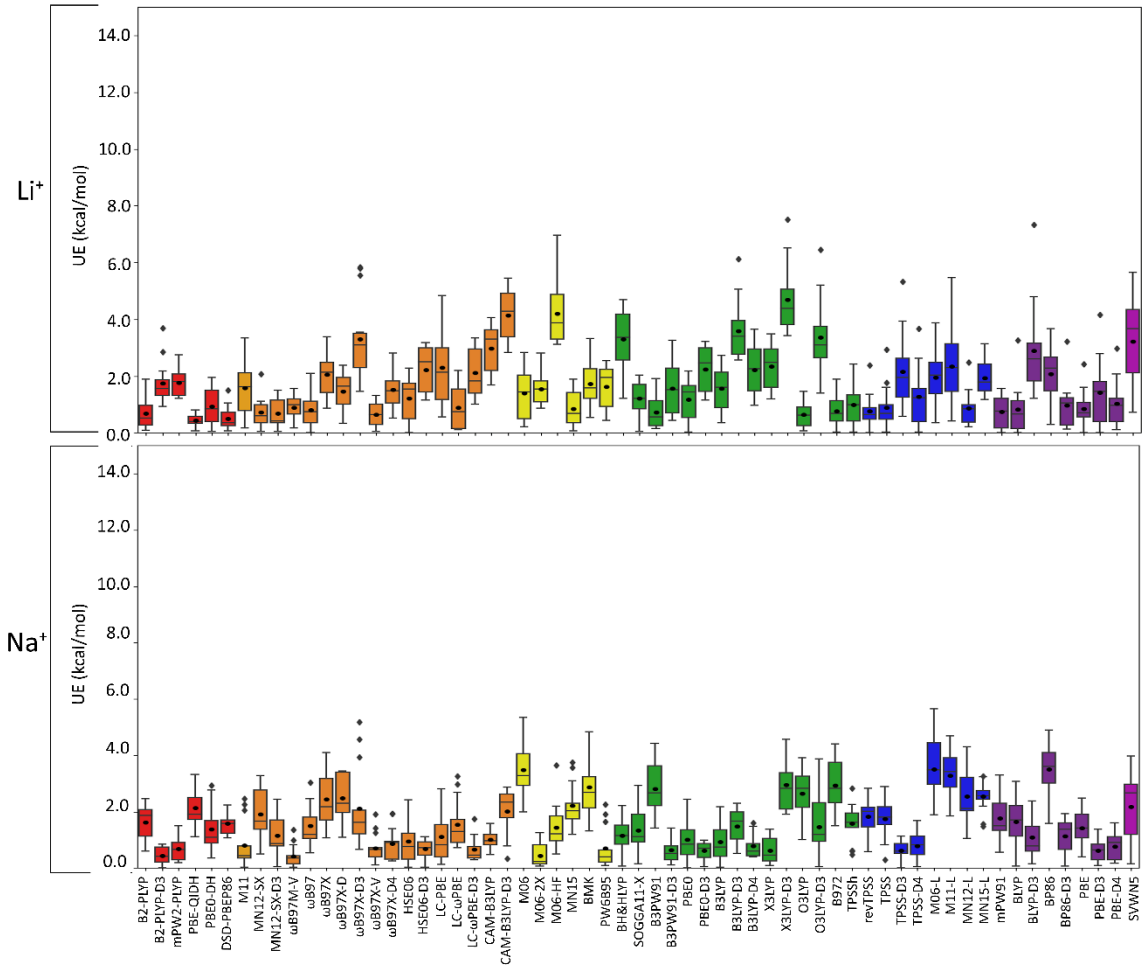


Figure B.2. Boxplots of the unsigned errors (kcal/mol) in Li^+ or Na^+–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

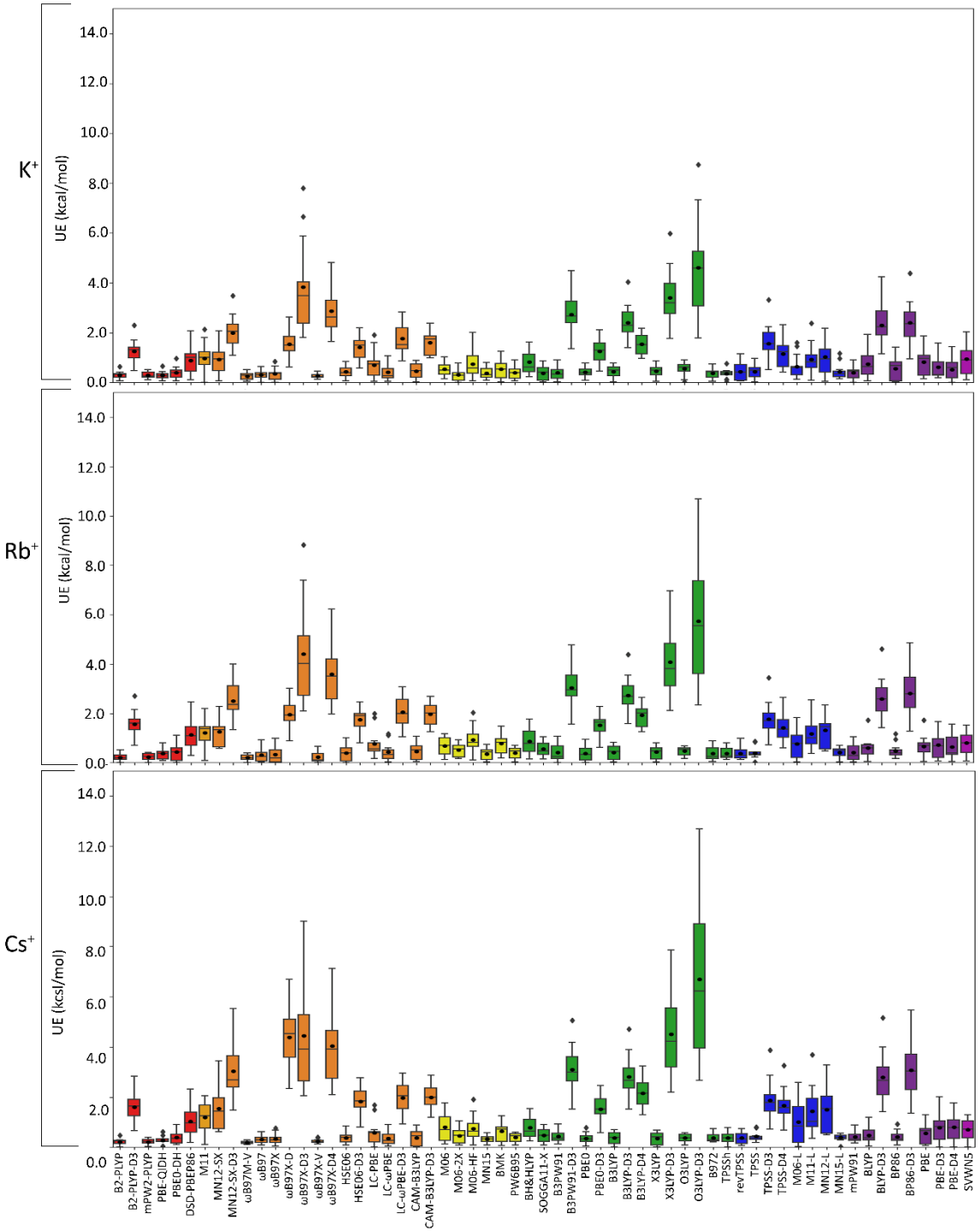


Figure B.3. Boxplots of the unsigned errors (kcal/mol) in K^+, Rb^+ or Cs^+–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

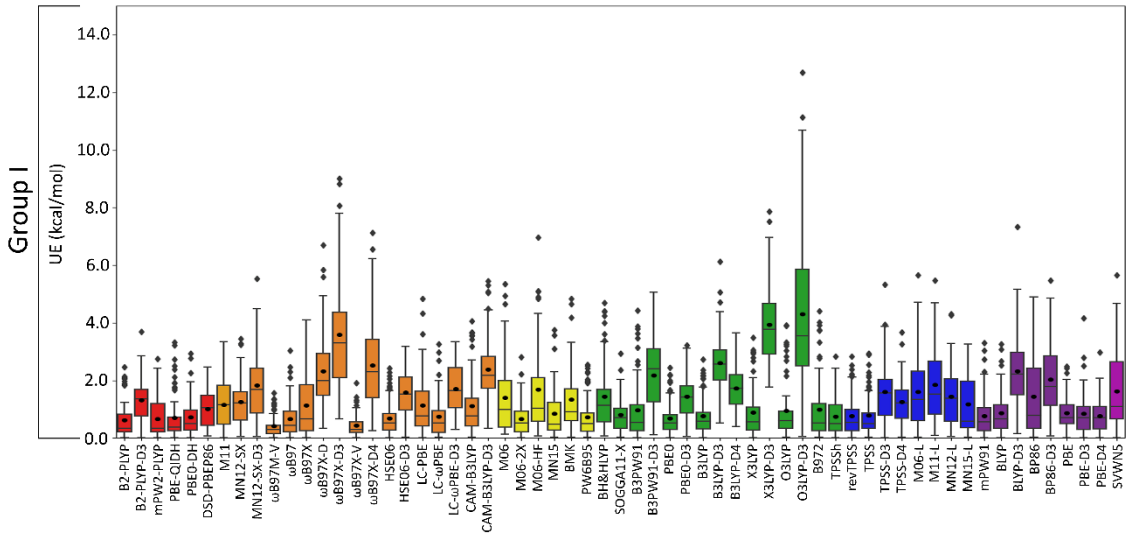


Figure B.4. Boxplot of the unsigned errors (kcal/mol) in group I metal–nucleic acid DFT binding energies relative to CCSD(T)/CBS reference values, with the functionals sorted according to double-hybrids (red), RSH (orange), GH meta-GGA (yellow), GH GGA (green), meta-GGA (blue), GGA (purple), and LDA (magenta).

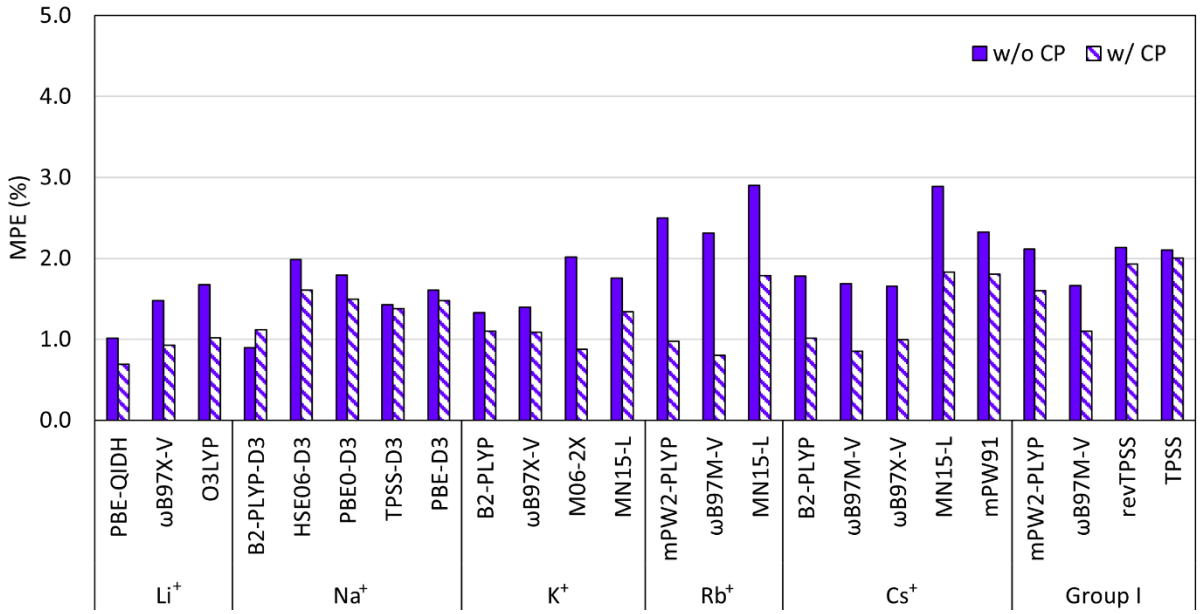


Figure B.5. MPEs (%) in the uncorrected (solid) and counterpoise-corrected (striped) DFT binding energies relative to the CCSD(T)/CBS reference values for the top-performing functionals for each metal and over all group I.

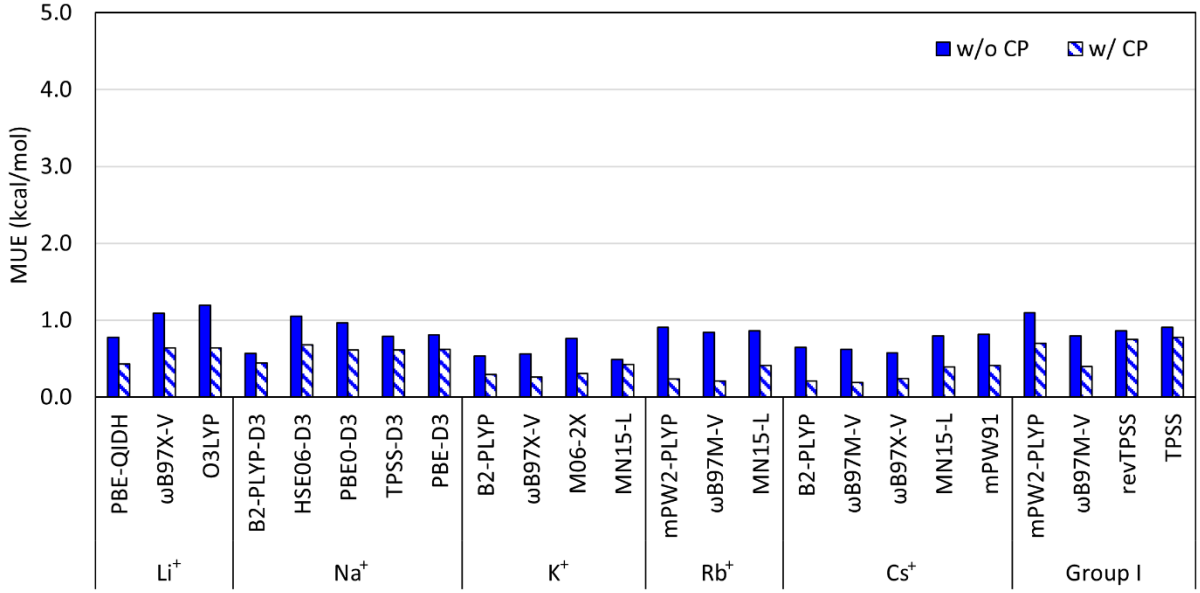


Figure B.6. MUEs (kcal/mol) in the uncorrected (solid) and counterpoise-corrected (striped) DFT binding energies relative to the CCSD(T)/CBS reference values for the top-performing functionals for each metal and over all group I.

Appendix C

Supplementary Information for Chapter 4: Building a Data Set of Group II Metal–Nucleic Acid Interactions and Evaluating the Performance of Density Functional Theory Methods

Contains Tables C.1–C.11 and Figures C.1–C.9

Table C.1. Coordination distances (Å) between each group II metal and nucleic acid component calculated at the MP2/def2-TZVPP level of theory.[a]

Binding Site		Be^{2+}	Mg^{2+}	Ca^{2+}	Sr^{2+}	Ba^{2+}
A(N6–N7)	X···N6	1.713	2.116	2.461	2.628	2.821
	X···N7	1.591	2.000	2.314	2.470	2.659
A(N1–N6)	X···N1	1.587	2.009	2.320	2.476	2.659
	X···N6	1.683	2.102	2.426	2.592	2.788
A(N3)	X···N3	1.515	1.933	2.219	2.370	2.548
G(O6–N7)	X···O6	1.547	1.945	2.213	2.350	2.509
	X···N7	1.649	2.041	2.382	2.547	2.746
G(N2–N3)	X···N2	1.686	2.102	2.430	2.595	2.786
	X···N3	1.568	1.995	2.309	2.470	2.660
C(O2–N3)	X···O2	1.532	1.917	2.160	2.286	2.432
	X···N3	1.657	2.082	2.441	2.630	2.863
C(N3–N4)	X···N3	1.572	1.984	2.278	2.413	
	X···N4	1.699	2.148	2.536	2.781	
T(O4)	X···O4	1.367	1.775	2.045	2.184	2.348
T(O2)	X···O2	1.372	1.779	2.055	2.198	2.365
U(O4)	X···O4	1.369	1.776	2.050	2.190	2.357
U(O2)	X···O2	1.374	1.782	2.062	2.206	2.377
P(2NB)	X···O(NB)	1.515	1.904	2.157	2.282	2.433
	X···O(NB)	1.515	1.904	2.156	2.282	2.433
P(1B1NB)	X···O(B)	1.476	1.919	2.235	2.383	2.593
	X···O(NB)	1.513	1.850	2.077	2.188	2.312
P(2B)	X···O(B)	1.477	1.866	2.409	2.569	2.768
	X···O(B)	1.477	1.866	2.408	2.569	2.768
	X···O(NB)[b]			2.144	2.255	2.368

[a]Definitions of each complex studied in the present work are shown in Figure 4.2. Only one complex could not be characterized, which corresponds to an empty space within the table. [b]For Ca^{2+}, Sr^{2+}, and Ba^{2+}, an additional coordination distance is reported to a non-bridging oxygen in the phosphate moiety (Figure C.1).

Table C.2. Binding energies (kcal/mol) of group II metal–nucleic acid complexes calculated at the CCSD(T)/CBS level of theory.[a]

Binding Site	Be^{2+}	Mg^{2+}	Ca^{2+}	Sr^{2+}	Ba^{2+}
A(N6–N7)	–277.7	–166.2	–98.5	–83.8	–70.7
A(N1–N6)	–271.6	–161.4	–96.2	–82.3	–70.0
A(N3)	–230.0	–133.4	–78.7	–68.2	–59.0
G(O6–N7)	–315.8	–203.3	–137.0	–121.8	–108.8
G(N2–N3)	–252.3	–141.1	–77.6	–64.3	–52.7
C(O2–N3)	–295.4	–186.8	–123.7	–110.4	–98.3
C(N3–N4)	–267.5	–158.6	–96.4	–83.8	
T(O4)	–237.5	–139.7	–91.0	–81.1	–71.6
T(O2)	–227.9	–133.6	–86.7	–77.1	–68.1
U(O4)	–234.8	–138.3	–90.3	–80.6	–71.3
U(O2)	–221.5	–128.7	–82.6	–73.2	–64.4
P(2NB)	–511.4	–376.8	–298.0	–280.8	–263.9
P(1B1NB)	–492.6	–361.7	–285.7	–269.5	–253.3
P(2B)	–463.4	–333.8	–283.4	–267.6	–252.6

[a]Definitions of each complex studied in present work as well as the relative energies between each complex are shown in Figure 4.2. Complexes that could not be characterized correspond to empty spaces within the table.

Table C.3. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for group II metal–nucleic acid complexes, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	Max	Outliers[a]
Double-Hybrid	B2-PLYP	0.8	0.5	0.1	2.0		1.2	1.0	0.1	5.0	Be(A(N3), P(2B)), Mg(P(2NB))
	B2-PLYP-D3	1.6	1.3	0.1	5.5	Ba(A/G(N3), T/U(O4))	2.1	1.4	0.2	6.7	Be(P(2B))
	mPW2-PLYP	0.8	0.6	0.0	2.2		1.4	1.3	0.0	5.7	Be(A/G(N3), P(1B1NB,2B), T/U(O4)
	PBE-QIDH	1.1	1.0	0.0	3.9	Mg(G(N3), U(O2))	1.6	1.5	0.0	5.6	Mg(G(N3), P(2B), T/U(O2))
	PBE0-DH	0.9	0.7	0.0	2.8	Mg(T/U(O2))	1.3	0.9	0.0	3.8	
	DSD-PBEP86	1.3	1.0	0.0	4.4	Ba(A/G(N3))	1.8	1.1	0.0	4.6	
RSH	M11	2.3	1.6	0.1	6.0		2.9	1.6	0.1	5.7	
	MN12-SX	4.7	3.1	0.1	13.2		5.9	2.3	0.2	12.0	

	MN12-SX-D3	5.6	4.2	0.2	16.0		6.6	3.1	0.5	14.4	
RSH meta-GGA	ωB97M-V	0.6	0.5	0.0	2.3	Ca/Sr(G(N3))	0.7	0.5	0.0	1.9	
RSH GGA	ωB97	1.8	1.2	0.0	5.3	Mg(G(N3))	3.2	2.7	0.0	9.0	
	ωB97X	2.2	1.8	0.0	6.9		4.3	4.4	0.0	12.9	
	ωB97X-D	3.4	2.1	0.1	9.2	Ba(A(N1,N3,N7)	5.3	3.5	0.1	11.3	
	ωB97X-D3	4.3	3.2	0.1	12.1		5.3	2.8	0.1	14.4	Sr/Ba(P(2B)), Ba(P(2NB))
	ωB97X-V	0.7	0.7	0.0	2.8	Ca(G(N3),U(O2)), Sr/Ba(G(N3))	1.0	0.7	0.0	2.9	Mg(T(O2))
	ωB97X-D4	3.2	3.0	0.0	11.8	Ba(T(O4)	3.4	2.5	0.0	10.0	
	HSE06	1.0	0.8	0.0	4.0	Ba(A(N3))	1.5	1.5	0.0	5.6	Be(A(N3), G(N7), P(2B))
	HSE06-D3	2.1	1.5	0.1	7.4	Ba(A(N3), T(O4))	2.9	1.7	0.1	7.2	
	LC-PBE	1.4	1.3	0.0	5.9	Ca/Sr/Ba(G(N3))	1.9	1.7	0.0	8.5	Be(C(O2), G(N7), P(2NB))
	LC-ωPBE	1.0	0.8	0.0	3.6	Sr/Ba(G(N3))	1.5	1.2	0.0	5.5	Be(G(N7), P(2NB))
	LC-ωPBE-D3	1.5	1.1	0.0	5.0	Ba(A(N3), T/U(O4))	2.2	1.6	0.0	6.9	Be(G(N7))
	CAM-B3LYP	1.4	0.9	0.0	3.6		2.5	2.3	0.0	8.3	Be(G(N7))
	CAM-B3LYP-D3	2.4	1.4	0.1	6.6	Ba(T/U(O4))	3.9	2.6	0.1	9.7	Be(A(N3), G(N7), P(1B1NB,2B), T/U(O4))
GH meta-GGA	M06	3.4	1.8	0.0	9.2	Ca/Sr/Ba(A(N3))	4.9	2.4	0.1	9.4	
	M06-2X	1.0	0.7	0.0	3.1	Ba(A(N3), U(O4))	1.5	1.0	0.0	3.9	
	M06-HF	1.7	1.0	0.1	4.0		3.3	2.8	0.0	10.5	
	MN15	1.6	1.2	0.0	4.7		2.5	1.9	0.1	7.5	
	BMK	1.4	1.1	0.0	4.4		2.3	2.0	0.0	6.4	
	PW6B95	1.1	0.7	0.0	2.9		1.8	1.8	0.0	7.0	Be(A(N3), G(N7), P(1B1NB,2B,2NB))
GH GGA	BH&HLYP	1.3	1.0	0.0	4.6	Ca/Sr/Ba(G(N3))	2.1	1.9	0.0	7.5	Be(C(O2), G(N7), P(2NB))
	SOGGA11-X	1.0	0.7	0.0	2.4		1.8	1.8	0.0	6.5	Be(A(N1), C(N3,O2), G(N7), P(1B1NB,2NB))
	B3PW91	1.1	1.0	0.0	5.2	Sr/Ba(A(N3))	1.7	1.2	0.0	4.0	
	B3PW91-D3	3.6	2.8	0.0	11.1	Ba(A(N3))	4.3	2.1	0.1	8.6	
	PBE0	1.0	0.8	0.0	3.8	Ba(A(N3))	1.6	1.6	0.0	5.9	Be(A(N1,N3), G(N3,N7), P(1B1NB,2B))
	PBE0-D3	1.9	1.3	0.1	6.7	Sr/Ba(A(N3)), Sr(T(O4))	2.8	1.8	0.1	7.3	

	B3LYP	1.3	1.0	0.0	4.2	Be/Mg/Ba(A(N3))	2.2	2.3	0.0	9.2	Be(A/G(N3), G(N7), P(1B1NB,2B), T(O2), T/U(O4)
	B3LYP-D3	4.0	2.1	1.0	10.4	Ba(A/G(N3), T(O4))	5.5	2.2	2.4	12.5	Be(A(N3), P(2B))
	B3LYP-D4	2.7	2.0	0.2	9.8	Ba(A(N3), T/U(O4))	3.7	2.1	0.7	9.8	Be(A(N3))
	X3LYP	1.5	1.1	0.0	4.7	Be/Mg(A(N3))	2.6	2.6	0.0	9.9	Be/Mg(A(N3)), Be(A(N1), G(N3,N7), P(1B1NB,2B), T/U(O2,O4)
	X3LYP-D3	5.1	2.7	1.6	12.5	Ba(A/G(N3), T/U(O4))	7.0	2.3	3.3	13.4	
	O3LYP	1.3	1.2	0.0	5.3	Be/Mg/Ba(A(N3))	2.3	2.6	0.0	10.0	Be(A/G(N3), P(1B1NB, 2B))
	O3LYP-D3	6.5	4.3	1.3	19.9	Ba(A/G(N3), T(O4))	8.3	3.3	2.7	18.6	Ba(P(1B1NB,2B,2NB))
	B97-2	1.1	1.0	0.0	4.5	Mg(U(O2)), Ba(A(N3))	1.6	1.4	0.0	5.1	
	TPSSh	1.4	1.2	0.0	5.8	Sr/Ba(A(N3))	2.2	2.2	0.0	9.6	Be(G(N3,N7), P(1B1NB,2B))
Local meta-GGA	revTPSS	1.4	1.3	0.1	6.6	Ca/Sr/Ba(A(N3)), Ba(G(N3))	2.2	2.2	0.1	10.1	Be(A/G(N3), P(2B))
	TPSS	1.7	1.5	0.0	7.0	Sr(A(N3)), Ba(A/G(N3))	2.7	2.8	0.0	12.5	Be(G(N3), P(1B1NB,2B))
	TPSS-D3	3.7	2.3	0.3	10.8	Ba(A/G(N3))	5.1	2.7	1.1	14.9	Be(A/G(N3), P(2B))
	TPSS-D4	3.0	2.2	0.2	9.4	Ba(A/G(N3), T(O4))	4.0	2.5	0.6	12.6	Be(A/G(N3), P(2B))
	M06-L	4.9	2.8	0.5	13.3	Ca/Ba(A(N3))	6.5	2.7	1.2	13.2	
	M11-L	6.2	4.1	0.3	18.9	Sr/Ba(A(N3))	7.9	3.2	0.8	16.5	Be(A(N3)), Ca/Sr/Ba(P(2NB), Sr(P(1B1NB))
	MN12-L	4.2	2.7	0.1	12.4	Ba(A(N3))	5.4	2.4	0.2	12.2	Mg(P(2B))
	MN15-L	2.2	1.6	0.1	6.5		3.8	3.4	0.1	11.8	
Local GGA	mPW91	1.0	0.8	0.0	4.1	Sr/Ba(A(N3)), Ba(G(N7))	1.6	1.2	0.0	4.7	
	BLYP	1.9	1.6	0.1	8.4	Mg/Ba(A(N3))	3.0	3.0	0.1	14.5	Be(G(N3), P(1B1NB, 2B), T/U(O4), U(O2)), Mg(A(N3))
	BLYP-D3	5.0	3.1	0.7	13.7	Ba(A/G(N3))	6.7	3.1	1.6	18.5	Be(G(N3), P(2B), T(O4)), Mg(A(N3))
	BP86	1.5	1.5	0.0	7.8	Ca/Sr/Ba(A(N3)), Ba(G(N3))	1.9	1.6	0.0	6.3	Be(A/G(N3), P(2B))
	BP86-D3	4.1	3.6	0.1	14.4	Ba(A/G(N3))	4.6	2.6	0.1	10.2	
	PBE	1.7	1.5	0.0	7.3	Mg/Sr/Ba(A(N3)), Mg(T(O2))	2.8	2.9	0.0	11.6	Be/Mg(A(N3)), Be(G(N3), P(2B). T(O2))
	PBE-D3	2.5	2.0	0.0	9.0	Mg/Sr/Ba(A(N3)), Ba(G(N3))	3.8	3.1	0.0	13.4	Be/Mg(A(N3)), Be(G(N3), P(2B))
	PBE-D4	2.2	1.9	0.0	8.3	Mg/Ba(A(N3)), Ba(G(N3), T(O4))	3.2	2.9	0.0	12.7	Be/Mg(A(N3), Be(G(N3). P(2B))
Local LDA	SVWN5	2.6	1.8	0.1	10.3	Mg(A(N3))	4.3	3.5	0.0	13.7	

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.4. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for group I and II metal–nucleic acid complexes, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	Max	Outliers[a]
Double-Hybrid	B2-PLYP	1.1	1.2	0.0	7.9	Na(A,G,T/U(O2))	0.9	0.9	0.0	5.0	Be(A/G(N3),P(2B),T(O4))
	B2-PLYP-D3	2.7	2.7	0.0	13.9	Cs(A),K/Rb(A(N3,N7)),K(G(N3))	1.7	1.2	0.0	6.7	Be(A/G(N3),T(O4)),Be/Ba(P(2B))
	mPW2-PLYP	1.2	1.0	0.0	5.3	Na/K(G(N3)),Li(A(N3),U(O4)),Na(A(N7))	1.0	1.1	0.0	5.7	Be(A,C,G,T,U)
	PBE-QIDH	1.4	1.7	0.0	12.7	Na(A),G(N3),U(O4))	1.1	1.3	0.0	5.6	Mg(A(N1,N7),G(N3),P,T/U))
	PBE0-DH	1.4	1.5	0.0	11.2	Na(A(N1,N7),G(N3)),K(A(N3))	1.0	0.9	0.0	3.8	Mg(T/U(O2),P(2B))
	DSD-PBEP86	2.1	2.0	0.0	9.7	K/Rb/Cs(A),Na/K(G(N3)),Na(A(N7))	1.4	1.0	0.0	4.6	Mg(G(N7),P(2NB),T/U(O2)),Ba(P(2B))
RSH	M11	2.7	2.1	0.0	9.4	Na(G(N3))	2.1	1.5	0.0	5.7	
	MN12-SX	4.1	2.9	0.1	13.2	Ba(A(N3))	3.7	2.9	0.0	12.0	Ba(P(2NB))
	MN12-SX-D3	5.7	4.5	0.1	19.4		4.3	3.4	0.1	14.4	
RSH meta-GGA	ωB97M-V	0.8	0.8	0.0	5.2	Na/K(G(N3)), Na(A(N7))	0.6	0.5	0.0	1.9	
RSH GGA	ωB97	1.8	1.6	0.0	11.6	Na(A(N1,N7)),Na/Mg(G(N3))	2.0	2.3	0.0	9.0	Be(A(N1,N7),C,G,T,U)),Be/Mg(P),Mg(C(N3))
	ωB97X	2.4	2.3	0.0	15.7	Li/Na(G(N3)),Na(A(N1,N7))	2.8	3.6	0.0	12.9	Be,Mg((A,G,C,T,P),U(O2))
	ωB97X-D	5.4	6.5	0.1	47.2	Na/K(G(N3),Rb(A(N3,N7),Cs((A,G(N7),C,T/U(O4),T(O2)))	3.9	3.0	0.1	11.3	
	ωB97X-D3	6.9	6.2	0.1	33.2	K(A(N1,N7),G(N3)),Rb/Cs(A)	4.5	2.6	0.1	14.4	Sr(P(2B)),Ba(P)
	ωB97X-V	0.9	1.0	0.0	7.3	Na/K(G(N3)),Na(A(N1,N7))	0.7	0.6	0.0	2.9	Mg(P(2B),T/U(O2))
	ωB97X-D4	5.3	5.7	0.0	29.1	K(G(N3)),Rb(A(N7)),Cs(A)	3.0	2.1	0.0	10.0	Ba(P,T(O4))
	HSE06	1.4	1.4	0.0	8.2	Na(A(N1,N7)),Na/K(G(N3)),Cs(A(N3))	1.1	1.2	0.0	5.6	Be(A(N1,N3),C(O2),G,P))
	HSE06-D3	3.3	2.8	0.1	15.2	K(A(N3),G(N3))Rb(A(N3,N7)),Cs(A,T(O4))	2.2	1.5	0.0	7.2	Be(A/G(N3),G(N7),P(1B1NB,2B))
	LC-PBE	2.1	2.0	0.0	10.8	Na/K(G(N3)),Li(A(N3),G(N7))	1.5	1.4	0.0	8.5	Li/Be/Mg(G(7)),Li/Be(C(O2)),Be(A(N7),P(2NB))
	LC-ωPBE	1.5	1.7	0.0	12.5	Na(A(N1,N7),G(N3))	1.1	1.1	0.0	5.5	Be(A/G(N7),C,P(1B1NB,2NB))
	LC-ωPBE-D3	3.2	3.1	0.0	14.9	K(A/G(N3)),Rb(A(N3,N7)),Cs(A,T(O4))	2.0	1.3	0.0	6.9	Be(A/G(N7),C(O2),P(1B1NB,2NB))
	CAM-B3LYP	2.0	1.6	0.0	7.6	Li(A(N3),T/U(O4)),K(G(N3))	1.8	2.0	0.0	8.3	Be(A(N3),C(O2),G(N7),P,T/U(O4))

	CAM-B3LYP-D3	4.2	2.9	0.1	14.5	Ba(T/U(O4))	3.1	2.2	0.1	9.7	Be(A(N3), G(N7), P(1B1NB,2B), T/U(O4))
GH meta-GGA	M06	3.4	2.8	0.0	18.9	Na(A(N1,N7),G(N3)),Sr/Ba(A(N3))	3.2	2.6	0.1	9.4	
	M06-2X	1.2	0.9	0.0	4.8	Li/Na/Rb/Cs/Ba(A(N3)),Li(U(O4)),Na(G(N3))	1.1	0.9	0.0	3.9	Ba(P(2NB))
	M06-HF	2.5	1.9	0.1	9.5	Li(A,C,G(N3),U(O2))	2.5	2.4	0.0	10.5	Be(A/G(N7),P(2NB))
	MN15	1.9	2.0	0.0	13.7	Na/K(G(N3),Na(A(N1,N7))	1.7	1.7	0.0	7.5	Mg(A(N1,N7),C/G(N3),P)
	BMK	2.3	2.5	0.0	17.8	Li(A(N1,N7)),Li/Na(G(N3))	1.8	1.7	0.0	6.4	Mg(A(N7),G(N3),P(2B))
	PW6B95	1.5	1.4	0.0	7.3	Na/K(A(N7),G(N3)),Na(A(N1))	1.3	1.5	0.0	7.0	Be(A,C,G,P,T/U(O4))
GH GGA	BH&HLYP	2.3	2.0	0.0	9.3	Na/K(G(N3)),Li(A(N3),T,U)	1.8	1.6	0.0	7.5	Be(A(N1,N3),C(O2),G(N7),P(1B1NB,2NB),T/U(O4))
	SOGGA11-X	1.5	1.4	0.0	11.2	Na(A(N1,N7),G(N3))	1.3	1.4	0.0	6.5	Be(A,C,G,P)
	B3PW91	1.8	2.2	0.0	14.9	Na(A(N1,N7),G(N3),T/U(O2)),Rb(A(N3))	1.3	1.2	0.0	4.4	
	B3PW91-D3	5.2	5.1	0.0	24.1	K/Rb(A(N3),Rb(A(N7)),Cs(A)	3.3	2.0	0.1	8.6	
	PBE0	1.4	1.4	0.0	8.4	Li(A(N3)),Na(A(N1,N)),Na/K(G(N3))	1.2	1.3	0.0	5.9	Be(A,C,G,P)
	PBE0-D3	2.9	2.4	0.1	12.7	K(A(N3),G(N3)),Rb(A(N3,N7),T(O4)),Cs(A,T(O4))	2.2	1.6	0.0	7.3	Be(A(N1,N7),C(O2),G,P)
	B3LYP	1.6	1.4	0.0	7.0	Na(A(N1,N7),G(N3))	1.5	1.9	0.0	9.2	Be(A(N1,N3),G(N3,N7),P(1B1NB,2B),T/U)
	B3LYP-D3	5.6	4.1	1.0	20.7	K(G(N3)),Rb(A(N3,N7)),Cs(A)	4.1	2.3	0.5	12.5	Be(A(N3),G,P(1B1NB,2B),T,U(O4))
	B3LYP-D4	3.9	3.1	0.2	16.0	Rb/Cs(A(N3)),Rb(A(N7)),Cs(T(O4))	2.7	1.9	0.4	9.8	Be(A(N3),G(N3),P(2B),T(O2)),Be/Ba(T/U(O4))
	X3LYP	1.8	1.5	0.0	6.3		1.8	2.2	0.0	9.9	Be,Mg(A(N3))
	X3LYP-D3	7.9	6.1	1.6	36.8	K(A(N3,N7),G(N3)),Rb(A),Cs(A,T(O4))	5.5	2.5	1.8	13.4	Be(A(N3),P(2B))
	O3LYP	1.8	2.0	0.0	11.6	Na(A(N1,N7),G(N3),T/U(O2))	1.6	2.1	0.0	10.0	Be/Sr/Ba(A(N3)),Na/K/Cs(A(N1,N7)),Na/K(G(N3)),Na(G(N7))
	O3LYP-D3	9.5	9.8	0.1	59.0	Rb/Cs(A),K(G(N3))	6.4	3.7	0.1	18.6	Ba(P)
	B97-2	1.8	2.2	0.0	15.4	Na(A(N1,N7),G(N3),T/U(O2))	1.3	1.3	0.0	5.1	Mg(A(N7),P,U(O2)),Na(A(N7))
	TPSSh	1.6	1.5	0.0	7.9	Na(A(N1,N7),G(N3)),Rb/Cs/Bs(A(N3))	1.5	1.8	0.0	9.6	Be(A,C,G,P,T/U(O4))
Local meta-GGA	revTPSS	1.7	1.6	0.0	7.8	Rb/Cs/Sr/Ba(A(N3)), Na(A(N1,N7),G(N3),T/U(O2))	1.5	1.8	0.0	10.1	Be(A,C(N3),G,P(1B1NB,2B))
	TPSS	1.8	1.6	0.0	8.0	Rb/Cs/Sr/Ba(A(N3)),Na(A(N1,N7),G(N3),U(O2))	1.8	2.3	0.0	12.5	Be(A,C/G(N3),G(N7),P(1B1NB,2B,T,U(O4))
	TPSS-D3	4.1	3.3	0.0	17.8	K(A/G(N3)),Rb(A(N3,N7)),Cs(A),Ba(A/G(N3))	3.4	2.7	0.0	14.9	Be(A/G(N3),P(1BNB,2B))
	TPSS-D4	3.4	2.8	0.0	15.1	K/Rb/Cs/Ba(A(N3)),K/Ba(G(N3)),Cs/Ba(T(O4))	2.7	2.4	0.0	12.6	Be(A/G(N3),P(2B)),Mg(A(N3))
	M06-L	4.3	3.3	0.1	18.0	Cs/Ca/Sr/Ba(A(N3)),Na(A(N1,N7),G(N3))	4.2	3.3	0.0	13.2	
	M11-L	5.4	3.8	0.3	18.9	Cs/Ca/Sr/Ba(A(N3)),Na/Ca/Sr/Ba(G(N3)),Sr/Ba(T(O4)),Na(A(N7)	5.0	3.9	0.1	16.5	
	MN12-L	4.1	2.9	0.1	16.5	Na/Ba(G(N3)),Na(A(N7))	3.5	2.7	0.1	12.2	Mg(P(2B))
	MN15-L	2.5	2.1	0.0	12.5	Na(A(N3,N7),G(N3),T/U(O2))	2.5	2.9	0.0	11.8	Be(P,T,U),Mg(A(N7),G,C,P,T/U(O2))
Local GGA	mPW91	1.5	1.7	0.0	11.6	Na(A(N1,N3),G(N3),U(O2)),K(G(N3)),Rb/Cs(A(N3))	1.2	1.1	0.0	4.7	Be(A(N3),G,P(1B1NB,2B))

