COMPUTATIONAL INVESTIGATION OF OXIDATIVE DAMAGE TO GUANINE: FORMATION, RECOGNITION AND REMOVAL BY DNA REPAIR ENZYMES IN HUMANS AND BACTERIA

SHAHIN SOWLATI-HASHJIN

Bachelor of Science, Shahid Beheshti University, 2006
Master of Science, Shahid Beheshti University, 2009
Master of Science, Saint Mary's University, 2013

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

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

COMPUTATIONAL INVESTIGATION OF OXIDATIVE DAMAGE TO GUANINE: FORMATION, RECOGNITION AND REMOVAL BY DNA REPAIR ENZYMES IN HUMANS AND BACTERIA

SHAHIN SOWLATI-HASHJIN

Date of Defence: January 19, 2018

Dr. S. Wetmore Supervisor	Professor	Ph.D.
Dr. S. Mosimann Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. P. Hazendonk Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. M. Roussel Internal-External Examiner	Professor	Ph.D.
Dr. E. Bushnell External Examiner Brandon University Brandon, Manitoba	Assistant Professor	Ph.D.
Dr. P.l Hayes Chair, Thesis Examination Committee	Professor	Ph.D.

Dedication

To my parents and brothers

Abstract

Formation of guanine oxidation products is among the most frequently occurring DNA damaging events, and has been suspected to be related to aging, Alzheimer's disease and various types of cancer. Guanine oxidative products are removed and replaced through the base excision repair (BER) pathway, which involves several enzymes including DNA glycosylases. BER is initiated when a DNA glycosylase detects a lesion and removes the damaged nucleobase. The present thesis employs computational chemistry to investigate the formation mechanisms of two guanine oxidation products, namely 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) and 7,8-dihydro-8-oxoguanine (OG). Moreover, several steps during the repair processes catalyzed by human (hOgg1) and bacterial (FPG) DNA glycosylases, including the lesion recognition, base removal, DNA backbone cleavage, and hydrolysis of the resulting DNA–enzyme complex, are investigated using quantum mechanical and molecular dynamics approaches. The atomic-level details provided by these calculations show the differences in the recognition and removal mechanisms catalyzed by human and bacterial glycosylases.

Acknowledgements

I am grateful to several people, without the help of whom, this thesis could not have been completed. First and foremost, I would like to thank my supervisor Dr. Stacey D. Wetmore who patiently and constantly encouraged me to find my way out of the maze. My special thanks go to my family for their dedication and support over all these years.

I would like to thank the Supervisory Committee Members, Dr. Steven Mosimann, Dr. Paul Hazendonk, and Dr. Brent Selinger for their useful comments, suggestions and helps during the program. I would like to thank Dr. Eric Bushnell who has agreed to be my external examiner. I would also like to thank my internal external examiner, Dr. Marc Roussel.

My thanks also go to the Wetmore group members, past and present: Dr. Purshotam Sharma, Dr. Mohammed Aboelnga, Katie Wilson, Stefan Lenz, Ryan Kung, Hanzala Hussain, Priya Bhutani, Rajwinder Kaur, and Preetleen Upneja.

I want to thank University of Lethbridge, School of Graduate Studies, and Alberta Innovates – Technology Futures (AI – TF) for funding.

My sincere thanks go to many of my friends, Javad Sheikhtaheri, Jarret McKinnon, Afshan Ardalan, Alya Jeraj, Chad Helfenbein, Rachael Wilbourn, and my soccer teammates!

Last but not least, I like to thank Gretsch, Zildjian, and Ahead companies, whose products together helped me make it through.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of Tables

Chapter 1

Chapter 2

Chapter 5

Chapter 6

List of Abbreviations

APE1	AP-endonuclease 1 (EC # 4.2.99.18)
AP-site	Apurinic/apyrimidinic-site (also abasic site)
B3LYP	Becke 88 exchange with Lee-Yang-Parr correlation density functional
B3LYP-D3	B3LYP with empirical dispersion
B. st.	Bacillus stearothermophilus
BER	Base excision repair
C. ac.	Clostridium acetobutylicum
DFT	Density functional theory
DNA	Deoxyribonucleic acid
E1cB	Elimination unimolecular conjugate base
EndoIII	Endonuclease III
FPG	Formamidopyrimidine–DNA glycosylase (also MutM)
hOgg1	Human 8-oxoguanine–DNA glycosylase
hUNG2	Human uracil–DNA glycosylase
IEF-PCM	Integral-equation formalism polarizable continuum method
L. lac.	Lactococcus lactis
M06-2X	Minnesota 06 with 56% HF exchange
MD	Molecular dynamics
M. jan.	Methanocaldococcus jannaschii
MM	Molecular mechanics
MutM	Formamidopyrimidine–DNA glycosylase (also FPG)
NEI	Endonuclease VIII
NEIL1	Endonuclease VIII-like glycosylase 1
NER	Nucleotide excision repair
NTH	Endonuclease III
PES	Potential energy surface
QM	Quantum mechanics
RMSD	Root-mean-square deviation
SMD	Density-based solvent model
H2TH	Helix-2turn-helix
HhH	Helix-hairpin-helix
WC	Watson-Crick
ωB97X-D	Head-Gordon type Becke function with empirical dispersion
ZPVE	Zero-point vibrational energy

Native and Modified Nucleobase and Nucleoside Abbreviations

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
dA	2′-deoxyadenosine
dC	2′-deoxycytidine
dG	2′-deoxyguanosine
dT	2′-deoxythymidine
5-UOH	5-hydroxy-2′-deoxyuridine
8-GOH	8-hydroxyguanine
8-oxoA	7,8-dihydro-8-oxoadenine
8-oxoG	7,8-dihydro-8-oxoguanine
DHT	Dihydrothymine
dOG	2′-deoxy-8-oxoguanosine
FapyA	4,6-diamino-5-formamidopyrimidine
FapyG	2,6-diamino-4-hydroxy-5-formamidopyrimidine
Gh	5-guanidinohydantoin
Ia	Iminoallantoin
N7-mFapyG	N7-methylFapyG
OG	8-oxoguanine
Sp	2-imino-5-5′-spirodihydantoin
Tg	Thymine glycol

One and Three-Letter Codes for the Amino Acids

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate, aspartic acid
E	Glu	Glutamate, glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Chapter 1: Introduction

1.1 General Overview

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), along with two other broad classes of polymers, namely proteins, lipids and carbohydrates, are the macromolecules that are essential for life. Specifically, DNA is responsible for encoding genetic information required for cell functions. Although DNA can be damaged due to exposure to various endogenous and exogenous factors, the enzymatic repair machinery in cells can combat the damage by replacing the damaged units with the native ones. Although experimental studies have provided invaluable insights into various aspects of DNA damage and repair processes, computational chemistry methods provide atomic-level detail of chemical processes. Moreover, computer modelling allows for the evaluation of the feasibility and/or probability of different proposed/hypothesized reaction mechanisms, as well as the proposal of new mechanisms. The present thesis employs computer modelling (quantum mechanical (QM) calculations and molecular dynamics (MD) simulations) to study the formation and enzymatic repair of a specific type of DNA damage, namely guanine oxidation damage.

1.2 DNA Structure

DNA generally exists in a double-stranded form, and is composed of four natural nucleobases, namely adenine (A) and guanine (G), which are known as purines, and cytosine (C) and thymine (T), which belong to the pyrimidine family (Figure 1.1). Each nucleobase is linked to a 2′-deoxyribose sugar by a so-called glycosidic bond, which forms a nucleoside. In addition, a phosphate group links to the nucleoside at the C5′ position to form a nucleotide, which joins with other nucleotides to form a DNA strand

(Figure 1.2). Two single DNA strands run anti-parallel with respect to each other to form the DNA double helix (duplex).

Purine nucleobases (G and A) pair opposite their matching pyrimidines (C and T, respectively) in the complementary strand using the Watson-Crick (WC) face to make hydrogen bonds (Figure 1.2), which results in an equal number of G and C, and A and T in the duplex. The WC interactions in the natural base pairs require the purine and pyrimidine to adopt the *anti* conformation (*i.e.,* χ $\angle$(O4′–C1′–N9–C4) for purines and $\angle$(O4′–C1′–N1–C2) for pyrimidines, equals 180 ± 90°, Figure 1.2). Indeed, there are two and three strong hydrogen bonds in the A:T and G:C base pairs, respectively (Figure 1.2). Since only one copy of DNA exists in each cell, maintaining the nucleobase sequence by the double-helical nature of DNA is of vital importance.