	BLYP	2.0	1.7	0.0	8.4	Mg/Ba(A(N3)),Na(N1,N7))	2.0	2.5	0.0	14.5	Be(A(N1,N7),C(N3),G,P(1B1NB,2B),T(O4),U),Mg(A(N3))
	BLYP-D3	6.0	4.8	0.3	26.3	K(A/G(N3)),Rb/Cs(A)	4.6	3.2	0.1	18.5	Be(G(N3), P(2B), T(O4),U(O2)),Mg(A(N3))
	BP86	2.4	2.6	0.0	13.3	Na(A(N1,N7),C,G(N3),T/U(O2)),Ba(A(N3))	1.7	1.5	0.0	6.3	Be(A (N3), P(2B)),Mg(P(2NB))
	BP86-D3	5.4	5.6	0.1	27.5	Na(U(O4)),K(A/G(N3)),Rb/Cs(A)	3.3	2.4	0.1	10.2	Ba(P(2B))
	PBE	2.0	1.7	0.0	7.3	Na/K(A(N7)),Mg/Ba(A(N3))	1.9	2.3	0.0	11.6	Be(A,G,P(1B1NB,2B),T,U),Mg(A(N3))
	PBE-D3	2.5	2.3	0.0	11.9	K(G(N3)),K/Rb/Mg/Sr/Ba(A(N3)),Ba(G(N3),T(O4))	2.4	2.7	0.0	13.4	Be(A(N1,N3),G,P(1B1NB,2B),T(O4)),Mg(A(N3))
	PBE-D4	2.2	2.2	0.0	11.9	K/Rb/Cs/Mg/Ba(A(N3)),K(G(N3)),Ba(T(O4))	2.0	2.5	0.0	12.7	Be (A(N1,N3),G,P(1B1NB,2B),T),Mg(A(N3))
Local LDA	SVWN5	3.2	2.5	0.0	10.5	Li/Na/Mg(A(N3)),K(A(N7))	3.0	3.0	0.0	13.7	Be(A(N3),G,P(2B)),Mg(A(N3),G(N7))

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.5. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the double-hybrid family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Metal	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Be^{2+}	B2-PLYP	0.8	0.4	0.1	1.6		2.3	1.2	0.4	5.0	P(2B)
	B2-PLYP-D3	1.3	0.5	0.3	2.3		3.6	1.4	1.5	6.7	P(2B)
	mPW2-PLYP	1.3	0.4	0.5	2.1		3.6	0.9	2.4	5.7	
	PBE-QIDH	0.5	0.4	0.1	1.5	U(O2)	1.3	1.0	0.2	3.3	U(O2)
	PBE0-DH	0.6	0.3	0.0	1.1		1.7	0.9	0.1	3.0	
	DSD-PBEP86	0.5	0.3	0.0	1.1		1.5	0.9	0.1	3.0	
Mg^{2+}	B2-PLYP	0.8	0.4	0.2	1.7	T(O2)	1.5	0.7	0.7	3.1	P(2NB)
	B2-PLYP-D3	0.5	0.4	0.1	1.9	A(N3)	0.8	0.6	0.3	2.5	A(N3), P(2B)
	mPW2-PLYP	0.3	0.4	0.0	1.5	A(N3), U(O2)	0.5	0.6	0.0	2.0	A(N3)
	PBE-QIDH	2.4	0.9	1.1	3.9		4.1	0.8	3.1	5.6	
	PBE0-DH	1.3	0.8	0.2	2.8		2.2	1.1	0.3	3.8	
	DSD-PBEP86	1.8	0.6	0.8	3.0		3.2	0.8	1.7	4.6	
Ca^{2+}	B2-PLYP	0.9	0.5	0.1	1.8		1.1	0.6	0.1	2.2	A(N3)
	B2-PLYP-D3	1.0	0.7	0.1	2.3		1.1	0.6	0.2	2.1	
	mPW2-PLYP	1.1	0.6	0.4	2.2		1.3	0.5	0.4	2.1	A(N3), U(O4)
	PBE-QIDH	0.9	0.9	0.2	2.6	T/U(O2)	0.9	0.6	0.2	2.2	T/U(O2)
	PBE0-DH	0.8	0.8	0.0	2.4		0.8	0.6	0.1	1.9	
	DSD-PBEP86	0.7	0.5	0.0	1.7		0.8	0.5	0.0	1.8	

Sr^{2+}	B2-PLYP	0.6	0.4	0.1	1.2		0.6	0.3	0.3	1.4	
	B2-PLYP-D3	1.9	1.0	0.5	3.7		1.9	0.6	1.0	2.9	
	mPW2-PLYP	0.8	0.6	0.2	1.8		0.8	0.4	0.2	1.4	
	PBE-QIDH	0.8	0.8	0.0	2.2		0.6	0.6	0.0	1.6	
	PBE0-DH	0.9	0.7	0.0	2.4		0.7	0.5	0.1	1.5	
	DSD-PBEP86	1.2	0.9	0.2	2.9		1.2	0.7	0.1	2.5	
Ba^{2+}	B2-PLYP	0.6	0.6	0.1	2.0	A(N3)	0.6	0.5	0.1	1.2	
	B2-PLYP-D3	3.4	1.5	1.2	5.5		3.1	0.8	2.1	4.5	
	mPW2-PLYP	0.6	0.4	0.1	1.3		0.6	0.3	0.2	1.0	
	PBE-QIDH	0.8	0.5	0.1	1.6		0.7	0.4	0.0	1.2	
	PBE0-DH	1.1	0.6	0.1	2.1		1.1	0.6	0.0	2.1	
	DSD-PBEP86	2.3	1.1	1.0	4.4		2.2	0.8	1.0	3.9	P(2B)

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.6. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the RSH family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Be^{2+}	M11	0.5	0.3	0.1	1.2	G(N3)	1.6	1.1	0.1	3.1	
	MN12-SX	1.2	0.5	0.1	2.0		3.9	2.3	0.2	9.4	P(1B1NB,2B)
	MN12-SX-D3	1.0	0.4	0.2	1.9		3.4	2.2	0.5	8.7	P(1B1NB,2B)
	ωB97M-V	0.3	0.2	0.1	0.6		1.0	0.5	0.2	1.7	
	ωB97	2.4	0.6	1.5	3.6	G(N3)	6.9	1.0	5.0	9.0	G(N3)
	ωB97X	3.8	0.8	2.4	5.1		10.9	1.4	8.2	12.9	
	ωB97X-D	3.4	0.7	2.2	4.5		9.8	1.3	6.9	11.3	A(N3)
	ωB97X-D3	1.3	0.4	0.6	2.1		3.8	1.3	1.4	5.5	
	ωB97X-V	0.3	0.2	0.0	0.7		0.8	0.5	0.0	1.5	
	ωB97X-D4	0.5	0.3	0.1	1.2	A(N3)	1.4	0.7	0.3	2.8	G/A(N3)
	HSE06	1.3	0.5	0.6	2.5		3.8	1.2	1.4	5.6	
	HSE06-D3	1.7	0.6	1.0	3.1	A(N3)	5.0	1.3	2.5	7.2	U(O2)
	LC-PBE	1.1	0.8	0.0	2.7		3.3	2.6	0.0	8.5	

	LC-ωPBE	0.9	0.5	0.0	1.7		2.8	1.6	0.0	5.5	
	LC-ωPBE-D3	1.5	0.5	0.6	2.2	A(N3)	4.4	1.6	1.4	6.9	
	CAM-B3LYP	2.3	0.5	1.2	3.0	P(2NB)	6.6	0.9	5.3	8.3	
	CAM-B3LYP-D3	2.8	0.7	1.5	3.8	P(2NB)	7.9	1.0	6.5	9.7	
	M11	1.5	1.0	0.3	3.5		2.8	1.6	0.5	5.0	
Mg^{2+}	MN12-SX	2.7	0.8	1.7	4.2		5.1	2.1	2.5	9.7	
	MN12-SX-D3	2.1	0.8	1.0	3.5		4.1	2.0	1.4	8.4	
	ωB97M-V	0.5	0.5	0.0	1.5	G(N3), T/U(O2)	0.8	0.6	0.1	1.9	
	ωB97	3.1	1.0	1.7	5.3		5.4	1.1	3.8	7.5	
	ωB97X	4.3	1.3	2.5	6.9		7.7	1.4	5.4	9.8	
	ωB97X-D	4.5	1.2	2.6	6.5		8.0	1.4	5.6	9.9	
	ωB97X-D3	1.2	0.6	0.1	2.1	T/U(O2)	2.4	1.3	0.1	3.9	
	ωB97X-V	0.9	0.6	0.2	2.2	T/U(O2)	1.6	0.7	0.4	2.9	
	ωB97X-D4	0.4	0.5	0.0	1.9	G(N3)	0.7	0.7	0.0	2.6	G(N3)
	HSE06	0.5	0.7	0.0	2.3	A(N3)	0.9	0.9	0.0	3.1	A(N3)
	HSE06-D3	1.0	0.8	0.1	3.5	A(N3)	1.7	1.1	0.1	4.6	T(O2), A(N3), G(N7)
	LC-PBE	0.8	0.7	0.2	2.4	G(N3,N7), C(O2)	1.5	1.4	0.3	5.0	G(N7)
	LC-ωPBE	1.1	0.7	0.0	2.3		1.8	0.9	0.1	3.1	
	LC-ωPBE-D3	0.5	0.4	0.0	1.2		0.9	0.6	0.0	2.5	G(N7)
	CAM-B3LYP	1.6	0.7	0.7	3.0		2.9	1.0	1.6	5.1	G(N7)
	CAM-B3LYP-D3	2.7	0.9	1.3	4.3	P(2NB)	4.8	1.0	3.3	6.9	G(N7)
	M11	2.6	1.2	1.1	4.5		3.2	1.4	1.0	5.0	
Ca^{2+}	MN12-SX	5.4	2.0	2.4	9.4	A(N3)	6.2	1.5	4.3	9.3	
	MN12-SX-D3	6.6	2.3	3.1	10.8	A(N3), P(2B)	7.6	1.6	5.7	10.8	
	ωB97M-V	0.7	0.7	0.1	2.3		0.8	0.5	0.1	1.8	
	ωB97	1.5	0.9	0.5	3.8	G(N3)	1.6	0.6	0.7	2.9	
	ωB97X	1.0	0.9	0.2	3.2	G(N3)	1.1	0.7	0.2	2.5	
	ωB97X-D	1.3	0.8	0.1	2.8		1.5	0.7	0.1	2.5	T/U(O2)
	ωB97X-D3	4.5	1.1	2.7	6.1		5.5	2.0	3.2	10.2	P(2B,1B1NB)
	ωB97X-V	1.0	0.8	0.3	2.7	G(N3)	1.1	0.5	0.4	2.1	
	ωB97X-D4	3.4	1.2	1.6	5.5		3.9	1.0	2.4	5.9	
	HSE06	0.8	0.7	0.0	2.1		0.8	0.6	0.0	1.7	
	HSE06-D3	1.4	1.1	0.4	4.0	A(N3)	1.6	0.9	0.3	3.1	

	LC-PBE	1.7	1.4	0.3	5.3	G(N3)	1.8	0.9	0.7	4.1	
	LC-ωPBE	1.1	1.0	0.2	3.2		1.1	0.7	0.3	2.5	
	LC-ωPBE-D3	1.2	0.7	0.2	2.3		1.5	0.8	0.1	2.5	
	CAM-B3LYP	1.1	1.0	0.0	3.6	G(N3)	1.2	0.8	0.0	2.8	
	CAM-B3LYP-D3	1.1	0.9	0.1	3.2	T/U(O4)	1.3	0.8	0.1	2.9	T(O4)
	M11	3.2	1.4	1.6	5.3		3.4	1.4	1.3	5.4	
Sr^{2+}	MN12-SX	6.8	2.4	3.0	11.8	A(N3)	7.0	1.8	4.8	10.8	
	MN12-SX-D3	8.4	2.9	3.8	13.8	P(2B,1B1NB)	8.6	1.9	6.6	12.7	
	ωB97M-V	0.5	0.6	0.0	2.0	G(N3)	0.4	0.4	0.0	1.3	
	ωB97	1.2	1.0	0.1	3.8	G(N3)	1.2	0.7	0.1	2.4	
	ωB97X	0.9	0.9	0.0	3.3	G(N3)	0.8	0.6	0.0	2.1	
	ωB97X-D	2.1	1.1	0.6	4.1		2.1	0.9	0.4	3.7	T/U(O2), G(N7)
	ωB97X-D3	6.3	1.6	3.5	8.2		6.8	2.3	4.2	12.2	P(2B,1B1NB,2NB)
	ωB97X-V	0.8	0.8	0.0	2.8	G(N3)	0.7	0.5	0.1	1.8	
	ωB97X-D4	4.2	1.5	1.9	7.0		4.4	1.1	2.7	6.4	
	HSE06	1.0	0.7	0.1	2.8	A(N3)	0.9	0.5	0.3	1.9	A(N3), G(N7)
	HSE06-D3	2.4	1.4	1.0	5.5	A(N3)	2.4	1.0	1.1	3.7	
	LC-PBE	1.7	1.5	0.2	5.9	G(N3)	1.6	0.9	0.4	3.8	G(N3)
	LC-ωPBE	1.1	1.0	0.0	3.6		1.0	0.6	0.0	2.3	
	LC-ωPBE-D3	1.7	0.9	0.3	3.2		1.8	0.9	0.2	3.2	
	CAM-B3LYP	1.0	0.9	0.1	3.5	G(N3)	0.9	0.6	0.1	2.2	G(N3)
	CAM-B3LYP-D3	2.0	1.3	0.3	4.5		2.0	1.0	0.2	3.7	T(O4)
	M11	3.8	1.5	1.8	6.0		3.7	1.5	1.4	5.7	
Ba^{2+}	MN12-SX	7.8	2.7	3.6	13.2		7.3	2.2	4.8	12.0	
	MN12-SX-D3	10.1	3.5	4.6	16.0		9.4	2.4	6.7	14.4	
	ωB97M-V	0.7	0.6	0.0	1.7		0.7	0.5	0.0	1.3	
	ωB97	0.8	0.9	0.0	3.3	G(N3)	0.6	0.5	0.0	1.7	
	ωB97X	0.8	0.7	0.1	2.8	G(N3)	0.6	0.4	0.1	1.5	G(N3,N7)
	ωB97X-D	5.7	2.6	2.1	9.2		5.0	1.3	2.5	6.9	T/U(O2)
	ωB97X-D3	8.6	2.5	4.5	12.1		8.3	2.7	5.7	14.4	P(2B,1B1NB,2NB)
	ωB97X-V	0.7	0.6	0.1	2.2	G(N3)	0.5	0.3	0.2	1.2	A(N3)
	ωB97X-D4	7.6	2.8	3.3	11.8		6.9	1.6	4.7	10.0	
	HSE06	1.3	1.2	0.0	4.0		1.3	1.0	0.0	2.7	

	HSE06-D3	3.9	1.7	1.9	7.4	A(N3), T(O4)	3.6	1.2	2.1	5.0	
	LC-PBE	1.5	1.7	0.0	5.8	G(N3)	1.2	1.0	0.0	3.1	
	LC-ωPBE	1.0	0.9	0.1	3.3		0.8	0.6	0.0	1.7	
	LC-ωPBE-D3	2.9	1.2	1.3	5.0		2.7	1.0	1.0	4.1	
	CAM-B3LYP	1.1	0.8	0.1	2.8		1.0	0.5	0.1	2.0	
	CAM-B3LYP-D3	3.5	1.7	1.4	6.6		3.2	1.1	1.2	4.7	T(O4)

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.7. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the GH meta-GGA family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[b]	MUE	SD	Min	MAX	Outliers[b]
Be^{2+}	M06	1.4	0.5	0.0	2.0	A(N3)	4.1	1.7	0.1	6.0	
	M06-2X	0.8	0.3	0.3	1.3		2.4	0.6	1.4	3.3	
	M06-HF	2.5	0.6	1.1	3.5	P(1B1NB,2B), A(N7)	7.3	1.7	5.1	10.5	
	MN15	1.0	0.4	0.2	1.5		3.0	1.1	0.5	5.0	
	BMK	1.3	0.5	0.6	2.5	T/U(O2)	3.6	1.0	2.3	5.6	
	PW6B95	1.7	0.4	1.0	2.4		4.9	1.3	2.6	7.0	
Mg^{2+}	M06	4.0	1.3	2.1	5.6		7.3	1.9	2.9	9.4	A(N3)
	M06-2X	0.3	0.2	0.0	0.7		0.5	0.4	0.0	1.4	
	M06-HF	2.7	0.8	1.1	4.0	P(2B)	4.9	1.5	3.0	7.7	
	MN15	3.1	1.0	1.7	4.7		5.5	1.3	3.7	7.5	
	BMK	2.9	1.0	1.4	4.4		5.1	0.9	3.7	6.4	
	PW6B95	0.5	0.5	0.0	2.0	A(N3)	0.9	0.7	0.0	2.7	A(N3), G(N7)
Ca^{2+}	M06	3.6	1.7	1.4	8.3	A(N3)	4.6	2.6	1.4	8.9	
	M06-2X	0.9	0.5	0.2	1.7		1.2	0.8	0.1	2.9	
	M06-HF	0.9	0.4	0.3	1.7		1.2	0.8	0.3	3.6	A(N1), P(2NB,1B1NB), C(O2)
	MN15	1.6	1.0	0.6	3.8		1.8	0.7	0.8	2.9	
	BMK	0.6	0.4	0.0	1.4		0.5	0.3	0.1	1.1	
	PW6B95	0.9	0.7	0.0	2.1		0.8	0.6	0.0	2.0	

Sr^{2+}	M06	3.8	1.9	1.7	9.0	A(N3)	4.3	2.3	1.4	8.2	
	M06-2X	1.2	0.6	0.5	2.5		1.4	0.8	0.3	3.1	
	M06-HF	1.0	0.4	0.4	1.9		1.2	0.7	0.3	2.9	P(2NB)
	MN15	1.4	1.1	0.0	3.8		1.3	0.8	0.1	2.4	
	BMK	0.8	0.6	0.1	1.7		0.7	0.5	0.1	1.7	
	PW6B95	1.0	0.6	0.2	1.9		1.0	0.4	0.2	1.6	C/T(O2)
Ba^{2+}	M06	4.2	1.9	2.2	9.2	A(N3)	4.3	2.3	1.6	8.1	
	M06-2X	1.8	0.7	1.1	3.1		1.9	1.0	0.6	3.9	
	M06-HF	1.5	0.8	0.1	2.8		1.6	1.0	0.0	3.7	P(2NB)
	MN15	0.9	0.9	0.0	3.0	G(N3)	0.6	0.5	0.1	1.6	
	BMK	1.3	0.9	0.1	2.7		1.4	0.9	0.0	2.7	
	PW6B95	1.3	0.8	0.1	2.9		1.4	0.9	0.0	2.7	

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.8. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the GH GGA family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Be^{2+}	BH&HLYP	1.8	0.5	0.8	2.5		5.2	1.5	2.0	7.5	
	SOGGA11-X	1.6	0.5	0.9	2.4		4.8	1.3	2.0	6.5	
	B3PW91	0.6	0.4	0.1	1.7	A(N7)	1.9	1.1	0.3	3.9	
	B3PW91-D3	1.6	0.6	0.8	3.2	A(N3)	4.5	1.4	2.1	7.3	
	PBE0	1.4	0.5	0.6	2.4		4.2	1.3	1.3	5.9	U(O2)
	PBE0-D3	1.9	0.6	1.1	3.2		5.5	1.3	2.6	7.3	U(O2)
	B3LYP	2.2	0.8	0.9	3.7		6.2	1.5	4.3	9.2	
	B3LYP-D3	3.1	1.0	1.4	5.1		8.7	1.8	6.2	12.5	P(2B)
	B3LYP-D4	2.3	0.9	0.9	4.3		6.5	1.7	4.4	9.8	
	X3LYP	2.5	0.8	1.1	4.1		7.2	1.4	5.4	9.9	
	X3LYP-D3	3.4	1.0	1.7	5.5		9.8	1.7	7.5	13.4	
	O3LYP	2.4	0.8	1.2	4.3	A(N3)	6.8	1.7	4.8	10.0	
	O3LYP-D3	3.4	1.0	1.9	5.9	A(N3)	9.8	2.1	7.3	14.1	
	B97-2	0.3	0.3	0.1	1.2	A(N3)	1.0	0.7	0.2	2.7	A(N3)

	TPSSh	2.0	0.7	0.9	3.4	G/A(N3)	5.8	1.9	3.0	9.6	
Mg^{2+}	BH&HLYP	0.9	0.6	0.2	1.9		1.6	1.1	0.3	3.8	
	SOGGA11-X	0.9	0.6	0.1	2.0		1.6	0.9	0.2	3.2	
	B3PW91	1.5	0.7	0.5	3.1	U(O2)	2.7	1.0	0.7	4.0	
	B3PW91-D3	0.7	0.7	0.0	2.8	A(N3)	1.1	1.0	0.1	3.7	A(N3)
	PBE0	0.6	0.7	0.0	2.1	A(N3), U(O2)	0.9	0.9	0.0	2.8	
	PBE0-D3	1.0	0.8	0.1	3.2	A(N3)	1.7	1.0	0.1	4.3	A(N3), G(N7), T(O2)
	B3LYP	1.0	1.0	0.0	4.0	A(N3)	1.6	1.3	0.0	5.3	A(N3)
	B3LYP-D3	2.9	1.2	1.0	6.1	A(N3)	5.0	1.3	3.3	8.2	A(N3)
	B3LYP-D4	1.7	1.2	0.2	5.3	A(N3)	2.8	1.5	0.9	7.0	A(N3)
	X3LYP	1.6	1.0	0.4	4.7	A(N3)	2.7	1.2	1.4	6.3	A(N3)
	X3LYP-D3	3.8	1.4	1.6	7.2	A(N3)	6.7	1.4	4.5	9.6	A(N3), P(2B)
	O3LYP	0.7	1.1	0.0	4.4	A(N3)	1.1	1.5	0.1	5.9	A(N3)
	O3LYP-D3	2.9	1.6	1.3	7.6	A(N3)	5.1	1.9	2.7	10.1	A(N3)
	B97-2	2.1	0.9	0.2	3.7		3.8	1.2	0.3	5.1	A(N7)
	TPSSh	0.6	0.8	0.0	3.0	A(N3)	1.0	1.0	0.0	4.0	A(N3)
Ca^{2+}	BH&HLYP	1.1	1.1	0.0	4.6	G(N3)	1.1	0.8	0.1	3.5	G(N3)
	SOGGA11-X	0.6	0.6	0.0	2.1	G(N3), U(O2)	0.5	0.4	0.0	1.6	G(N3), U(O2)
	B3PW91	0.7	0.8	0.0	2.9	A(N3)	0.7	0.6	0.0	2.3	A(N3)
	B3PW91-D3	3.9	1.6	1.6	7.0		4.4	1.0	2.7	5.9	
	PBE0	0.8	0.8	0.0	2.2		0.8	0.6	0.0	1.8	
	PBE0-D3	1.3	1.0	0.2	3.7	A(N3)	1.5	0.9	0.1	2.9	
	B3LYP	0.8	0.5	0.1	1.6		0.8	0.4	0.2	1.6	
	B3LYP-D3	3.2	1.5	1.0	5.9		3.5	0.9	2.4	5.4	
	B3LYP-D4	2.0	1.5	0.4	5.1		2.1	1.2	0.7	4.7	T(O4)
	X3LYP	0.9	0.6	0.0	1.9		0.9	0.5	0.0	1.9	
	X3LYP-D3	3.9	1.5	1.6	6.6		4.5	1.0	3.3	6.5	
	O3LYP	0.9	0.9	0.0	2.7		0.8	0.7	0.0	2.1	
	O3LYP-D3	5.6	2.0	3.0	9.6		6.6	2.0	3.8	10.9	P(2B)
	B97-2	0.7	0.6	0.1	2.5	A(N3)	0.7	0.5	0.1	1.9	A(N3)
	TPSSh	0.8	0.8	0.0	3.4	A(N3)	0.9	0.7	0.1	2.7	A(N3)
Sr^{2+}	BH&HLYP	1.1	1.2	0.0	4.6	G(N3)	1.0	0.9	0.0	3.0	
	SOGGA11-X	0.8	0.5	0.1	1.7	G(N3), U(O2)	0.7	0.4	0.1	1.4	
	B3PW91	1.0	1.0	0.1	4.0	A(N3)	1.1	0.8	0.1	2.8	
	B3PW91-D3	5.0	2.0	2.1	8.8		5.1	1.1	3.4	6.7	

	PBE0	0.9	0.7	0.1	2.7		0.8	0.5	0.2	1.8	
	PBE0-D3	2.1	1.3	1.0	5.0	A(N3)	2.2	1.0	0.8	3.4	
	B3LYP	0.8	0.9	0.0	2.8		0.7	0.6	0.0	1.9	
	B3LYP-D3	4.4	1.9	1.6	7.6		4.4	1.0	3.3	6.2	
	B3LYP-D4	2.7	1.8	0.7	6.4		2.6	1.2	0.9	5.2	
	X3LYP	0.9	0.8	0.0	2.3		0.7	0.6	0.0	1.6	
	X3LYP-D3	5.9	2.1	2.5	9.0		6.0	1.3	4.6	8.4	
	O3LYP	0.9	1.0	0.2	4.0		0.9	0.6	0.2	2.7	A(N3)
	O3LYP-D3	8.5	2.9	4.5	13.6		8.9	2.5	5.6	14.1	P(2B,1B1NB,2NB)
	B97-2	0.9	0.9	0.1	3.5	A(N3)	1.0	0.7	0.1	2.4	A(N3)
	TPSSh	1.2	1.2	0.0	4.6	A(N3)	1.3	0.9	0.0	3.1	A(N3)
	BH&HLYP	1.6	1.2	0.1	4.0		1.5	1.1	0.1	3.1	
	SOGGA11-X	1.3	0.8	0.1	2.3		1.3	0.8	0.1	2.6	
	B3PW91	1.9	1.3	0.3	5.2		1.9	1.2	0.2	3.3	
	B3PW91-D3	6.9	2.7	3.0	11.1		6.3	1.4	4.5	8.6	
	PBE0	1.2	1.1	0.1	3.8		1.3	0.9	0.1	2.5	
	PBE0-D3	3.4	1.6	1.7	6.7	A(N3)	3.3	1.2	1.7	4.7	
	B3LYP	1.6	1.2	0.0	4.2		1.6	1.0	0.0	2.8	
Ba^{2+}	B3LYP-D3	6.5	2.6	2.6	10.4		5.8	1.3	4.3	7.9	
	B3LYP-D4	5.0	2.5	1.8	9.8		4.4	1.3	2.5	7.0	
	X3LYP	1.4	1.2	0.0	3.6		1.4	0.9	0.0	2.4	
	X3LYP-D3	8.6	3.1	3.8	12.5		7.9	1.8	6.1	11.4	
	O3LYP	1.5	1.4	0.2	5.3	A(N3)	1.6	1.2	0.1	3.2	
	O3LYP-D3	12.4	4.4	6.4	19.9		11.6	3.5	7.7	18.6	P(2B,1B1NB,2NB)
	B97-2	1.5	1.3	0.0	4.5		1.7	1.2	0.0	3.4	
	TPSSh	2.2	1.4	0.7	5.8		2.2	1.2	0.5	4.0	

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.9. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the meta-GGA family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Be^{2+}	revTPSS	1.9	0.9	0.6	3.8	G/A(N3)	5.5	2.4	2.7	10.1	
	TPSS	2.5	1.0	0.8	4.4	P(2NB), G(N3)	7.1	2.5	4.2	12.5	P(2B)
	TPSS-D3	3.1	1.1	1.2	5.4	P(2NB), G/A(N3)	9.0	2.7	5.9	14.9	P(2B)
	TPSS-D4	2.6	1.0	0.8	4.8	P(2NB), G/A(N3)	7.4	2.5	4.3	12.6	P(2B)
	M06-L	1.6	0.7	0.5	2.6		4.5	1.7	1.2	7.2	
	M11-L	2.1	0.8	0.3	3.5		6.2	2.4	0.8	9.7	
	MN12-L	0.9	0.6	0.1	2.0		2.9	2.6	0.2	9.4	P(2B)
	MN15-L	2.5	0.8	1.6	4.1		7.3	2.1	4.2	11.4	
Mg^{2+}	revTPSS	0.8	1.0	0.1	3.7	A(N3)	1.4	1.3	0.2	4.9	A(N3)
	TPSS	1.0	1.3	0.0	4.9	A(N3)	1.6	1.8	0.0	6.5	A(N3)
	TPSS-D3	2.3	1.6	0.3	6.5	A/G(N3)	3.8	2.1	1.1	8.7	A(N3), P(2B)
	TPSS-D4	1.5	1.5	0.2	6.0	A(N3)	2.5	2.0	0.6	8.0	A/G(N3)
	M06-L	4.4	1.3	2.3	6.6		7.8	1.8	4.5	10.9	A(N3)
	M11-L	3.2	1.2	1.1	5.3		5.8	2.1	1.5	9.6	A(N3), T/U(O4), P(2B), A(N7)
	MN12-L	3.5	1.0	2.2	5.0		6.6	2.4	3.8	12.2	
	MN15-L	4.4	1.1	2.5	6.5		7.9	1.8	5.1	11.8	P(2B)
Ca^{2+}	revTPSS	1.0	1.1	0.1	4.0	A(N3)	1.0	0.8	0.1	3.1	A(N3)
	TPSS	1.0	1.1	0.2	4.3	A/G(N3)	1.1	0.9	0.2	3.4	A(N3)
	TPSS-D3	3.1	1.8	0.8	7.1	A(N3)	3.4	1.3	1.7	5.6	
	TPSS-D4	2.5	1.7	0.4	5.9		2.6	1.2	1.1	4.7	
	M06-L	6.7	2.6	3.9	13.3	A(N3)	8.0	2.9	4.5	13.2	
	M11-L	7.9	3.2	4.2	16.0	A(N3)	9.3	2.8	4.9	15.0	
	MN12-L	4.4	1.8	1.7	8.7	A(N3)	5.0	1.4	2.8	7.3	
	MN15-L	0.8	0.8	0.2	2.4	A(N3), T(O2)	0.9	0.8	0.2	2.3	
Sr^{2+}	revTPSS	1.3	1.4	0.1	5.3	A(N3)	1.2	0.9	0.2	3.6	A(N3)
	TPSS	1.5	1.5	0.3	5.7	A/G(N3)	1.4	1.1	0.2	3.9	A(N3)
	TPSS-D3	4.2	2.2	1.3	8.8	A(N3)	4.1	1.3	2.5	6.2	
	TPSS-D4	3.2	1.9	0.8	7.1		3.1	1.2	1.6	4.9	
	M06-L	6.0	2.6	3.6	12.8	A(N3)	6.5	2.7	3.3	11.0	
	M11-L	8.9	3.7	4.4	18.3	A(N3)	9.3	3.2	4.6	16.1	

	MN12-L	5.7	2.2	2.3	10.8	A(N3)	5.8	1.7	3.4	9.1	
	MN15-L	1.1	0.9	0.2	3.7	A(N3), T(O2)	1.0	0.6	0.4	2.6	A(N3)
Ba^{2+}	revTPSS	2.3	1.8	0.3	6.6	A(N3)	2.1	1.3	0.2	4.0	
	TPSS	2.6	1.8	0.9	7.0	A(N3)	2.5	1.4	0.6	4.6	
	TPSS-D3	5.9	2.8	2.3	10.8		5.3	1.5	3.5	8.3	
	TPSS-D4	5.3	2.6	1.9	9.4		4.7	1.4	2.8	6.9	
	M06-L	5.9	2.5	3.6	12.3		5.8	2.6	2.8	10.0	
	M11-L	9.2	4.0	5.0	18.9	A(N3)	8.9	3.7	3.6	16.5	
	MN12-L	6.9	2.5	3.1	12.4	A(N3)	6.5	2.2	3.6	10.9	
	MN15-L	2.0	1.5	0.1	5.4		1.9	1.0	0.1	3.6	

[a]Outliers in the boxplot statistics for a given complex are provided with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.10. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the GGA family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Be^{2+}	mPW91	1.1	0.4	0.3	2.0		3.1	1.2	0.6	4.7	U(O2)
	BLYP	2.4	1.3	0.5	4.5		6.9	3.5	1.9	14.5	
	BLYP-D3	3.5	1.4	1.1	5.6		10.0	3.7	5.7	18.5	
	BP86	0.7	0.8	0.0	2.8	G/A(N3)	2.0	2.2	0.1	6.3	A(N3), P(2B)
	BP86-D3	1.4	1.1	0.1	4.2	A(N3)	3.8	2.9	0.2	9.6	A(N3), P(2B)
	PBE	2.5	1.1	0.6	4.4	P(2NB), G/A(N3)	7.1	2.5	3.2	11.6	
	PBE-D3	2.9	1.2	0.9	5.3		8.3	2.6	4.6	13.4	A/G(N3), P(2B)
	PBE-D4	2.6	1.2	0.6	5.5		7.3	2.7	3.3	12.7	A/G(N3), P(2NB,2B)
Mg^{2+}	mPW91	0.9	0.6	0.2	2.3	T/U(O2)	1.6	0.7	0.3	3.0	
	BLYP	2.2	2.1	0.3	8.4	A(N3)	3.5	2.7	0.6	11.2	A(N3)
	BLYP-D3	4.1	2.5	0.7	10.2	A(N3)	6.9	3.1	2.1	13.7	A(N3), T(O2), P(2B, 2NB)
	BP86	1.6	0.8	0.3	3.3		2.9	1.4	0.6	6.0	
	BP86-D3	1.1	1.5	0.1	5.7	A/G(N3)	1.9	2.0	0.1	7.6	A(N3)
	PBE	1.8	2.1	0.2	7.3	G/A(N3), T(O2)	2.8	2.7	0.7	9.8	A(N3)

	PBE-D3	2.6	2.1	0.2	8.5	A/G(N3)	4.2	2.7	0.6	11.3	A(N3)
	PBE-D4	2.0	2.0	0.2	8.1	A/G(N3)	3.2	2.6	0.6	10.8	A/G(N3)
Ca^{2+}	mPW91	0.7	0.6	0.0	2.0		0.7	0.5	0.1	1.6	
	BLYP	1.1	0.8	0.2	3.4	A(N3)	1.2	0.8	0.2	2.9	P(2NB)
	BLYP-D3	3.9	2.3	0.7	7.9		4.1	1.5	1.6	6.4	
	BP86	1.0	1.3	0.0	4.8	A/G(N3)	1.0	1.0	0.1	3.8	A(N3)
	BP86-D3	4.2	2.3	1.1	9.1		4.5	1.4	2.6	7.1	
	PBE	1.2	1.1	0.0	3.8		1.3	0.9	0.0	3.0	
	PBE-D3	1.3	1.6	0.0	5.6	A(N3)	1.3	1.2	0.0	4.4	
	PBE-D4	1.1	1.4	0.0	5.0	A(N3)	1.1	1.1	0.0	3.9	A(N3)
Sr^{2+}	mPW91	0.9	0.8	0.1	3.0		0.9	0.6	0.1	2.0	A(N3)
	BLYP	1.3	1.4	0.1	5.2	A(N3)	1.1	0.9	0.1	3.5	A(N3)
	BLYP-D3	5.6	2.9	1.5	10.5		5.3	1.4	2.9	7.4	
	BP86	1.4	1.8	0.0	6.4	A/G(N3)	1.3	1.2	0.0	4.3	A(N3)
	BP86-D3	5.8	2.9	1.9	11.7		5.6	1.3	3.8	7.9	
	PBE	1.2	1.3	0.2	5.2		1.0	0.8	0.4	3.5	A(N3)
	PBE-D3	2.2	2.0	0.4	7.3	A(N3)	2.0	1.3	0.6	5.0	
	PBE-D4	1.7	1.9	0.1	6.3	A(N3)	1.5	1.3	0.1	4.3	
Ba^{2+}	mPW91	1.3	1.2	0.1	4.1		1.5	1.1	0.0	2.9	
	BLYP	2.4	1.9	0.4	6.8		2.1	1.2	0.3	4.0	
	BLYP-D3	8.0	3.8	2.7	13.7		7.0	1.6	4.4	9.7	
	BP86	2.6	2.1	0.7	7.8	A(N3)	2.4	1.3	0.4	4.6	
	BP86-D3	8.2	3.8	3.0	14.4		7.2	1.6	5.2	10.2	
	PBE	1.7	1.8	0.2	6.5	A(N3)	1.5	1.2	0.1	3.8	
	PBE-D3	3.7	2.4	1.3	9.0		3.3	1.4	1.5	5.5	
	PBE-D4	3.5	2.4	1.1	8.3		3.1	1.4	0.9	4.9	

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

Table C.11. Calculated errors in functional performance reported as MPEs (%) and MUEs (kcal/mol) as well as standard (SD), minimum (Min), and maximum (Max) deviations evaluated for the LDA family for each group II metal–nucleic acid complex, in addition to complexes that resulted in outliers according to boxplot statistics.

Family	Functional	MPE	SD	Min	Max	Outliers[a]	MUE	SD	Min	MAX	Outliers[a]
Be^{2+}	SVWN5	3.0	1.1	1.1	5.3		8.5	2.3	5.6	12.8	
Mg^{2+}	SVWN5	4.2	2.2	1.4	10.3	A(N3)	7.2	2.7	4.0	13.7	A(N3), G(N7)
Ca^{2+}	SVWN5	2.0	1.3	0.4	4.0		2.2	1.2	0.4	4.0	
Sr^{2+}	SVWN5	1.8	1.4	0.2	4.7		1.6	1.0	0.4	3.3	
Ba^{2+}	SVWN5	1.9	1.7	0.1	6.2	A(N3)	1.5	1.1	0.0	3.7	A(N3), G(N7)

[a]Outliers in the boxplot statistics for a given complex are provided, with the nucleic acid component (A, C, T, G, U, or P) listed and the corresponding binding site in parentheses.

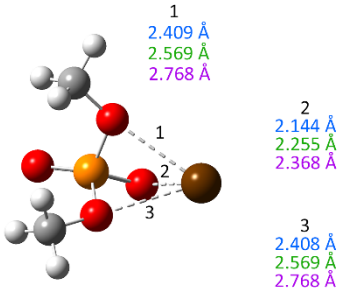


Figure C.1. The distinct structural motif formed between Ca^{2+} (blue), Sr^{2+} (green), or Ba^{2+} (purple) and the phosphate moiety at the 2 bridging oxygens, with an additional contact formed with a non-bridging oxygen.

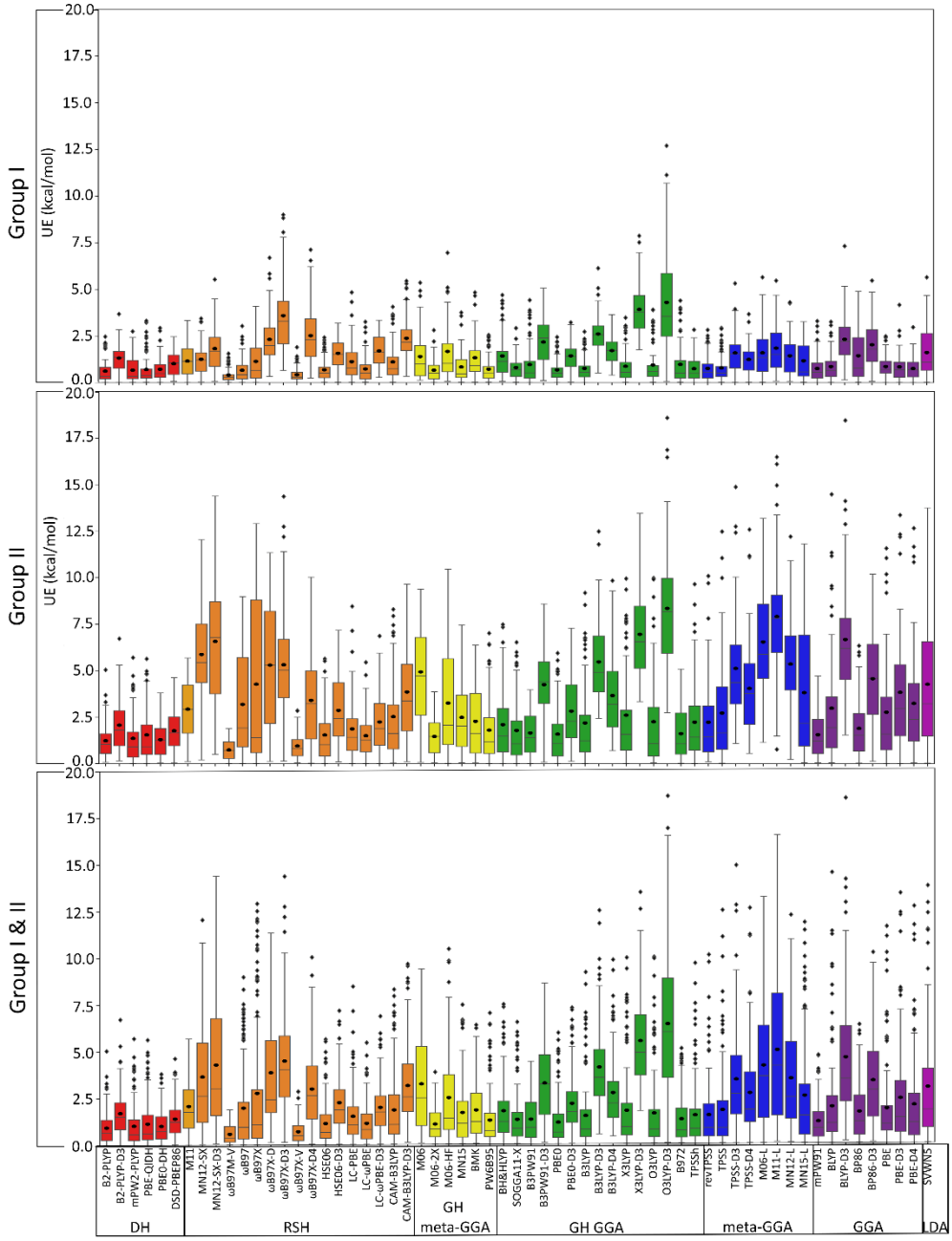


Figure C.2. Boxplots of the unsigned errors (kcal/mol) for group I metals (top), group II metals (middle), and group I and II metals (bottom). Each color represents a functional family according to the rungs of Jacob's Ladder. The boxplot statistics for group I metal–nucleic acid interactions were adapted from our previous study.[1]

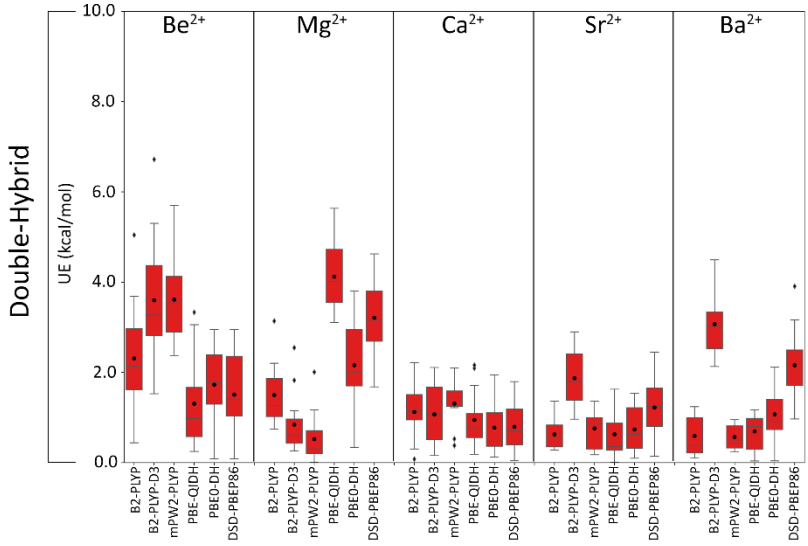


Figure C.3. Boxplot of the unsigned errors (kcal/mol) for the double-hybrid family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

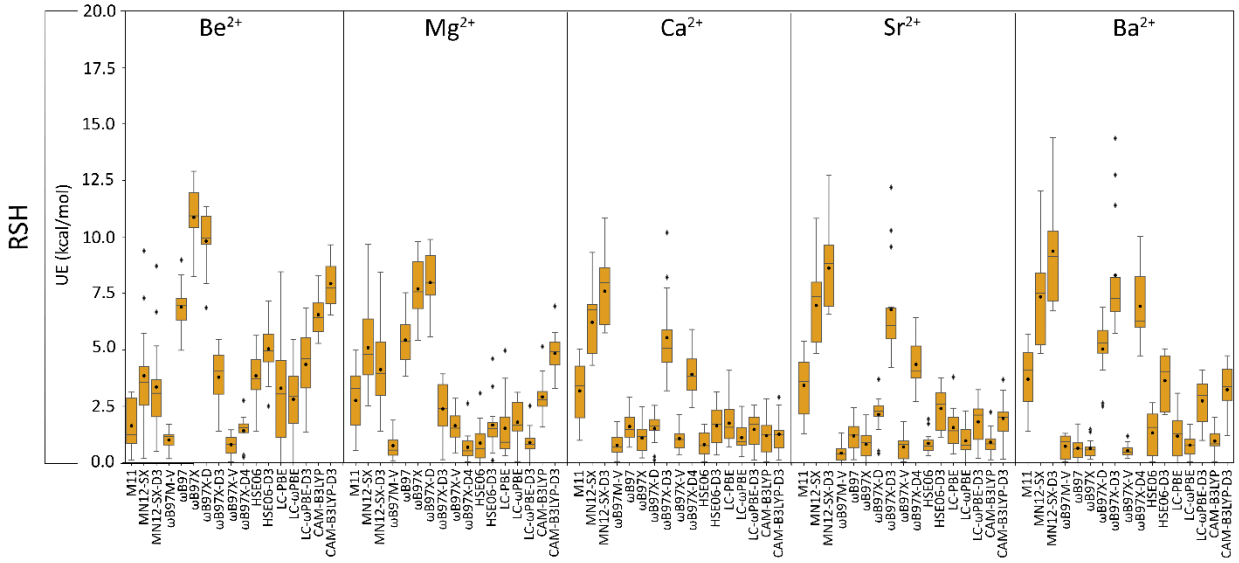


Figure C.4. Boxplot of the unsigned errors (kcal/mol) for the range-separated hybrid family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

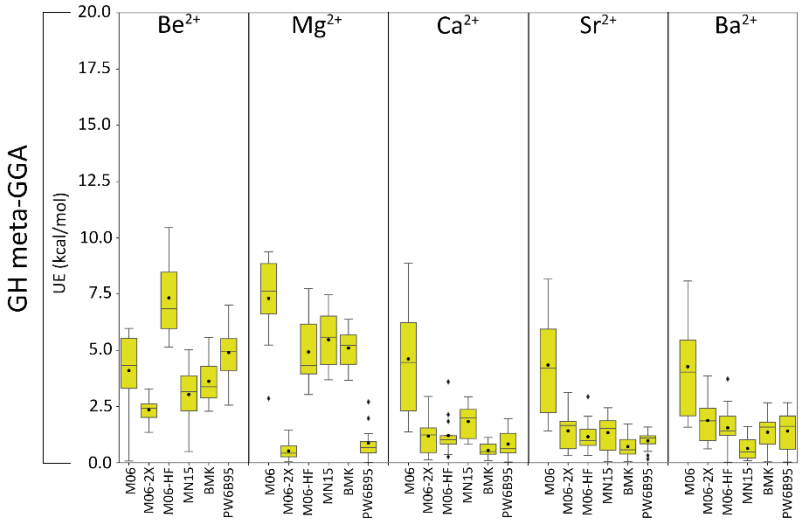


Figure C.5. Boxplot of the unsigned errors (kcal/mol) for the global hybrid meta-GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

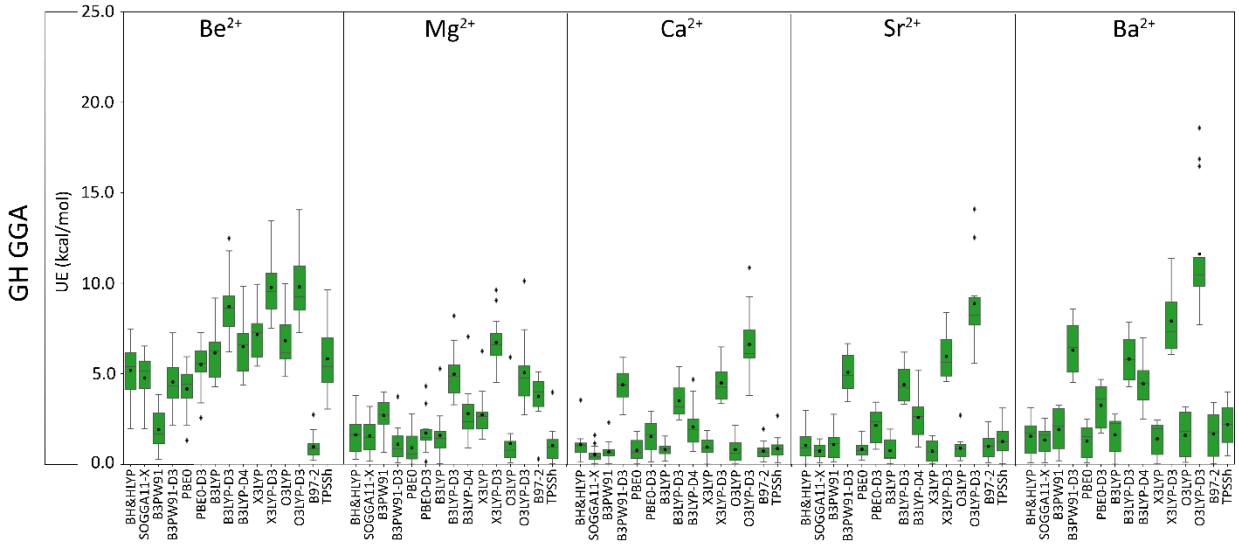


Figure C.6. Boxplot of the unsigned errors (kcal/mol) for the global hybrid GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

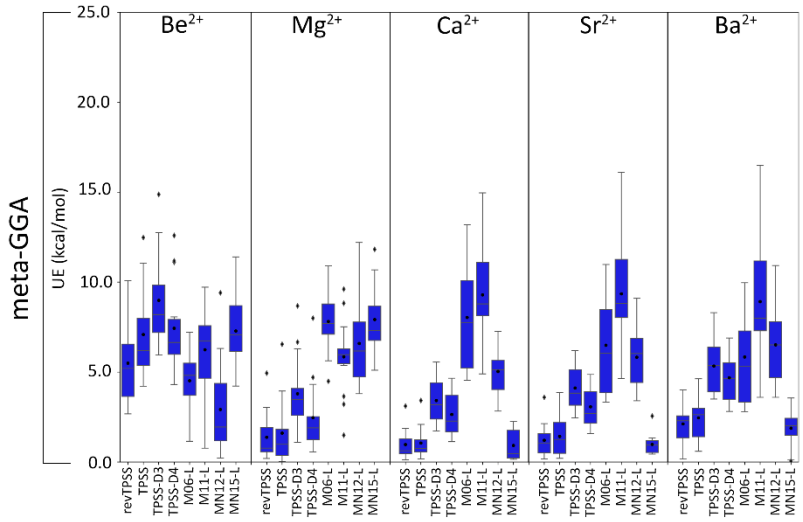


Figure C.7. Boxplot of the unsigned errors (kcal/mol) for the meta-GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

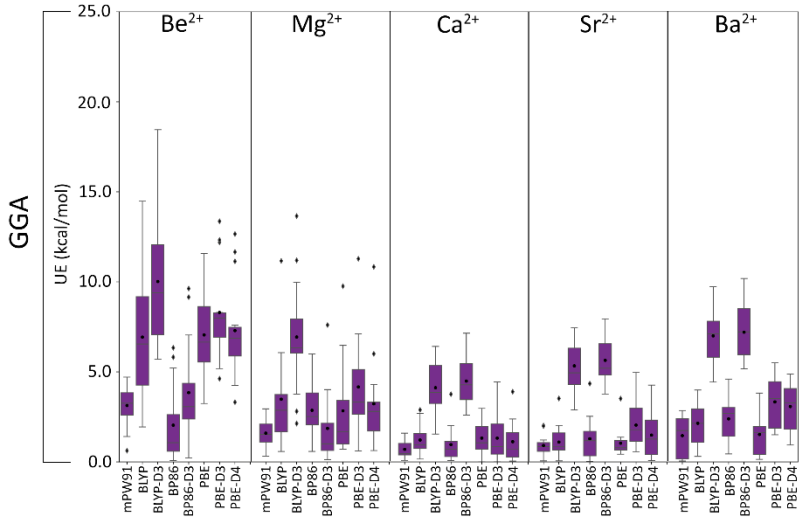


Figure C.8. Boxplot of the unsigned errors (kcal/mol) for the GGA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

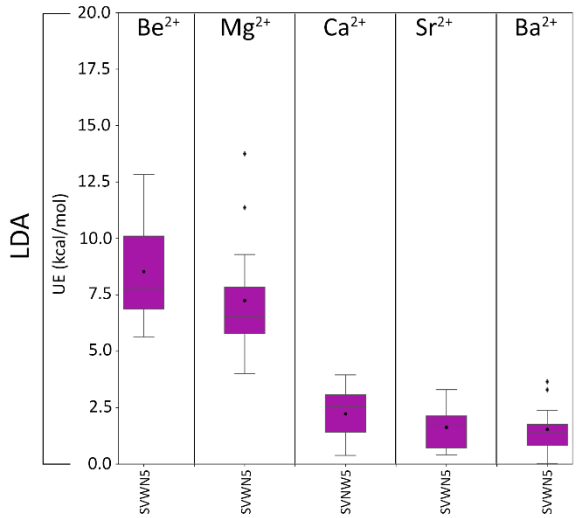


Figure C.9. Boxplot of the unsigned errors (kcal/mol) for the LDA family with respect to CCSD(T)/CBS reference values grouped according to group II metal.

Appendix D

Supplementary Information for Chapter 5: Assessment of DFT Methods for the Structural Prediction of Transition and Post-Transition Metal–Nucleic Acid Complexes

Contains Tables D.1–D.23 and Figures D.1–D.35

Section D.1. Description of Analysis Metrics:

The analysis metrics used throughout the Results and Discussion include the all-heavy-atom and inner-shell root-mean-square-deviations (RMSDs), average RMSDs (ARMSDs), maximum RMSDs (Max RMSDs), as well as inner-shell mean percent errors (MPEs) and maximum percent errors (Max PEs). Note that the term 'heavy-atom' refers to all atoms excluding hydrogen.

All-heavy-atom and inner-shell RMSD were calculated as:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}$$

wherein N is the number of heavy atoms in the complex for the all-heavy-atom RMSD or the number of heavy atoms within the coordination sphere for the inner-shell RMSD, and δ_i is the distance between the position of atom i in the calculated structure and the reference crystallographic structure. The calculated ARMSD is the average RMSD over all complexes for a given functional or DFT functionals for a given complex. The Max RMSD is the largest deviation in the inner-shell or global structure over all complexes in the test set.

The inner-shell MPE was calculated as:

$$MPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{x_i}\right|$$

wherein N is the number of coordination distances, x_i is the crystallographic coordination distance, and y_i is the calculated coordination distance. The Max PE corresponds to the largest deviation in a coordination distance over all complexes in the test set.

Section D.2. Fractional Occupation Number Weighted Density Analysis Results:

Although DFT can accurately describe the ground-state properties of molecules, the single-Slater determinant can limit the reliable DFT description of systems containing strong static electron correlation. The first-row transition metals are particularly susceptible to multi-reference behavior since partially filled d-orbitals lead to multiple near-degenerate electronic configurations and therefore can experience greater static electron correlation effects. Additionally, multi-reference complexes with large delocalization are considered notoriously difficult systems to describe by DFT.[1] Although multireference configuration interaction (MR-CI) is generally recommended for the chemical properties of multireference systems,[2] such methods are prohibitively computationally expensive for predicting the electronic structure of large biomolecules. Therefore, it is important to assess the ability of DFT functionals to offer a reasonable description of the geometry of first-row transition metal–nucleic acid complexes. Nevertheless, each system must be carefully assessed for static correlation effects. In this light, we used the fractional occupation number weighted density (FOD) analysis[1-3] as a qualitative diagnostic tool to identify and visualize the multi-reference character in the 53 nucleic acid–metal complexes investigated in the present study (Figure D.5–D.8).

As anticipated, the first-row transition metals suffer the most from strong static correlation effects. Indeed, Cr^{3+}, Mn^{2+}, Co^{2+}, Fe^{2+}, and Cu^{2+} generally have the largest NFOD values (> 0.7), while Ni^{2+} and Zn^{2+} have NFOD values of ~0.05–0.52 (Figure D.5). Among the complexes involving the heavier metals, only DOMQAZ (Au^{2+}) and 4I1G (Pt^{2+}) exhibit NFOD values > 0.7, while the remaining heavier metal–nucleic acid complexes show minimal static electron correlation effects (NFOD ~0.08–0.50; Figure D.6). Despite the first-row metals showing a high degree of static correlation, the effects are localized on

the metal center for many complexes (i.e., those involving Fe^{2+}, Cr^{3+}, or Mn^{2+}). There is a higher degree of delocalization about the metal center for Cu^{2+} complexes and select Co^{2+} complexes (1FD5 and QAFLOZ).

To ensure the general trends in functional performance prevail regardless of NFOD value, we divided our 41 test set into two groups, namely complexes with low multi-reference character (NFOD < 0.7) and complexes with high multi-reference character (NFOD > 0.7), and reanalyzed the reliability of functionals for each subgroup. Overall, 19 complexes (Figure D.5, Appendix D) that span different metals contain high multi-reference character is treated as a separate test set (denoted 19 subset), which includes the Cr^{3+} (XOSLUO), Mn^{2+} (2OEU, 4HIG, 4DWY), Fe^{2+} (HOPJAY, MORHEJ, WAWPUF, FOJPOM), Co^{2+} (QAFLOZ, AZINAA, 1FD5, 5FHJ), Cu^{2+} (FUDQUR, 1D39, AMPCUS, AMADCS, JUNGIK), Pt^{2+} (4I1G), and Au^{2+} (DOMQAZ) complexes. The remaining complexes (Figure D.6) with low multi-reference character comprise the so-called 22 subset. The most reliable functionals for geometries across the 19 subset are ωB97X-V, ωB97X-D3(BJ), and MN15, which have the smallest MPEs (~7–8%) and Max PEs (~46–47%), in addition to small all-heavy-atom ARMSDs (~0.28–0.31 Å) and Max RMSDs (~0.56–0.70 Å; Figure D.8; Table D.4). For the 22 subset with low multi-reference character, the methods with the lowest MPEs (< 3%), all-heavy-atom ARMSDs (0.332–0.343 Å), and fairly low Max PEs (~8–11%) and RMSDs (~0.750–0.780 Å) are ωB97X-V, ωB97X-D3(BJ), and MN15, although all functionals have highly similar MPEs (up to ~4%), Max PEs (up to ~14%), all-heavy-atom ARMSDs (up to ~0.39 Å), and fluctuating Max RMSDs (up to ~1.1 Å).

Overall, all functionals describe the global geometries of the metal–nucleic acid complexes in the two subsets similarly, with average all-heavy-atom ARMSDs of 0.353 Å

(22 subset) and 0.364 Å (19 subset), and average inner-shell MPEs of 3.3% (22 subset) and 9.5% (19 subset). Most importantly, regardless of whether complexes with the greatest multi-reference character are excluded, ωB97X-V, ωB97X-D3(BJ), and MN15 remain as the top three performing methods for accurately reproducing the geometries as found over the 41 complex test set (Figure D.8). In fact, regardless of the multireference character, there are DFT methods that can reliably describe the geometries of metal–nucleic acid complexes and we discuss our entire 41 test set in the main text. Furthermore, for the 12 challenging complexes that could not be reproduced with any DFT method (Figure D.7), the NFOD values range between 0.076 and 1.652 and therefore there is no correlation between the ability to optimize nucleic acid–metal complexes and the degree of multi-reference character. Nevertheless, although reasonable geometries can be calculated for all complexes in the 41 test set using single-reference DFT methods, multi-reference methods should be used to accurately describe the energetics of complexes showing significant multi-reference character.[4]

Table D.1. Comparison of inner-shell MPEs (%), Max PEs (%), and RMSDs (Å), as well as all-heavy-atom RMSDs (Å) for 12 representative complexes from all-electron gas-phase MN15 optimizations with def2-SVP or def2-TZVP.[a]

	def2-SVP				def2-TZVP				Absolute Differences			
	Inner Shell			All Heavy Atom	Inner Shell			All Heavy Atom	Inner Shell			All Heavy Atom
Complex	MPE	Max PE	RMSD	RMSD	MPE	Max PE	RMSD	RMSD	MPE	Max PE	RMSD	RMSD
4HIG	18.7	37.5	0.338	0.250	19.3	38.3	0.347	0.257	0.6	0.8	0.009	0.007
1FD5	8.6	14.3	0.262	0.318	8.9	15.0	0.282	0.290	0.3	0.7	0.020	0.028
JENGEO	0.6	1.8	0.050	0.091	0.8	2.2	0.054	0.093	0.2	0.4	0.004	0.002
AWUBAW	5.8	8.8	0.223	0.389	7.5	11.3	0.240	0.393	1.7	2.5	0.017	0.004
5IX7	4.3	4.5	0.083	0.613	4.4	4.6	0.085	0.610	0.1	0.1	0.002	0.003
BORFOE	2.9	5.4	0.154	0.662	3.9	6.0	0.250	0.313	1.0	0.6	0.096	0.349
MELTON	3.5	5.1	0.265	0.447	4.0	5.7	0.260	0.446	0.5	0.6	0.005	0.001
HOPJAY	6.4	8.9	0.157	0.192	6.3	8.6	0.152	0.184	0.1	0.3	0.005	0.008
AHOFEK	0.7	1.3	0.025	0.032	0.8	0.9	0.017	0.027	0.1	0.4	0.008	0.005
GIKTIE	1.6	1.8	0.049	0.369	1.2	1.4	0.045	0.349	0.4	0.4	0.004	0.020
FUDQUR	3.3	5.2	0.083	0.147	3.4	7.8	0.089	0.140	0.1	2.6	0.003	0.007
PUBPOR	2.1	3.4	0.174	0.292	1.6	3.2	0.166	0.289	0.5	0.2	0.008	0.003
Average[b]	4.9		0.155	0.317	5.2		0.166	0.283	0.5		0.015	0.036

[a]Models for the complexes explored are shown in Figure 5.1. [b]Average MPE or RMSD calculated across all complexes.

Table D.2. Comparison of inner-shell MPEs (%), Max PEs (%), and RMSDs (Å), as well as all-heavy-atom RMSDs (Å) for the 2 negatively charged complexes in the 41 test set from all-electron gas-phase MN15 optimizations with def2-SVP or def2-SVPD.[a]

	def2-SVP				def2-SVPD				Absolute Differences			
	Inner Shell			All Heavy Atom	Inner Shell			All Heavy Atom	Inner Shell			All Heavy Atom
Complex	MPE	Max PE	RMSD	RMSD	MPE	Max PE	RMSD	RMSD	MPE	Max PE	RMSD	RMSD
XOSLUO	1.0	1.5	0.052	0.113	0.8	1.4	0.050	0.112	0.2	0.1	0.002	0.001
5IX7	4.3	4.5	0.083	0.613	4.1	4.3	0.079	0.611	0.2	0.2	0.004	0.002
Average[b]	2.7		0.068	0.363	2.4		0.064	0.361	0.2	0.2	0.003	0.002

[a]Models for the complexes explored are shown in Figure 5.1. [b]Average MPE or RMSD calculated across all complexes.

Table D.3. Comparison of inner-shell MPEs (%), Max PEs (%), and RMSDs (Å), as well as all-heavy-atom RMSDs (Å) for 10 representative complexes from all-electron gas-phase MN15 optimizations with a scalar-relativistic (SDD) or relativistic (ZORA) basis set.[a]

	SDD				ZORA				Absolute Differences			
	Inner Shell			All Heavy Atom	Inner Shell			All Heavy Atom	Inner Shell			All Heavy Atom
Complex	MPE	Max PE	RMSD	RMSD	MPE	Max PE	RMSD	RMSD	MPE	Max PE	RMSD	RMSD
AHOFEK	0.7	1.3	0.025	0.032	0.6	0.8	0.020	0.029	0.1	0.5	0.005	0.003
BORFOE	2.9	5.4	0.066	0.123	2.8	5.2	0.157	0.673	0.1	0.2	0.091	0.550
BUDDUA	1.7	2.1	0.064	0.461	1.9	2.4	0.053	0.433	0.2	0.3	0.011	0.028
CODLAL	1.9	5.2	0.249	0.287	1.9	6.1	0.251	0.288	0.0	0.9	0.022	0.001
DEGJUU	2.0	3.5	0.069	0.197	1.3	2.6	0.061	0.190	0.7	0.9	0.008	0.007
DOMQAZ	1.3	2.9	0.063	0.270	0.9	1.8	0.055	0.258	0.4	1.1	0.008	0.012
GIKTIE	1.6	1.8	0.049	0.369	0.8	1.0	0.041	0.358	0.8	0.8	0.008	0.011
MEHSEY	3.0	6.2	0.032	0.341	3.2	6.6	0.032	0.671	0.2	0.4	0.000	0.330
PUBPOR	2.1	3.4	0.174	0.292	1.8	2.6	0.135	0.253	0.3	0.8	0.039	0.039
WIPRUL	3.1	6.4	0.103	0.152	2.6	5.5	0.091	0.149	0.5	0.9	0.012	0.003
Average[b]	2.0		0.089	0.252	1.8		0.090	0.330	0.3		0.018	0.098

[a]Models for the complexes explored are shown in Figure 5.1. [b]Average MPE or RMSD calculated across all complexes.

Table D.4. Comparison of all-heavy-atom ARMSDs (Å) and Max RMSDs (Å), and inner-shell MPEs (%), and Max PEs (%) for each functional from all-electron gas-phase optimizations on the 41, 19, and 22 test set devised based on the fractional occupation density (FOD) analysis.[a]

	41 Complexes[a]				19 Complexes[b]				22 Complexes[c]			
	All Heavy Atom		Inner Shell		All Heavy Atom		Inner Shell		All Heavy Atom		Inner Shell	
Functional	ARMSD	Max RMSD	MPE	Max PE	ARMSD	Max RMSD	MPE	Max PE	ARMSD	Max RMSD	MPE	Max PE
B3LYP	0.373	1.038	6.6	103.8	0.391	1.038	10.1	103.8	0.358	0.692	3.6	10.5
B3LYP-D3	0.339	0.982	6.0	106.0	0.331	0.982	9.0	106.0	0.346	0.780	3.4	10.0
B97D3	0.388	1.092	6.5	75.2	0.386	0.796	9.8	75.2	0.390	1.092	3.7	14.5
BLYP	0.388	1.138	7.5	94.2	0.411	1.138	11.2	94.2	0.369	0.697	4.3	12.9
BLYP-D3	0.371	0.849	6.7	72.2	0.387	0.849	10.1	72.2	0.357	0.806	3.8	11.3
BP86	0.393	1.113	7.0	89.3	0.428	1.113	11.0	89.3	0.363	0.686	3.6	12.2
BP86-D3	0.378	0.848	6.5	71.7	0.383	0.848	10.1	71.7	0.361	0.801	3.5	11.5
MN15	0.310	0.777	5.2	46.2	0.283	0.561	8.0	46.2	0.332	0.777	2.8	8.8
MN15-L	0.338	0.764	6.1	45.2	0.310	0.764	9.0	45.2	0.362	0.739	3.6	11.2
PW6B95	0.346	0.971	5.8	95.8	0.352	0.971	9.3	95.8	0.340	0.756	2.7	10.3
PBE	0.415	1.149	7.0	89.3	0.469	1.149	11.0	89.3	0.369	0.700	3.6	10.7
PBE-D3	0.364	0.851	6.6	72.5	0.390	0.851	10.2	72.5	0.341	0.703	3.4	10.6
PBEh-3c	0.333	0.712	5.5	51.0	0.336	0.712	8.4	51.0	0.330	0.689	2.9	12.4
TPSS	0.385	1.139	6.6	85.7	0.391	0.871	10.4	85.7	0.381	1.139	3.4	11.1
TPSS-D3	0.363	0.946	6.5	102.8	0.386	0.946	10.4	102.8	0.345	0.676	3.2	9.6
TPSSh	0.387	1.083	6.4	118.7	0.408	1.083	10.5	118.7	0.368	0.791	2.8	10.2
ωB97X	0.321	1.000	5.4	109.9	0.306	1.000	7.9	109.9	0.334	0.759	3.2	11.6
ωB97X-D	0.328	0.938	5.3	92.8	0.315	0.938	7.8	92.8	0.339	0.757	3.1	10.3
ωB97X-D3	0.329	0.773	5.1	45.7	0.314	0.700	7.9	45.7	0.343	0.773	2.7	10.4
ωB97X-V	0.321	0.754	4.7	46.5	0.311	0.688	7.1	46.5	0.330	0.754	2.7	11.7
Average[d]	0.359		6.2		0.364		9.5		0.353		3.3	

[a]Evaluated over all 41 complexes in our test set (highlighted in black in Figure 5.1). [b]Evaluated over the 19 complexes in our test set with greater multi-reference character (NFOD > 0.7, Figure D.6). [c]Evaluated over the 22 complexes in our test set with small multi-reference character (NFOD < 0.7, Figure D.7). [d]Average ARMSD or MPE calculated across all functionals.

Table D.5. All-heavy-atom RMSDs (Å) and ARMSDs (Å) from all-electron gas-phase DFT optimizations relative to the crystal structure for our entire 53-complex test set.[a]

		Functional															
Metal	Structure ID	B3LYP	B3LYP-D3	B97D3	BLYP	BLYP-D3	BP86	BP86-D3	MN15	MN15-L	PBE	PBE-D3	TPSS	TPSS-D3	ωB97X	ωB97X-D	ARMSD[b]
Cr	XOSLUO	0.122	0.111	0.113	0.120	0.116	0.114	0.114	0.113	0.114	0.113	0.113	0.113	0.112	0.110	0.110	0.114
Mn	4DWY	0.441	0.553	0.508	0.438	0.505	0.471	0.498	0.542	0.764	0.468	0.579	0.479	0.489	0.536	0.532	0.520
	4HIG	0.253	0.245	0.227	0.237	0.229	0.230	0.226	0.250	0.243	0.230	0.227	0.234	0.229	0.232	0.232	0.235
	2OEU	0.877	0.633	0.557	0.950	0.558	0.666	0.675	0.561	0.532	1.000	0.489	0.696	0.541	0.578	0.790	0.674
	3PDR-MN5	1.010	0.992	0.959	1.014	0.982	1.004	0.996	0.916	0.890	1.019	1.004	1.016	0.993	0.990	0.985	0.985
	3PDR-MN176	1.951	1.663	1.325	1.410	1.356	1.419	1.366	1.689	1.077	1.326	1.354	1.767	1.359	1.662	1.672	1.493
Fe	WAWPUF	0.173	0.178	0.234	0.207	0.219	0.245	0.238	0.202	0.247	0.249	0.223	0.228	0.247	0.171	0.176	0.216
	HOPJAY	0.201	0.179	0.185	0.172	0.181	0.197	0.197	0.192	0.183	0.200	0.196	0.190	0.191	0.205	0.204	0.192
	MORHEJ	0.249	0.290	0.324	0.268	0.313	0.293	0.340	0.320	0.374	0.296	0.321	0.287	0.323	0.256	0.281	0.302
	FOJPOM	0.103	0.084	0.080	0.089	0.077	0.075	0.085	0.075	0.049	0.068	0.072	0.072	0.075	0.107	0.103	0.081
Co	QAFLOZ	0.166	0.143	0.282	0.184	0.266	0.222	0.466	0.122	0.324	0.719	0.745	0.304	0.736	0.167	0.139	0.332
	1FD5	0.269	0.272	0.274	0.261	0.339	0.734	0.313	0.318	0.265	0.357	0.323	0.334	0.358	0.271	0.274	0.331
	AZINAA	0.380	0.355	0.368	0.423	0.396	0.432	0.410	0.237	0.228	0.430	0.417	0.412	0.260	0.211	0.247	0.347
	5FHJ	0.463	0.458	0.436	0.367	0.395	0.386	0.413	0.407	0.386	0.386	0.402	0.380	0.394	0.458	0.454	0.412
Ni	ZITWOO	0.220	0.222	0.234	0.242	0.240	0.244	0.246	0.201	0.196	0.242	0.244	0.236	0.238	0.212	0.211	0.229
	JEXYOD	0.197	0.190	0.193	0.206	0.199	0.199	0.199	0.188	0.284	0.197	0.192	0.197	0.192	0.192	0.187	0.201
	MELTON	0.442	0.445	0.443	0.444	0.445	0.445	0.449	0.447	0.445	0.446	0.448	0.444	0.445	0.444	0.443	0.445
	JENGEO	0.096	0.092	0.160	0.132	0.146	0.126	0.137	0.091	0.429	0.107	0.109	0.130	0.154	0.093	0.092	0.140
Cu	1D39	0.896	0.285	0.691	0.681	0.678	0.685	0.657	0.309	0.301	0.675	0.668	0.684	0.666	0.306	0.310	0.566
	AMPCUS	1.038	0.982	0.796	0.849	0.849	0.846	0.848	0.413	0.356	1.107	0.851	0.871	0.946	1.000	0.938	0.846
	JUNGIK	0.409	0.339	0.707	1.138	0.687	1.113	0.661	0.199	0.280	1.149	0.664	0.704	0.623	0.338	0.332	0.623
	AMADCS	0.289	0.338	0.313	0.318	0.326	0.324	0.347	0.313	0.212	0.328	0.336	0.325	0.343	0.336	0.311	0.317
	FUDQUR	0.156	0.154	0.170	0.181	0.143	0.142	0.130	0.147	0.175	0.163	0.161	0.159	0.159	0.151	0.149	0.156
	MOHJEZ	0.893	0.750	0.753	0.943	0.759	0.890	0.756	0.702	0.704	0.872	0.753	0.895	1.068	0.870	1.041	0.843
Zn	FAJZAT	0.211	0.166	1.092	0.237	0.170	0.212	0.147	0.172	0.486	0.213	0.176	0.206	0.152	0.204	0.147	0.266
	AWUBAW	0.586	0.604	0.469	0.579	0.603	0.577	0.598	0.389	0.377	0.594	0.605	0.596	0.610	0.444	0.600	0.549
	5JVW	0.388	0.379	0.400	0.413	0.404	0.410	0.402	0.352	0.309	0.412	0.406	0.413	0.402	0.367	0.354	0.387
	1NLC	0.946	0.885	0.741	0.922	0.769	0.951	0.904	0.736	0.782	0.941	0.908	0.948	0.882	1.129	1.167	0.907
Pd	FODDIN	0.414	0.096	0.123	0.416	0.111	0.410	0.157	0.092	0.118	0.409	0.162	0.411	0.165	0.104	0.101	0.219
	PUBPOR	0.435	0.325	0.356	0.434	0.366	0.394	0.468	0.292	0.512	0.420	0.256	0.383	0.287	0.295	0.270	0.366
Ag	5IX7	0.612	0.613	0.613	0.618	0.618	0.618	0.619	0.613	0.612	0.617	0.617	0.618	0.620	0.610	0.610	0.615

	COLDAL	0.311	0.286	0.295	0.315	0.294	0.292	0.296	0.287	0.295	0.288	0.289	0.295	0.287	0.287	0.300	0.294
	BUDDUA	0.484	0.490	0.502	0.475	0.479	0.466	0.470	0.461	0.504	0.465	0.466	0.494	0.506	0.484	0.507	0.484
	6LBW	0.342	0.362	0.356	0.336	0.355	0.348	0.368	0.355	0.489	0.347	0.359	0.355	0.373	0.356	0.359	0.364
	5AY2	1.729	1.731	1.728	1.730	1.738	1.729	1.734	1.725	1.717	1.730	1.732	1.727	1.733	1.729	1.729	1.729
Cd	BORFOE	0.625	0.631	0.633	0.657	0.686	0.653	0.661	0.662	0.331	0.700	0.703	0.671	0.676	0.624	0.612	0.635
	FAJYIA	0.390	0.417	0.415	0.401	0.425	0.394	0.434	0.426	0.739	0.397	0.413	0.389	0.415	0.406	0.420	0.432
	WIPRUL	0.154	0.149	0.153	0.166	0.151	0.155	0.149	0.152	0.173	0.148	0.145	0.148	0.143	0.150	0.156	0.153
	6CB3	0.878	1.014	0.893	0.890	0.814	0.885	0.870	0.902	1.165	0.891	0.820	1.708	0.888	0.844	0.859	0.955
Pt	4I1G	0.385	0.386	0.374	0.368	0.370	0.387	0.389	0.394	0.392	0.378	0.379	0.383	0.384	0.127	0.126	0.348
	DEGJUU	0.206	0.211	0.214	0.217	0.221	0.215	0.215	0.197	0.183	0.215	0.216	0.209	0.211	0.203	0.205	0.209
	SUTTAC	0.692	0.780	0.815	0.697	0.806	0.686	0.801	0.777	0.218	0.691	0.603	1.139	0.608	0.759	0.757	0.722
	GIKTIE	0.332	0.352	0.344	0.341	0.361	0.329	0.348	0.369	0.397	0.335	0.347	0.330	0.344	0.343	0.348	0.348
	1IHH	2.522	2.060	2.520	2.089	2.533	2.532	2.525	2.493	0.888	2.525	2.521	2.074	2.518	2.060	2.504	2.291
Au	ZOBTAO	0.360	0.371	0.369	0.364	0.375	0.361	0.376	0.415	0.412	0.361	0.370	0.368	0.379	0.368	0.362	0.374
	DOMQAZ	0.566	0.296	0.695	0.555	0.711	0.563	0.261	0.270	0.457	0.595	0.249	0.567	0.249	0.256	0.271	0.437
	AHOFEK	0.054	0.048	0.059	0.079	0.071	0.057	0.052	0.032	0.052	0.054	0.051	0.053	0.050	0.037	0.040	0.053
Hg	GOFYIJ	1.646	1.755	1.744	1.707	1.774	1.678	1.787	1.706	1.809	1.643	1.727	2.043	1.720	1.725	1.721	1.746
	4R6M	0.991	0.991	0.993	0.996	0.996	0.997	0.997	0.991	0.991	0.996	0.996	0.995	0.995	0.992	0.993	0.994
Pb	MEHSEY	0.326	0.390	0.348	0.351	0.331	0.393	0.340	0.341	0.413	0.455	0.329	0.291	0.329	0.358	0.381	0.358
	TPMTPB	2.839	2.833	2.933	2.860	2.935	3.102	2.933	2.859	2.614	2.839	2.840	2.826	2.831	2.839	2.828	2.861
	6A85	0.710	0.686	0.659	0.676	1.213	0.769	0.620	1.019	1.276	0.632	0.575	0.873	0.820	0.682	1.132	0.823
ARMSD[b]		0.604	0.563	0.599	0.599	0.598	0.617	0.590	0.547	0.534	0.624	0.580	0.628	0.591	0.552	0.578	0.588

[a]See Figure 5.1 for models of all complexes included in our test set. 5WSQ was excluded due to a significant change in the metal coordination site across all functionals. [b]Average RMSD calculated across all functionals for a given complex or across all complexes for a given functional.