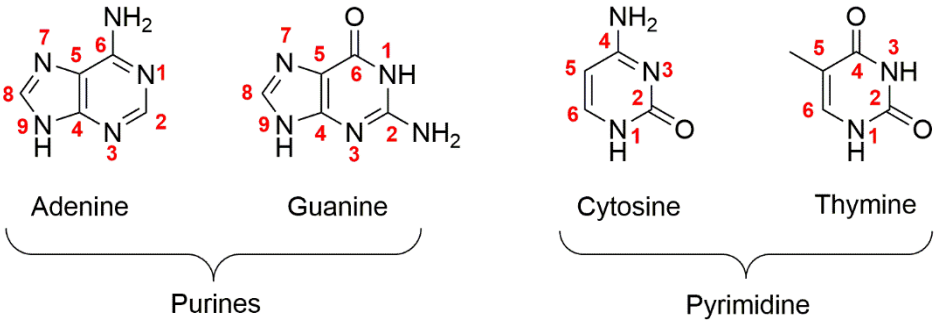

Figure 1.1. Four natural DNA nucleobases classified into the purines and pyrimidines.

1.3 DNA Damage

DNA undergoes various damaging events, through exposure to toxic substances, high-energy radiation, and several other sources.[1-8] Although damage may occur to the sugar-phosphate backbone, damage to the nucleobases is more frequent. Specifically, nucleobase modifications arising from oxidation,[1] deamination,[8] depurination,[8] and alkylation[8] are particularly common. Nucleobase damage can be divided into two broad categories, known as bulky and non-bulky damage. Non-bulky damage typically changes

only a few atoms in a canonical nucleobase, and is usually carcinogenic or cytotoxic.[9-12] One of the main subjects of this thesis is a specific type of non-bulky damage, namely oxidation of guanine.

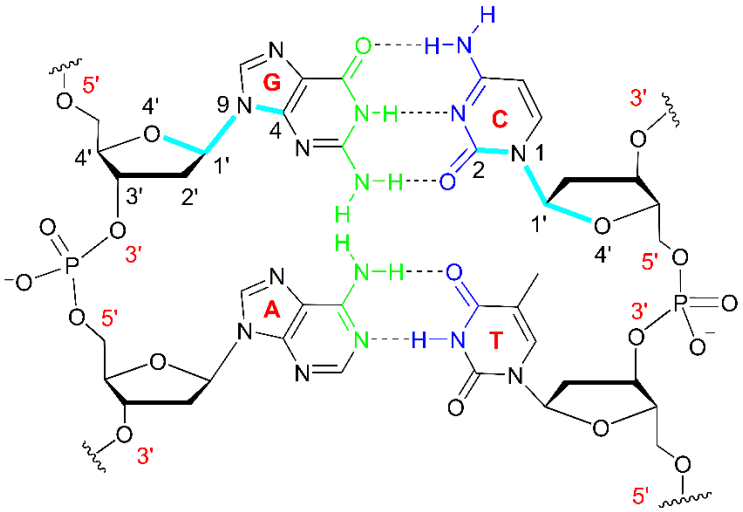

Figure 1.2. Pairing patterns of the purine and pyrimidine nucleotides. Watson-Crick faces are highlighted in green and blue for the purines and pyrimidines, respectively. The χ is highlighted in cyan for guanine (∠(O4′–C1′–N9–C4)) and cytosine (∠(O4′–C1′–N1–C2)).

The oxidation of G to form 7,8-dihydroxy-8-oxoguanine (8oxoG or OG) and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) is the most frequent non-bulky nucleobase damage pathway (Figure 1.3).[13-17] In fact, oxidation of guanine is the most frequent damaging event, with the formation of OG occurring ~ 1000 to 2800 times per day per cell.[13-17] This is mainly due to the lower one-electron oxidation potential and lower ionization potential of guanine relative to other canonical nucleobases.[4, 5, 18] The unsaturated N7=C8 bond (Figure 1.1), which is not involved in hydrogen-bonding interactions with the opposing C is more easily accessible to oxidizing agents, such as hydrogen peroxide (H_2O_2) and the hydroxyl radical (•OH).[6, 19] In fact, guanine-rich non-coding regions of DNA are suggested to be a means of protecting DNA against oxidation.

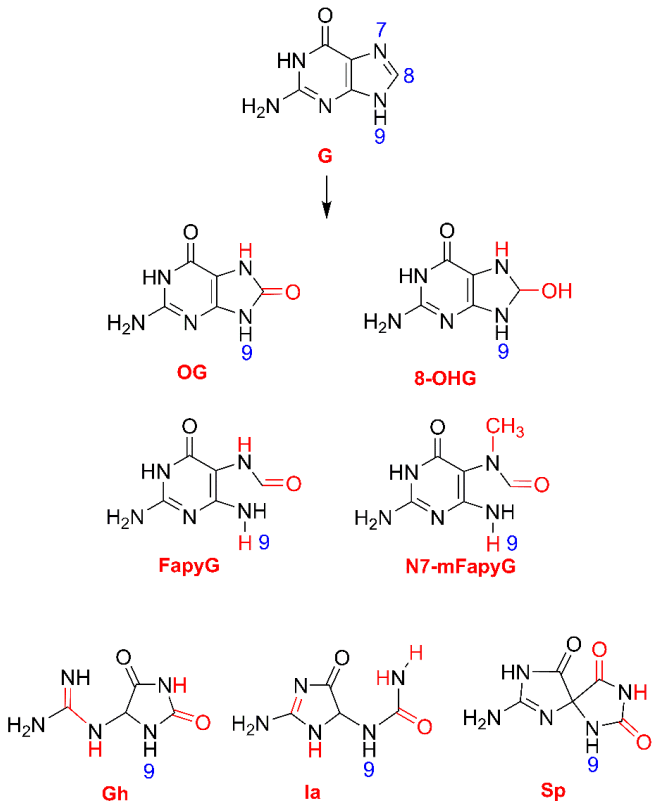

Figure 1.3. Guanine and its several common oxidation products. Structural changes to G to generate the oxidized products are highlighted in red.

The products of guanine oxidation are not limited to OG and FapyG, but also include 7,8-dihydro-8-hydroxyguanine (8-OHG), 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine (N7-mFapyG), spiroiminodihydantoin (Sp), guanidinohydantoin (Gh), and iminoallantoin (Ia, Figure 1.3).[14, 20, 21] Most of these products cause G:C to T:A transversion mutations.[22-26] As an example, OG leads to these mutations mainly due to the preference to rotate about the C1′–N9 bond (*i.e.*, the glycosidic bond) to adopt the *syn* conformation ($\chi = 0 \pm 90°$, Figure 1.4), which reduces the electrostatic repulsion between O8 of the damaged nucleobase, and O4′ and O5′ of the sugar-phosphate backbone. Consequently, the Hoogsteen face of OG (which includes O6, N7–H, and N2–H, Figure 1.4, blue and Figure 1.5) can be mistaken as T by

replicative polymerases, which insert A opposite the lesion and leads to a G:C to T:A mutation after the second round of DNA replication.[27] The formation of guanine oxidation products has been proposed to be mutagenic,[28-31] and lead to cancer,[32, 33] Alzheimer's disease[34] and aging.[35] Specifically, the G:C to T:A transversion mutation is among the most common mutations in cells associated with lung, breast, ovarian, colorectal and renal cancer.[10]

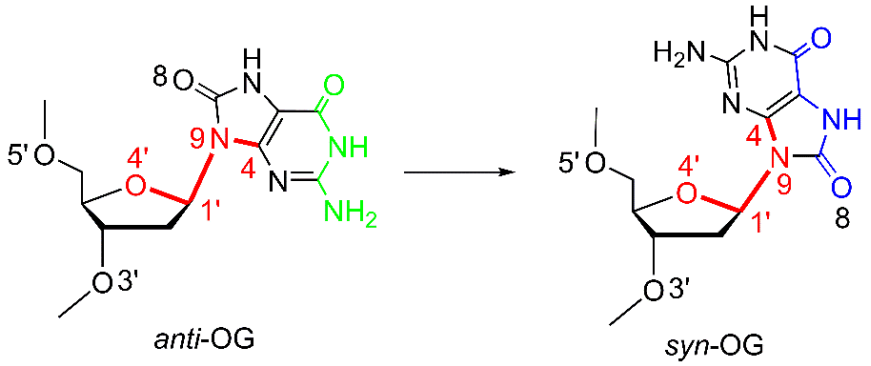

Figure 1.4. The *anti* and *syn* conformations of OG, with the WC and Hoogsteen faces highlighted in green and blue, respectively. The χ dihedral angle (∠(O4′–C1′–N9–C4)) is highlighted in red.

The mechanism of guanine oxidation, the characteristics of oxidative products, and the pathway for the repair of the damaged products have been the subject of numerous experimental studies.[4, 7, 36-39] Due to the complementary atomic level information provided by computer simulations, computational techniques have also been used to explore the impact of guanine oxidation on the genome. For example, previous computational studies have investigated the oxidation/reduction potential of G[40-43] and the associated oxidation products (OG, FapyG, Sp, and Gh),[44] the sequence dependency of oxidation,[45, 46] the pK_a and isomerization of Gh and Sp,[47] and the mutagenicity of FapyG.[48] Other studies have focused on the recognition[49-51] and removal of guanine oxidation products within the context of base excision repair in bacteria[52-55] and human

cells.[56-63] Together, these experimental and computational studies have heightened our understanding of damaging events and repair processes in living organisms.