Table D.6. Comparison of all-heavy-atom ARMSDs (Å) and Max RMSDs (Å) for all-electron gas-phase and implicit water DFT optimizations of 4 complexes.[a]

	Gas Phase		Implicit Water	
Functional	ARMSD	Max RMSD	ARMSD	Max RMSD
B3LYP	1.123	1.729	0.536	0.837
B3LYP-D3	1.121	1.731	0.583	1.008
B97D3	1.092	1.728	0.592	1.091
BLYP	1.140	1.730	0.577	0.962
BLYP-D3	1.077	1.738	0.573	1.087
BP86	1.125	1.729	0.456	0.658
BP86-D3	1.089	1.734	0.574	0.888
MN15	1.080	1.725	0.398	0.634
MN15-L	1.144	1.717	0.560	1.132
PBE	1.122	1.730	0.485	0.831
PBE-D3	1.075	1.732	0.566	1.135
TPSS	1.331	1.727	0.594	1.210
TPSS-D3	1.171	1.733	0.519	0.828
ωB97X	1.109	1.729	0.481	0.833
ωB97X-D	1.155	1.729	0.472	0.840
Average[b]	1.130		0.531	

[a]The crystallographic structures included in this subset are 5AY2, MOHJEZ, 4R6M, and 6CB3. See Figure 5.1 for models of all complexes included in our test set. [b]Average ARMSD calculated across all functionals.

Table D.7. Comparison of all-heavy-atom RMSDs (Å) and ARMSDs (Å) for all-electron gas-phase and implicit water DFT optimizations of 4 complexes.[a]

	Gas-Phase				Implicit Water			
Functional	5AY2	MOHJEZ	4R6M	6CB3	5AY2	MOHJEZ	4R6M	6CB3
B3LYP	1.729	0.893	0.991	0.878	0.061	0.837	0.409	0.837
B3LYP-D3	1.731	0.750	0.991	1.014	0.076	1.008	0.410	0.837
B97D3	1.728	0.753	0.993	0.893	0.088	1.091	0.413	0.775
BLYP	1.730	0.943	0.996	0.890	0.079	0.852	0.415	0.962
BLYP-D3	1.738	0.759	0.996	0.814	0.090	1.087	0.413	0.703
BP86	1.729	0.890	0.997	0.885	0.110	0.625	0.430	0.658
BP86-D3	1.734	0.756	0.997	0.870	0.231	0.888	0.429	0.747
MN15	1.725	0.702	0.991	0.902	0.078	0.554	0.327	0.634
MN15-L	1.717	0.704	0.991	1.165	0.064	0.636	0.410	1.132
PBE	1.730	0.872	0.996	0.891	0.097	0.831	0.431	0.580
PBE-D3	1.732	0.753	0.996	0.820	0.104	1.135	0.431	0.596
TPSS	1.727	0.895	0.995	1.708	0.089	1.210	0.412	0.665
TPSS-D3	1.733	1.068	0.995	0.888	0.096	0.828	0.432	0.722
ωB97X	1.729	0.870	0.992	0.844	0.028	0.833	0.328	0.737
ωB97X-D	1.729	1.041	0.993	0.859	0.024	0.840	0.409	0.617
ARMSD[b]	1.729	0.843	0.994	0.955	0.088	0.884	0.407	0.747

[a]The crystallographic structures included in this subset are 5AY2, MOHJEZ, 4R6M, and 6CB3. See Figure 5.1 for models of all complexes included in our test set. [b]Average RMSD calculated across all functionals.

Table D.8. Comparison of all-heavy-atom ARMSDs (Å) and Max RMSDs (Å) for all-electron gas-phase and implicit water optimizations of 5 complexes.[a]

	Gas Phase				Implicit Water			
	Unconstrained		Constrained[b]		Unconstrained		Constrained[b]	
Functional	ARMSD	Max RMSD	ARMSD	Max RMSD	ARMSD	Max RMSD	ARMSD	Max RMSD
B3LYP	1.243	1.951	0.712	1.365	0.783	1.206	0.677	1.304
B3LYP-D3	1.013	1.663	0.723	1.325	0.770	1.201	0.517	1.122
B97D3	0.910	1.325	0.691	1.322	0.683	1.337	0.523	1.130
BLYP	1.131	1.631	0.736	1.396	0.817	1.255	0.619	0.849
BLYP-D3	1.035	1.356	0.716	1.346	0.759	1.178	0.513	0.811
BP86	1.151	1.614	0.792	1.379	0.758	1.180	0.547	0.854
BP86-D3	0.944	1.366	0.746	1.328	0.795	1.360	0.417	0.695
MN15	1.035	1.689	0.611	0.925	0.779	1.239	0.383	0.568
MN15-L	0.963	1.276	0.582	0.900	0.888	1.079	0.410	0.573
PBE	1.106	1.611	0.718	1.369	0.780	1.230	0.506	0.814
PBE-D3	0.945	1.354	0.704	1.344	0.772	1.323	0.432	0.722
TPSS	1.240	1.767	0.725	1.347	0.858	1.565	0.512	0.825
TPSS-D3	0.981	1.359	0.786	1.321	0.814	1.386	0.436	0.702
ωB97X	1.067	1.662	0.692	1.312	0.747	1.404	0.482	0.740
ωB97X-D	1.158	1.672	0.686	1.287	0.725	1.422	0.425	0.702
Average[c]	1.061		0.708		0.782		0.493	

[a]The crystallographic structures included in this subset are 3PDR-MN5, 3PDR-MN176, 1NLC, 6A85, and 1IHH. See Figure 5.1 for models of all complexes included in our test set. [b]Model was truncated and constrained to crystallographic coordinates at C1′ of nucleobase ligands, and C5′ and C3′ of phosphate models. [c]Average ARMSD calculated across all functionals.

Table D.9. Comparison of all-heavy-atom RMSDs (Å) and ARMSDs (Å) for all-electron gas-phase and implicit water DFT optimizations of 5 complexes.[a]

	Gas Phase										Implicit Water									
	Unconstrained					Constrained[b]					Unconstrained					Constrained[b]				
Functional	3PDR-MN176	3PDR-MN5	1NLC	6A85	1IHH	3PDR-MN176	3PDR-MN5	1NLC	6A85	1IHH	3PDR-MN176	3PDR-MN5	1NLC	6A85	1IHH	3PDR-MN176	3PDR-MN5	1NLC	6A85	1IHH
B3LYP	1.951	1.010	0.946	0.710	2.522	0.395	0.984	0.415	0.402	1.365	1.206	0.602	0.375	0.590	1.142	0.364	1.105	0.271	0.341	1.304
B3LYP-D3	1.663	0.992	0.885	0.686	2.060	0.426	1.107	0.427	0.332	1.325	1.201	0.635	0.404	0.470	1.142	0.382	1.122	0.414	0.281	0.387
B97D3	1.325	0.959	0.741	0.659	2.520	0.372	0.981	0.429	0.353	1.322	1.337	0.609	0.4056	0.268	0.793	0.378	1.130	0.432	0.281	0.394
BLYP	1.410	1.014	0.922	0.676	2.089	0.399	0.987	0.429	0.472	1.396	1.255	0.611	0.4575	0.585	1.178	0.391	0.674	0.849	0.566	0.615
BLYP-D3	1.356	0.982	0.769	1.213	2.533	0.411	0.993	0.443	0.388	1.346	1.156	0.634	0.4188	0.409	1.178	0.397	0.682	0.811	0.314	0.362
BP86	1.419	1.004	0.951	0.769	2.532	0.390	1.138	0.470	0.583	1.379	1.180	0.619	0.4751	0.365	1.152	0.384	0.690	0.325	0.484	0.854
BP86-D3	1.366	0.996	0.904	0.620	2.525	0.391	1.232	0.440	0.339	1.328	1.360	0.634	0.4687	0.360	1.152	0.409	0.695	0.352	0.254	0.376
MN15	1.689	0.916	0.736	1.019	2.493	0.392	0.925	0.849	0.404	0.484	1.239	0.897	0.4068	0.631	0.719	0.343	0.568	0.391	0.261	0.353
MN15-L	1.077	0.890	0.782	1.276	0.888	0.441	0.900	0.627	0.607	0.338	1.079	0.858	0.7789	1.007	0.718	0.312	0.573	0.375	0.452	0.338
PBE	1.326	1.019	0.941	0.632	2.525	0.393	1.004	0.432	0.391	1.369	1.230	0.623	0.4524	0.572	1.021	0.388	0.684	0.291	0.353	0.814
PBE-D3	1.354	1.004	0.908	0.575	2.521	0.393	0.998	0.440	0.346	1.344	1.323	0.636	0.4222	0.459	1.021	0.392	0.722	0.308	0.261	0.478
TPSS	1.767	1.016	0.948	0.873	2.074	0.402	1.015	0.465	0.399	1.347	1.565	0.617	0.4619	0.525	1.123	0.395	0.675	0.339	0.330	0.825
TPSS-D3	1.359	0.993	0.882	0.820	2.518	0.414	1.004	0.436	0.758	1.321	1.386	0.622	0.4423	0.497	1.123	0.385	0.702	0.425	0.259	0.409
ωB97X	1.662	0.990	1.129	0.682	2.060	0.411	0.993	0.408	0.335	1.312	1.404	0.614	0.3526	0.502	0.860	0.369	0.691	0.301	0.311	0.740
ωB97X-D	1.672	0.985	1.167	1.132	2.504	0.407	0.682	0.408	0.647	1.287	1.422	0.582	0.4244	0.403	0.793	0.365	0.702	0.391	0.282	0.384
ARMSD[b]	1.493	0.985	0.907	0.823	2.291	0.402	0.996	0.475	0.450	1.218	1.290	0.653	0.450	0.510	1.008	0.377	0.761	0.418	0.335	0.576

[a]The crystallographic structures included in this subset are 3PDR-MN5, 3PDR-MN176, 1NLC, 6A85, and 1IHH. See Figure 5.1 for models of all complexes included in our test set. [b]Model was truncated and constrained to crystallographic coordinates at C1′ of nucleobase ligands, and C5′ and C3′ of phosphate models. [b]Average RMSD calculated across all functionals.

Table D.10. Comparison of all-heavy-atom RMSDs (Å) and ARMSDs (Å) for each functional from all-electron gas-phase and implicit water optimizations of 2 complexes.[a]

	Gas Phase			Implicit Water		
	Unconstrained		Constrained[b]	Unconstrained		Constrained[b]
Functional	GOFYIJ	TPMTPB[c]	GOFYIJ	GOFYIJ	TPMTPB[c]	GOFYIJ
B3LYP	1.646	2.839	1.634	0.980	2.847	1.791
B3LYP-D3	1.755	2.833	1.429	0.955	2.819	1.386
B97D3	1.744	2.933	1.379	1.669	2.758	1.551
BLYP	1.707	2.860	1.834	1.671	2.990	1.954
BLYP-D3	1.774	2.935	1.368	1.512	2.814	1.525
BP86	1.678	3.102	1.604	1.438	2.985	1.766
BP86-D3	1.787	2.933	2.221	1.512	2.754	2.089
MN15	1.706	2.859	1.374	1.857	2.688	1.554
MN15-L	1.809	2.614	1.296	1.065	2.701	1.364
PBE	1.643	2.839	1.630	1.232	2.832	1.836
PBE-D3	1.727	2.840	1.349	0.867	2.800	1.518
TPSS	2.043	2.826	1.618	0.846	3.019	1.965
TPSS-D3	1.720	2.831	1.342	1.542	2.934	1.440
ωB97X	1.725	2.839	1.394	1.258	2.749	1.436
ωB97X-D	1.721	2.828	1.399	1.191	2.826	1.393
ARMSD[d]	1.746	2.861	1.525	1.387	2.834	1.638

[a]See Figure 5.1 for models of all complexes included in our test set. [b]Model was truncated and constrained to crystallographic coordinates at C1′ of nucleobase ligands. [c]Constraints were not considered for this complex due to the lack of a suitable constraint point (see main text for details). [d]Average RMSD calculated across all functionals.

Table D.11. All-heavy-atom RMSDs (Å) and ARMSDs (Å) from all-electron gas-phase DFT optimizations relative to the crystal structure for our 41 test set.[a]

		Functional																				
Metal	Structure ID	B3LYP	B3LYP-D3	B97D3	BLYP	BLYP-D3	BP86	BP86-D3	MN15	MN15-L	PW6B95	PBE	PBE-D3	PBEh-3c	TPSS	TPSS-D3	TPSSh	ωB97X	ωB97X-D	ωB97X-D3	ωB97X-V	ARMSD[b]
Cr	XOSLUO	0.122	0.111	0.113	0.120	0.116	0.114	0.114	0.113	0.114	0.109	0.113	0.113	0.151	0.113	0.112	0.111	0.110	0.110	0.126	0.128	0.117
Mn	4DWY	0.441	0.553	0.508	0.438	0.505	0.471	0.498	0.542	0.764	0.735	0.468	0.579	0.672	0.479	0.489	0.481	0.536	0.532	0.700	0.688	0.554
	4HIG	0.253	0.245	0.227	0.237	0.229	0.230	0.226	0.250	0.243	0.228	0.230	0.227	0.239	0.234	0.229	0.246	0.232	0.232	0.242	0.241	0.236
	2OEU	0.877	0.633	0.557	0.950	0.558	0.666	0.675	0.561	0.532	0.601	1.000	0.489	0.712	0.696	0.541	0.879	0.578	0.790	0.483	0.481	0.663
Fe	WAWPUF	0.173	0.178	0.234	0.207	0.219	0.245	0.238	0.202	0.247	0.180	0.249	0.223	0.157	0.228	0.247	0.201	0.171	0.176	0.167	0.169	0.206
	HOPJAY	0.201	0.179	0.185	0.172	0.181	0.197	0.197	0.192	0.183	0.200	0.200	0.196	0.182	0.190	0.191	0.186	0.205	0.204	0.191	0.188	0.191
	MORHEJ	0.249	0.290	0.324	0.268	0.313	0.293	0.340	0.320	0.374	0.271	0.296	0.321	0.246	0.287	0.323	0.276	0.256	0.281	0.276	0.301	0.295
	FOJPOM	0.103	0.084	0.080	0.089	0.077	0.075	0.085	0.075	0.049	0.103	0.068	0.072	0.111	0.072	0.075	0.078	0.107	0.103	0.106	0.110	0.086
Co	QAFLOZ	0.166	0.143	0.282	0.184	0.266	0.222	0.466	0.122	0.324	0.142	0.719	0.745	0.163	0.304	0.736	0.172	0.167	0.139	0.153	0.155	0.289
	1FD5	0.269	0.272	0.274	0.261	0.339	0.734	0.313	0.318	0.265	0.273	0.357	0.323	0.247	0.334	0.358	0.335	0.271	0.274	0.272	0.274	0.318
	AZINAA	0.380	0.355	0.368	0.423	0.396	0.432	0.410	0.237	0.228	0.241	0.430	0.417	0.229	0.412	0.260	0.391	0.211	0.247	0.226	0.226	0.326
	5FHJ	0.463	0.458	0.436	0.367	0.395	0.386	0.413	0.407	0.386	0.428	0.386	0.402	0.333	0.380	0.394	0.371	0.458	0.454	0.434	0.436	0.409
Ni	ZITWOO	0.220	0.222	0.234	0.242	0.240	0.244	0.246	0.201	0.196	0.211	0.242	0.244	0.205	0.236	0.238	0.226	0.212	0.211	0.218	0.217	0.225
	JEXYOD	0.197	0.190	0.193	0.206	0.199	0.199	0.199	0.188	0.284	0.189	0.197	0.192	0.194	0.197	0.192	0.194	0.192	0.187	0.178	0.178	0.197
	MELTON	0.442	0.445	0.443	0.444	0.445	0.445	0.449	0.447	0.445	0.446	0.446	0.448	0.451	0.444	0.445	0.444	0.444	0.443	0.448	0.447	0.446
	JENGEO	0.096	0.092	0.160	0.132	0.146	0.126	0.137	0.091	0.429	0.090	0.107	0.109	0.112	0.130	0.154	0.107	0.093	0.092	0.097	0.096	0.130
Cu	1D39	0.896	0.285	0.691	0.681	0.678	0.685	0.657	0.309	0.301	0.319	0.675	0.668	0.252	0.684	0.666	0.816	0.306	0.310	0.310	0.309	0.525
	AMPCUS	1.038	0.982	0.796	0.849	0.849	0.846	0.848	0.413	0.356	0.971	1.107	0.851	0.404	0.871	0.946	1.083	1.000	0.938	0.455	0.470	0.804
	JUNGIK	0.409	0.339	0.707	1.138	0.687	1.113	0.661	0.199	0.280	0.753	1.149	0.664	0.549	0.704	0.623	0.709	0.338	0.332	0.375	0.544	0.614
	AMADCS	0.289	0.338	0.313	0.318	0.326	0.324	0.347	0.313	0.212	0.353	0.328	0.336	0.658	0.325	0.343	0.310	0.336	0.311	0.329	0.337	0.337
	FUDQUR	0.156	0.154	0.170	0.181	0.143	0.142	0.130	0.147	0.175	0.140	0.163	0.161	0.089	0.159	0.159	0.151	0.151	0.149	0.097	0.098	0.146
Zn	FAJZAT	0.211	0.166	1.092	0.237	0.170	0.212	0.147	0.172	0.486	0.202	0.213	0.176	0.155	0.206	0.152	0.195	0.204	0.147	0.119	0.137	0.240
	AWUBAW	0.586	0.604	0.469	0.579	0.603	0.577	0.598	0.389	0.377	0.435	0.594	0.605	0.389	0.596	0.610	0.598	0.444	0.600	0.392	0.468	0.526
	5JVW	0.388	0.379	0.400	0.413	0.404	0.410	0.402	0.352	0.309	0.351	0.412	0.406	0.369	0.413	0.402	0.400	0.367	0.354	0.373	0.372	0.384
Pd	FODDIN	0.414	0.096	0.123	0.416	0.111	0.410	0.157	0.092	0.118	0.089	0.409	0.162	0.155	0.411	0.165	0.409	0.104	0.101	0.095	0.100	0.207
	PUBPOR	0.435	0.325	0.356	0.434	0.366	0.394	0.468	0.292	0.512	0.257	0.420	0.256	0.233	0.383	0.287	0.379	0.295	0.270	0.340	0.241	0.347
Ag	5IX7	0.612	0.613	0.613	0.618	0.618	0.618	0.619	0.613	0.612	0.620	0.617	0.617	0.539	0.618	0.620	0.616	0.610	0.610	0.532	0.519	0.603
	COLDAL	0.311	0.286	0.295	0.315	0.294	0.292	0.296	0.287	0.295	0.301	0.288	0.289	0.390	0.295	0.287	0.297	0.287	0.300	0.265	0.236	0.295
	BUDDUA	0.484	0.490	0.502	0.475	0.479	0.466	0.470	0.461	0.504	0.489	0.465	0.466	0.516	0.494	0.506	0.499	0.484	0.507	0.480	0.462	0.485

	6LBW	0.342	0.362	0.356	0.336	0.355	0.348	0.368	0.355	0.489	0.355	0.347	0.359	0.421	0.355	0.373	0.356	0.356	0.359	0.457	0.421	0.374
Cd	BORFOE	0.625	0.631	0.633	0.657	0.686	0.653	0.661	0.662	0.331	0.634	0.700	0.703	0.550	0.671	0.676	0.638	0.624	0.612	0.632	0.588	0.628
	FAJYIA	0.390	0.417	0.415	0.401	0.425	0.394	0.434	0.426	0.739	0.389	0.397	0.413	0.390	0.389	0.415	0.386	0.406	0.420	0.371	0.395	0.421
	WIPRUL	0.154	0.149	0.153	0.166	0.151	0.155	0.149	0.152	0.173	0.138	0.148	0.145	0.159	0.148	0.143	0.148	0.150	0.156	0.176	0.178	0.155
Pt	4I1G	0.385	0.386	0.374	0.368	0.370	0.387	0.389	0.394	0.392	0.390	0.378	0.379	0.395	0.383	0.384	0.388	0.127	0.126	0.390	0.129	0.346
	DEGJUU	0.206	0.211	0.214	0.217	0.221	0.215	0.215	0.197	0.183	0.191	0.215	0.216	0.171	0.209	0.211	0.205	0.203	0.205	0.198	0.200	0.205
	SUTTAC	0.692	0.780	0.815	0.697	0.806	0.686	0.801	0.777	0.218	0.756	0.691	0.603	0.689	1.139	0.608	0.791	0.759	0.757	0.674	0.754	0.725
	GIKTIE	0.332	0.352	0.344	0.341	0.361	0.329	0.348	0.369	0.397	0.341	0.335	0.347	0.339	0.330	0.344	0.327	0.343	0.348	0.413	0.407	0.352
Au	ZOBTAO	0.360	0.371	0.369	0.364	0.375	0.361	0.376	0.415	0.412	0.367	0.361	0.370	0.305	0.368	0.379	0.369	0.368	0.362	0.275	0.270	0.360
	DOMQAZ	0.566	0.296	0.695	0.555	0.711	0.563	0.261	0.270	0.457	0.255	0.595	0.249	0.586	0.567	0.249	0.573	0.256	0.271	0.640	0.627	0.462
	AHOFEK	0.054	0.048	0.059	0.079	0.071	0.057	0.052	0.032	0.052	0.038	0.054	0.051	0.032	0.053	0.050	0.044	0.037	0.040	0.031	0.034	0.048
Pb	MEHSEY	0.326	0.390	0.348	0.351	0.331	0.393	0.340	0.341	0.413	0.595	0.455	0.329	0.500	0.291	0.329	0.473	0.358	0.381	0.773	0.540	0.413
ARMSD[b]		0.373	0.339	0.388	0.388	0.371	0.393	0.371	0.310	0.338	0.346	0.415	0.364	0.333	0.385	0.364	0.387	0.321	0.328	0.329	0.321	0.358

[a]See Figure 5.1 for models of all complexes included in our test set. [b]Average RMSD calculated across all functionals for a given complex or across all complexes for a given functional.

Table D.12. Comparison of all-heavy-atom and inner-shell ARMSDs (Å) over 20 functionals relative to crystal structure references for each metal in the gas phase.[a]

Metal	All Heavy Atom	Inner shell
Cr^{3+}	0.116	0.049
Mn^{2+}	0.484	0.336
Fe^{2+}	0.194	0.138
Co^{2+}	0.339	0.229
Ni^{2+}	0.249	0.180
Cu^{2+}	0.485	0.498
Zn^{2+}	0.383	0.174
Pd^{2+}	0.277	0.085
Ag^{+}	0.458	0.121
Cd^{2+}	0.401	0.148
Pt^{2+}	0.407	0.203
$Au^{+/2+}$	0.290	0.052
Pb^{2+}	0.413	0.036

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1.

Table D.13. All-heavy-atom ARMSDs (Å) for each functional–metal combination from all-electron DFT optimizations in the gas phase.[a]

Functional	Cr	Mn	Fe	Co	Ni	Cu	Zn	Pd	Ag	Cd	Pt	Au	Pb	Average[b]
B3LYP	0.112	0.523	0.181	0.320	0.239	0.558	0.395	0.425	0.469	0.390	0.404	0.326	0.326	0.359
B3LYP-D3	0.111	0.477	0.183	0.307	0.237	0.419	0.383	0.211	0.463	0.399	0.432	0.239	0.390	0.327
B97D3	0.113	0.431	0.206	0.340	0.257	0.535	0.654	0.239	0.470	0.400	0.437	0.374	0.348	0.370
BLYP	0.120	0.542	0.184	0.309	0.256	0.633	0.409	0.425	0.469	0.408	0.406	0.333	0.351	0.373
BLYP-D3	0.116	0.431	0.197	0.349	0.257	0.536	0.393	0.238	0.464	0.421	0.439	0.386	0.331	0.351
BP86	0.114	0.455	0.203	0.444	0.254	0.622	0.400	0.402	0.459	0.401	0.404	0.327	0.393	0.375
BP86-D3	0.114	0.466	0.215	0.471	0.258	0.529	0.382	0.312	0.462	0.415	0.438	0.230	0.340	0.356
MN15	0.113	0.451	0.197	0.271	0.232	0.276	0.304	0.192	0.454	0.413	0.434	0.239	0.341	0.301
MN15-L	0.114	0.513	0.213	0.301	0.338	0.265	0.391	0.315	0.470	0.414	0.297	0.307	0.413	0.335
PW6B95	0.109	0.521	0.189	0.271	0.234	0.507	0.329	0.174	0.441	0.387	0.420	0.220	0.595	0.338
PBE	0.113	0.566	0.203	0.473	0.248	0.684	0.406	0.415	0.457	0.415	0.405	0.337	0.455	0.398
PBE-D3	0.113	0.432	0.203	0.472	0.248	0.536	0.396	0.209	0.457	0.420	0.386	0.223	0.329	0.340
PBEh-3c	0.151	0.541	0.174	0.243	0.240	0.390	0.304	0.194	0.467	0.366	0.399	0.307	0.500	0.329
TPSS	0.113	0.470	0.194	0.357	0.252	0.548	0.405	0.397	0.469	0.403	0.515	0.329	0.291	0.365
TPSS-D3	0.112	0.420	0.209	0.437	0.257	0.547	0.388	0.226	0.471	0.412	0.387	0.226	0.329	0.340
TPSSh	0.111	0.535	0.185	0.317	0.243	0.614	0.398	0.394	0.442	0.390	0.428	0.329	0.473	0.374
ωB97X	0.110	0.449	0.185	0.277	0.235	0.426	0.338	0.200	0.461	0.393	0.358	0.220	0.358	0.308
ωB97X-D	0.110	0.516	0.191	0.278	0.233	0.408	0.367	0.185	0.472	0.396	0.359	0.224	0.381	0.317
ωB97X-D3	0.126	0.475	0.185	0.271	0.235	0.313	0.294	0.218	0.433	0.393	0.419	0.315	0.773	0.342
ωB97X-V	0.128	0.470	0.192	0.272	0.234	0.352	0.326	0.171	0.415	0.387	0.372	0.310	0.540	0.321
Average[b]	0.116	0.484	0.194	0.339	0.249	0.485	0.383	0.277	0.458	0.401	0.407	0.290	0.413	0.346

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average ARMSD calculated across all metals for a given functional or across all functionals for a given metal.

Table D.14. Inner-shell ARMSDs (Å) from all-electron DFT optimizations for each functional–metal combination in the gas phase.[a]

Functional	Cr	Mn	Fe	Co	Ni	Cu	Zn	Pd	Ag	Cd	Pt	Au	Pb	Average[b]
B3LYP	0.046	0.323	0.136	0.218	0.178	0.612	0.165	0.045	0.121	0.142	0.229	0.059	0.035	0.178
B3LYP-D3	0.045	0.341	0.131	0.196	0.176	0.375	0.154	0.119	0.123	0.136	0.229	0.056	0.034	0.163
B97D3	0.047	0.334	0.141	0.231	0.179	0.506	0.206	0.127	0.123	0.147	0.221	0.060	0.039	0.182
BLYP	0.048	0.335	0.134	0.240	0.190	0.706	0.187	0.074	0.120	0.153	0.227	0.084	0.045	0.196
BLYP-D3	0.046	0.339	0.137	0.239	0.185	0.597	0.171	0.132	0.123	0.155	0.226	0.070	0.043	0.189
BP86	0.046	0.320	0.144	0.414	0.186	0.688	0.177	0.044	0.127	0.143	0.219	0.061	0.041	0.201
BP86-D3	0.049	0.323	0.148	0.256	0.190	0.578	0.162	0.138	0.133	0.139	0.223	0.056	0.041	0.187
MN15	0.052	0.340	0.141	0.194	0.168	0.223	0.158	0.095	0.132	0.133	0.226	0.036	0.032	0.148
MN15-L	0.047	0.347	0.139	0.232	0.204	0.281	0.185	0.143	0.129	0.241	0.223	0.047	0.041	0.174
PW6B95	0.045	0.346	0.140	0.190	0.177	0.476	0.189	0.066	0.105	0.144	0.227	0.048	0.028	0.168
PBE	0.045	0.332	0.145	0.263	0.186	0.778	0.177	0.043	0.127	0.166	0.217	0.057	0.040	0.198
PBE-D3	0.047	0.341	0.148	0.248	0.185	0.589	0.169	0.079	0.130	0.162	0.220	0.057	0.040	0.186
PBEh-3c	0.066	0.355	0.120	0.191	0.168	0.293	0.175	0.066	0.114	0.151	0.225	0.029	0.027	0.152
TPSS	0.046	0.331	0.142	0.239	0.182	0.587	0.169	0.038	0.126	0.148	0.230	0.058	0.039	0.180
TPSS-D3	0.048	0.337	0.148	0.242	0.179	0.614	0.157	0.083	0.130	0.144	0.224	0.057	0.038	0.185
TPSSh	0.046	0.321	0.138	0.222	0.175	0.740	0.161	0.029	0.108	0.136	0.229	0.050	0.035	0.184
ωB97X	0.045	0.344	0.137	0.196	0.174	0.398	0.191	0.084	0.122	0.129	0.078	0.047	0.030	0.152
ωB97X-D	0.045	0.323	0.135	0.187	0.169	0.362	0.145	0.104	0.122	0.128	0.075	0.047	0.030	0.144
ωB97X-D3	0.057	0.346	0.123	0.186	0.178	0.260	0.181	0.103	0.102	0.132	0.223	0.031	0.032	0.150
ωB97X-V	0.058	0.344	0.132	0.186	0.175	0.302	0.194	0.083	0.098	0.136	0.086	0.034	0.032	0.143
Average[b]	0.049	0.336	0.138	0.229	0.180	0.498	0.174	0.085	0.121	0.148	0.203	0.052	0.036	0.172

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average ARMSD calculated across all metals for a given functional or across all functionals for a given metal.

Table D.15. Comparison of all-heavy-atom ARMSDs (Å) and Max RMSDs (Å), and inner-shell ARMSDs (Å), Max RMSDs (Å), MPEs (%), and Max PEs (%) for each functional from all-electron gas-phase and implicit water DFT optimizations.[a]

	Gas Phase						Implicit Water					
	All Heavy Atom		Inner Shell				All Heavy Atom		Inner Shell			
Functional	ARMSD	Max RMSD	ARMSD	Max RMSD	MPE	Max PE	ARMSD	Max RMSD	ARMSD	Max RMSD	MPE	Max PE
B3LYP	0.373	1.038	0.213	1.234	6.6	103.8	0.343	0.733	0.176	0.682	5.1	39.8
B3LYP-D3	0.339	0.982	0.185	0.918	6.0	106.0	0.328	0.816	0.177	0.681	4.9	38.2
B97D3	0.388	1.092	0.210	0.955	6.5	75.2	0.360	0.902	0.185	0.647	5.6	38.7
BLYP	0.388	1.138	0.235	1.449	7.5	94.2	0.373	1.078	0.219	1.268	6.3	76.8
BLYP-D3	0.371	0.849	0.222	0.940	6.7	72.2	0.362	0.875	0.198	0.804	5.7	53.2
BP86	0.393	1.113	0.244	1.414	7.0	89.3	0.373	1.199	0.226	1.267	6.1	75.2
BP86-D3	0.378	0.848	0.220	0.918	6.5	71.7	0.400	0.918	0.213	0.963	5.7	70.1
MN15	0.310	0.777	0.165	0.700	5.2	46.2	0.301	1.074	0.171	0.700	4.5	36.0
MN15-L	0.338	0.764	0.192	0.684	6.1	45.2	0.339	1.049	0.199	0.684	5.6	54.2
PW6B95	0.346	0.971	0.198	0.923	5.8	95.8	0.296	1.211	0.174	0.966	4.8	89.3
PBE	0.415	1.149	0.243	1.447	7.0	89.3	0.377	1.206	0.225	1.248	6.0	73.5
PBE-D3	0.364	0.851	0.221	0.927	6.6	72.5	0.383	1.312	0.218	0.972	5.8	83.0
PBEh-3c	0.333	0.712	0.173	0.700	5.5	51.0	0.273	0.573	0.155	0.691	4.9	47.2
TPSS	0.385	1.139	0.215	0.922	6.6	85.7	0.349	0.833	0.191	0.71	5.5	58.7
TPSS-D3	0.363	0.946	0.220	0.932	6.5	102.8	0.369	0.838	0.192	0.671	5.3	32.6
TPSSh	0.387	1.083	0.227	1.390	6.4	118.7	0.351	1.323	0.181	1.074	5.2	77.0
ωB97X	0.321	1.000	0.175	0.942	5.4	109.9	0.334	1.063	0.172	0.899	5.4	56.3
ωB97X-D	0.328	0.938	0.165	0.809	5.3	92.8	0.332	0.884	0.159	0.575	4.8	37.9
ωB97X-D3	0.329	0.773	0.169	0.693	5.1	45.7	0.300	0.649	0.158	0.682	4.8	45.1
ωB97X-V	0.322	0.754	0.161	0.605	4.7	46.5	0.286	0.637	0.158	0.684	4.8	48.1
Average[b]	0.359		0.203		6.2		0.273		0.187		5.3	

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average ARMSD or MPE calculated across all functionals.

Table D.16. Inner-shell MPEs (%) and AMPEs (%) for each functional from all-electron DFT optimizations in the gas phase.[a]

Functional	Cr	Mn	Fe	Co	Ni	Cu	Zn	Pd	Ag	Cd	Pt	Au	Pb	AMPE[b]
B3LYP	0.6	15.0	5.1	7.6	3.4	12.9	7.8	2.6	2.0	3.1	8.7	2.9	5.6	6.0
B3LYP-D3	0.4	15.0	5.0	6.8	3.3	9.6	7.5	2.5	2.1	2.6	8.4	2.5	5.6	5.5
B97D3	0.3	14.8	5.5	8.8	3.9	11.0	8.4	2.6	2.0	3.3	8.1	3.1	5.7	6.0
BLYP	0.9	15.0	5.3	8.4	3.9	16.2	9.0	3.6	1.7	4.3	9.0	4.4	6.5	6.8
BLYP-D3	0.4	14.9	5.2	8.6	3.8	12.3	8.3	3.1	2.0	3.3	8.5	3.7	5.1	6.1
BP86	0.4	14.5	5.8	8.7	3.4	15.3	8.2	2.2	2.5	3.3	8.3	2.9	4.5	6.2
BP86-D3	0.6	14.6	5.9	8.8	3.8	11.8	7.9	1.9	3.0	2.8	8.0	2.4	2.9	5.7
MN15	1.0	15.0	5.5	6.2	2.7	6.2	6.4	1.3	2.9	2.4	8.1	1.0	3.0	4.7
MN15-L	0.7	15.6	5.3	8.2	4.1	7.8	8.4	2.2	2.3	3.5	8.2	2.2	4.5	5.6
PW6B95	0.5	13.6	5.0	6.6	3.0	11.8	5.0	1.7	2.2	2.7	8.3	1.8	3.6	5.1
PBE	0.5	14.9	5.8	8.9	3.4	15.0	8.1	2.0	2.5	3.4	8.2	2.7	5.0	6.2
PBE-D3	0.5	15.0	6.0	8.8	3.5	11.9	7.9	1.7	2.7	3.1	8.1	2.5	4.0	5.8
PBEh-3c	0.7	15.0	4.6	6.6	4.4	8.1	5.8	1.0	2.4	2.6	8.2	1.0	3.9	4.9
TPSS	0.5	14.8	5.6	8.4	3.2	13.1	7.7	2.0	2.4	3.0	8.5	2.7	4.3	5.9
TPSS-D3	0.6	14.7	6.0	8.3	3.2	13.3	7.4	1.5	2.7	2.6	8.2	2.4	3.1	5.7
TPSSh	0.5	13.8	5.2	8.0	3.0	15.0	5.1	1.7	2.3	2.3	8.4	2.3	4.3	5.5
ωB97X	0.4	14.8	5.2	6.6	3.1	9.7	7.3	2.0	2.3	2.4	3.9	1.7	4.6	4.9
ωB97X-D	0.4	15.1	5.1	6.4	3.1	9.2	7.3	1.7	2.2	2.3	3.7	1.8	4.4	4.8
ωB97X-D3	0.4	13.9	4.8	6.4	3.5	7.0	5.3	1.3	2.5	2.4	8.1	1.2	2.9	4.6
ωB97X-V	0.5	13.7	4.9	6.3	3.1	7.6	5.1	1.4	2.4	2.3	3.6	1.3	3.3	4.3
AMPE[b]	0.5	14.7	5.3	7.7	3.4	11.2	7.2	2.0	2.4	2.9	7.6	2.3	4.3	5.5

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average MPE calculated across all metals for a given functional or across all functionals for a given metal.

Table D.17. Inner-shell MPEs (%) and AMPEs (%) from all-electron DFT optimizations in the gas phase.[a]

		MPE																				AMPE[b]
Metal	Structure ID	B3LYP	B3LYP-D3	B97D3	BLYP	BLYP-D3	BP86	BP86-D3	MN15	MN15-L	PW6B95	PBE	PBE-D3	PBEh-3c	TPSS	TPSS-D3	TPSSh	ωB97X	ωB97X-D	ωB97X-D3	ωB97X-V	
Cr	XOSLUO	0.6	0.4	0.3	0.9	0.4	0.4	0.6	1.0	0.7	0.4	0.4	0.5	0.7	0.5	0.6	0.5	0.4	0.4	0.4	0.5	0.5
Mn	4DWY	15.4	16.3	16.0	15.6	16.2	15.7	16.1	15.6	16.9	12.8	15.9	16.7	12.1	15.7	16.1	12.2	16.0	16.4	12.5	12.4	15.1
	4HIG	18.1	17.5	17.1	17.7	17.2	16.9	16.6	18.7	18.0	17.0	17.1	16.8	20.4	17.1	16.7	17.2	17.3	17.6	17.5	17.2	17.5
	2OEU	11.4	11.4	11.4	11.7	11.4	11.1	11.3	10.5	11.8	10.9	11.6	11.5	12.5	11.6	11.3	11.9	11.2	11.2	11.6	11.5	11.4
Fe	WAWPUF	6.1	6.0	7.4	6.9	7.0	8.3	7.6	7.6	7.2	5.6	8.4	8.5	4.2	8.1	8.5	6.3	6.0	6.2	4.7	5.3	7.3
	HOPJAY	5.7	6.0	6.5	6.1	6.3	6.7	6.6	6.4	6.6	6.0	6.7	6.6	5.8	6.5	6.5	6.3	5.8	5.9	6.1	6.1	6.3
	MORHEJ	5.7	6.5	7.7	6.3	7.2	7.7	8.4	7.4	6.9	6.2	7.8	8.2	6.0	7.7	8.2	7.2	5.6	6.0	6.0	6.1	6.9
	FOJPOM	2.9	1.6	0.3	1.8	0.4	0.3	1.0	0.6	0.4	2.1	0.3	0.8	2.3	0.2	0.6	0.8	3.2	2.3	2.3	2.3	1.3
Co	QAFLOZ	6.7	4.6	9.1	7.7	9.1	8.6	9.4	2.6	9.9	3.9	9.5	9.4	4.9	8.4	9.2	6.6	6.2	3.3	4.1	4.1	6.9
	1FD5	8.9	8.7	9.4	9.6	9.3	8.9	8.7	8.6	9.4	8.6	9.0	8.8	10.0	8.5	8.3	9.6	8.6	8.5	9.0	8.7	9.0
	AZINAA	7.3	6.0	7.7	8.1	7.4	8.5	7.9	6.6	4.8	6.4	8.3	7.9	4.0	7.9	6.7	7.6	4.2	6.3	4.6	4.8	6.7
	5FHJ	7.5	7.8	8.8	8.4	8.5	8.8	9.0	7.0	8.8	7.5	8.8	9.0	7.5	8.7	8.9	8.4	7.4	7.6	7.7	7.7	8.2
Ni	ZITWOO	4.9	4.9	5.7	5.8	5.5	5.3	5.4	4.2	6.2	4.8	5.4	5.4	5.9	4.9	5.2	4.9	4.6	4.7	5.4	4.9	5.2
	JEXYOD	2.9	2.4	3.7	3.2	3.5	2.7	3.9	2.3	3.5	2.3	2.8	2.7	3.6	2.6	2.6	2.5	2.5	2.5	2.5	2.5	2.9
	MELTON	3.9	4.6	5.1	4.7	5.1	4.5	5.4	3.5	5.0	3.8	4.6	5.2	4.7	4.1	4.3	3.8	4.0	4.1	4.9	3.9	4.5
	JENGEO	2.0	1.3	1.3	1.9	1.3	1.1	0.7	0.6	1.7	1.2	1.0	0.7	3.3	1.0	0.6	0.8	1.4	1.4	1.2	1.2	1.3
Cu	1D39	19.5	8.6	18.1	21.1	17.7	20.5	17.7	8.1	9.9	8.7	20.5	17.7	8.8	20.4	17.9	19.3	8.9	8.6	8.5	8.3	14.4
	AMPCUS	25.1	25.5	6.4	17.1	16.6	16.9	16.6	12.2	10.6	23.9	15.7	16.4	9.4	17.0	24.9	27.9	26.2	23.6	12.2	12.9	17.9
	JUNGIK	9.9	7.6	18.1	29.5	16.1	27.9	15.0	4.4	7.7	19.6	27.8	15.4	12.1	17.7	14.3	17.7	7.2	7.5	7.4	10.3	14.7
	AMADCS	6.2	3.4	8.1	7.9	7.3	6.9	6.8	3.4	6.5	3.3	6.8	6.8	6.7	6.5	6.4	6.1	3.4	3.5	3.6	3.4	5.7
	FUDQUR	3.8	3.0	4.3	5.3	3.9	4.1	3.1	2.8	4.1	3.7	4.1	3.4	3.7	3.9	3.2	4.1	3.0	2.8	3.2	3.2	3.6
Zn	FAJZAT	2.5	2.1	2.1	3.0	2.3	2.7	2.5	2.0	2.8	1.9	2.6	2.4	2.2	2.5	2.4	2.0	2.0	2.1	2.2	2.0	2.3
	AWUBAW	7.1	7.0	8.2	8.0	7.8	7.5	7.5	5.8	7.2	6.7	7.5	7.4	7.9	7.0	6.9	6.7	6.9	6.9	7.1	6.9	7.2
	5JVW	13.8	13.2	14.7	16.1	14.7	14.5	13.7	11.5	15.0	6.4	14.2	13.9	7.3	13.5	13.0	6.5	12.9	13.0	6.6	6.4	12.0
Pd	FODDIN	1.6	1.0	1.1	2.4	1.7	1.1	0.6	0.5	0.8	0.8	1.0	0.7	0.3	1.1	0.7	0.9	1.2	0.8	0.4	0.6	1.0
	PUBPOR	3.7	4.0	4.0	4.7	4.4	3.2	3.2	2.1	3.5	2.7	3.0	2.7	1.8	2.9	2.4	2.6	2.7	2.6	2.1	2.2	3.1
Ag	5IX7	2.7	3.1	3.0	2.5	3.0	3.9	4.3	4.3	3.5	3.2	3.8	4.1	3.7	3.7	4.0	3.6	3.0	3.0	3.7	3.5	3.5
	CODLAL	2.7	1.9	1.7	2.0	1.8	1.4	1.8	1.9	2.1	2.4	1.5	1.5	2.3	1.8	1.8	1.8	3.2	2.9	2.1	2.6	2.1
	BUDDUA	0.4	0.7	0.5	0.4	0.6	1.4	1.9	1.7	0.7	0.5	1.3	1.6	0.7	1.1	1.5	1.0	0.5	0.4	1.0	0.7	0.9
	6LBW	2.3	2.7	2.6	2.1	2.6	3.4	3.9	3.8	2.9	2.7	3.3	3.6	2.7	3.1	3.5	3.0	2.4	2.5	3.0	2.8	2.9
Cd	BORFOE	3.3	3.0	3.7	4.6	3.9	3.8	3.4	2.9	4.0	3.3	4.2	4.0	3.0	3.5	3.2	3.0	2.7	2.7	2.9	2.7	3.4

	FAJYIA	2.7	1.6	2.3	3.9	2.3	2.7	1.5	1.1	2.3	1.7	2.6	1.9	1.6	2.4	1.6	1.2	1.6	1.3	1.4	1.6	2.0
	WIPRUL	3.2	3.2	3.7	4.4	3.6	3.3	3.3	3.1	4.1	3.1	3.3	3.3	3.1	3.0	3.0	2.8	2.9	3.0	3.0	2.8	3.3
Pt	4I1G	26.5	26.6	25.5	24.7	24.8	26.0	26.1	27.6	27.3	27.2	26.1	26.1	27.8	26.4	26.5	27.0	8.6	8.4	27.2	8.9	23.8
	DEGJUU	3.3	3.0	3.1	4.3	3.9	3.0	2.7	2.0	2.4	2.4	2.9	2.7	2.5	3.1	2.9	2.7	2.8	2.7	2.2	2.4	2.9
	SUTTAC	1.7	1.3	1.3	2.6	2.0	1.4	1.1	1.3	1.3	1.3	1.3	1.2	0.9	1.6	1.3	1.3	1.4	1.2	1.2	1.3	1.4
	GIKTIE	3.3	2.6	2.7	4.3	3.2	2.8	2.2	1.6	2.0	2.3	2.6	2.2	1.4	2.9	2.4	2.6	2.7	2.3	1.9	2.0	2.5
Au	ZOBTAO	2.6	2.2	2.1	3.2	2.6	1.9	1.6	0.9	1.8	1.8	1.8	1.6	1.2	1.9	1.6	1.8	2.3	2.2	1.4	1.7	1.9
	DOMQAZ	3.7	3.1	4.3	6.1	5.0	4.2	3.5	1.3	2.9	2.2	4.0	3.5	1.2	4.0	3.4	3.1	1.6	1.6	1.4	1.4	3.1
	AHOFEK	2.6	2.3	2.9	4.0	3.6	2.6	2.3	0.7	2.0	1.5	2.4	2.2	0.6	2.4	2.2	1.9	1.3	1.5	0.7	0.9	2.0
Pb	MEHSEY	5.6	5.6	5.7	6.5	5.1	4.5	2.9	3.0	4.5	3.6	5.0	4.0	3.9	4.3	3.1	4.3	4.6	4.4	2.9	3.3	4.3
AMPE[b]		6.6	6.0	6.5	7.5	6.7	7.0	6.5	5.2	6.1	5.8	7.0	6.6	5.5	6.6	6.5	6.4	5.4	5.3	5.1	4.7	6.2

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average MPE calculated across all functionals for a given complex or across all complexes for a given functional.

Table D.18. Max PEs (%) from all-electron DFT optimizations in the gas phase for the 41 test set, complexes that do not include Cu^{2+}, and only Cu^{2+} complexes.[a]

	Max PE		
Functional	41 test set	w/o Cu^{2+}	Cu^{2+}
B3LYP	103.8	45.0	103.8
B3LYP-D3	106.0	44.9	106.0
B97D3	75.2	42.7	75.2
BLYP	94.2	42.3	94.2
BLYP-D3	72.2	42.3	72.2
BP86	89.3	43.5	89.3
BP86-D3	71.7	43.5	71.7
MN15	37.5	46.2	27.2
MN15-L	35.0	45.2	34.9
PW6B95	95.8	45.6	95.8
PBE	89.3	43.5	89.3
PBE-D3	72.5	43.5	72.5
PBEh-3c	51.0	46.2	51.0
TPSS	85.7	44.3	85.7
TPSS-D3	102.8	44.3	102.8
TPSSh	118.7	45.2	118.7
ωB97X	109.9	35.0	109.9
ωB97X-D	92.8	35.0	92.8
ωB97X-D3	34.6	45.7	28.3
ωB97X-V	46.5	34.3	46.5

[a]Models for the complexes explored are shown in Figure 5.1.

Table D.19. Inner-shell MPEs (%) and AMPEs (%) from all-electron DFT optimizations in implicit water.[a]

		MPE																				AMPE[b]
Metal	Structure ID	B3LYP	B3LYP-D3	B97D3	BLYP	BLYP-D3	BP86	BP86-D3	MN15	MN15-L	PW6B95	PBE	PBE-D3	PBEh-3c	TPSS	TPSS-D3	TPSSh	ωB97X	ωB97X-D	ωB97X-D3	ωB97X-V	
Cr	XOSLUO	1.6	2.3	2.3	2.2	2.2	1.1	0.7	0.5	1.2	0.5	1.1	0.8	0.6	0.9	0.6	0.5	1.2	1.0	0.4	0.5	1.1
Mn	4DWY	16.2	16.4	16.5	16.4	16.6	16.5	16.7	16.3	16.5	16.7	16.6	16.7	15.9	16.5	16.7	16.4	16.1	16.2	16.7	16.6	16.5
	4HIG	17.4	17.0	16.6	17.2	11.7	16.8	16.6	18.2	17.1	16.4	16.7	16.5	19	16.7	16.6	17.5	16.7	17.0	16.8	16.5	16.8
	2OEU	11.8	11.6	11.4	11.8	11.6	11.3	11.3	11.8	11.5	11.1	11.4	11.8	12.4	11.7	11.5	12.2	11.3	11.4	11.5	11.4	11.6
Fe	WAWPUF	6.7	7.2	8.4	7.6	7.0	8.8	9.2	8.2	7.6	6.4	8.9	9.1	4.8	8.6	9.0	7.1	6.6	6.7	4.7	6.1	7.4
	HOPJAY	6.0	6.2	6.8	6.4	6.6	6.9	7.0	6.5	6.8	6.2	6.9	7.1	5.9	6.8	6.9	6.5	6.0	6.1	6.1	6.1	6.5
	MORHEJ	6.3	7.1	8.3	7.0	7.8	8.3	8.9	8.0	7.5	6.8	8.4	8.8	6.2	8.2	8.7	7.8	6.1	6.5	6.7	6.8	7.5
	FOJPOM	2.5	1.3	0.5	1.5	0.2	0.3	1.2	0.4	0.2	1.9	0.4	1.0	2.1	0.1	0.8	0.5	2.9	2.0	2.0	2.0	1.2
Co	QAFLOZ	4.5	3.0	6.6	3.8	5.9	8.5	5.7	4.6	8.8	4.4	8.0	7.0	3.8	8.1	7.9	3.2	4.1	4.0	1.9	2.5	5.3
	1FD5	8.6	8.4	9.2	9.5	9.2	8.8	8.6	7.8	9.1	8.5	8.8	8.7	9.5	8.3	8.2	8.4	8.3	8.5	8.7	8.5	8.7
	AZINAA	6.6	6.4	7.1	7.3	6.6	7.7	7.5	6.3	5.1	6.2	7.6	7.4	3.0	7.2	6.7	7.0	5.7	6.1	3.7	3.7	6.2
	5FHJ	7.8	8.2	9.0	8.6	8.7	9.0	9.2	6.6	8.8	8.3	9.0	9.2	6.6	8.9	9.1	8.5	8.0	7.9	8.3	8.2	8.4
Ni	ZITWOO	4.1	4.8	5.2	5.0	5.2	4.5	5.3	3.5	5.1	3.9	4.6	5.2	4.8	4.2	5.1	4.2	3.8	3.9	5.1	4.0	4.6
	JEXYOD	2.4	2.0	3.7	3.0	3.5	2.3	2.3	2.0	2.8	2.0	2.4	2.4	3.4	2.0	2.1	1.9	2.1	2.1	1.9	2.0	2.4
	MELTON	3.4	4.1	4.6	4.2	4.7	4.1	5.0	3.0	4.2	3.2	4.1	4.8	3.3	3.7	3.8	3.3	3.4	3.5	4.1	3.1	3.9
	JENGEO	1.2	0.7	0.8	1.7	0.9	0.6	0.2	0.6	0.6	0.7	0.5	0.3	2.3	0.5	0.3	0.6	0.8	0.8	0.8	0.8	0.8
Cu	1D39	8.6	8.4	10.8	20.3	10.7	20.3	18.9	7.8	10.5	8.0	19.6	10.6	8.4	10.5	9.2	8.7	8.2	8.3	10.6	10.8	11.5
	AMPCUS	5.3	7.6	6.7	4.0	5.2	4.1	5.5	8.6	8.6	20.6	4.8	5.3	7.3	5.2	6.8	6.1	6.3	9.4	8.4	8.0	7.2
	JUNGIK	9.1	8.4	9.7	10.6	9.2	9.2	8.3	3.4	4.5	5.7	9.3	8.8	10.3	8.9	7.8	8.6	8.5	8.4	8.9	10.6	8.4
	AMADCS	4.2	4.6	5.9	5.8	5.4	5.2	5.2	3.8	4.2	3.9	5.2	5.1	4.2	4.8	4.8	4.1	4.2	4.3	4.1	3.9	4.6
	FUDQUR	2.8	2.0	2.9	3.9	2.8	2.8	2.0	1.9	3.0	2.4	2.9	2.3	2.5	2.8	2.1	2.6	2.1	1.9	2.0	2.1	2.5
Zn	FAJZAT	1.8	1.6	1.9	2.7	2.0	1.9	2.0	1.5	2.5	1.2	1.9	1.8	1.6	1.7	1.8	1.3	1.4	1.4	1.6	1.3	1.7
	AWUBAW	5.5	5.5	6.2	6.2	6.3	5.8	6.0	4.4	5.7	5.1	5.9	18.2	5.2	5.4	7.9	5.3	4.8	5.1	4.8	4.3	6.2
	5JVW	12.4	11.4	13.6	14.9	13.1	12.8	12.1	9.4	15.2	5.6	12.8	12.3	6.5	12.0	11.5	5.8	11.0	11.1	5.6	5.4	10.7
Pd	FODDIN	1.7	1.0	0.9	2.4	1.5	1.1	0.4	0.4	0.7	0.8	0.9	0.6	0.3	1.1	0.5	0.9	1.2	0.8	0.2	0.3	0.9
	PUBPOR	3.0	3.7	3.7	3.9	4.0	2.5	3.4	1.7	3.4	2.0	2.3	2.0	0.8	2.3	1.8	2.1	2.2	2.4	1.2	1.3	2.5
Ag	5IX7	3.4	3.8	3.7	3.2	3.7	4.6	5.0	5.0	4.1	3.9	4.5	4.7	4.7	4.4	4.6	4.2	3.6	3.6	4.5	4.4	4.2
	COLDAL	2.3	1.2	1.6	1.7	1.3	0.9	1.0	1.3	2.5	2.0	0.9	0.8	1.8	1.3	1.1	1.5	2.6	2.4	1.2	2.6	1.6
	BUDDUA	1.3	0.7	0.7	1.1	0.5	0.5	1.1	0.5	0.8	1.2	0.5	0.7	0.4	0.5	0.5	0.5	1.3	1.3	0.4	0.5	0.8
	6LBW	2.3	2.7	2.6	2.1	2.6	3.4	3.9	3.8	2.9	0.8	3.3	3.6	1.7	3.1	3.5	1.7	2.4	2.5	2.1	1.8	2.6
Cd	BORFOE	5.4	4.2	6.4	15.1	19.4	23.8	13.1	3.7	13.9	4.6	22.3	13.9	2	14.6	8.1	24.6	11.9	3.8	1.9	1.7	10.7
	FAJYIA	4.8	2.9	3.7	6.1	3.6	4.2	2.8	2.5	4.9	2.8	4.3	3.1	1.1	3.8	2.9	3.6	3.1	2.7	1.0	1.1	3.3

	WIPRUL	4.4	3.9	4.5	5.5	4.5	4.4	3.7	3.1	4.5	3.8	4.4	4.0	1.9	4.2	3.8	4	3.7	3.6	1.7	1.8	3.8
Pt	4I1G	6.5	7.0	9.8	9.3	10.3	10.6	10.9	8.3	9.9	7.2	10.8	11.0	28.1	9.6	10.3	9.2	26.8	7.5	27.4	27.4	12.9
	DEGJUU	2.8	2.4	2.4	3.6	3.2	2.4	2.1	1.5	1.9	2.0	2.2	2.0	1.7	2.5	2.2	2.2	2.4	2.2	1.4	1.6	2.2
	SUTTAC	1.4	1.3	1.1	2.3	1.7	1.1	1.0	1.3	1.1	1.3	1.2	1.1	0.8	1.3	1.1	1.1	1.3	1.1	1.3	1.3	1.3
	GIKTIE	2.4	1.7	1.6	3.3	2.5	1.8	1.0	0.9	1.0	1.5	1.6	1.2	0.5	1.9	1.3	1.6	2.0	1.5	0.7	1.0	1.6
Au	ZOBTAO	2.5	2.1	2.0	3.1	2.5	1.8	1.5	0.8	1.7	1.7	1.7	1.5	1.1	1.8	1.5	1.7	2.2	2.0	1.3	1.6	1.8
	DOMQAZ	2.6	2.3	3.2	4.6	3.8	3.0	2.6	1.2	2.0	1.5	2.9	2.6	1.0	2.9	2.5	2.2	1.3	1.3	1.1	1.2	2.3
	AHOFEK	2.5	2.3	3.0	4.0	3.6	2.6	2.4	0.7	2.0	1.4	2.5	2.3	0.5	2.5	2.3	2	1.3	1.5	0.7	0.8	2.
Pb	MEHSEY	3.4	3.1	3.5	4.2	3.7	2.8	4.3	2.2	3.7	2.3	2.9	2.7	1.9	2.5	1.8	2.4	2.7	2.7	1.1	1.6	2.8
AMPE[b]		5.1	4.9	5.6	6.3	5.7	6.1	5.7	4.5	5.6	4.8	6.0	5.8	4.9	5.5	5.3	5.2	5.4	4.8	4.7	4.8	5.3

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average MPE calculated across all functionals for a given complex or across all complexes for a given functional.

Table D.20. Comparison of inner-shell MPEs (%), Max PEs (%), and ARMSDs (Å), and all-heavy-atom ARMSDs (Å), and Max RMSDs (Å) from all-electron DFT optimizations in the gas phase and implicit water over the first-row transition metal complexes.[a]

	Gas Phase						Implicit Water					
	Inner Shell				All Heavy Atom		Inner Shell				All Heavy Atom	
Functional	MPE	Max PE	ARMSD	Max RMSD	ARMSD	Max RMSD	MPE	Max PE	ARMSD	Max RMSD	ARMSD	Max RMSD
B3LYP	8.3	103.8	0.279	1.235	0.359	1.038	6.4	39.2	0.203	0.581	0.332	0.733
B3LYP-D3	7.3	106.0	0.226	0.918	0.321	0.982	6.3	36.8	0.202	0.578	0.321	0.816
B97D3	8.2	75.2	0.267	0.955	0.386	1.092	6.9	37.9	0.211	0.587	0.361	0.902
BLYP	9.3	94.2	0.308	1.449	0.381	1.138	7.0	41.3	0.251	1.268	0.355	0.951
BLYP-D3	8.4	72.2	0.284	0.940	0.353	0.849	6.8	34.5	0.216	0.572	0.343	0.749
BP86	9.0	89.3	0.331	1.414	0.391	1.113	6.6	38.3	0.254	1.267	0.386	0.948
BP86-D3	8.4	71.7	0.282	0.918	0.378	0.848	6.7	32.4	0.244	0.963	0.380	0.786
MN15	6.4	37.5	0.195	0.382	0.273	0.561	5.7	33.6	0.194	0.387	0.292	1.074
MN15-L	7.6	35.0	0.223	0.395	0.315	0.764	6.9	33.4	0.206	0.382	0.296	0.719
PW6B95	7.3	95.8	0.252	0.923	0.332	0.971	6.5	89.3	0.232	0.966	0.308	1.211
PBE	9.0	89.3	0.327	1.447	0.423	1.149	6.7	37.6	0.256	1.248	0.381	0.935
PBE-D3	8.4	72.5	0.285	0.927	0.374	0.851	6.7	33.1	0.252	0.972	0.392	1.312
PBEh-3c	6.9	51.0	0.210	0.622	0.303	0.712	6.2	47.2	0.192	0.588	0.260	0.573
TPSS	8.4	85.7	0.280	0.922	0.362	0.871	6.5	37.4	0.215	0.583	0.353	0.722
TPSS-D3	8.5	102.8	0.287	0.932	0.370	0.946	6.5	34.6	0.223	0.571	0.348	0.790
TPSSh	8.2	118.7	0.305	1.390	0.373	1.083	6.2	39.1	0.204	0.577	0.321	0.773
ωB97X	7.3	109.9	0.236	0.942	0.308	1.000	6.0	38.3	0.201	0.575	0.317	0.810
ωB97X-D	7.2	92.8	0.218	0.809	0.317	0.938	6.1	37.5	0.202	0.575	0.331	0.884
ωB97X-D3	6.3	34.6	0.204	0.406	0.282	0.700	6.1	39.4	0.191	0.585	0.273	0.603
ωB97X-V	6.3	46.5	0.215	0.605	0.294	0.688	6.0	48.1	0.191	0.614	0.269	0.590
Average[b]	7.7		0.254		0.341		6.5		0.217		0.331	

[a]Models for the complexes explored are shown in Figure 5.1. [b]Average ARMSD or MPE calculated across all functionals.

Table D.21. Comparison of inner-shell MPEs (%), Max PEs (%), and ARMSDs (Å), and all-heavy-atom ARMSDs (Å), and Max RMSDs (Å) from ECP DFT optimizations in the gas phase and implicit water over the first-row transition metal complexes.[a]

	Gas Phase						Implicit Water					
	Inner Shell				All Heavy Atom		Inner Shell				All Heavy Atom	
Functional	MPE	Max PE	ARMSD	Max RMSD	ARMSD	Max RMSD	MPE	Max PE	ARMSD	Max RMSD	ARMSD	Max RMSD
B3LYP	7.4	104.9	0.231	0.924	0.320	0.944	6.4	39.2	0.198	0.585	0.323	0.810
B3LYP-D3	7.3	106.7	0.226	0.921	0.314	0.992	6.3	36.8	0.199	0.573	0.310	0.821
B97D3	8.6	103.8	0.285	0.958	0.384	1.104	6.9	37.9	0.205	0.585	0.355	0.868
BLYP	8.1	87.7	0.255	0.958	0.345	0.794	7.0	41.3	0.209	0.577	0.325	0.735
BLYP-D3	7.4	72.7	0.227	0.837	0.346	0.828	6.8	34.5	0.207	0.564	0.328	0.759
BP86	7.9	86.5	0.252	0.901	0.338	0.811	6.6	38.3	0.206	0.570	0.348	0.734
BP86-D3	7.9	81.4	0.251	0.893	0.364	0.831	6.7	32.4	0.208	0.556	0.351	0.791
MN15	6.5	35.5	0.201	0.583	0.310	0.919	5.7	33.6	0.180	0.350	0.256	0.562
MN15-L	7.4	35.6	0.212	0.419	0.312	0.787	6.9	33.4	0.204	0.378	0.304	0.742
PW6B95	6.7	82.8	0.221	0.934	0.307	0.758	6.3	89.1	0.231	0.985	0.312	1.197
PBE	7.8	85.0	0.253	0.881	0.359	0.809	6.7	37.6	0.210	0.577	0.348	0.735
PBE-D3	7.5	74.6	0.231	0.890	0.354	0.813	6.7	33.1	0.212	0.559	0.362	0.853
PBEh-3c	6.8	49.4	0.204	0.603	0.298	0.685	6.0	43.8	0.207	0.868	0.301	1.449
TPSS	8.5	105.2	0.276	0.947	0.363	0.941	6.5	37.4	0.202	0.570	0.351	0.908
TPSS-D3	8.3	101.5	0.283	0.951	0.384	1.111	6.5	34.6	0.208	0.576	0.319	0.786
TPSSh	7.1	103.6	0.230	0.923	0.334	1.029	6.0	36.2	0.199	0.571	0.314	0.814
ωB97X	6.3	35.6	0.196	0.395	0.273	0.542	6.0	38.3	0.192	0.576	0.334	1.131
ωB97X-D	6.5	35.6	0.195	0.390	0.271	0.518	6.1	37.5	0.193	0.576	0.328	1.059
ωB97X-D3	6.3	43.4	0.203	0.572	0.285	0.688	5.9	38.5	0.186	0.575	0.276	0.671
ωB97X-V	6.1	41.7	0.197	0.532	0.283	0.678	5.8	38.6	0.185	0.570	0.272	0.666
Average[b]	7.3		0.231		0.327		6.4		0.202		0.321	

[a]Models for the complexes explored shown in Figure 5.1. [b]Average ARMSD or MPE calculated across all functionals.

Table D.22. Inner-shell MPEs (%) and AMPEs (%) from ECP DFT optimizations in the gas phase.[a]

		MPE																				AMPE[b]
Metal	Structure ID	B3LYP	B3LYP-D3	B97D3	BLYP	BLYP-D3	BP86	BP86-D3	MN15	MN15-L	PW6B95	PBE	PBE-D3	PBEh-3c	TPSS	TPSS-D3	TPSSh	ωB97X	ωB97X-D	ωB97X-D3	ωB97X-V	
Cr	XOSLUO	1.6	1.3	2.3	2.0	2.2	1.0	0.6	0.4	1.1	0.8	0.9	0.7	0.6	0.8	0.5	0.8	1.2	1.0	0.9	0.8	1.1
Mn	4DWY	15.3	16.0	15.9	15.6	16.3	15.8	16.2	14.9	17.0	15.5	16.0	16.3	15.4	15.5	16.0	15.2	15.7	16.2	15.7	15.6	15.8
	4HIG	18.6	18.1	17.4	18.4	18.0	18.1	17.5	17.2	18.1	17.6	18.2	17.9	20.4	18.6	18.3	18.8	17.8	17.9	17.8	17.6	18.1
	2OEU	11.7	11.7	11.4	11.9	11.6	12.2	12.0	11.3	11.8	11.3	12.3	11.8	12.4	12.1	12.5	11.9	11.7	11.9	11.7	11.1	11.8
Fe	WAWPUF	5.8	5.9	7.1	6.3	6.6	7.5	7.2	7.7	8.0	6.0	7.9	8.0	4.8	7.4	8.0	6.7	5.7	5.8	4.7	5.8	6.6
	HOPJAY	5.6	5.9	6.3	5.9	6.2	6.4	6.5	6.4	6.7	5.9	6.5	6.4	5.7	6.3	6.3	6.1	5.6	5.7	6.0	5.9	6.1
	MORHEJ	5.4	6.1	7.4	5.8	6.7	7.3	7.9	7.5	7.5	6.0	7.4	7.8	5.5	7.2	7.7	6.8	5.2	5.6	5.4	5.8	6.6
	FOJPOM	3.3	2.0	0.1	2.4	1.1	0.6	0.4	0.5	0.2	2.5	0.4	0.2	2.5	0.7	0.2	1.2	3.7	2.9	2.6	2.6	1.5
Co	QAFLOZ	4.0	5.5	8.1	6.5	7.8	7.0	8.2	2.9	9.3	4.1	7.8	7.8	4.2	6.6	7.7	5.6	3.7	5.3	4.3	4.1	6.0
	1FD5	8.9	8.7	9.4	9.6	9.3	8.9	8.7	8.6	9.4	8.4	9.0	8.8	9.6	8.5	8.3	8.4	8.6	8.5	8.7	8.5	8.8
	AZINAA	5.8	5.4	6.3	6.0	5.9	6.3	6.3	5.2	4.7	5.3	6.2	6.2	3.3	5.8	5.0	5.6	3.6	4.6	4.2	4.2	5.3
	5FHJ	7.0	7.3	8.2	7.7	7.9	8.0	8.1	7.5	8.4	7.1	8.1	8.2	6.8	7.8	7.9	7.2	7.0	7.2	7.3	7.2	7.6
Ni	ZITWOO	4.6	4.8	5.3	5.3	5.1	4.8	5.1	3.7	5.6	4.4	4.9	5.1	5.5	4.4	4.9	4.5	4.3	4.4	5.2	4.6	4.8
	JEXYOD	2.7	3.5	3.5	3.0	3.5	2.4	2.4	2.0	3.1	2.3	2.4	2.4	3.4	2.3	2.2	2.3	2.5	2.4	2.4	2.3	2.7
	MELTON	3.5	4.3	4.6	4.1	4.6	3.8	4.0	2.7	4.5	3.3	3.9	4.0	4.1	3.5	3.6	3.3	3.5	3.6	4.5	3.4	3.8
	JENGEO	1.8	1.2	1.0	1.9	1.1	0.8	0.3	0.4	1.1	1.1	0.8	0.4	2.5	0.7	0.3	1.0	1.4	1.3	1.1	1.1	1.1
Cu	1D39	8.4	8.3	18.1	20.9	9.4	20.6	20.0	8.6	9.8	8.1	20.4	10.4	8.5	20.6	18.6	8.2	9.3	9.4	8.5	8.4	12.7
	AMPCUS	25.0	25.3	24.2	6.5	7.2	6.8	7.6	8.5	10.6	10.5	6.5	6.8	8.8	24.9	24.4	24.9	9.2	9.1	9.4	8.9	13.3
	JUNGIK	9.5	6.9	16.9	19.3	15.0	17.6	19.4	8.7	6.8	19.5	16.4	19.6	11.8	19.5	17.7	7.5	6.5	6.8	10.1	9.4	13.2
	AMADCS	4.0	3.8	7.9	7.7	7.1	6.8	6.6	12.6	5.8	3.7	6.8	6.6	7.8	6.4	6.2	5.7	3.8	3.9	3.9	3.8	6.0
	FUDQUR	4.0	3.2	4.4	5.3	4.1	4.2	3.4	3.3	4.0	4.2	4.2	3.7	3.8	4.0	3.4	4.2	3.3	3.0	3.5	3.4	3.8
Zn	FAJZAT	2.2	2.1	1.8	2.6	2.1	2.4	2.4	2.1	2.8	1.9	2.3	2.2	2.2	2.3	2.2	1.9	1.9	1.9	2.2	0.1	2.1
	AWUBAW	6.7	6.5	7.2	7.0	6.9	6.5	6.5	4.1	6.9	5.9	6.6	6.5	7.0	6.0	6.6	5.8	6.0	6.2	6.4	6.1	6.4
	5JVW	12.8	12.0	13.4	14.8	13.4	13.0	12.4	8.8	13.7	5.7	12.9	12.5	6.7	12.1	11.5	5.8	11.7	11.9	6.0	5.8	10.8
AMPE[b]		7.4	7.3	8.7	8.2	7.5	7.9	7.9	6.5	7.4	6.7	7.9	7.5	6.8	8.5	8.3	7.1	6.4	6.5	6.4	6.1	7.3

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average MPE calculated across all functionals for a given complex or across all complexes for a given functional.

Table D.23. Inner-shell MPEs (%) and AMPEs (%) from ECP DFT optimizations in implicit water.[a]

		MPE																				AMPE[b]
Metal	Structure ID	B3LYP	B3LYP-D3	B97D3	BLYP	BLYP-D3	BP86	BP86-D3	MN15	MN15-L	PW6B95	PBE	PBE-D3	PBEh-3c	TPSS	TPSS-D3	TPSSh	ωB97X	ωB97X-D	ωB97X-D3	ωB97X-V	
Cr	XOSLUO	1.6	2.3	2.3	2.2	2.2	1.1	0.7	0.5	1.2	0.8	1.1	0.8	0.5	0.9	0.6	0.8	1.2	1.0	0.9	0.8	1.2
Mn	4DWY	16.0	16.2	15.5	16.3	16.5	16.5	16.7	15.5	16.7	15.9	16.7	16.8	14.8	16.3	16.5	15.9	15.9	16.0	16.5	16.5	16.2
	4HIG	17.9	17.5	17.1	18.3	17.8	17.6	17.3	16.9	17.6	17.0	17.5	17.4	19.2	18.3	18.1	18.4	17.1	17.4	17.3	17.0	17.6
	2OEU	11.6	12.1	11.8	12.1	11.9	12.1	11.7	10.9	11.5	11.5	12.0	12.0	12.3	12.3	12.0	12.2	11.5	11.5	11.5	11.5	11.8
Fe	WAWPUF	6.4	5.8	8.0	7.0	6.6	8.1	8.6	8.2	7.6	6.0	8.4	8.6	4.8	8.2	8.5	6.7	6.2	6.8	4.7	5.8	7.1
	HOPJAY	5.8	6.0	6.5	6.1	6.3	6.6	6.6	6.4	6.9	6.0	6.6	6.8	5.8	6.5	6.5	6.3	5.7	5.8	5.9	5.9	6.3
	MORHEJ	6.0	6.6	8.0	6.4	7.2	6.4	8.4	8.0	8.1	6.5	8.0	8.3	6.2	7.8	8.2	7.4	5.7	6.1	6.4	6.4	7.1
	FOJPOM	3.0	1.8	0.3	2.1	0.9	0.4	0.6	0.3	0.4	2.2	0.2	0.4	2.2	0.5	0.4	1.0	3.5	2.7	2.4	2.4	1.4
Co	QAFLOZ	3.5	3.5	5.5	6.2	6.2	6.7	6.6	2.4	8.3	2.7	6.4	6.4	3.2	6.7	6.7	3.2	2.1	2.1	1.5	1.8	4.6
	1FD5	8.6	8.4	9.2	9.5	9.2	8.8	8.6	7.8	9.1	8.2	8.8	8.7	8.8	8.3	8.2	8.0	8.3	8.5	8.5	8.3	8.6
	AZINAA	5.3	5.3	5.8	5.4	5.5	5.8	5.9	5.5	4.6	5.0	5.7	5.7	2.7	5.4	5.5	5.1	3.4	3.8	3.4	3.5	4.9
	5FHJ	7.1	7.5	8.4	7.9	8.1	8.1	8.3	6.2	8.4	7.2	8.2	8.3	6.1	7.9	8.1	7.3	7.1	7.3	7.5	7.5	7.6
Ni	ZITWOO	3.9	4.5	5.1	4.7	4.9	4.2	5.1	3.1	4.7	3.7	4.2	5.0	4.6	3.9	4.9	3.8	3.6	3.7	4.9	3.9	4.3
	JEXYOD	2.3	1.9	3.6	2.7	3.5	2.1	2.0	1.4	2.4	2.0	2.1	2.0	2.4	1.9	1.9	1.9	2.1	2.0	2.0	2.0	2.2
	MELTON	3.0	3.8	4.2	3.7	4.2	3.5	3.6	2.3	3.7	2.8	3.6	3.7	3.0	3.1	3.2	2.9	3.0	3.1	3.1	2.8	3.3
	JENGEO	1.2	0.7	0.8	1.7	0.9	0.6	0.2	0.6	0.6	0.7	0.5	0.3	2.1	0.5	0.3	0.6	0.8	0.8	0.7	0.8	0.8
Cu	1D39	8.3	8.2	9.3	9.6	9.2	9.4	8.9	8.4	10.3	8.0	9.1	9.0	8.0	8.5	8.5	8.0	9.1	9.2	9.5	9.4	8.9
	AMPCUS	7.0	7.7	6.9	4.8	6.3	5.6	7.0	8.6	10.2	20.5	5.7	6.9	7.4	6.4	7.0	7.3	7.5	7.9	8.0	7.9	7.8
	JUNGIK	9.5	8.4	9.5	10.4	9.0	9.0	8.3	5.8	4.0	8.2	9.1	8.5	9.9	8.6	8.2	8.2	8.6	8.5	9.0	9.0	8.5
	AMADCS	4.3	4.4	5.9	5.9	5.2	5.2	5.1	4.3	4.3	4.2	4.8	4.9	4.4	4.3	4.6	4.2	4.4	4.6	4.2	4.1	4.7
	FUDQUR	3.1	2.3	3.1	4.1	3.0	3.1	2.3	2.5	2.9	2.4	3.1	2.6	2.7	3.0	2.4	2.9	2.4	2.2	2.4	2.4	2.7
Zn	FAJZAT	1.5	1.6	1.8	2.3	1.7	1.7	1.9	1.8	2.6	1.3	1.6	1.8	1.6	1.6	1.8	1.3	1.3	1.5	1.6	1.3	1.7
	AWUBAW	4.8	4.8	5.4	5.5	5.5	5.3	5.2	2.7	5.5	4.5	5.3	5.2	4.7	4.9	4.3	4.5	4.7	4.8	4.4	4.0	4.8
	5JVW	11.3	10.4	12.5	13.5	12.2	11.7	11.0	7.2	14.0	5.0	11.6	11.4	6.0	10.9	10.4	5.1	10.0	10.2	5.3	5.1	9.7
AMPE[b]		6.4	6.3	6.9	7.0	6.8	6.7	6.7	5.7	6.9	6.3	6.7	6.7	6.0	6.5	6.5	6.0	6.1	6.1	5.9	5.8	6.4

[a]Models for the 41 complexes explored are highlighted in black in Figure 5.1. [b]Average MPE calculated across all functionals for a given complex or across all complexes for a given functional.