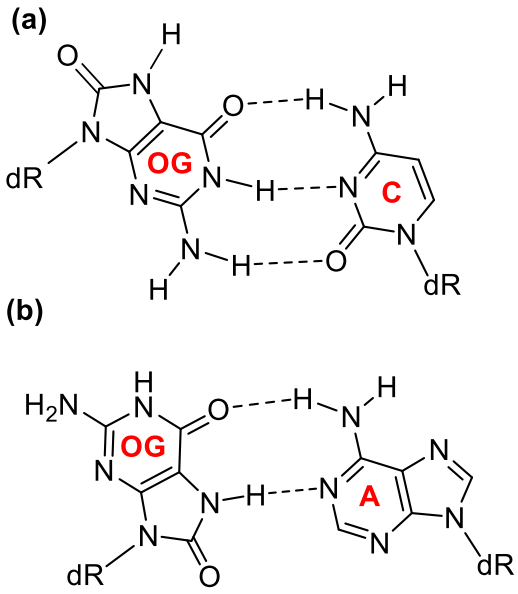

Figure 1.5. (a) Watson-Crick base pair between *anti*-OG and C, and (b) Hoogsteen base pair between *syn*-OG and A.

Despite providing key insights into the characteristics of the guanine oxidation pathways, previous computational studies have typically employed gas-phase optimizations, with environmental effects included through single-point calculations in aqueous solution. However, a recent *ab initio* study has shown that the structures and energy barriers predicted in the gas phase and water for the imidazole ring opening along the FapyG formation pathway can be significantly different.[64] Furthermore, despite the abundance of water in cells, and the fact that the nucleobase edge may be solvent exposed in the DNA duplex, the environment in which the oxidation process occurs may not be well represented by the dielectric constant of water. Additionally, although previous studies showed that the N9 substituent on the nucleobase model (*i.e.*, –H, $–CH_2OH$, or $–CH(CH_3)OCH_3$) has a negligible effect on the relative reaction enthalpies,[65] the potential

impact of the explicit DNA environment (*i.e.,* the sugar–phosphate backbone, base opposing the lesion and flanking bases) has been neglected. To address the importance of the model, the computational method, and the environment, Chapter 2 considers various models and methods to set the ground for future computational studies of reactions relevant to DNA oxidation damage and repair processes.

1.4 DNA Repair

Considering the number of base pairs (more than 10^9) in a human cell,[66] DNA damage seems inevitable. However, several repair mechanisms have been developed to process the damage. For example, nucleotide excision repair (NER)[67, 68] removes a section of DNA that includes the damage, while base excision repair (BER)[67, 69-71] removes only the damaged nucleoside. NER typically targets bulky forms of damage, while non-bulky damage, including guanine oxidation products, are processed through the BER pathway.

Several enzymes are involved in BER, each of which catalyze different steps using some of the 20 different building blocks known as amino acids (Figure 1.6). In the first BER step, a DNA glycosylase detects the lesion, flips the damaged nucleotide into the active site, and cleaves the glycosidic bond between the nucleobase and 2′-deoxyribose.[70, 72-74] A monofunctional glycosylase only catalyzes glycosidic bond cleavage,[70, 72, 74] while the deglycosylation step facilitated by a bifunctional glycosylase occurs before or after cleavage of the C1′–O4′ bond in the 2′-deoxyribose ring (sugar-ring opening).[73, 75] Moreover, all bifunctional glycosylases process the DNA backbone on the 3′-side of the damaged site by breaking the C3′–O3′ bond (β-lyase activity), while some bifunctional glycosylases are also capable of nicking the DNA backbone on the 5′-side of

the lesion (C5′–O5′ bond, δ-lyase activity, Figure 1.7). The products of the process facilitated by a bifunctional glycosylase with β-lyase activity is a 5′-monophosphate and an apurinic/apyrimidinic site (AP-site or abasic site, Figure 1.7). After the glycosylase processes the substrate, an endonuclease nicks the 3′- and 5′-sides of the lesion in the case of a monofunctional glycosylase or the 5′-side in the case of a bifunctional glycosylase that only exhibits β-lyase activity.[76] If the lesion is processed by a bifunctional glycosylase with β- and δ-lyase activity, the 3′-phosphate of the +1 nucleotide with respect to lesion is removed by a 3′-phosphotase. The BER process is complete when a repair polymerase inserts the complementary canonical nucleotide opposite the remaining base, and a ligase seals the DNA backbone.[77]

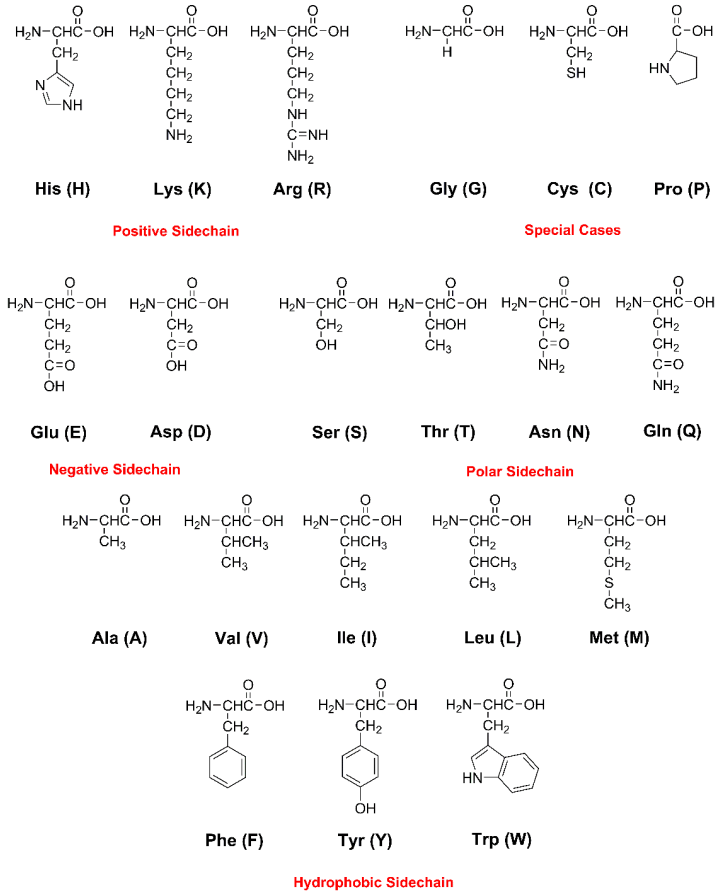

Figure 1.6. Chemical structures of 20 naturally-occurring amino acids classified based on the electronic property of their sidechain, with the corresponding 3- and 1-letter codes.

A few examples of DNA glycosylases that target oxidative DNA damage are provided in Table 1.1. Another subject of the present thesis is the function of DNA glycosylases that are specialized for removing guanine oxidation products in human cells (human oxoguanine glycosylase, hOgg1) and in bacteria (formamidopyrimidine glycosylase, FPG) during the initial steps of the BER pathway. Interestingly, despite sharing the guanine major oxidation products (OG and FapyG) as substrates, hOgg1 and FPG belong to two different structural families. In the following sections, the function and substrate specificity of hOgg1 and FPG are introduced in detail.

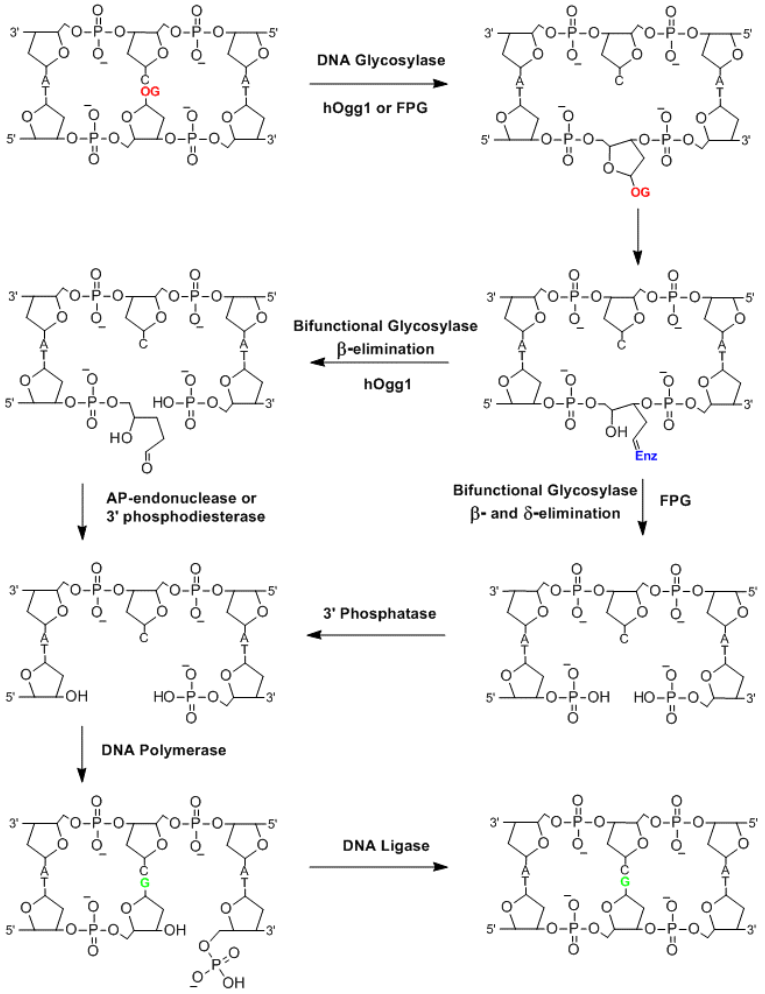

Figure 1.7. Stages of the base excision repair pathway involved in repairing guanine oxidative damage (OG in red).

1.4.1 Human oxoguanine glycosylase (hOgg1)

hOgg1 belongs to the endonuclease III Nth-like superfamily, which is characterized by a Helix-hairpin-Helix (HhH) motif followed by a Gly, Pro or Val-rich loop and contains an absolutely conserved catalytic Asp residue (Asp268, Figure 1.8).[75, 78] In addition, the HhH motif contains a Lys residue (Lys249) that serves as the catalytic nucleophile for bifunctional members of this family (Figure 1.8). hOgg1 is primarily responsible for the removal of OG and FapyG,[79, 80] but has been shown to possess some activity toward a few other substrates, such as N7-mFapyG, 4,6-diamino-5-formamidopyrimidine (FapyA), 8-oxo-7,8-dihydroadenine (8-oxoA), and 8-hydroxyguanine (8-OHG).[81, 82] The β-lyase activity of hOgg1 is much slower than the deglycosylation reaction[82] and, although the enzyme has been classified as a bifunctional glycosylase, a few studies have proposed the monofunctional nature of hOgg1.[63, 83]

Table 1.1. Classification of several DNA glycosylases that target oxidative DNA damage based on the substrate preference or structural features.

	Substrate Preference	
Substrate Examples	Tg,[a] DHT[b]	OG, FapyG
Glycosylase	Nth,[c] Nei[d]	hOgg1, FPG
	Structural Feature	
Motif	Helix-hairpin-Helix (HhH)	Helix-2Turn-Helix (H2TH)
Glycosylase	hOgg1, Nth	FPG, Nei

[a] Thymine glycol. [b] Dihydrothymine. [c] Endonuclease III. [d] Endonuclease VIII.

In addition to discrimination of OG or FapyG from canonical G, hOgg1 shows a preference for the nucleobase opposite the lesion (Figure 1.9). hOgg1 removes OG and FapyG from the OG:C and FapyG:C pairs at similar rates.[82] However, FapyG is removed from FapyG:A with a rate constant that is 60-fold greater than for OG from the OG:A base pair.[82] Moreover, FapyG removal by hOgg1 from FapyG:C is 47-fold more faster than from the corresponding FapyG:A pair.[82]

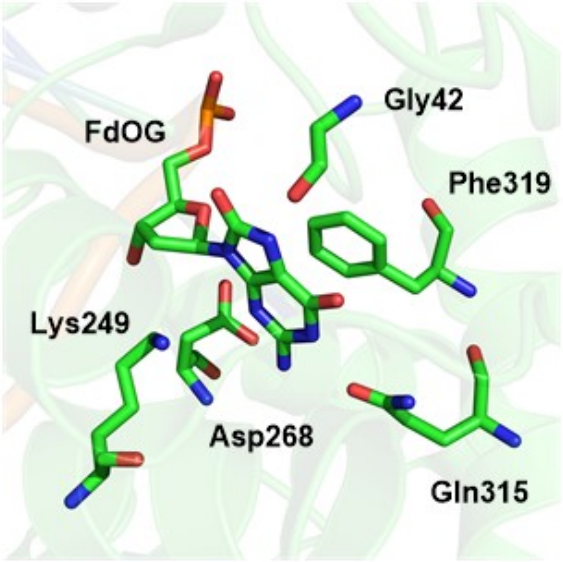

Figure 1.8. X-ray crystal structure of 2′-fluoro-dOG (FdOG) bound in the hOgg1 active site (PDB ID: 3KTU).

The rate constant for base removal from the OG:C or FapyG:C pairs is ~ 247- and 833-fold greater than the combined rate constant of the deglycosylation and β-elimination reactions, respectively.[82] The magnitude of the effect of the opposing base for FapyG removal is reduced compared to the efficiency of hOgg1 on the corresponding OG-containing duplexes, where there is a 3000-fold preference of hOgg1 for the OG:C over the OG:A base pair (Figure 1.9).[82]

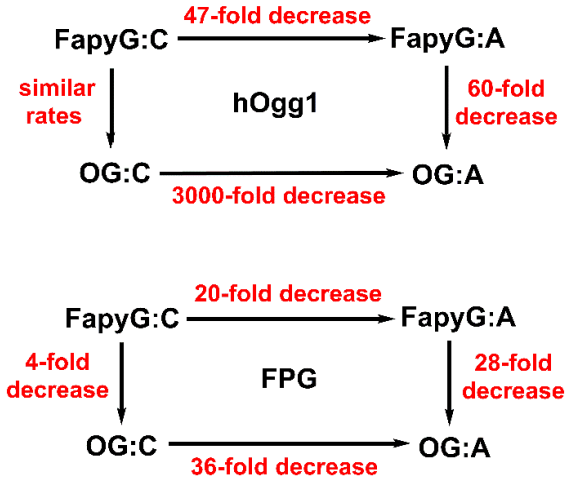

Figure 1.9. Relative rates of OG and FapyG removal when paired opposite C or A by hOgg1 (top) or FPG (bottom).[82]

The hOgg1 specificity for OG:C is proposed to be the result of five hydrogen bonds between the enzyme (Arg204, Asn149, and Arg154) and the opposing C, with the replacement of Arg154 with histidine eliminating the specificity.[84] This is further supported by analyzing the structure of an Ogg orthologue from the bacterium *Clostridium acetobutylicum* (*C. ac.*) Ogg, which can excise OG opposite all four natural nucleobases.[85] *C. ac.* Ogg maintains the fold and general DNA interactions of hOgg1, but Met132 replaces Arg154 of hOgg1, and two of the five hydrogen bonds with the opposing base are absent.[85, 86] Interestingly, the opposite-base preference is also observed in the AP-lyase activity of hOgg1 and the yeast Ogg1 homologue (hOgh1).[82, 87] Specifically, the preferential cleavage of an AP-site:C pairing suggests that hOgh1 interacts with the base in the complementary strand or discriminates between structural differences of the AP-site opposite different bases.[87]

hOgg1 is inactive toward natural guanine despite G differing from OG at only the N7 and C8 positions. Crystal structures have revealed the presence of a hydrogen bond between the backbone carbonyl of Gly42 and N7–H of OG.[84] Since there is no N7–H in G, it has been suggested that this hydrogen bond may be responsible for the observed substrate specificity.[48] However, excision of N7-mFapyG by hOgg1 indicates that N7–H···Gly42 might not be crucial for the recognition or removal. Other than the hydrogen bond between N7–H and Gly42, alignment of other active site residues is also important. Lys249Arg and Lys249Gln mutants have been shown to be catalytically impaired, supporting that Lys249 is the nucleophile.[88] Lys249 has also been proposed to play a role in OG recognition. Specifically, the interaction between the $Lys249^{+}$–$Cys253^{-}$ dipole and the local dipole between O8 and N7–H of OG has been proposed to be a factor in discrimination between G and OG.[89] However, a study on the Lys249Cys/Cys253Lys

mutant (inverted dipole) has shown that the $Lys249^{+}$–$Cys253^{-}$ dipole does not substantially affect processing of OG.[83] Moreover, Lys249 must be neutral to serve as the nucleophile, and quantum mechanical calculations have shown that the $Lys249^{+}$–$Cys253^{-}$ pair destabilizes the binding pocket, whereas the neutral counterpart stabilizes the complex.[60]

Several mutational studies have been conducted to disclose the roles of active site residues in the lesion recognition and/or binding steps. For example, replacing His270 with alanine (His270Ala) or leucine (His270Leu) greatly diminishes the ability of the enzyme to bind OG:C-containing DNA. Hence, the OG excision rate is decreased, presumably due to disruption of the hydrogen bond between His270 and the 5′-phosphate of OG.[90] Similarly, mutating Phe319, which stacks against the lesion (Figure 1.8) to Ala decreases both OG binding and excision.[84] Asp268 has been shown to be a critical residue, likely contributing to the stability and function of hOgg1. Specifically, Asp268Glu and Asp268Gln mutants show lower thermal stability compared to the WT enzyme.[91] Although mutating Asp268 to Ala or Asn diminishes the enzymatic activity, the lesion recognition is not affected.[92, 93]