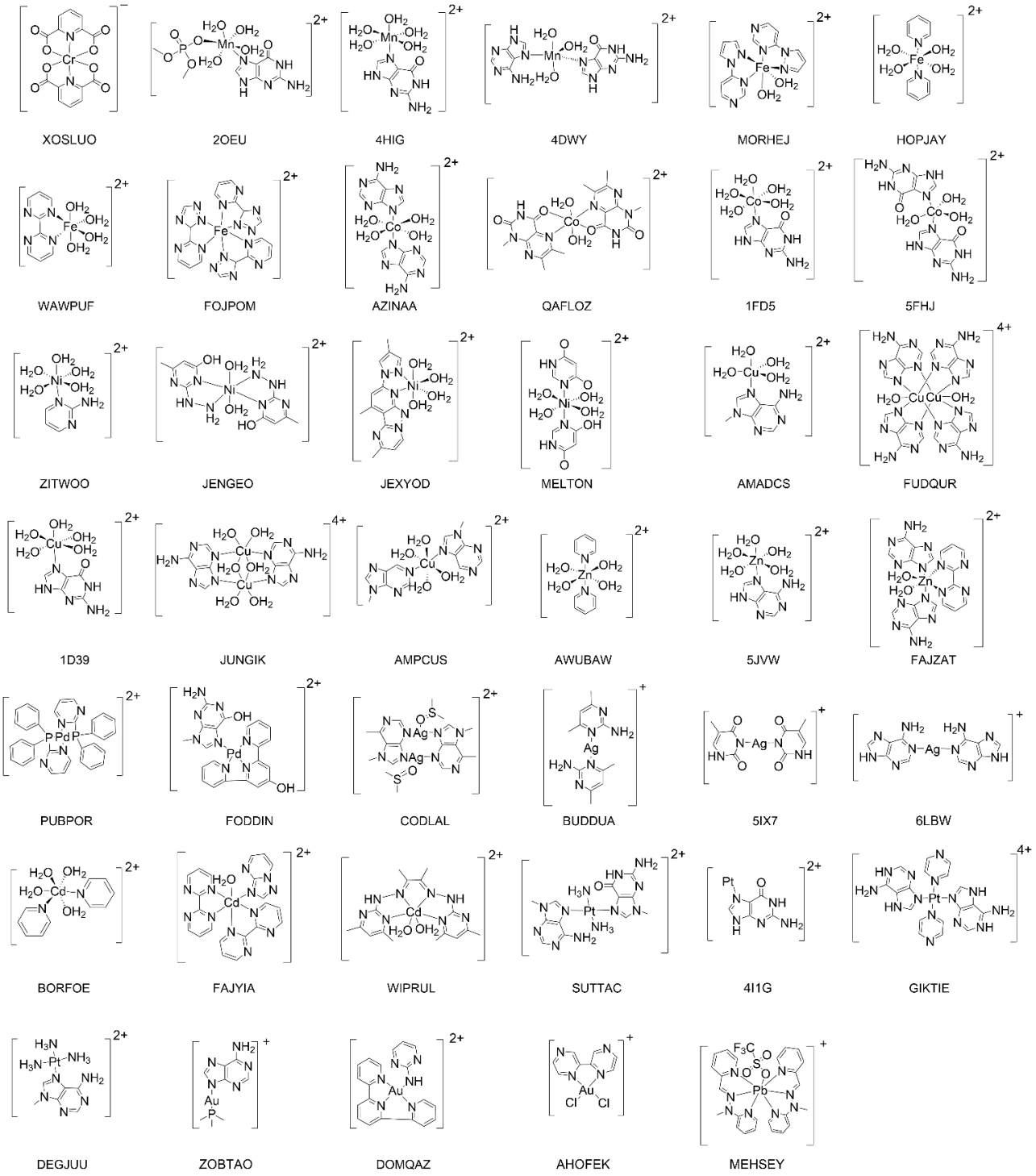


Figure D.1. Models of representative transition and post-transition metal–nucleic acid complexes in the 41-test set from the PDB and CSD investigated in the present work.

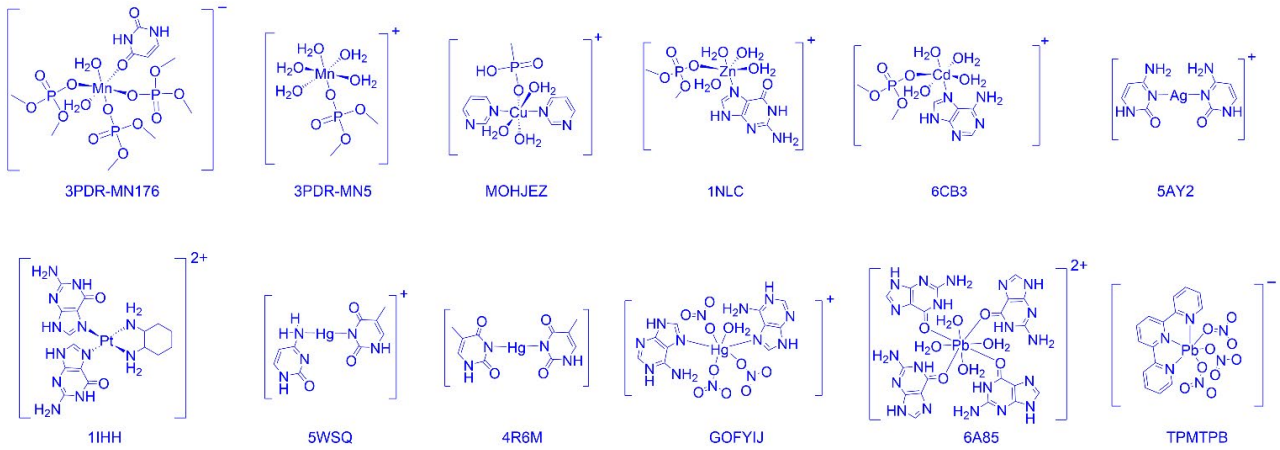


Figure D.2. Models of representative transition and post-transition metal–nucleic acid complexes in the 12 test set from the PDB and CSD that were analyzed separately due to the inability of all functionals to reproduce experimental geometries.

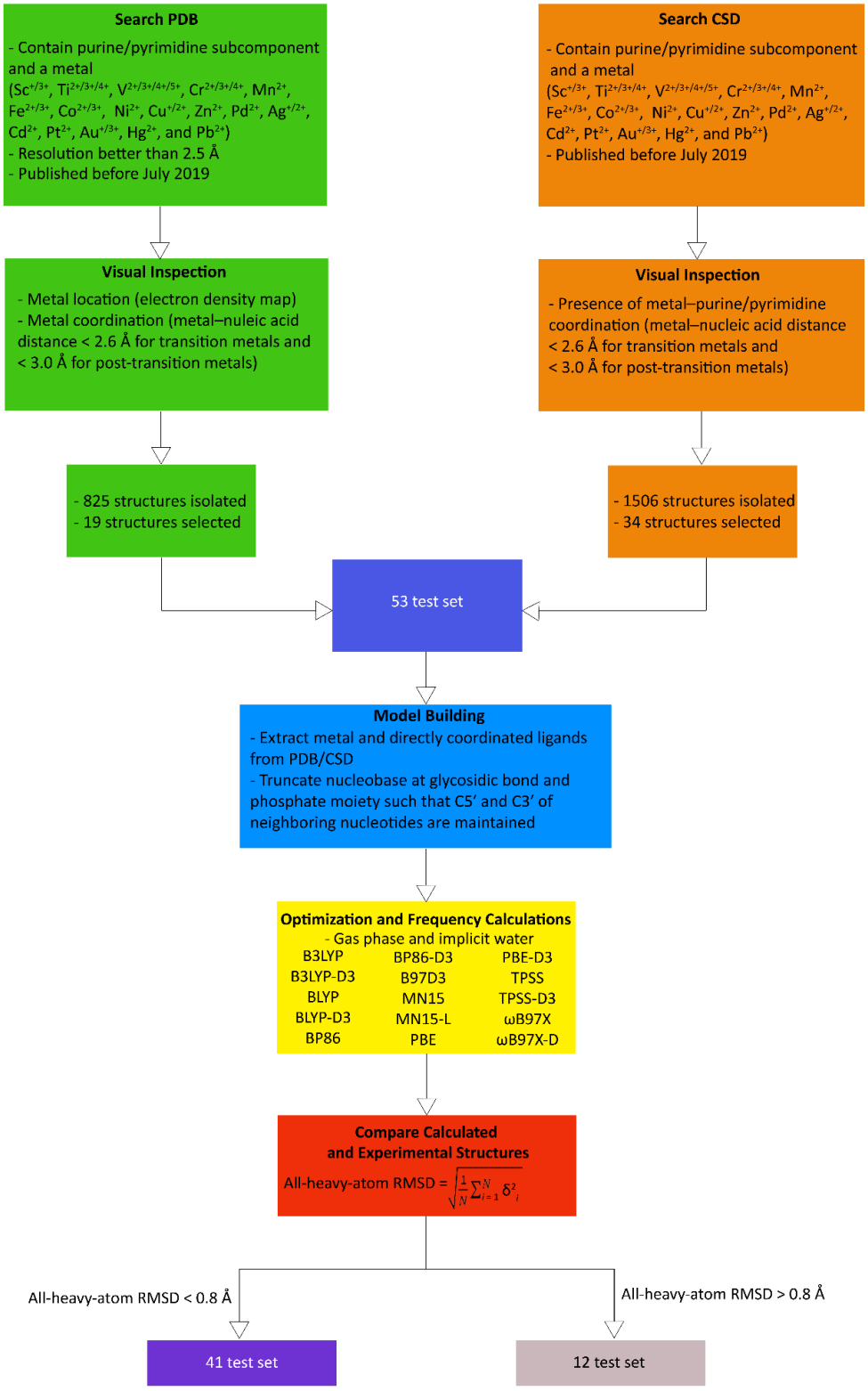


Figure D.3. A schematic of the general workflow for the generation of our test sets.

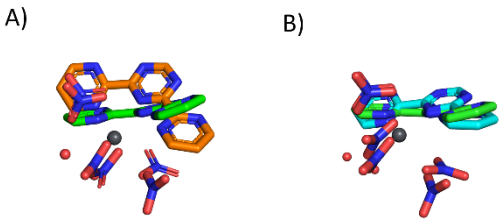


Figure D.4. Overlay of the Pb^{2+} complex (TPMTPB) from the reference crystal structure (green) and a gas-phase MN15 A) def2-SVP (orange) or B) def2-SVPD (cyan) optimization, which demonstrates that the crystal structure geometry cannot be reproduced regardless of basis set.

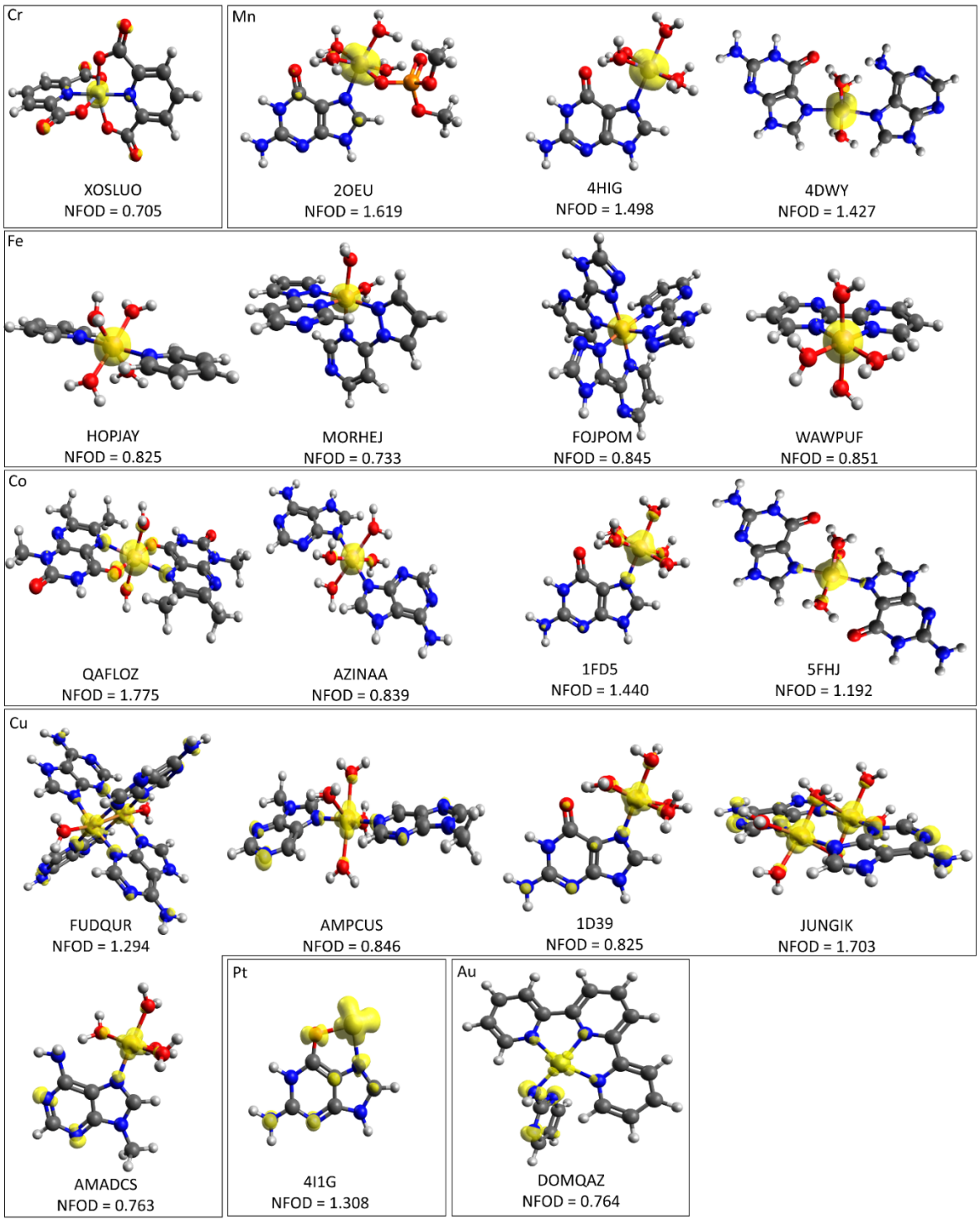


Figure D.5. FOD analysis (σ = 0.005 e Bohr^{-3}, TPSS/def2-TZVP, T_{el} = 5000 K) of 19 complexes in our test set that have the strongest static electron correlation effects (NFOD > 0.7). Areas highlighted in yellow represent static electron correlation.

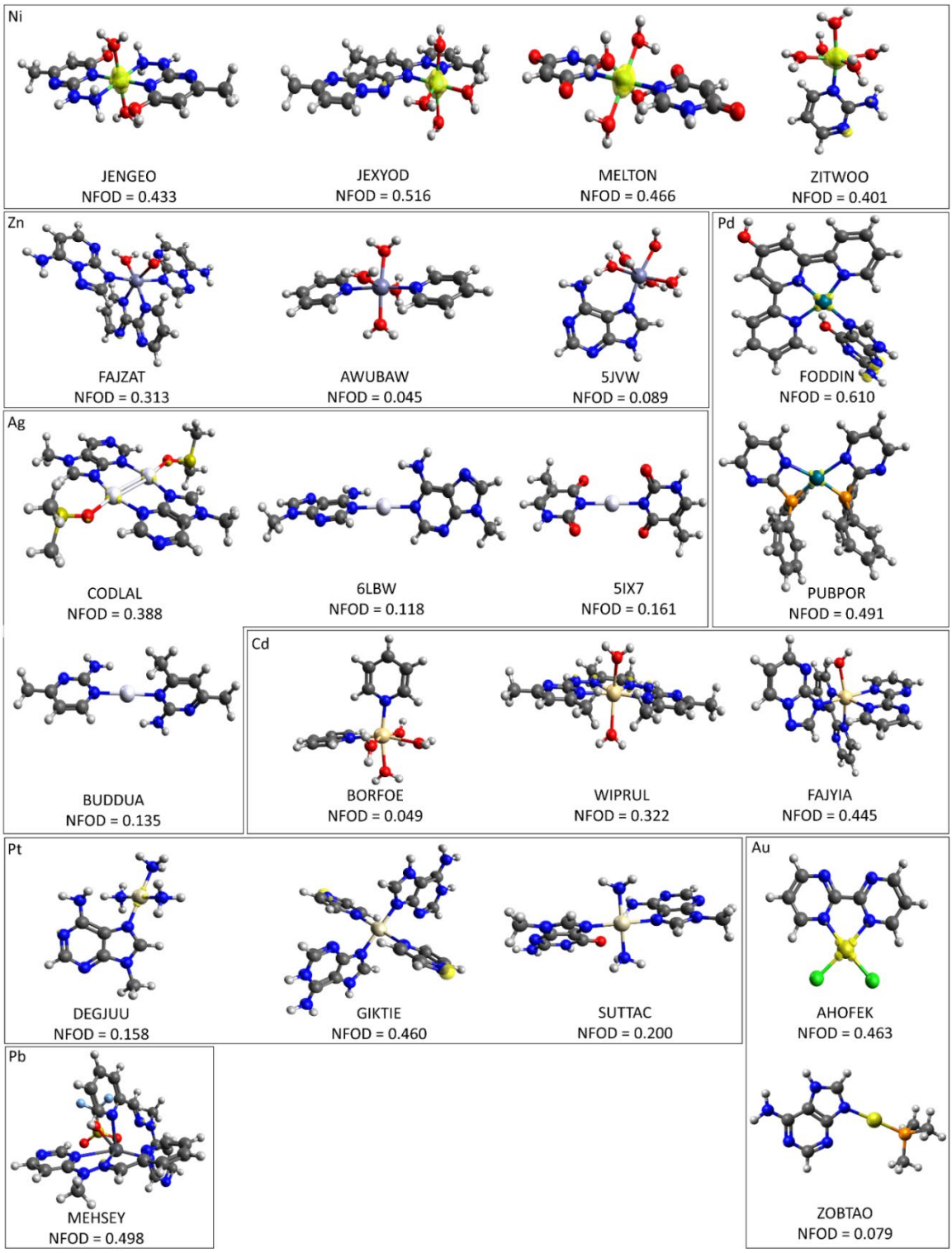


Figure D.6. FOD analysis (σ = 0.005 e Bohr^{-3}, TPSS/def2-TZVP, T_{el} = 5000 K) of 22 complexes in our test set identified to have the weakest static electron correlation effects (NFOD < 0.7). Areas highlighted in yellow represent static electron correlation.

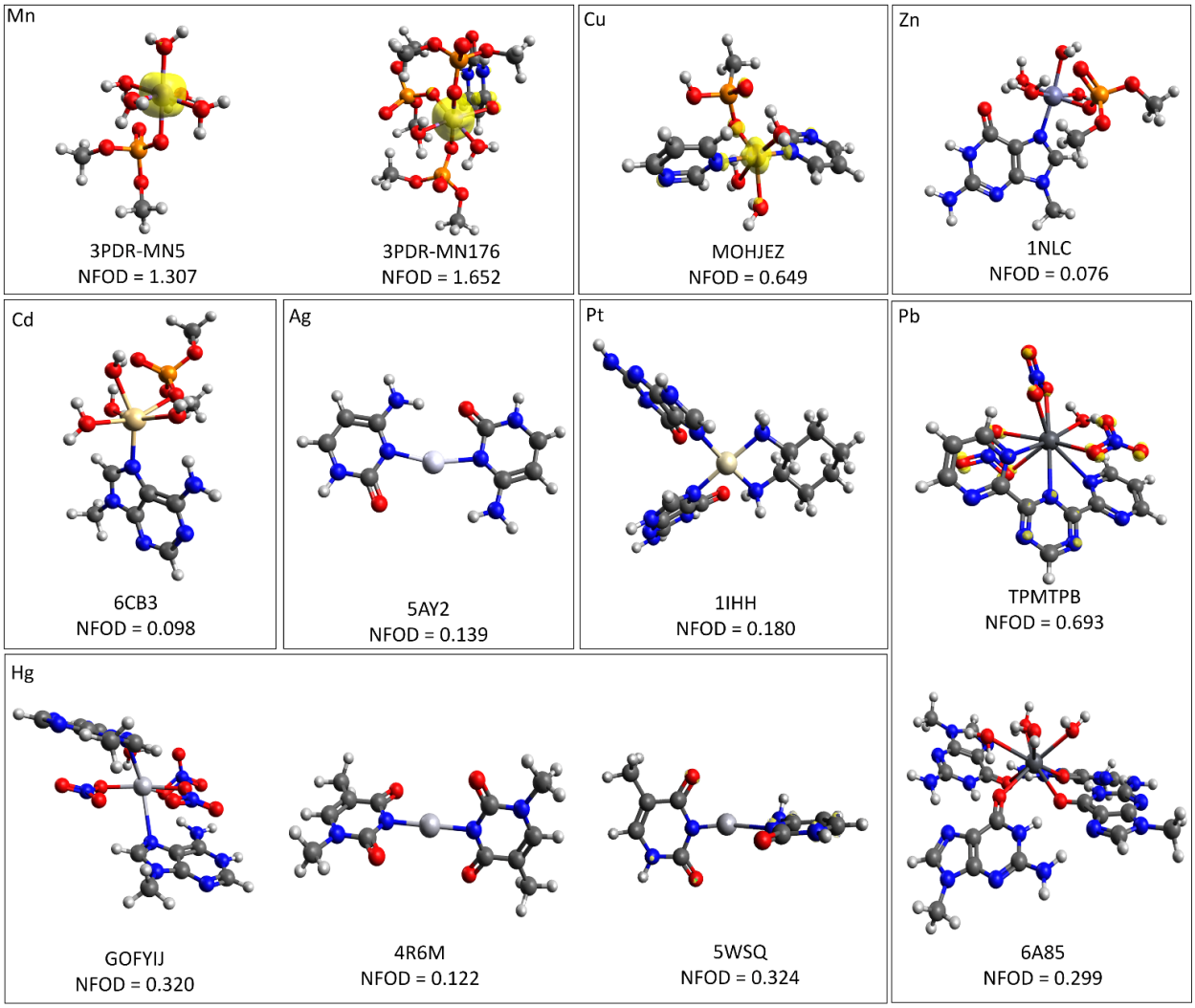


Figure D.7. FOD analysis (σ = 0.005 e Bohr^{-3}, TPSS/def2-TZVP, T_{el} = 5000 K) of the 12 complexes in our 53 test set that DFT cannot reproduce the crystallographic structures, which are also highlighted in blue in Figure 1. Areas highlighted in yellow represent static electron correlation.

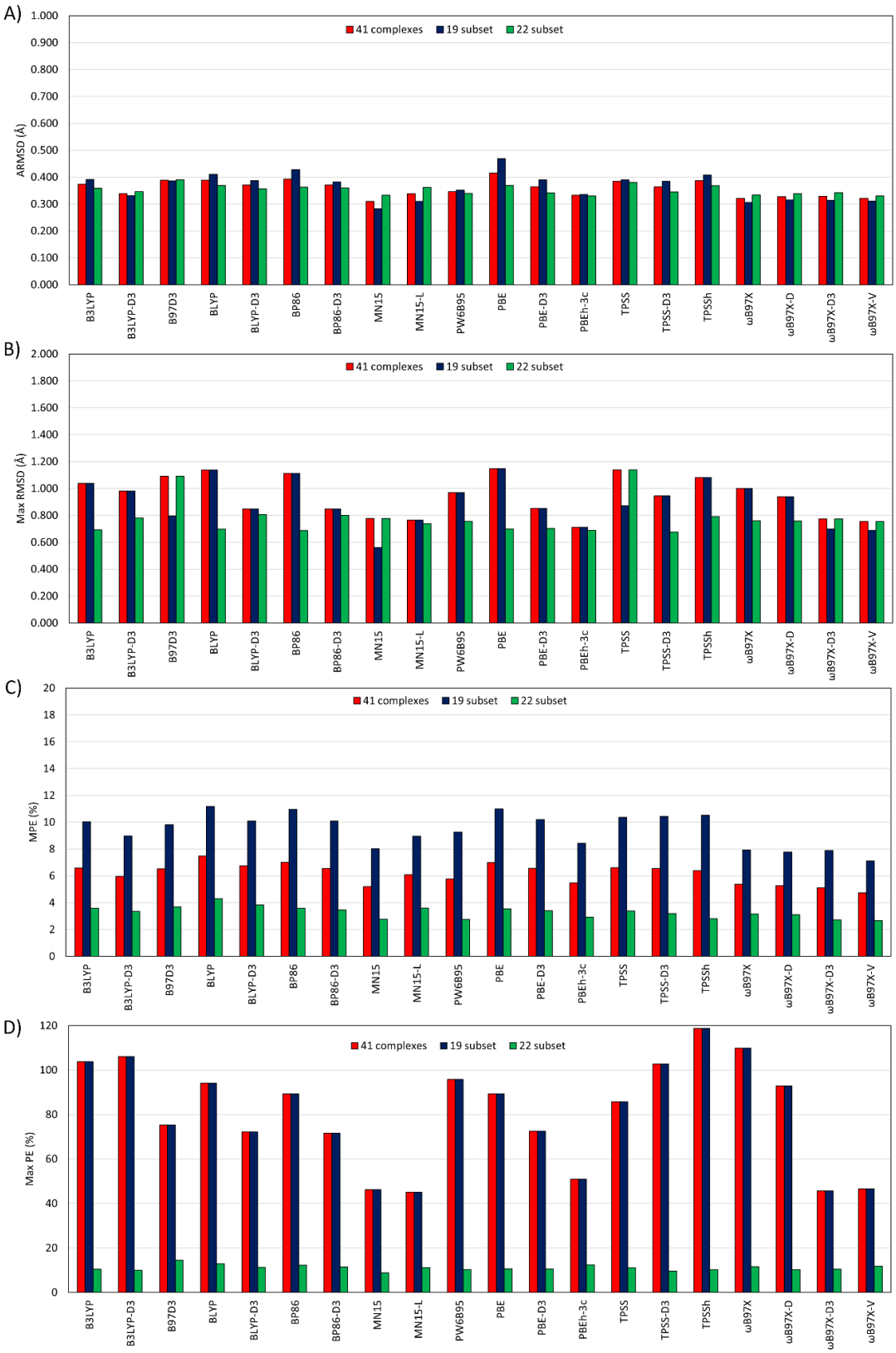


Figure D.8. All-heavy-atom A) average RMSDs (Å) and B) maximum RMSDs (Å), and inner-shell C) mean percent errors (%) and D) maximum percent errors (%) relative to

crystal structure references for each functional from all-electron gas-phase optimizations over the 41 test set, the 19 subset of complexes with high multi-reference character (NFOD > 0.7), and the 22 subset of complexes with low multi-reference character (NFOD < 0.7).

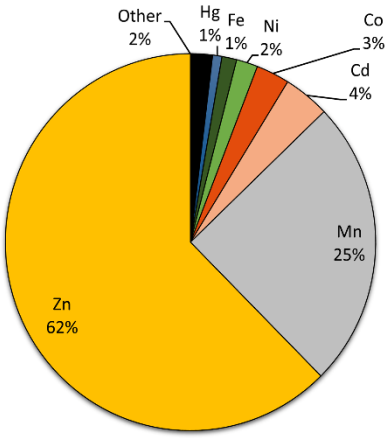


Figure D.9. Metal distribution across all nucleic acid complexes in the PDB (2076 structures). 'Other' represents the combined total of $V^{2+/3+/4+/5+}$, $Cr^{2+/3+/4+}$, $Au^{+/3+}$, $Ag^{+/2}$, Pb^{2+}, $Cu^{+/2+}$, and Pt^{2+}.

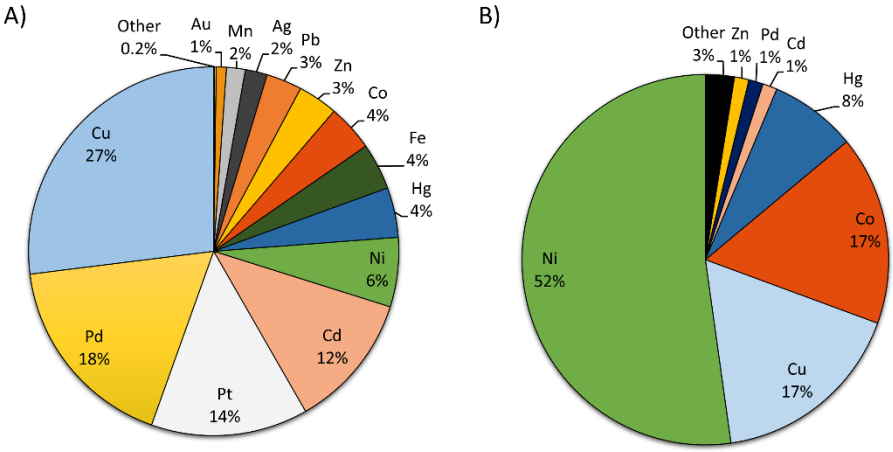


Figure D.10. Metal distribution of A) 1349 pyrimidine complexes, with 'Other' representing the combined total of $Cr^{2+/3+/4+}$ and $V^{2+/3+/4+/5+}$, or B) 157 purine complexes, with 'Other' representing the combined total of Mn^{2+}, $Ag^{+/2+}$, $Fe^{2+/3+}$, and Pt^{2+} in the CSD.

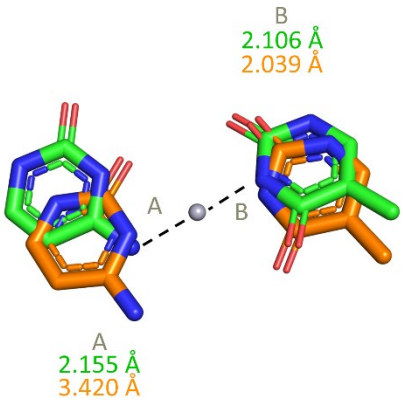


Figure D.11. Overlay of the Hg^{2+} complex (5WSQ) from the reference crystal structure (green) and a MN15 gas-phase optimization (orange), which demonstrates a change in coordination from N4 to N3 of C.

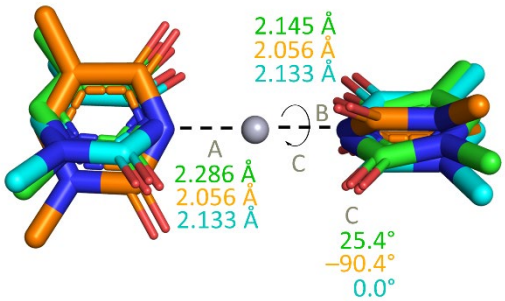


Figure D.12. Overlay of the Hg^{2+} complex (4R6M) from the reference structure (green) and a B3LYP gas-phase optimization (orange) or optimization in implicit water (cyan), which demonstrates implicit solvent is required to describe the ligand co-planarity. The dihedral angle corresponds to ∠(C2N3N3C2).

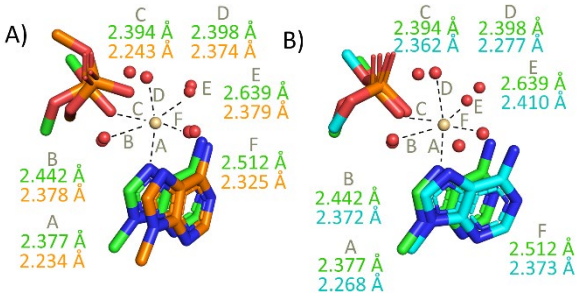


Figure D.13. Overlay of the Cd^{2+} complex (6CB3) from the reference crystal structure (green) and a MN15 A) gas-phase optimization (orange) or B) optimization in implicit water (cyan), which demonstrates implicit solvent is required to maintain the relative ligand orientation about the metal center.

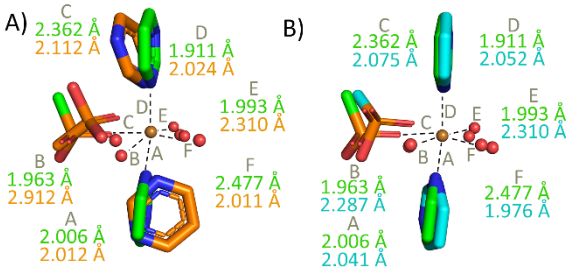


Figure D.14. Overlay of the Cu^{2+} complex (MOHJEZ) from the reference crystal structure (green) and a MN15 A) gas-phase optimization (orange) or B) optimization in implicit water (cyan), which demonstrates implicit solvent is required to maintain the relative ligand orientation about the metal center and inner shell coordination.

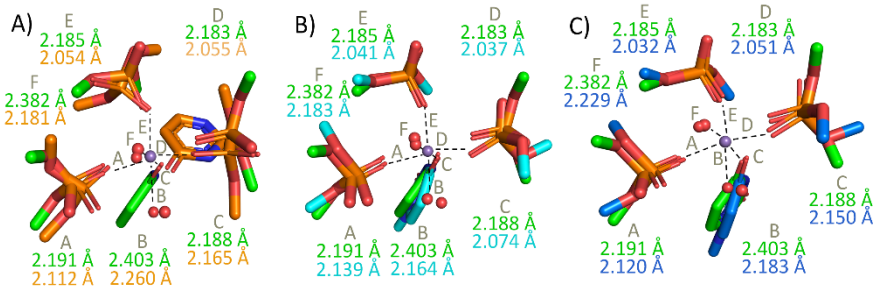


Figure D.15. Overlay of the Mn^{2+} complex (3PDR-MN176) from the reference crystal structure (green) and a B3LYP A) gas-phase optimization (orange), B) constrained gas-phase optimization (cyan) or C) constrained optimization in implicit water (blue), which demonstrates that constraints are required to maintain the relative ligand orientation about the metal center regardless of environment.

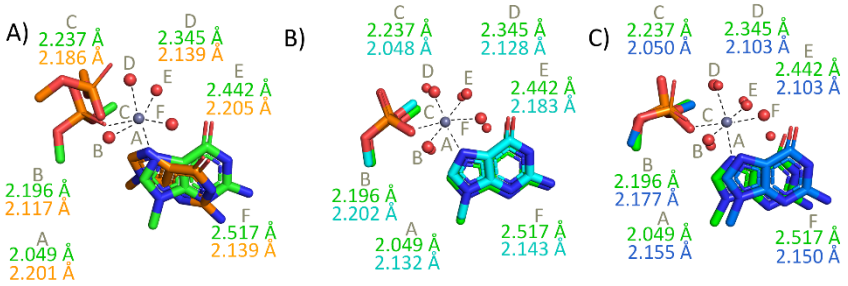


Figure D.16. Overlay of the Zn^{2+} complex (1NLC) from the reference crystal structure (green) and a B3LYP A) gas-phase optimization (orange), B) constrained gas-phase optimization (cyan) or C) constrained optimization in implicit water (blue), which demonstrates that constraints are required to maintain the relative ligand orientation about the metal center regardless of environment.

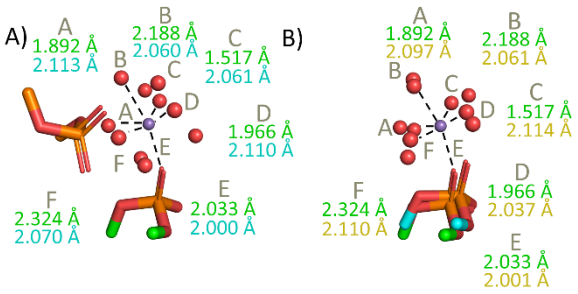


Figure D.17. Overlay of the Mn^{2+} complex (3PDR-MN5) from the reference crystal structure (green) and a MN15-L A) gas-phase optimization (orange) or B) constrained optimization in implicit water (cyan), which demonstrates that combining an implicit water environment with constraints is required to maintain the relative ligand orientation about the metal center.

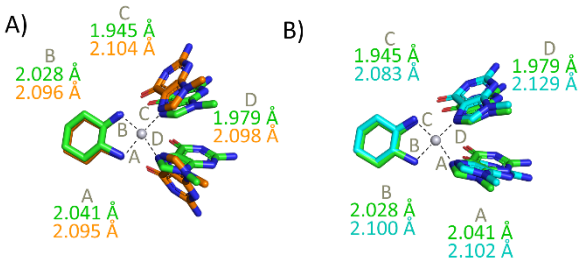


Figure D.18. Overlay of the Pt^{2+} complex (1IHH) from the reference crystal structure (green) and a BLYP constrained optimization in A) the gas phase (orange) or B) implicit water (cyan), which demonstrates that combining an implicit water environment with constraints is required to maintain the relative ligand orientation about the metal center.

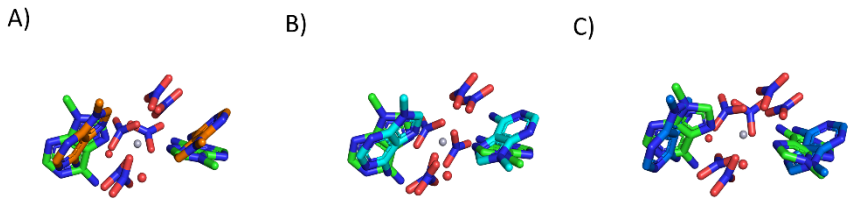


Figure D.19. Overlay of the Hg^{2+} complex (GOFYIJ) from the reference crystal structure (green) and a MN15 A) gas-phase optimization (orange), B) implicit water optimization (cyan) or C) constrained optimization in implicit water (blue), which demonstrates that the crystal structure geometry cannot be reproduced by DFT methods regardless of environment or constraints.

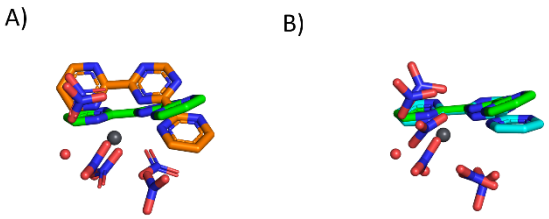


Figure D.20. Overlay of the Pb^{2+} complex (TPMTPB) from the reference crystal structure (green) and a MN15 A) gas-phase (orange) or B) implicit water (cyan) optimization, which demonstrates that the crystal structure geometry cannot be reproduced by DFT methods regardless of environment.