1.4.2 Formamidopyrimidine glycosylase (FPG)

Formamidopyrimidine glycosylase (FPG) belongs to another DNA glycosylase family, which includes bacterial FPG (or MutM),[94] endonuclease VIII (Nei)[95, 96] and the mammalian homologs known as Neil1 (Nei-like 1).[97, 98] Members of this family share a Helix-2Turn-Helix (H2TH) motif.[73] FPG uses Arg258 to bind the minor groove of DNA with only the tip of the zinc finger, which enters from the major groove.[99] FPG uses a terminal proline (Pro2) as the nucleophile and a catalytic Glu (Figure 1.10). Similar to hOgg1, FPG processes 8-OHG and AP-sites, but has a broader range of substrates,

including 4-6-diamino-5-formamidopyrimidine (FapyA), 5-hydroxy-2′-deoxyuridine (5-OHU), oxazolone, oxidative derivatives of ethenoadenine, and several other types of oxidative damage.[95, 100-103] Interestingly, FPG does not remove the 8-oxoA lesion, which is structurally similar to OG from the DNA duplex.[100]

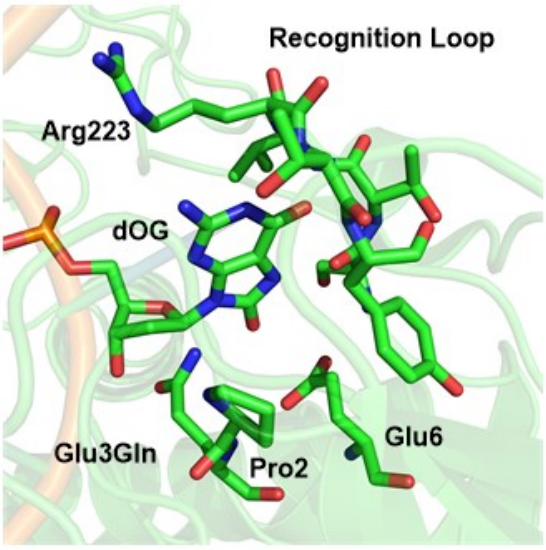

Figure 1.10. X-ray crystal structure of dOG bound in the active site of the Glu3Gln FPG mutant (PDB ID: 1R2Y).

As for hOgg1, FPG has been shown to exhibit an activity preference based on the nucleobase opposite OG (Figure 1.9).[82, 100] Specifically, FPG more efficiently processes OG opposite C than A, but only at an 18- to 35-fold greater rate,[82, 100] leading to appreciable activity toward OG regardless of the opposing (C or A) nucleobase. Moreover, OG opposite G, T, or an abasic site in DNA is readily cleaved by FPG.[82] Since OG does not form stable base pairs with G or T, the lack of a stable base pair at the lesion site may facilitate recognition and removal processes.[100] However, when the opposing base is A, the lesion is stabilized thermally (T_m) and thermodynamically (ΔG) relative to the G:A pair.[100, 104] This has been suggested to relate to the specificity constants of 110 and 270 for OG excision by FPG from OG:G and OG:T, respectively, compared to the

excision rate from the OG:C pair.[100] On the other hand, the lower specificity constant for the OG:A pair (18-fold lower than when OG is opposite C) has been proposed to arise from a combination of the high thermodynamic stability of the OG:A Hoogsteen pairing and weaker substrate binding.[105]

At the recognition step, Arg258 is the only residue that directly interacts with the WC face of OG.[99] However, since FPG excises Fapy moieties with either a hydrogen atom (*e.g.*, FapyG) or a methyl group at N7 (*e.g.*, N7-mFapyG), recognition through hydrogen bonding with N7 is unlikely. His89 and Arg109 interact with the strand opposing the lesion, and are suggested to be involved in kinking the DNA duplex.[73] Within the active site pocket, Arg223 or Lys217 (in Bacillus stearothermophilus (*B. st.*) and Escherichia coli (*E. coli*) FPG, respectively) has been proposed to participate in OG recognition.[106] Another residue that has been proposed to be involved in the recognition step is Met73 (*E. coli* FPG); however, since the interaction occurs between the backbone carbonyl and the lesion, mutational studies would not provide additional information.[106]

1.5 Research Questions for BER

Despite exhibiting a preference for the lesion and opposing base,[82] sharing several steps in the early stages of BER (Figure 1. 7), and targeting major guanine oxidation products (*e.g.*, OG and FapyG), hOgg1 and FPG are different in several ways. Specifically, hOgg1 and FPG belong to different structural families. Furthermore, hOgg1 and FPG have different substrate specificities.[73, 82, 100, 101] hOgg1 only efficiently removes OG and FapyG, while FPG efficiently targets a wide spectrum of damaged purines and pyrimidines.[73, 82] Moreover, the opposing base preference in the two enzymes is different.[82] hOgg1 is basically inert toward OG:A, whereas FPG excises OG from the OG:A pair at an appreciable rate. Interestingly, both enzymes show reduced opposing

base preference for FapyG. Particularly, although FapyG is removed from the FapyG:C pair at a greater rate by both enzymes, the enzymatic activities toward FapyG:A are not negligible. Finally, although FPG possesses β- and δ-lyase activities, the β-elimination reaction is weakly catalyzed by hOgg1. In fact, the bifunctional mode of hOgg1 has been doubted by both experimental and computational studies.[63, 83] As mentioned above, the previous studies exploring the recognition and catalytic mechanisms employed by hOgg1 and FPG provide a wealth of information. However, the actual mechanism of substrate recognition, opposing base preferences, and catalysis that lead to the differential enzymatic functions and catalytic activities are not completely understood. In the following sections, facts and open questions about the lesion recognition stage and the chemical steps catalyzed by hOgg1 and FPG are discussed in detail.

1.5.1 Recognition

Although factors early in the recognition step have been proposed to play a role in dictating the observed hOgg1 and FPG activity dependence on the opposing base (*i.e.*, differences in the base pair strengths,[100, 104] direct contacts between the enzyme and the opposing base,[84, 86, 107] or disruptions of the DNA–enzyme interface that destabilize the pre-catalytic complex[49]), differences in the glycosidic orientation of OG paired opposite C and A may influence how the substrate binds in the active site and therefore the enzymatic activity. Specifically, crystal structures of hOgg1 bound to OG-containing DNA reveal that *anti*-OG binds in the active site when paired opposite C in the DNA duplex.[84, 89, 108] Moreover, crystal structures show that OG from OG:C and OG:A pairs in DNA binds in the *anti* conformation in the active site of *Methanocaldococcus jannaschii* (*M. jan.*) and *C. ac.* Ogg.[86, 109] In contrast, an available crystal structure of FPG suggests that OG paired opposite C binds in the *syn* conformation in the active site.[110]

Nevertheless, FapyG paired opposite C binds in the *anti* conformation in the FPG active site.[111] These differences in the OG binding conformation could reflect the ability of OG to rotate about the glycosidic bond during the base flipping step or in the active-site pocket, and/or inherent differences in the capability of these enzymes to bind different OG glycosidic conformations.

Molecular dynamics (MD) studies on hOgg1 include comparison of the eversion of *anti*-OG and G from the DNA duplex into the active site,[112] and the dynamics of wild type[113] (WT) and mutant[114] hOgg1 bound to DNA containing *anti*-OG. These studies have provided valuable insights into the lesion recognition step and the potential role of active site residues in the recognition and catalysis steps. Nevertheless, the ability of the hOgg1 active site to accommodate different OG glycosidic conformations has not been addressed. Moreover, the potential effect of the OG glycosidic conformer on the hOgg1 active site and its role in enzymatic activity are not clear.

Despite MD simulations being used to consider how *anti*- and/or *syn*-OG are bound in the FPG active site,[49-51, 106, 115] and identify potential catalytically important active site residues,[50, 51, 106, 115] the preferred OG binding conformation in FPG is a matter of increasing debate. Specifically, two separate MD studies concluded that the *syn* conformer is better accommodated in the active site than *anti*-OG,[51, 115] while another MD study suggests that *anti*-OG binding is preferred.[106] However, the work by Perlow-Poehnelt *et al.* proposed that *syn*-OG can rotate into an *anti* conformation in the FPG active site.[50] To complement these works, a recent study reveals that initiating MD simulations from a *high anti* OG conformation (*i.e.,* $\chi = -64°$) in the Lactococcus lactic (*L. lac.*) FPG active site results in spontaneous drift towards the *anti*-OG orientation.[49] Thus, the actual dependence of FPG lesion binding on the glycosidic conformation has

not been entirely resolved. The effects of the OG and FapyG glycosidic conformations on the function of hOgg1 and FPG is investigated in this thesis.

1.5.2 Removal

The first chemical step during BER involves nucleophilic attack (by a lysine (Lys249) in hOgg1 and a proline (Pro2) in FPG), which cleaves the glycosidic bond and opens the 2′-deoxyribose ring at C1′–O4′ (Figure 1.11). The resulting intermediate is a DNA–enzyme crosslink with a double bond between the nucleophilic nitrogen and the C1′ atom of the ring-opened sugar, which is known as a Schiff base. Next, a β-elimination reaction occurs through proton abstraction from C2′ by a general base, which subsequently leads to the C3′–O3′ bond cleavage (*i.e.*, 3′-phosphate elimination). It is assumed that once this step is completed, hOgg1 dissociates from DNA, while FPG catalyzes the δ-elimination reaction through which the C4′ proton is abstracted and the 5′-phosphate departs. The final step recovers the enzyme through hydrolysis of the Schiff base and leaves the nicked DNA strand for further processing (Figure 1.11).

Despite the knowledge about the general chemical steps of the catalytic process, several aspects of the reactions catalyzed by hOgg1 and FPG have not been completely explained. For example, the order of the glycosidic bond cleavage and ring-opening reaction is not clear and has been a subject of controversy. In fact, besides the proposed S_N1 mechanisms for the deglycosylation step,[61, 91, 116-118] two S_N2 mechanisms have been proposed for the first chemical step.[73, 75] In the first mechanism, nucleophilic attack at C1′ of 2′-deoxyribose leads to nucleobase departure (Figure 1.12a). In the second mechanism, an acid (Asp268 in hOgg1 and Glu3 in FPG, Figures 1.8 and 1.10, respectively) facilitates cleavage of the C1′–O4′ bond of 2′-deoxyribose upon nucleophilic attack (Figure 1.12b). Although quantum mechanical calculations have predicted that the second mechanism

leads to a lower barrier for OG excision by both enzymes,[53, 60] more structural evidence is required to unequivocally identify the favored pathway, a goal of the present thesis.

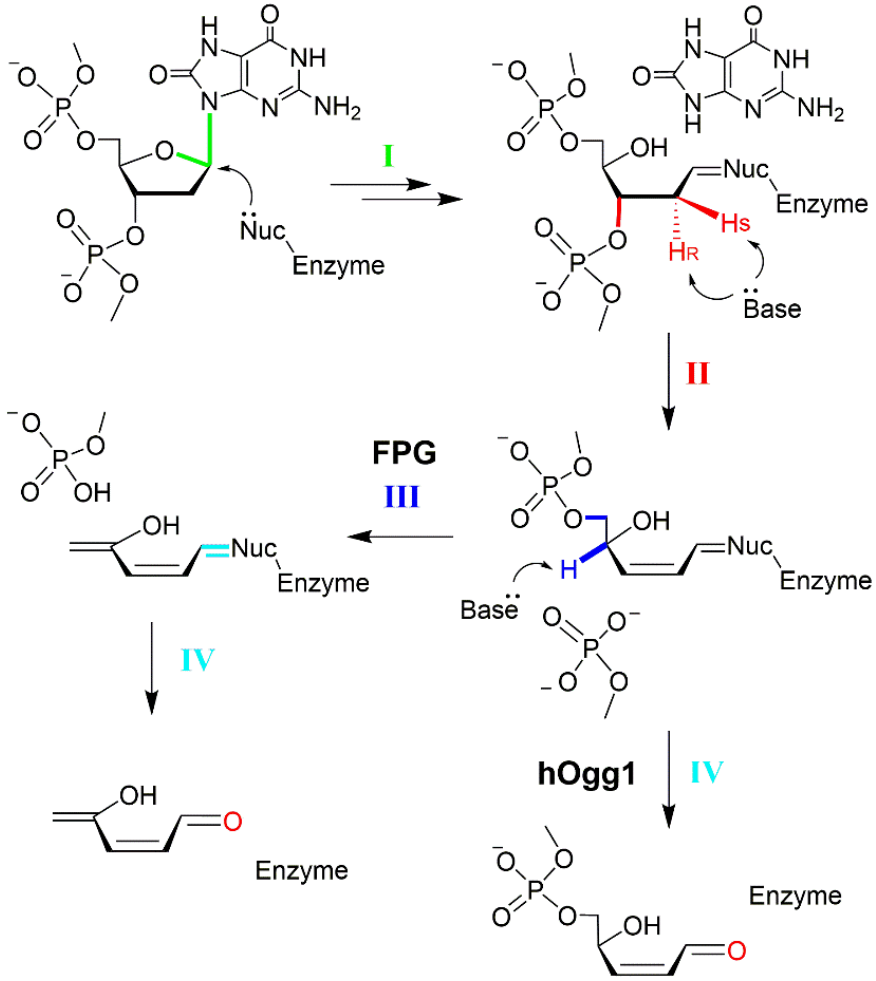

Figure 1.11. Proposed chemical steps catalyzed by hOgg1 and FPG during BER: deglycosylation/ring opening (I), β-elimination (II), δ-elimination (for FPG only) (III) and crosslink hydrolysis (IV).

Furthermore, due to discrepancies in the proposed mechanisms, the exact roles of the catalytic Asp268 in hOgg1 and Glu3 in FPG remain elusive. Specifically, Asp268 has been proposed to initiate nucleoside hydrolysis[83] or stabilize the Schiff base intermediate.[91] Both Asp268[84, 91] and Glu3[73, 119] have also been suggested to deprotonate the nucleophile, and transfer a proton to 2′-deoxyribose of the damaged nucleotide to facilitate the ring-opening step.

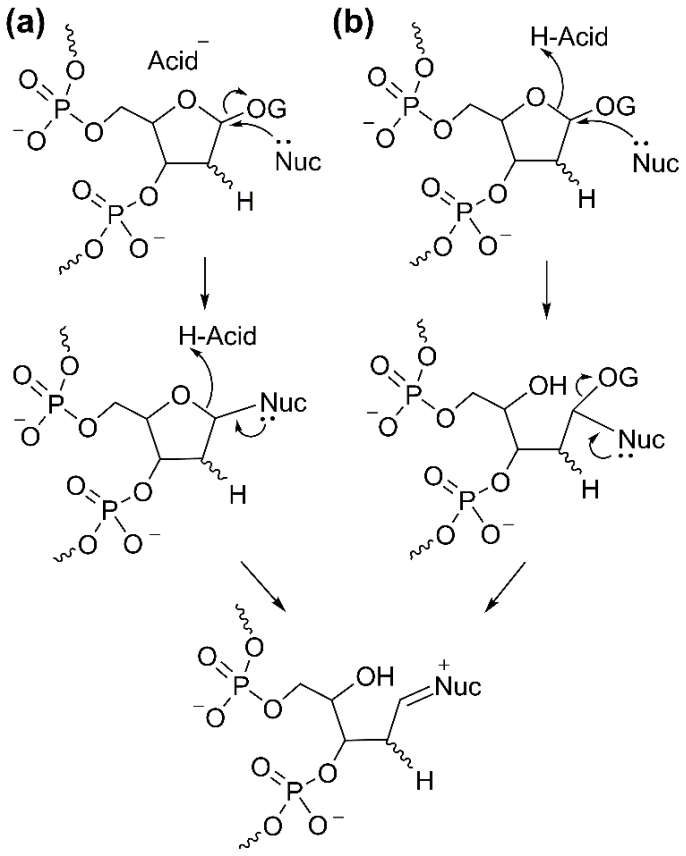

Figure 1.12. Proposed mechanisms for the formation of the DNA–enzyme crosslink intermediate in the first BER chemical step facilitated by a bifunctional glycosylase, which involve deglycosylation (a) prior to or (b) after ring opening.

Furthermore, despite some assumptions, the identity of the general base for proton abstraction (C2′–H or C4′–H during β- and δ-elimination catalysis, respectively), as well as the identity of the C2′ proton (*pro*-S or *pro*-R, also H_S or H_R, Figure 1.11) that promotes the β-elimination reaction, are unknown for both enzymes. Although the pK_a of C2′–H (~ 18) and C4′–H (~ 50) are lower in the Schiff base than the nucleotide,[120, 121] strong bases are required to abstract the protons.[73] It has been proposed that a crystallographic water in the proximity of C2′ (3.5 Å) plays this role in *B. st.* FPG during the β-lyase phase.[105] However, this molecule was not observed in related enzymes, including *E. coli* FPG. The excised OG has also been proposed (primarily in the case of

hOgg1)[84, 122] to facilitate proton abstraction; however, this proposal does not explain the β-lyase activity of the enzymes on an abasic site.[73] In contrast, OG has been suggested to inhibit the elimination step in hOgg1.[123] Additionally, since OG is not present in a crystal structure of the borohydride-trapped DNA–FPG crosslink, it has been proposed that excised OG diffuses away after cleavage.[105] Finally, in addition to unknowns surrounding C2′ proton abstraction, there are currently no proposals for the identity of the general base that abstracts the C4′ proton in the δ-elimination reaction catalyzed by FPG.