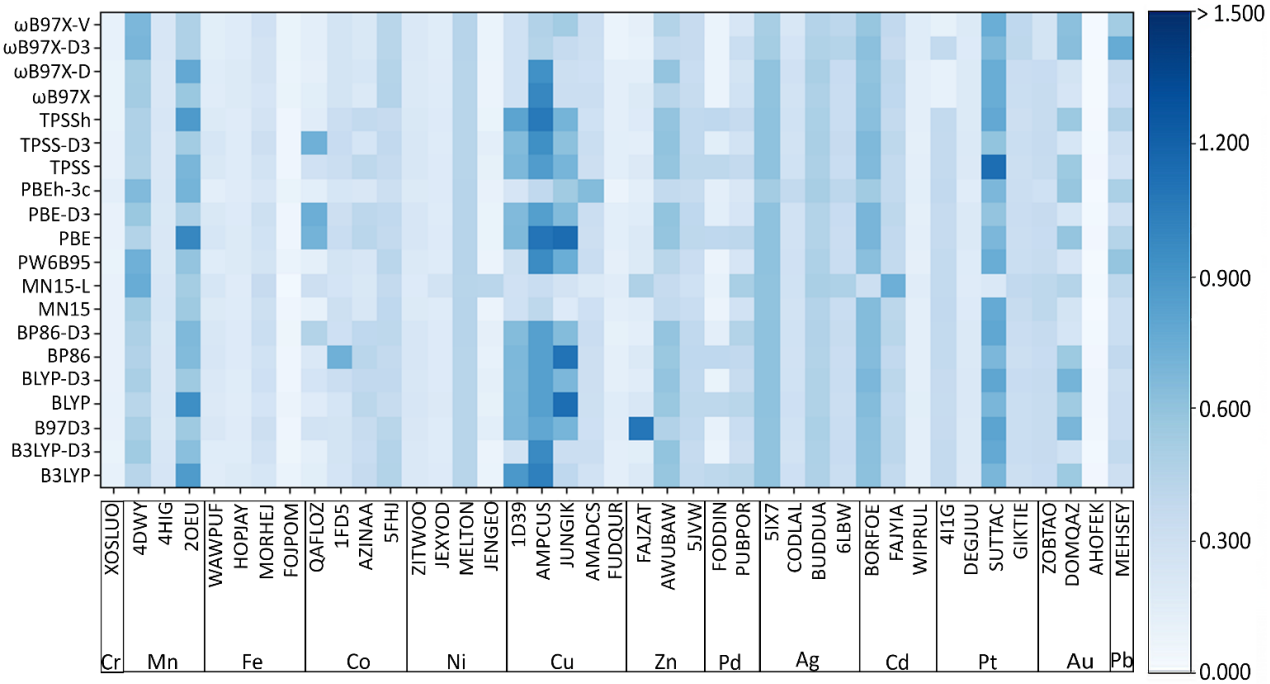


Figure D.21. All-heavy-atom RMSDs (Å) from all-electron gas-phase DFT optimizations with 20 functionals relative to crystal structure references for 41-representative metal–nucleic acid complexes considered in the present work.

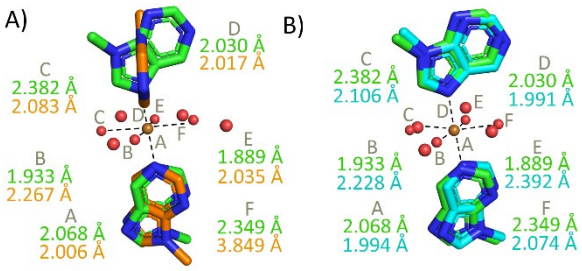


Figure D.22. Overlay of the Cu^{2+} complex (AMPCUS) from the reference crystal structure (green) and a A) B3LYP (orange) or B) MN15 (cyan) gas-phase optimization, which highlights the ability of MN15 to maintain the coordination geometry of Cu^{2+} complexes.

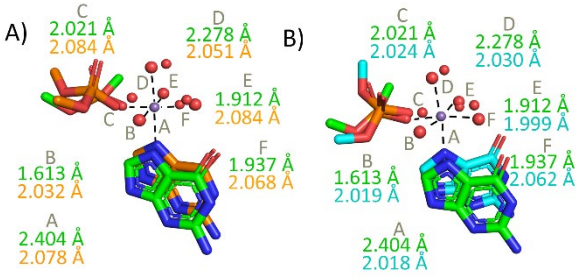


Figure D.23. Overlay of the Mn^{2+} complex (2OEU) from the reference crystal structure (green) and a A) B3LYP (orange) or B) MN15 (cyan) gas-phase optimization, which highlights the relative shift of the phosphate and nucleobase ligands.

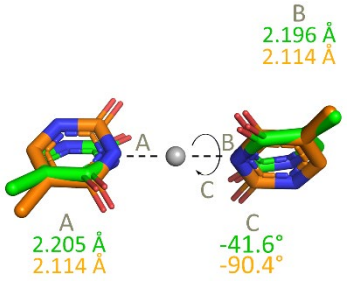


Figure D.24. Overlay of the Ag^{+} complex (5IX7) from the reference crystal structure (green) and a BP86 gas-phase optimization (orange), which highlights a change in ligand co-planarity. The dihedral angle corresponds ∠(C2N3N3C2).

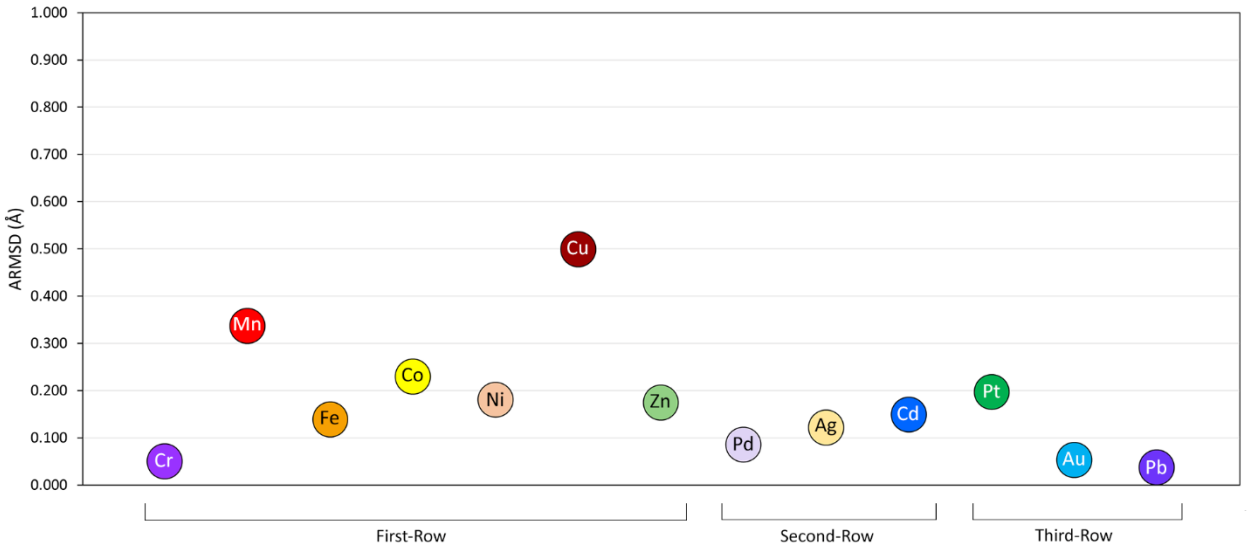


Figure D.25. Inner-shell ARMSDs (Å) across 20 functionals and all complexes relative to crystal structure references for each metal from all-electron gas-phase optimizations.

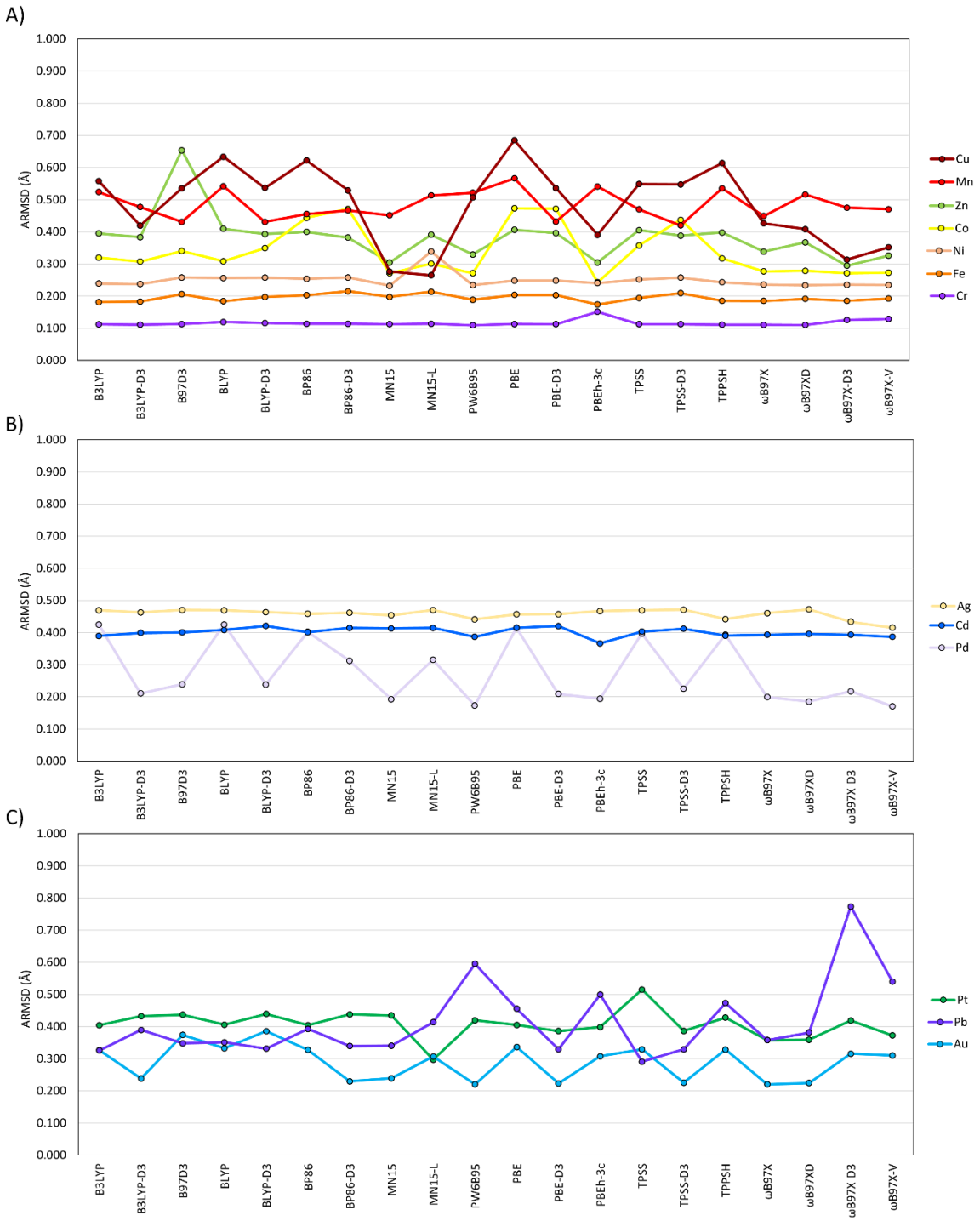


Figure D.26. All-heavy-atom ARMSDs (Å) across all complexes relative to crystal structure references for A) first-row, B) second-row, or C) third-row metals and each functional from all-electron gas-phase optimizations.

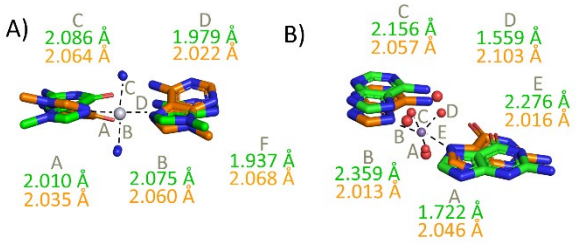


Figure D.27. Overlay of the A) MN15 gas-phase optimized Pt^{2+} complex (SUTTAC, orange) or B) MN15-L gas-phase optimized Mn^{2+} complex (4DWY, orange) and the reference crystal structure (green), which highlights minimal changes in relative ligand orientations for these functional–complex combinations.

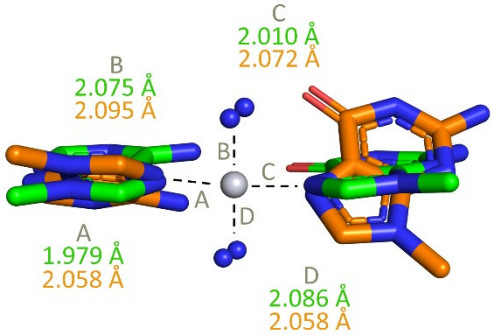


Figure D.28. Overlay of the Pt^{2+} complex (SUTTAC) from the reference crystal structure (green) and a TPSS gas-phase optimization (orange), which highlights the change in ligand co-planarity.

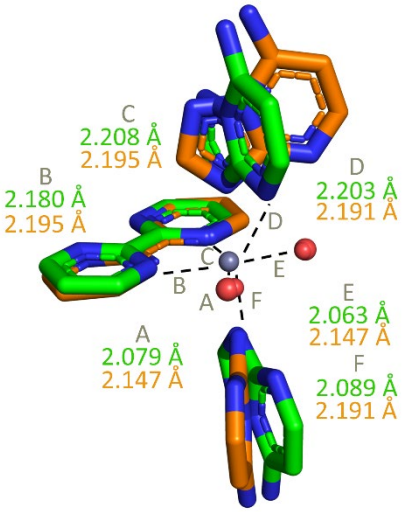


Figure D.29. Overlay of the Zn^{2+} complex (FAJZAT) from the reference crystal structure (green) and a B97D3 gas-phase optimization (orange), which highlights the change in ligand co-planarity.

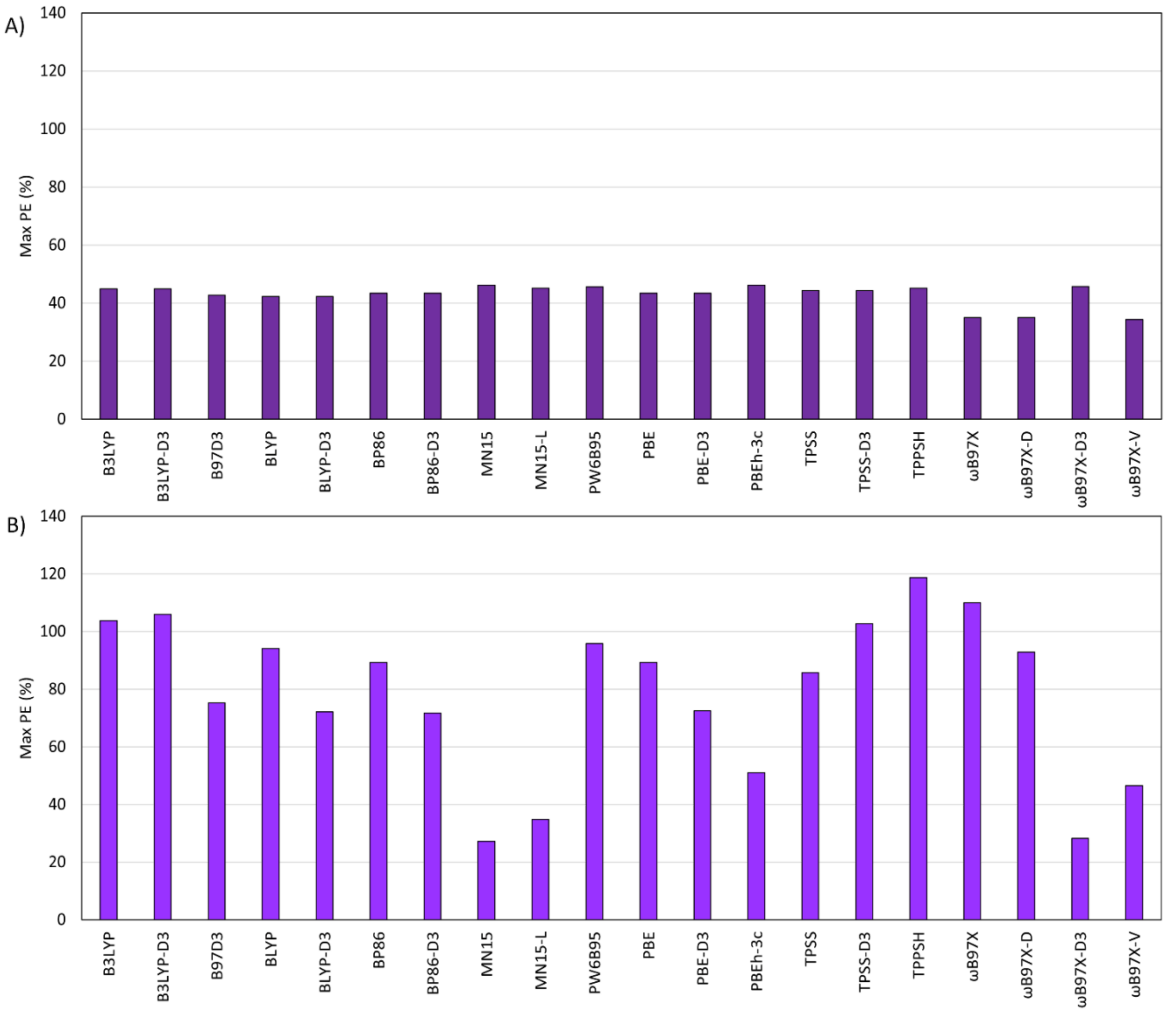


Figure D.30. Inner-shell Max PEs (%) evaluated for A) all complexes that do not include Cu^{2+} (36 systems) or B) only Cu^{2+} complexes from all-electron gas-phase optimizations.

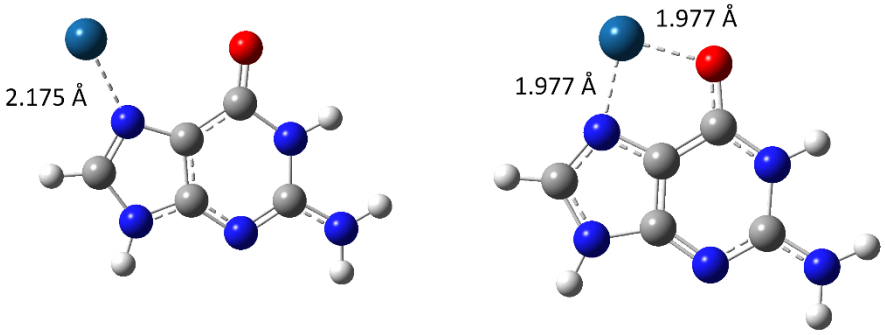


Figure D.31. Crystal structure reference (left) and a MN15 optimized structure (right) of the Pt^{2+}–complex (4I1G), which is a representative example of most functionals incorrectly describing the coordination geometry (with the exception of ωB97X and ωB97X-D).

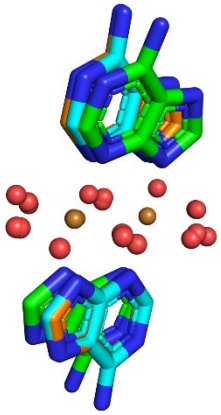


Figure D.32. Overlay of the Cu^{2+} complex (JUNGIK) from the reference crystal structure (green) and a A) ωB97X-V (orange) or B) PBEh-3c (cyan) gas-phase optimization, which highlights the reorientation of water ligands about the metal centers.

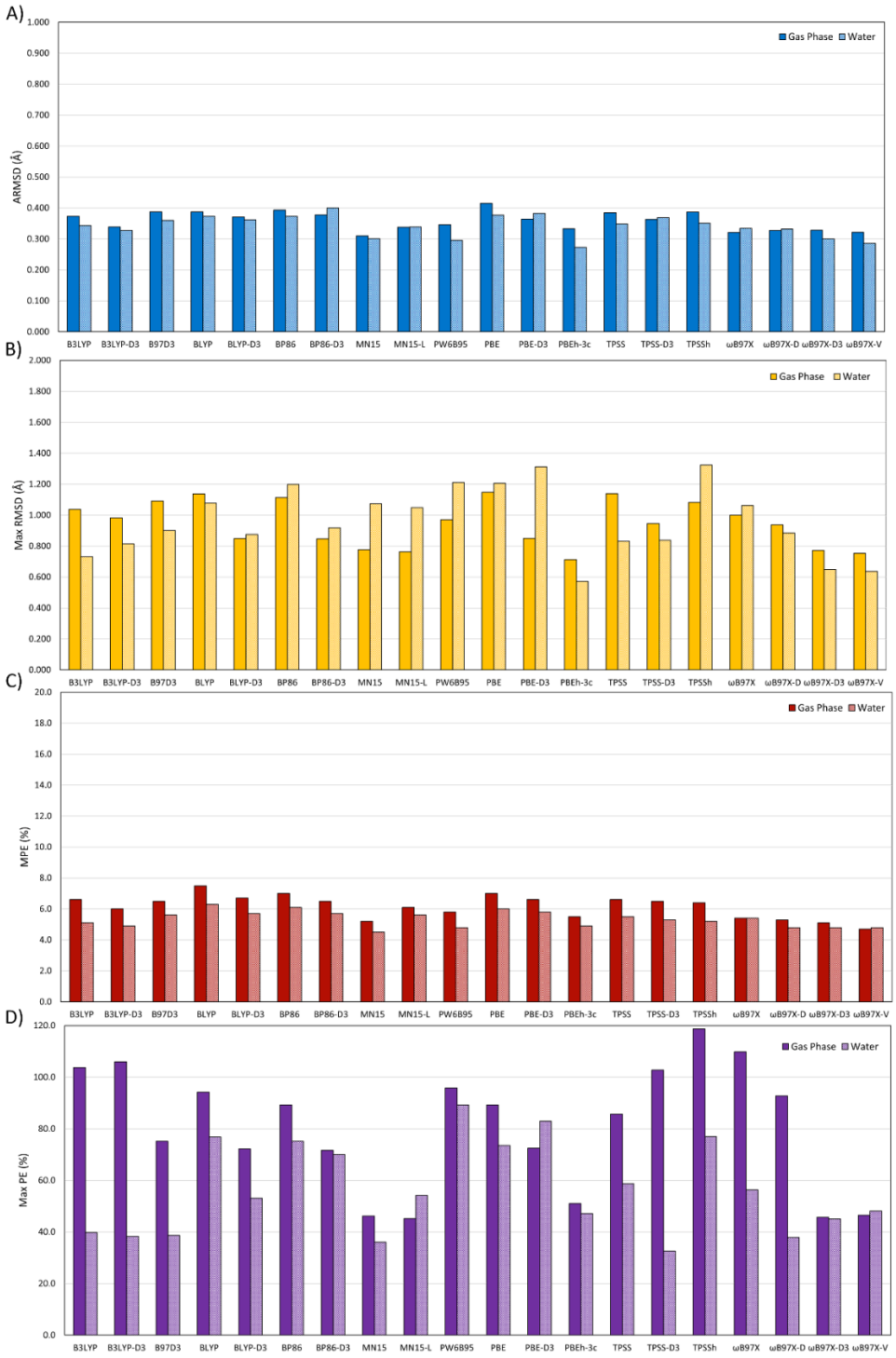


Figure D.33. All-heavy-atom A) ARMSDs (Å) and B) Max RMSDs (Å), and inner-shell C) MPEs (%) and D) Max PEs (%) across all complexes relative to crystal structure references for each functional from all-electron gas-phase and implicit water optimizations.

Figure D.34. All-heavy-atom A) ARMSDs (Å) and B) Max RMSDs (Å), and inner-shell C) MPEs (%) and D) Max PEs across all complexes relative to crystal structure references for each functional from gas-phase all-electron and ECP optimizations.

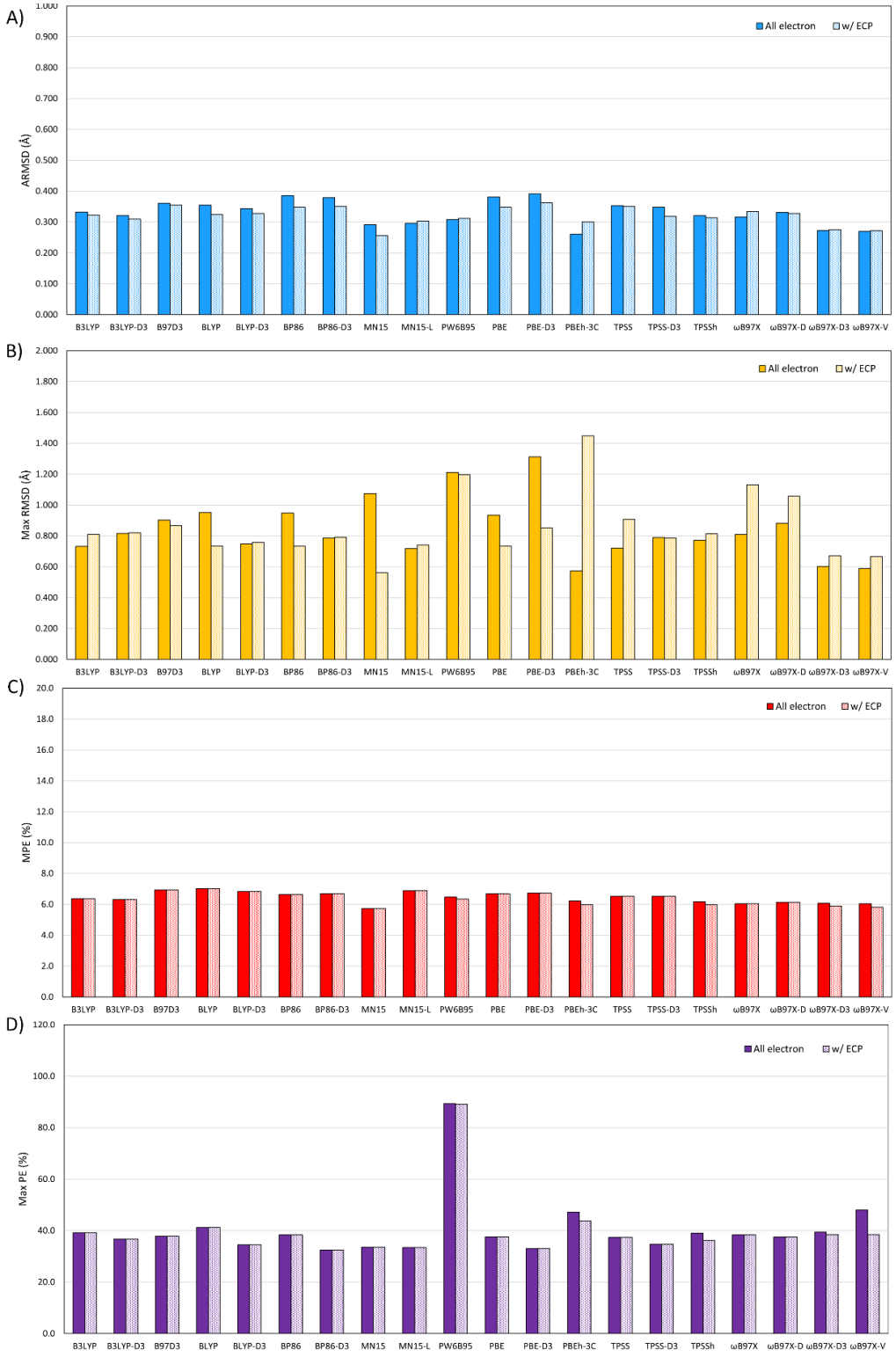


Figure D.35. All-heavy-atom A) ARMSDs (Å) and B) Max RMSDs (Å), and inner-shell C) MPEs (%) and B) Max PEs (%) across all complexes relative to crystal structure references for each functional from all-electron and ECP optimizations in implicit water.

Appendix E

Supplementary Information for Chapter 6: A DFT Investigation into the Structure and Strength of Hydrated Li^+–Nucleic Acid Interactions

Contains Tables E.1–E.6 and Figures E.1–E.8

Table E.1. BEs (kcal/mol) evaluated in the gas phase and various implicit environments for Li^+ binding to all positions in G or dimethyl phosphate.[a]

Implicit Environment	G(O6–N7)	G(N2–N3)	P(2NB)	P(1B1NB)	P(2B)
Gas	–78.7	–52.3	–172.1	–162.4	–148.3
Benzene	–37.6	–21.6	–79.3	–73.0	–63.1
diethylether	–22.3	–11.3	–46.3	–41.6	–34.5
1-bromopropane[b]	–14.4	–3.1	–28.8	–25.0	–20.1
acetone	–9.1	–2.6	–17.2	–15.0	–10.9
1,2-ethanediol	–7.4	–2.8	–13.7	–11.9	–8.1
water	–6.6	–2.7	–11.9	–10.3	–6.8

[a]All geometries were optimized in each solvent using IEF-PCM-MN15/aug-cc-pVDZ and single-point calculations were carried out with IEF-PCM-MN15/aug-cc-pVTZ. See Figure 1 for definitions of complexes considered in present work. The data reported does not contain deformation energy. The G models are capped at the N9 position with a hydrogen. [b]1-bromopropane mimics the dielectric of a DNA environment.

Table E.2. BEs (kcal/mol) of hydrated Li^+–G and Li^+–P complexes in the gas phase.[a]

	Li^+–G(N7)[b]		Li^+–G(O6)[b]		Li^+–P(2NB)	
# of Waters	BE		BE		BE	
0	–78.7		–78.7		–172.1	
1	–66.3		–66.3		–164.0	
2	–55.0		–55.0		–159.9	
3	–51.8		–50.9		–156.4	
4	–48.9		–47.9		–162.4	
5	–47.6		–45.8		–159.6	

[a]All geometries were optimized with MN15/aug-cc-pVDZ and single-point calculations were carried out using MN15/aug-cc-pVTZ. The data reported does not contain deformation energy. The G models are capped at the N9 position with a hydrogen. [b]Li^+ is coordinated to both N7 and O6 for 1 and 2 waters, while monodentate coordination to either O6 or N7 is preferred when 3 or more waters are present.

Table E.3. BEs (kcal/mol) for explicitly hydrated Li^+–G and Li^+–P complexes in implicit water.[a]

	Li^+–G(N7)[b]		Li^+–G(O6)[b]		Li^+–Phosphate(2NB)	
# of Waters	BE		BE		BE	
0	–6.6		–6.6		–11.9	
1	–8.1		–8.1		–19.8	
2	–9.9		–9.9		–26.9	
3	–17.4		–18.5		–32.7	
4	–18.7		–17.6		–38.7	
5	–19.8		–19.2		–38.0	

[a]All geometries were optimized with IEF-PCM-MN15/aug-cc-pVDZ and single-point calculations were carried out using IEF-PCM-MN15/aug-cc-pVTZ. The data reported does not contain deformation energy. The G models are capped at the N9 position with a hydrogen. [b]Li^+ is coordinated to both N7 and O6 for 1 and 2 waters, while monodentate coordination to either O6 or N7 is preferred when 3 waters are present.

Table E.4. BEs (kcal/mol) for explicitly hydrated Li^+–G and Li^+–P complexes in the gas phase and implicit water.[a]

	Li^+–G(N7)[b]			Li^+–G(O6)[b]			Li^+–Phosphate(2NB)	
# of Waters	Gas Phase	Water		Gas Phase	Water		Gas Phase	Water
0	–78.7	–6.6		–78.7	–6.6		–172.1	–11.9
1	–66.3	–8.1		–66.3	–8.1		–164.0	–19.8
2	–55.0	–9.9		–55.0	–9.9		–159.9	–26.9
3	–51.8	–17.4		–50.9	–18.5		–156.4	–32.7
4	–48.9	–18.7		–47.9	–17.6		–162.4	–38.7
5	–47.6	–19.8		–45.8	–19.2		–159.6	–38.0

[a]All geometries were optimized with IEF-PCM-MN15/aug-cc-pVDZ and single-point calculations were carried out using IEF-PCM-MN15/aug-cc-pVTZ in each respective solvent. The data reported does not include deformation energy. The G models are capped at the N9 position with a hydrogen. [b]Li^+ is coordinated to both N7 and O6 for 1 and 2 waters, while monodentate coordination to either O6 or N7 is preferred when 3 waters are present.

Table E.5. BEs (kcal/mol) evaluated in 1-bromopropane and water for all explicitly hydrated Li^+–nucleobase complexes involving 3 or 5 waters.[a]

	n = 3		n = 5		Difference[b]	
Complex	1-bromopropane	Water	1-bromopropane	Water	1-bromopropane	Water
A(N1)	–12.7	–10.2	–14.0	–11.7	–1.3	–1.5
A(N3)	–10.7	–8.1	–9.8	–8.2	0.9	–0.1
A(N7)	–13.4	–11.6	–19.1	–12.7	–5.7	–1.1
C(N3)	–25.8	–20.9	–27.9	–23.4	–2.1	–2.5
C(O2)	–25.7	–20.3	–25.1	–20.9	0.6	–0.6
G(N3)	–8.5	–7.1	–	–	–	–
G(N7)	–22.5	–17.2	–23.8	–20.0	–1.3	–2.8
G(O6)	–23.3	–17.9	–23.6	–19.6	–0.3	–1.7
T(O2)	–13.3	–10.4	–14.5	–12.2	–1.2	–1.8
T(O4)	–12.9	–10.2	–17.1	–13.5	–4.2	3.3
U(O2)	–11.2	–10.4	–13.7	–11.8	–2.5	–1.4
U(O4)	–13.0	–9.8	–15.7	–13.1	–2.7	–3.3

[a]All geometries were optimized with IEF-PCM-MN15/aug-cc-pVDZ and single-point calculations were carried out using IEF-PCM-MN15/aug-cc-pVTZ in each respective solvent. The data reported does not contain deformation energy. All nucleobase models are capped with a methyl group at the N9 and N1 position in purines and pyrimidines, respectively. [b]The differences in BEs are calculated as the hydrated Li^+–nucleic acid complex with 3 waters subtracted from complexes with 5 Li^+-bound waters in 1-bromopropane or water.

Table E.6. BEs (kcal/mol) of the WCF base pairing in hydrated Li^+–nucleobase pair complexes compared to canonical WCF base pairs in the absence of Li^+.[a]

	n = 3						n = 5					
	1-bromopropane			Water			1-bromopropane			Water		
Complex	BE_{Li^+B-P}	BE_{BP}	Diff.[b]	BE_{Li^+B-P}	BE_{BP}	Diff.[b]	BE_{Li^+B-P}	BE_{BP}	Diff.[b]	BE_{Li^+B-P}	BE_{BP}	Diff.[b]
A(N7):T	–10.8	–11.2	0.4	–10.1	–10.4	0.3	–10.9	–11.2	0.3	–10.2	–10.4	0.2
A(N3):T	–11.2		0.0	–10.4		0.0	–		–	–		–
A:T(O2)	–11.4		–0.2	–10.5		–0.1	–10.6		0.6	–10.0		0.4
A:T(O4)	–10.7		0.5	–10.4		0.0	–10.9		0.3	–9.1		1.3
A(N7):U	–10.9	–11.4	0.5	–10.2	–10.5	0.3	–11.0	–11.4	0.4	–10.2	–10.5	0.3
A(N3):U	–11.3		0.1	–10.5		0.0	–11.3		0.1	–		–
A:U(O2)	–11.4		0.0	–10.6		–0.1	–11.0		0.4	–10.3		0.2
A:U(O4)	–10.7		0.7	–10.1		0.4	–10.8		0.6	–10.0		0.5
G(N7):C	–20.8	–19.1	–1.7	–17.1	–16.6	–0.5	–19.8	–19.1	–0.7	–17.1	–16.6	–0.5
G(O6):C	–19.4		–0.3	–16.8		–0.2	–19.1		0.0	–16.6		0.0
G(N3):C	–19.2		–0.1	–16.7		–0.1	–		–	–		–
G:C(O2)	–18.3		0.8	–15.9		0.7	–		–	–		–

[a]All geometries were optimized with IEF-PCM-MN15/aug-cc-pVDZ and single-point calculations were carried out using IEF-PCM-MN15/aug-cc-pVTZ in each respective solvent. The data reported does not include deformation energy. All nucleobase models are capped with a methyl group at the N9 and N1 position in purines and pyrimidines, respectively. [b]The differences are calculated as the hydrated Li^+–base pair BE subtracted by the hydrated Li^+–nucleobase BE in each solvent for 3 or 5 waters.

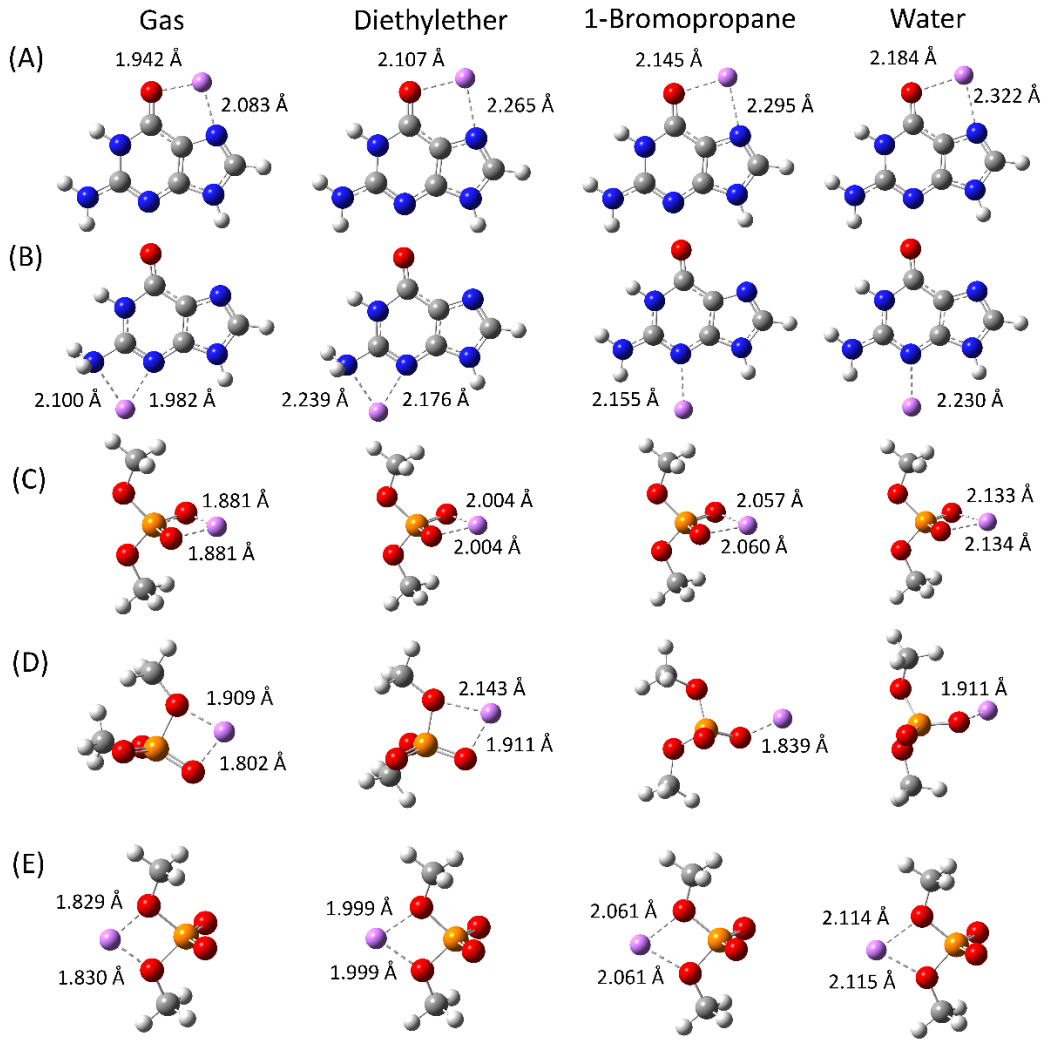


Figure E.1. Structures of (A) Li^+–G(O6–N7), (B) Li^+–G(N2–N3), (C) Li^+–P(2NB), (D) Li^+–P(1B1NB), and (E) Li^+–P(2B) complexes in different dielectric mediums highlighting the G and P complexes that undergo a change in coordination geometry. The ∠(OPO) angles and ∠(N2C2N3) angles (°) are provided.