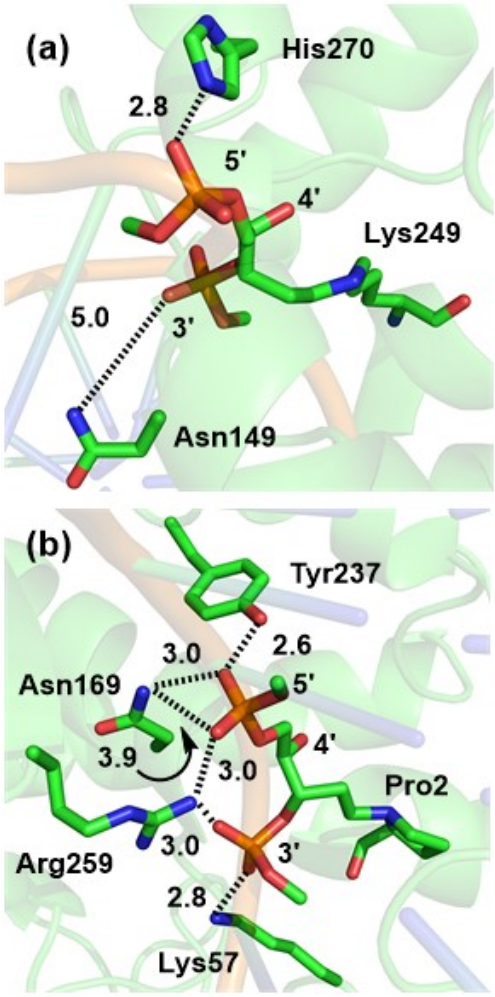

Figure 1.13. Interactions between active site residues, and the 3′- and 5′-phosphate groups in a) hOgg1 (PDB ID: 1LWY) and b) FPG (PDB ID: 1L1Z). Select distances are provided (Å). Main chain residues and water molecules are not shown for clarity.

Although there has been some speculation about the general base that abstracts the C2′ proton in these enzymes, even less is known about the phosphate elimination step. A crystal structure of the Schiff base intermediate in the hOgg1 active site (PDB ID: 1LWY)[122] shows that the 3′- or 5′-phosphates of the crosslink only interact with the Asn149 backbone or the His270 sidechain, respectively (Figure 1.13a). On the other hand, crystal structures of FPG bound to various substrates reveal that the 3′- and 5′-phosphates are in close proximity to conserved residues in FPG.[105, 124] Specifically, the 3′-phosphate interacts with Lys57 and Arg259, while the 5′-phosphate interacts with Asn169, Arg259 and Tyr242 (Figure 1.13b). It has been proposed that Lys57 and Arg259 protonate the 3′- and 5′-phosphates, respectively, while the other residues further stabilize the increasing negative charge formed upon phosphate departure through hydrogen-bonding contacts. Although fewer stabilizing contacts between hOgg1 and 3′-phosphate might to some extent explain the observed lower β-elimination rate in hOgg1, it cannot explain the absolute lack of δ-lyase activity in hOgg1. This thesis will strive to address the intrinsic mechanisms of these reactions.

1.6 Thesis Overview

The main focus of this thesis is to provide a deeper understanding of guanine oxidation damage, recognition, and removal processes facilitated by hOgg1 and FPG. Specifically, in Chapter 2, the formation mechanisms of FapyG and OG are considered using several quantum mechanical methods (density functional theory in combination with various basis sets) and model sizes (from nucleobase to trimer of nucleoside pairs that involve ~ 25 to 200 atoms). In this chapter, the importance of the computational

model (*i.e.*, combination of model size, environment, and method/basis set) that provides the best accuracy-computational cost balance is highlighted.

In Chapters 3 and 4, molecular dynamic (MD) simulations are employed to study the lesion–active site interactions between different *anti* and *syn* OG and FapyG glycosidic conformers. This work reveals how hOgg1 and FPG distinguish between the glycosidic conformers of OG and FapyG.

In the remaining chapters, the chemical steps involved in the catalytic mechanisms of damage removal by both enzymes are investigated in detail using quantum mechanical methods. Specifically, Chapter 5 models the first three BER chemical steps, namely deglycosylation, ring opening and β-elimination, and investigates the potential role of the cleaved OG as the cofactor for FPG. The effects of the nucleophile and subsequent Schiff base intermediate on the energy barriers are also evaluated by comparing to a previous study for hOgg1.[63] Chapter 6 studies the β- and δ-elimination reactions, and compares the inherent chemistry of the two reactions, providing a rationale for the lack of δ-elimination catalysis (δ-lyase activity) by hOgg1. Chapter 7 takes a closer look at the β- and δ-elimination reactions, as well as crosslink hydrolysis, by examining various mechanisms using different Schiff base models and general acid/bases. A new pathway for the β-elimination facilitated by hOgg1 is proposed. Similarly, new pathways for the β- and δ-elimination reactions catalyzed by FPG are suggested.

Finally, in Chapter 8, general conclusions regarding the computational approach/method, model size, and catalytic mechanisms of the two DNA glycosylases, and potential directions for future work are presented. The present thesis highlights the complementary and irreplaceable/unique role of computer modelling in expanding our

knowledge of DNA damage and repair processes by unveiling atomic-level detail of the formation, recognition, and removal of oxidative DNA damage.

1.7 References

1. A. Klungland, I. Rosewell, S. Hollenbach, E. Larsen, G. Daly, B. Epe, E. Seeberg, T. Lindahl and D. E. Barnes, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, 96, 13300–13305.

2. M. A. Kalam, K. Haraguchi, S. Chandani, E. L. Loechler, M. Moriya, M. M. Greenberg and A. K. Basu, *Nucleic Acids Research*, 2006, 34, 2305–2315.

3. M. D. Evans, M. Dizdaroglu and M. S. Cooke, *Mutat. Res.*, 2004, 567, 1–61.

4. C. J. Burrows and J. G. Muller, *Chem. Rev. (Washington, DC, U. S.)*, 1998, 98, 1109–1151.

5. S. Steenken, *Chem. Rev. (Washington, DC, U. S.)*, 1989, 89, 503–520.

6. R. N. Barnett, A. Bongiorno, C. L. Cleveland, A. Joy, U. Landman and G. B. Schuster, *J. Am. Chem. Soc.*, 2006, 128, 10795–10800.

7. A. P. Breen and J. A. Murphy, *Free Radic. Biol. Med.*, 1995, 18, 1033–1077.

8. T. Lindahl and D. E. Barnes, *Cold Spring Harb Symp Quant Biol*, 2000, 65, 127–133.

9. V. A. Bohr, O. P. Ottersen and T. Tonjum, *Neuroscience*, 2007, 145, 1183–1186.

10. C. Greenman, P. Stephens, R. Smith, G. L. Dalgliesh, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, S. Edkins, S. O/'Meara, I. Vastrik, E. E. Schmidt, T. Avis, S. Barthorpe, G. Bhamra, G. Buck, B. Choudhury, J. Clements, J. Cole, E. Dicks, S. Forbes, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, A. Jenkinson, D. Jones, A. Menzies, T. Mironenko, J. Perry, K. Raine, D. Richardson, R. Shepherd, A. Small, C. Tofts, J. Varian, T. Webb, S. West, S. Widaa, A. Yates, D. P. Cahill, D. N. Louis, P. Goldstraw, A. G. Nicholson, F. Brasseur, L. Looijenga, B. L. Weber, Y.-E. Chiew, A. deFazio, M. F. Greaves, A. R. Green, P. Campbell, E. Birney, D. F. Easton, G. Chenevix-Trench, M.-H. Tan, S. K. Khoo, B. T. Teh, S. T. Yuen, S. Y. Leung, R. Wooster, P. A. Futreal and M. R. Stratton, *Nature (London, U. K.)*, 2007, 446, 153–158.

11. T. B. Kryston, A. B. Georgiev, P. Pissis and A. G. Georgakilas, *Mutat Res*, 2011, 711, 193–201.

12. D. Ziech, R. Franco, A. Pappa and M. I. Panayiotidis, *Mutat Res*, 2011, 711, 167–173.

13. J. Cadet, T. Douki and J.-L. Ravanat, *Free Radic. Biol. Med.*, 2010, 49, 9–21.

14. S. Kanvah, J. Joseph, G. B. Schuster, R. N. Barnett, C. L. Cleveland and U. Landman, *Acc. Chem. Res.*, 2010, 43, 280–287.

15. B. van Loon, E. Markkanen and U. Hübscher, *DNA Repair*, 2010, 9, 604–616.

16. B. Tudek, A. Winczura, J. Janik, A. Siomek, M. Foksinski and R. Olinski, *Am. J. Transl. Res.*, 2010, 2, 254–284.

17. M. Foksinski, R. Rozalski, J. Guz, B. Ruszkowska, P. Sztukowska, M. Piwowarski, A. Klungland and R. Olinski, *Free Radic Biol Med*, 2004, 37, 1449–1454.

18. S. D. Wetmore, R. J. Boyd and L. A. Eriksson, *Chem. Phys. Lett.*, 2000, 322, 129–135.

19. A. Yadav and P. C. Mishra, *Chem. Phys.*, 2012, 405, 76–88.

20. G. Pratviel and B. Meunier, *Chemistry*, 2006, 12, 6018–6030.

21. M. S. Cooke, M. D. Evans, M. Dizdaroglu and J. Lunec, *Faseb J*, 2003, 17, 1195–1214.

22. S. Shibutani, M. Takeshita and A. P. Grollman, *Nature (London, U. K.)*, 1991, 349, 431–434.

23. V. Duarte, J. G. Muller and C. J. Burrows, *Nucleic Acids Res*, 1999, 27, 496–502.

24. P. T. Henderson, J. C. Delaney, J. G. Muller, W. L. Neeley, S. R. Tannenbaum, C. J. Burrows and J. M. Essigmann, *Biochemistry*, 2003, 42, 9257–9262.