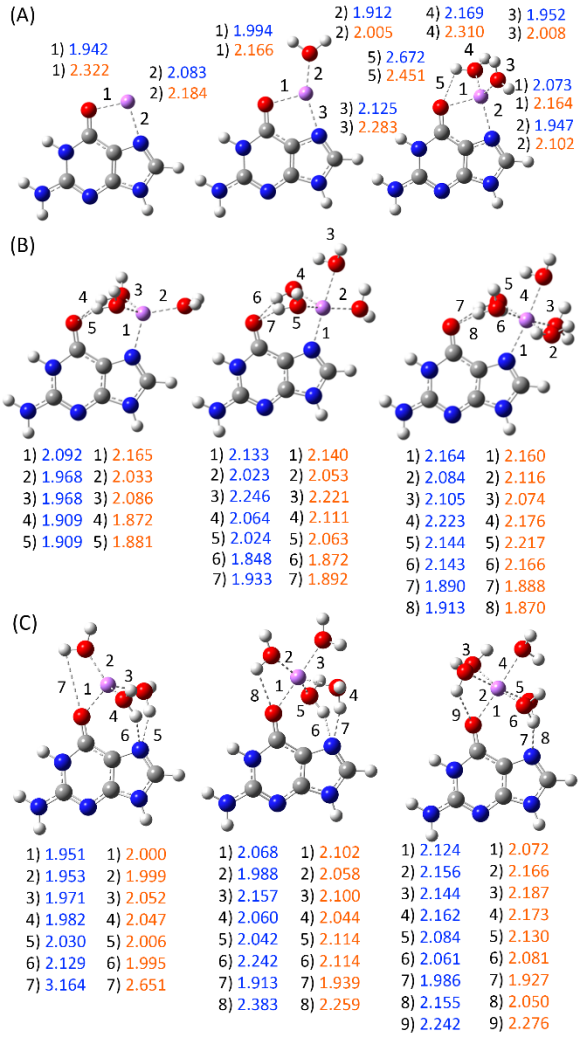


Figure E.2. Key distances (Å) in explicitly hydrated Li^+–G complexes in the gas phase (blue) and implicit water (orange) involving (A) 0 to 2 waters with bidentate Li^+–G(O6–N7) coordination, (B) 3 to 5 waters with monodentate Li^+–G(N7) coordination, and (C) 3 to 5 waters with monodentate Li^+–G(O6) coordination.

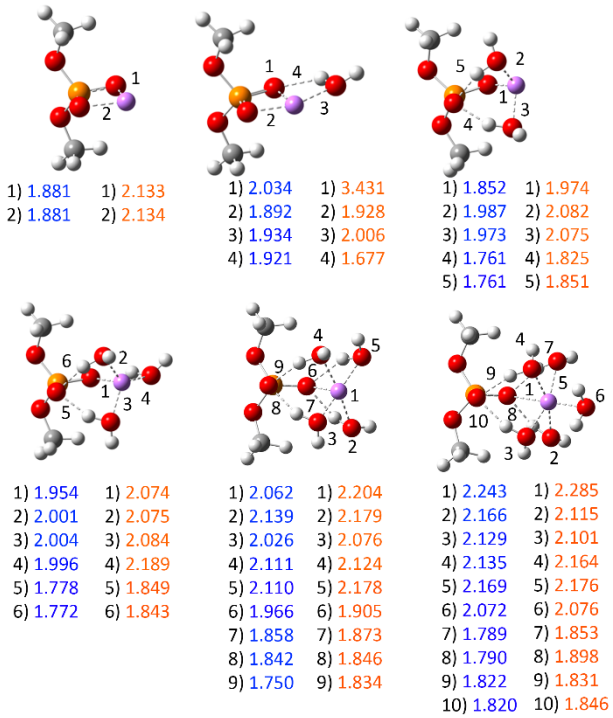


Figure E.3. Key distances (Å) in explicitly hydrated Li^+–P complexes in the gas phase (blue) and implicit water (orange) involving 0 to 5 waters with a transition from bidentate (Li^+–P(2NB)) to monodentate (Li^+–P(1NB)) coordination.

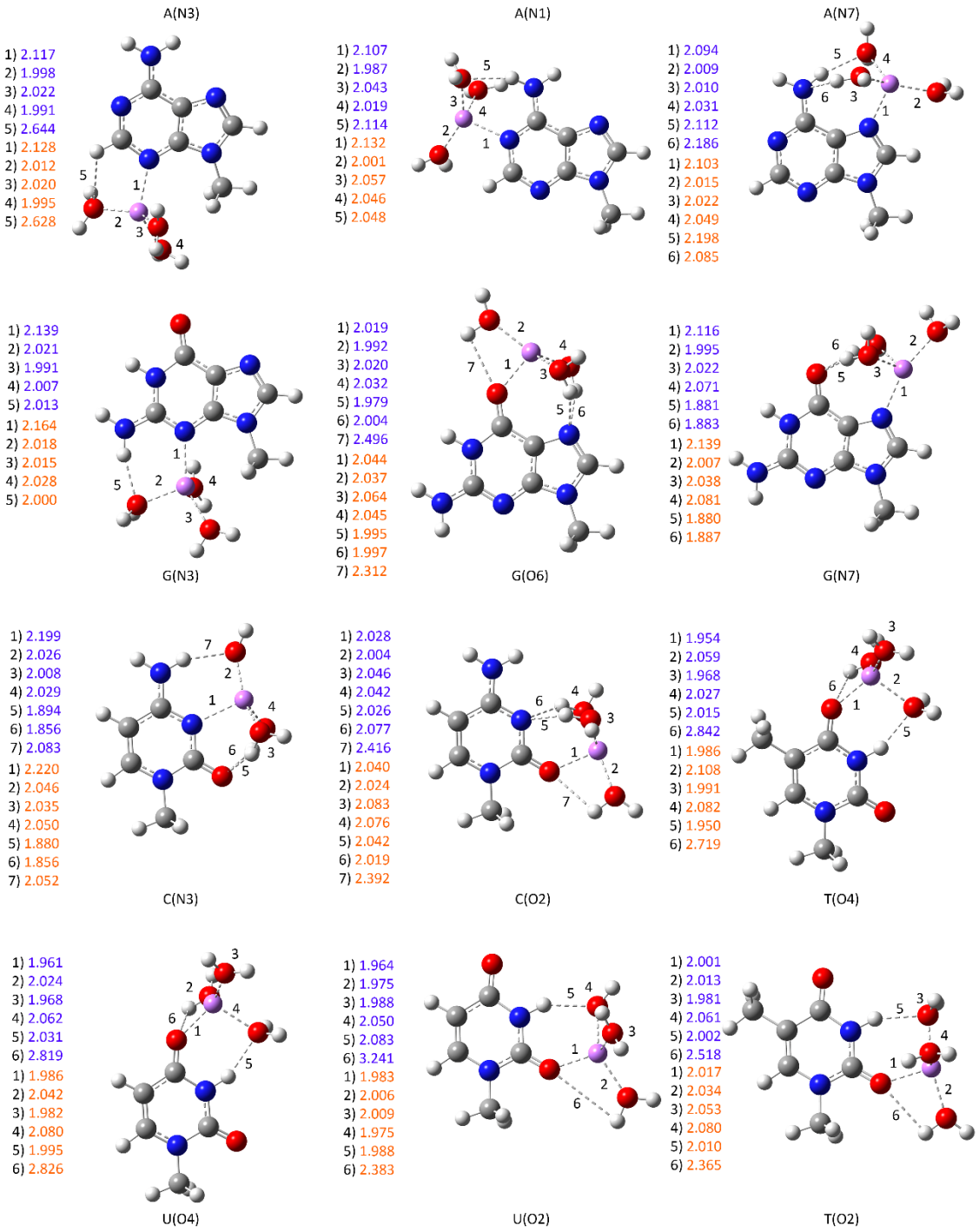


Figure E.4. Key distances (Å) in explicitly hydrated Li^+–nucleobase complexes involving 3 waters in 1-bromopropane (blue) and implicit water (orange).

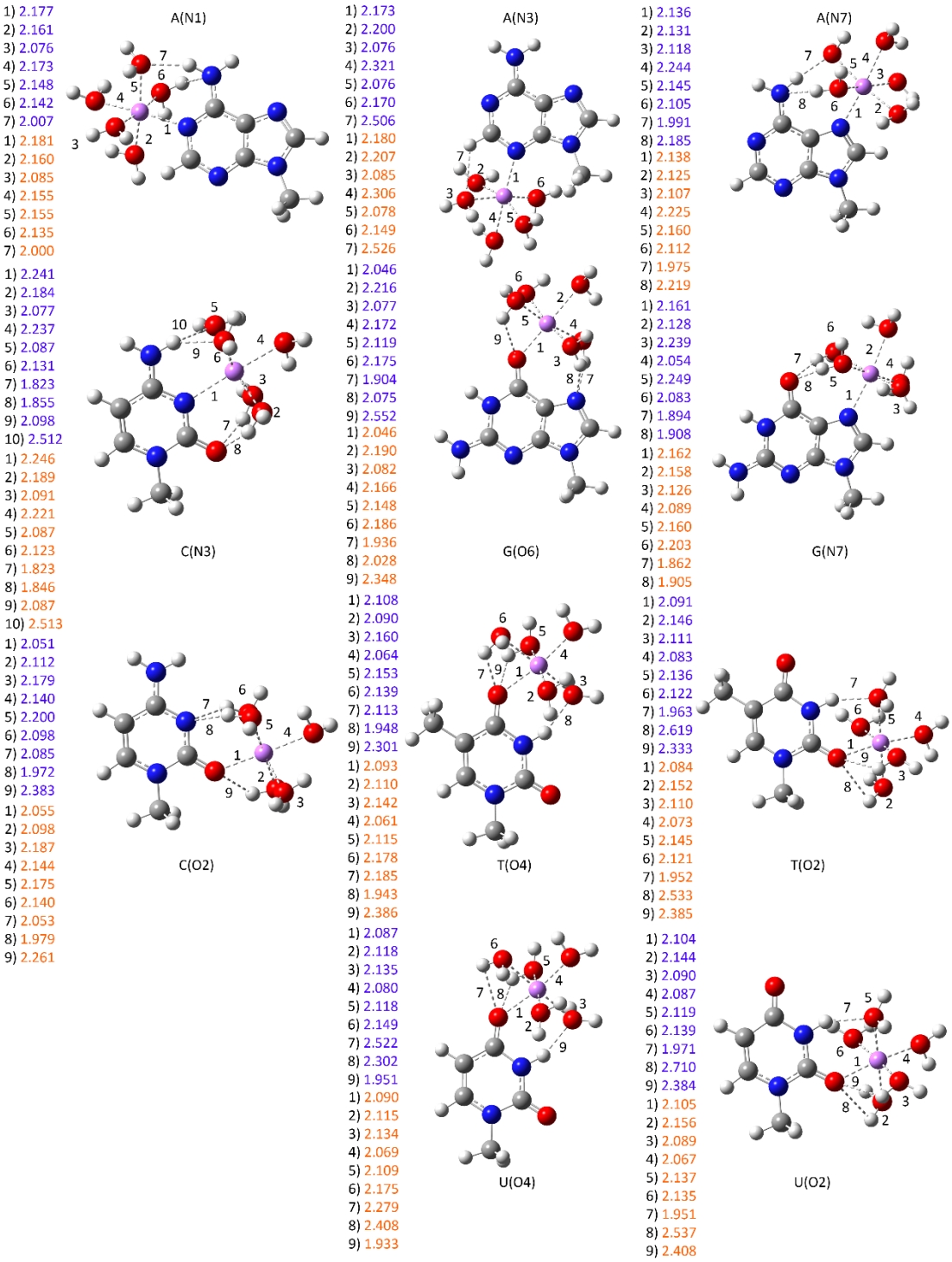


Figure E.5. Key distances (Å) in explicitly hydrated Li^{+}–nucleobase complexes involving 5 waters in 1-bromopropane (blue) and implicit water (orange).

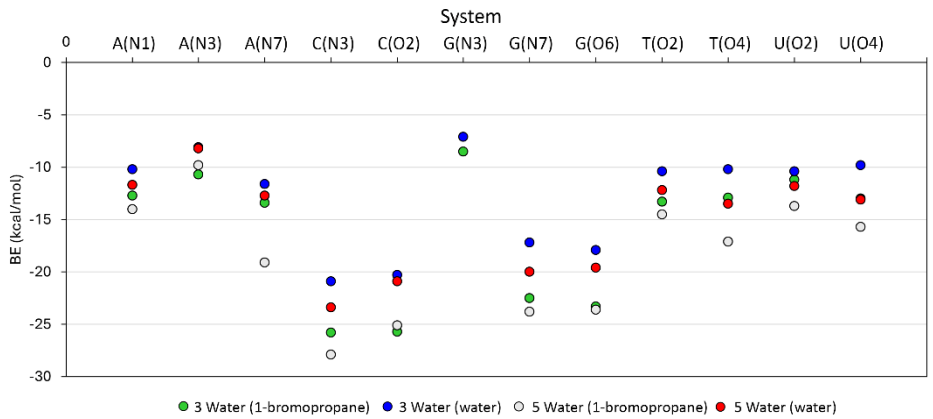


Figure E.6. BEs (kcal/mol) for explicitly hydrated Li^+–nucleobase complexes involving 3 or 5 waters in 1-bromopropane ($\varepsilon = 8$) and implicit water ($\varepsilon = 78$).

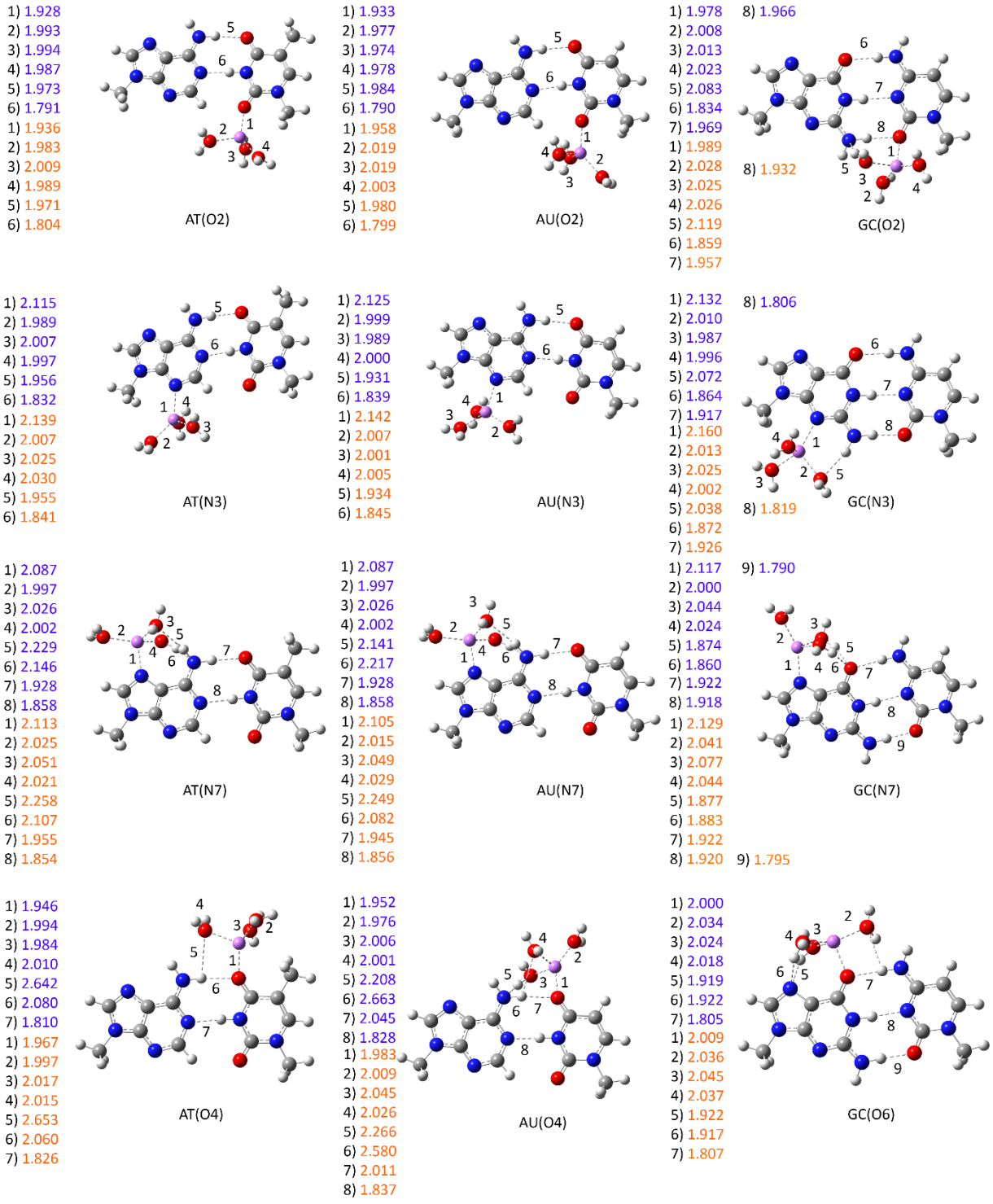


Figure E.7. Key distances (Å) in explicitly hydrated Li^+–nucleobase interactions within a WCF base pair involving 3 waters in 1-bromopropane (blue) and implicit water (orange).

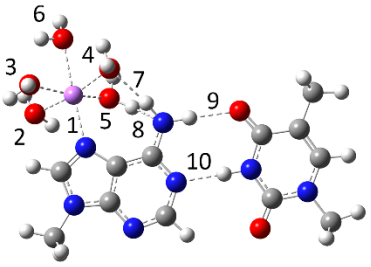
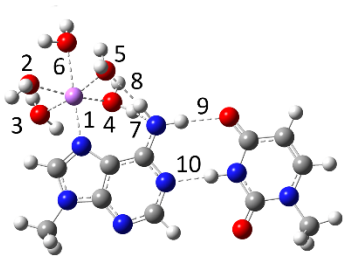
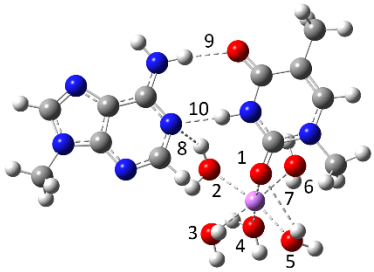
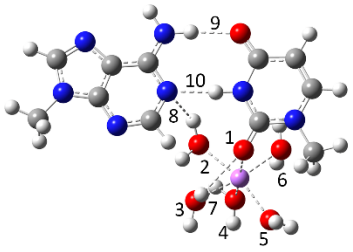
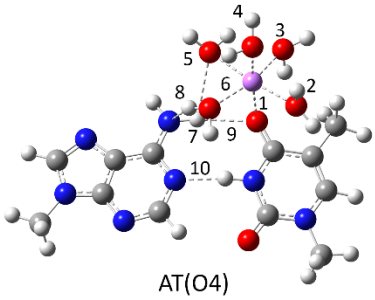
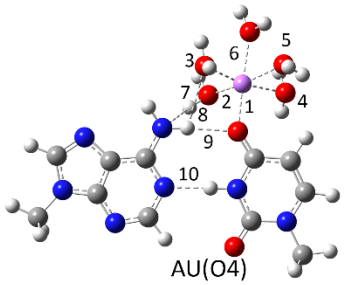
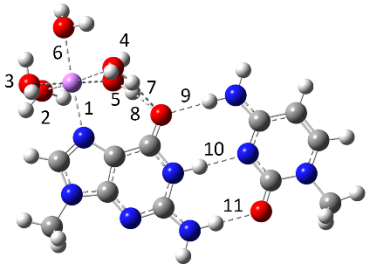
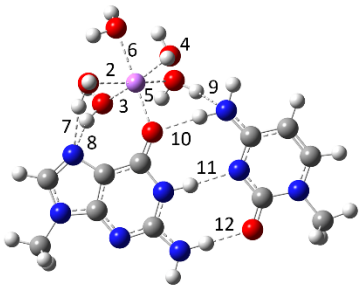
1) 2.134	1) 2.145		1) 2.137	1) 2.137	
2) 2.140	2) 2.135		2) 2.117	2) 2.103	
3) 2.119	3) 2.088		3) 2.139	3) 2.127	
4) 2.126	4) 2.144		4) 2.091	4) 2.010	
5) 2.113	5) 2.230		5) 2.136	5) 2.155	
6) 2.242	6) 2.216		6) 2.259	6) 2.249	
7) 2.243	7) 2.339		7) 2.275	7) 2.275	
8) 2.003	8) 1.990		8) 2.000	8) 2.000	
9) 1.936	9) 1.955		9) 1.943	9) 1.957	
10) 1.857	10) 1.855	AT(N7)	10) 1.850	10) 1.851	AU(N7)
1) 2.139	1) 2.190		1) 2.117	1) 2.118	
2) 2.133	2) 2.162		2) 2.200	2) 2.201	
3) 2.149	3) 2.164		3) 2.136	3) 2.139	
4) 2.087	4) 2.070		4) 2.118	4) 2.111	
5) 2.091	5) 2.101		5) 2.061	5) 2.060	
6) 2.136	6) 2.096		6) 2.142	6) 2.139	
7) 2.393	7) 2.523		7) 2.358	7) 2.405	
8) 2.029	8) 2.051		8) 2.358	8) 2.376	
9) 1.947	9) 1.898		9) 1.934	9) 1.936	
10) 1.867	10) 2.051	AT(O2)	10) 1.873	10) 1.873	AU(O2)
1) 2.151	1) 2.138		1) 2.022	1) 2.024	
2) 2.085	2) 2.053		2) 2.082	2) 2.180	
3) 2.165	3) 2.159		3) 2.088	3) 2.115	
4) 2.111	4) 2.106		4) 2.178	4) 2.137	
5) 2.158	5) 2.262		5) 2.157	5) 2.109	
6) 2.101	6) 2.067		6) 2.229	6) 2.193	
7) 2.566	7) 2.554		7) 2.186	7) 2.193	
8) 2.211	8) 2.253		8) 2.578	8) 2.587	
9) 2.175	9) 2.157		9) 1.991	9) 1.965	
10) 1.829	10) 1.830	AT(O4)	10) 1.827	10) 1.847	AU(O4)
1) 2.162	1) 2.167		1) 2.033	1) 2.027	
2) 2.130	2) 2.142		2) 2.171	2) 2.181	
3) 2.086	3) 2.074		3) 2.136	3) 2.084	
4) 2.196	4) 2.200		4) 2.128	4) 2.101	
5) 2.173	5) 2.178		5) 2.114	5) 2.208	
6) 2.139	6) 2.142		6) 2.197	6) 2.210	
7) 1.856	7) 1.859		7) 1.939	7) 1.933	
8) 1.897	8) 1.900		8) 2.016	8) 2.028	
9) 1.922	9) 1.921		9) 2.513	9) 2.862	
10) 1.914	10) 1.917		10) 1.899	10) 1.912	
11) 1.794	11) 1.799	GC(N7)	11) 1.920	11) 1.917	GC(O6)
			12) 1.817	12) 1.820	

Figure E.8. Key distances (Å) in explicitly hydrated Li^+–nucleobase complexes within a WCF base pair involving 5 waters in 1-bromopropane (blue) and implicit water (orange).

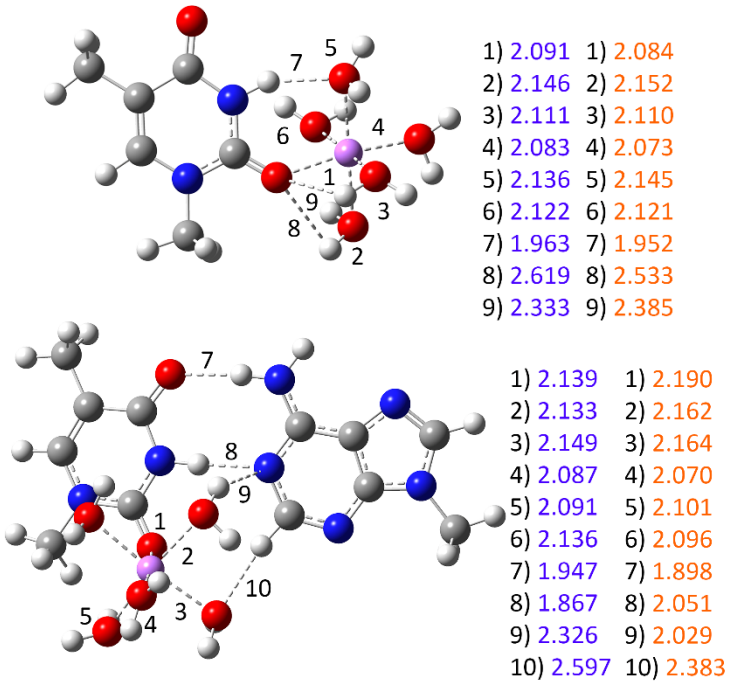


Figure E.9. Example structures and key distances (Å) of explicitly hydrated Li^+–T complex (top) and hydrated Li^+–AT complex (bottom) involving 5 waters in 1-bromopropane (blue) and implicit water (orange).

Appendix G

Supplementary Information for Chapter 8: Designing a Crown Ether–Nucleic Acid Adduct for Targeting Lithium: A Density Functional Theory and Ab Initio Molecular Dynamics Approach

Contains Tables G.1–G.4 and Figures G.1–G.8

Table G.1. Coordination distances (Å) and BEs (kcal/mol) of Li^+–CE complex in the gas phase and water.

Complex	Distance (Å)							
	Li^+–O^1	Li^+–O^2	Li^+–O^3	Li^+–O^4	Li^+–O^5	Ave		BE (kcal/mol)
Li^+–CE (gas)	1.955	1.993	2.032	2.019	2.169	2.034		–101.0
Li^+–CE (water)	1.998	2.076	2.151	2.055	2.218	2.100		–18.2

Table G.2. Coordination distances (Å) and BEs (kcal/mol) of Li^+–CEC and Li^+–CECBP where Li^+ is bound to the crown ether center in water.

Complex	Distance (Å)							
	Li^+–O^1	Li^+–O^2	Li^+–O^3	Li^+–O^4	Li^+–O^5	Ave		BE (kcal/mol)
Li^+–CEC	2.039	2.075	2.128	2.049	2.184	2.095		–16.1
Li^+–CEC BP	2.039	2.092	2.140	2.040	2.179	2.098		–16.4

Table G.3. Coordination distances (Å) and BEs (kcal/mol) of Li^+–CEC and Li^+–CECBP where Li^+ is bound to the O6–N7 position of G in water.

Complex	Distance (Å)				
	Li^+–N7	Li^+–O6	Ave		BE (kcal/mol)
Li^+–G	2.318	2.080	2.199		–5.7
Li^+–CEC BP	2.039	2.092	2.098		–6.0

Table G.4. Average WCF base pair distances (Å) and of Li^+–CECBP over the AIMD simulations.

Complex	Distance (Å)		
	G(O6)–C(H-N)	G(N-H)–C(N3)	G(N-H)–C(O2)
Li^+–CEC	1.899±0.164	1.929±0.122	1.946±0.182
Li^+–G	1.951±0.238	2.466±0.760	3.019±1.170

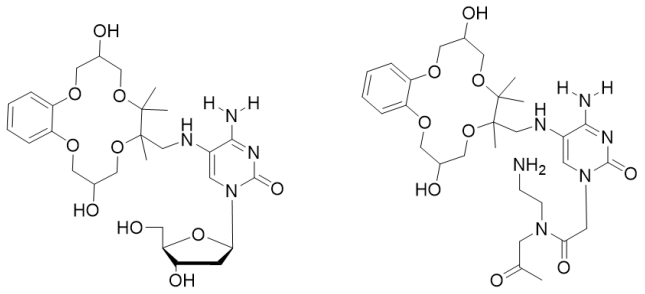


Figure G.1. Diagram of the nucleoside (left) and peptide (right) connection in the CECNBP and CECPBP models.

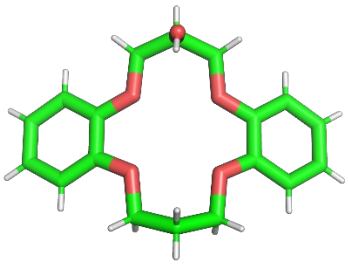


Figure G.2. The BEGVAL crystal structure used in present work to construct the CE 3d variant.[1]

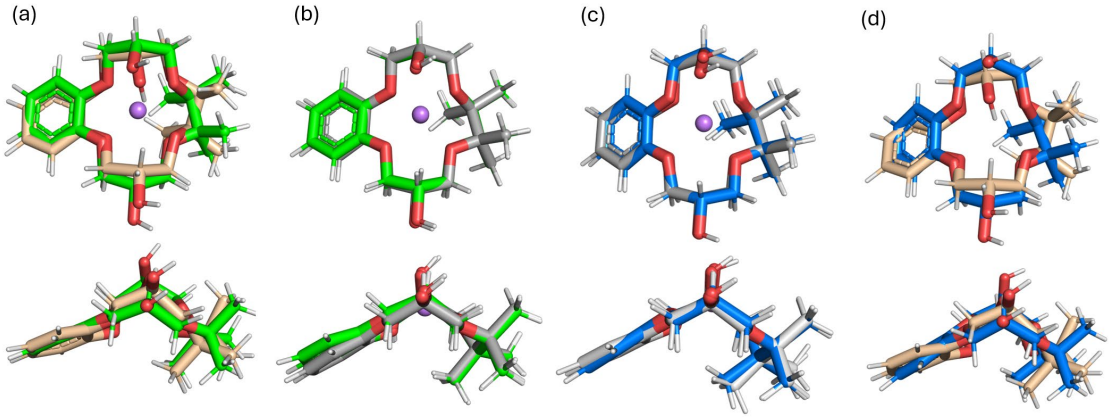


Figure G.3. Overlays of DFT-optimized structures between (a) gas-phase (green) and water (grey) Li^+–CE complex, (b) water Li^+–CE (grey) and CE monomer (blue), (c) gas-phase Li^+–CE (green) and CE monomer (yellow), and (d) water (blue) and gas-phase (yellow) CE monomers.

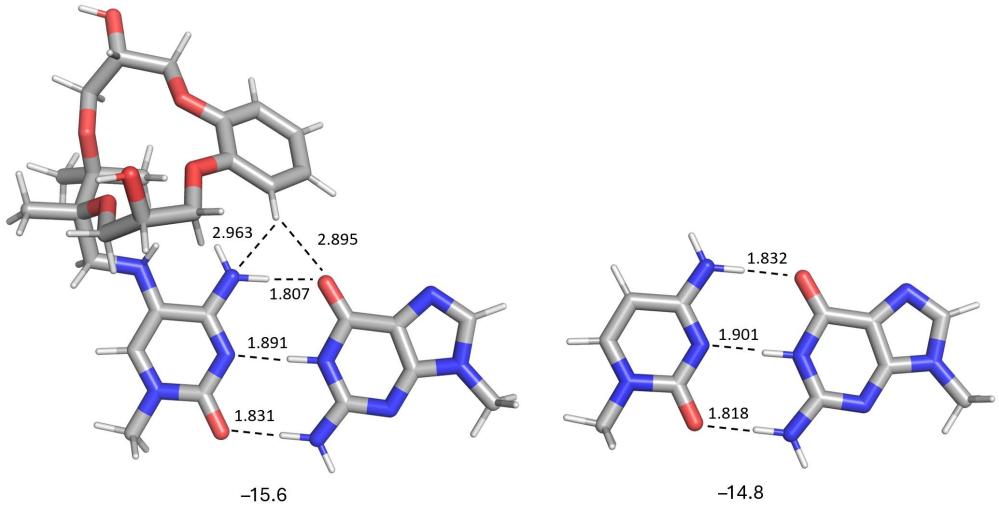


Figure G.4. DFT-optimized structures in implicit water of (a) CECBP and (b) GC, with WCF distances (Å) and calculated binding energies (kcal/mol).

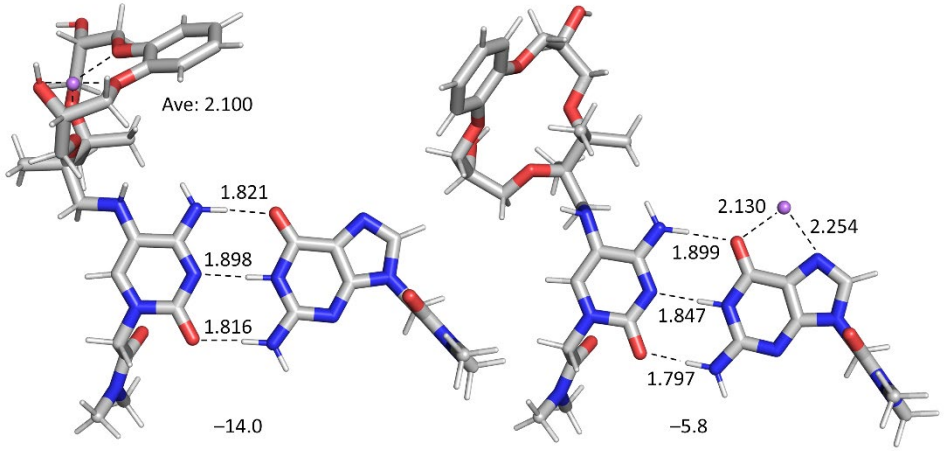


Figure G.5. DFT-optimized structures in implicit water of peptide linkage models (PNA) with Li^+ interacting at the CE center (left) or G(O6–N7) (right).

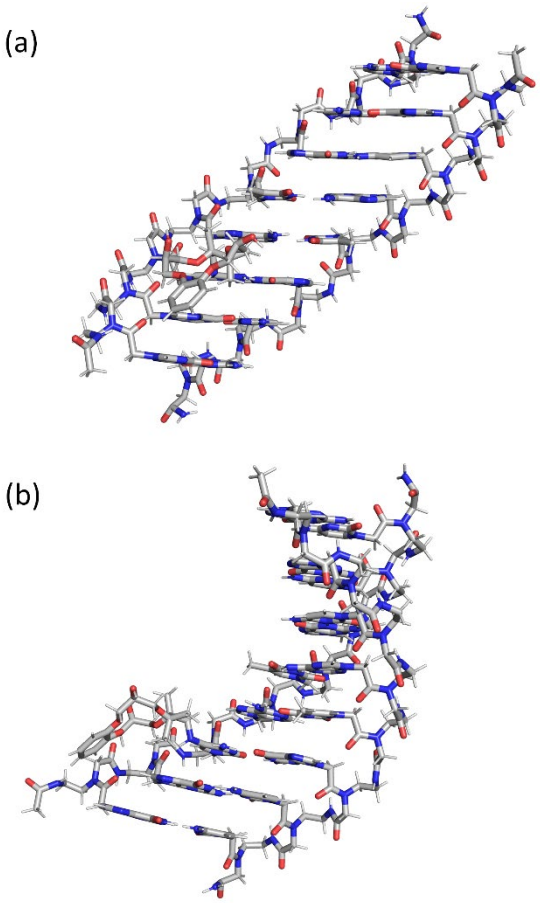


Figure G.6. GFN2-xTB-optimized structure of a CE-modified PNA duplex in implicit water with the (a) front view and (b) side view of PNA.

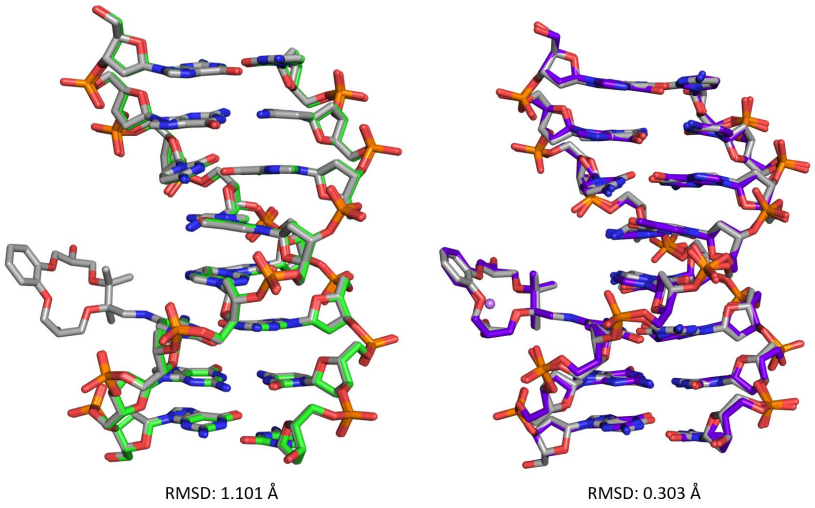


Figure G.7. Overlay of the unmodified DNA duplex (green) with a CE-modified DNA duplex (grey) (left), and an overlay of a CE-modified DNA duplex without Li^+ (grey) and with Li^+ (purple; right). All structures were optimized in implicit water using GFN2-xTB.

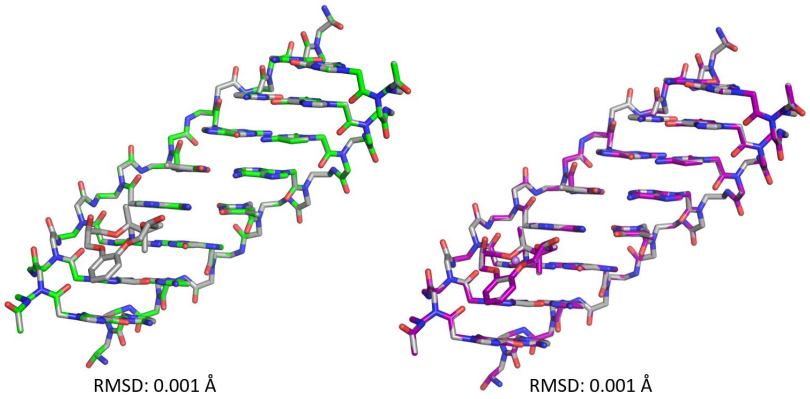


Figure G.8. Overlay of the unmodified PNA duplex (green) with a CE-modified PNA duplex (grey) (left), and an overlay of a CE-modified PNA duplex without Li^+ (grey) and with Li^+ (purple; right). All structures were optimized in implicit water using GFN2-xTB.