25. O. Kornyushyna, A. M. Berges, J. G. Muller and C. J. Burrows, *Biochemistry*, 2002, 41, 15304–15314.

26. P. T. Henderson, J. C. Delaney, F. Gu, S. R. Tannenbaum and J. M. Essigmann, *Biochemistry*, 2002, 41, 914–921.

27. G. A. Locatelli, H. Pospiech, N. Tanguy Le Gac, B. van Loon, U. Hubscher, S. Parkkinen, J. E. Syvaoja and G. Villani, *Biochem J*, 2010, 429, 573–582.

28. N. R. Jena, *J Biosci*, 2012, 37, 503–517.

29. N. R. Jena and P. C. Mishra, *Free Radic Biol Med*, 2012, 53, 81–94.

30. B. Tudek, *J Biochem Mol Biol*, 2003, 36, 12–19.

31. M. Dizdaroglu, G. Kirkali and P. Jaruga, *Free Radic. Biol. Med.*, 2008, 45, 1610–1621.

32. D. C. Malins and R. Haimanot, *Cancer Res.*, 1991, 51, 5430–5432.

33. D. C. Malins, N. L. Polissar and S. J. Gunselman, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, 93, 2557–2563.

34. W. R. Markesbery and M. A. Lovell, *Antioxid Redox Signal*, 2006, 8, 2039–2045.

35. E. A. Kemeleva, O. I. Sinitsyna, K. A. Conlon, M. Berrios, N. G. Kolosova, D. O. Zharkov, E. A. Vasyunina and G. A. Nevinsky, *Biochemistry*, 2006, 71, 612–618.

36. F. L. Muller, M. S. Lustgarten, Y. Jang, A. Richardson and H. Van Remmen, *Free Radic. Biol. Med.*, 2007, 43, 477–503.

37. W. L. Neeley and J. M. Essigmann, *Chem Res Toxicol*, 2006, 19, 491–505.

38. J. S. Stover, M. Ciobanu, D. E. Cliffel and C. J. Rizzo, *J. Am. Chem. Soc.*, 2007, 129, 2074–2081.

39. T. Gimisis and C. Cismaş, *Eur. J. Org. Chem.*, 2006, 2006, 1351–1378.

40. K. Lewis, K. Copeland and G. Hill, *Int. J. Quantum Chem.*, 2014, 114, 1678–1684.

41. M. Puiatti, D. M. Vera and A. B. Pierini, *Phys Chem Chem Phys*, 2008, 10, 1394–1399.

42. J. Gu, Y. Xie and H. F. Schaefer, *Nucleic Acids Research*, 2007, 35, 5165–5172.

43. B. T. Psciuk, R. L. Lord, B. H. Munk and H. B. Schlegel, *J. Chem. Theory Comput.*, 2012, 8, 5107–5123.

44. B. T. Psciuk and H. B. Schlegel, *J. Phys. Chem. B*, 2013, 117, 9518–9531.

45. I. Saito, M. Takayama, H. Sugiyama, K. Nakatani, A. Tsuchida and M. Yamamoto, *J. Am. Chem. Soc.*, 1995, 117, 6406–6407.

46. H. Sugiyama and I. Saito, *J. Am. Chem. Soc.*, 1996, 118, 7063–7068.

47. V. Verdolino, R. Cammi, B. H. Munk and H. B. Schlegel, *J. Phys. Chem. B*, 2008, 112, 16860–16873.

48. N. R. Jena and P. C. Mishra, *ChemPhysChem*, 2013, 14, 3263–3270.

49. A. V. Popov, A. V. Endutkin, Y. N. Vorobjev and D. O. Zharkov, *BMC Struct. Biol.*, 2017, 17, 5.

50. R. A. Perlow-Poehnelt, D. O. Zharkov, A. P. Grollman and S. Broyde, *Biochemistry*, 2004, 43, 16092–16105.

51. P. Amara, L. Serre, B. Castaing and A. Thomas, *Protein Sci.*, 2004, 13, 2009–2021.

52. I. D. Blank, K. Sadeghian and C. Ochsenfeld, *Scientific Reports*, 2015, 5, 10369.

53. K. Sadeghian, D. Flaig, I. D. Blank, S. Schneider, R. Strasser, D. Stathis, M. Winnacker, T. Carell and C. Ochsenfeld, *Angew. Chem., Int. Ed. Engl.*, 2014, 53, 10044–10048.

54. S. Sowlati-Hashjin and S. D. Wetmore, *J. Phys. Chem. B*, 2014, 118, 14566–14577.

55. S. Sowlati-Hashjin and S. D. Wetmore, *Phys. Chem. Chem. Phys.*, 2015, 17, 24696–24706.

56. J. Sebera, Y. Hattori, D. Sato, D. Reha, R. Nencka, T. Kohno, C. Kojima, Y. Tanaka and V. Sychrovsky, *Nucleic Acids Res*, 2017, 45, 5231–5242.

57. J. Sebera, L. Trantirek, Y. Tanaka, R. Nencka, J. Fukal and V. Sychrovsky, *RSC Advances*, 2014, 4, 44043–44051.

58. J. Sebera, L. Trantirek, J. Fukal, Y. Tanaka and V. Sychrovsky, *Chemicke Listy*, 2014, 108, 364–367.

59. J. Sebera, L. Trantírek, Y. Tanaka and V. Sychrovsky, *J. Phys. Chem. B*, 2012, 116, 12535–12544.

60. K. Sadeghian and C. Ochsenfeld, *J. Am. Chem. Soc.*, 2015, 137, 9824-9831.

61. P. Schyman, J. Danielsson, M. Pinak and A. Laaksonen, *J. Phys. Chem. A*, 2005, 109, 1713–1719.

62. M. Calvaresi, A. Bottoni and M. Garavelli, *J. Phys. Chem. B*, 2007, 111, 6557–6570.

63. J. L. Kellie and S. D. Wetmore, *J. Phys. Chem. B*, 2012, 116, 10786–10797.

64. P. Liu, Q. Wang, M. Niu and D. Wang, *Scientific Reports*, 2017, 7, 7798.

65. B. H. Munk, C. J. Burrows and H. B. Schlegel, *Chem. Res. Toxicol.*, 2007, 20, 432–444.

66. G. M. Blackburn and M. Gait, J, eds., *Nucleic Acids in Chemistry and Biology*, Oxford University Press, New York, 1996.

67. R. D. Wood, M. Mitchell, J. Sgouros and T. Lindahl, *Science (Washington, DC, U. S.)*, 2001, 291, 1284–1289.

68. T. Nouspikel, *Cell Mol Life Sci*, 2009, 66, 994–1009.

69. B. Sedgwick, P. A. Bates, J. Paik, S. C. Jacobs and T. Lindahl, *DNA Repair*, 2007, 6, 429–442.

70. H. E. Krokan and M. Bjørås, *Cold Spring Harbor Perspectives in Biology*, 2013, 5, a012583.

71. D. O. Zharkov, *Cell. Mol. Life Sci.*, 2008, 65, 1544–1565.

72. J. T. Stivers and Y. L. Jiang, *Chem. Rev. (Washington, DC, U. S.)*, 2003, 103, 2729–2759.

73. D. O. Zharkov, G. Shoham and A. P. Grollman, *DNA Repair*, 2003, 2, 839–862.

74. P. J. Berti and J. A. B. McCann, *Chem. Rev. (Washington, DC, U. S.)*, 2006, 106, 506–555.

75. H. M. Nash, S. D. Bruner, O. D. Scharer, T. Kawate, T. A. Addona, E. Spooner, W. S. Lane and G. L. Verdine, *Curr Biol*, 1996, 6, 968–980.

76. D. M. Wilson and D. Barsky, *Mutat. Res., DNA Repair*, 2001, 485, 283–307.

77. G. L. Dianov, S. L. Allinson, H. Budworth and K. Sleeth, *Eukaryotic DNA Damage Surveill. Repair*, 2004, 27–48.

78. M. M. Thayer, H. Ahern, D. Xing, R. P. Cunningham and J. A. Tainer, *Embo J*, 1995, 14, 4108–4120.

79. R. Lu, H. M. Nash and G. L. Verdine, *Curr Biol*, 1997, 7, 397–407.

80. J. Hu, N. C. de Souza-Pinto, K. Haraguchi, B. A. Hogue, P. Jaruga, M. M. Greenberg, M. Dizdaroglu and V. A. Bohr, *J. Biol. Chem.*, 2005, 280, 40544–40551.

81. A. Jensen, G. Calvayrac, B. Karahalil, V. A. Bohr and T. Stevnsner, *J. Biol. Chem.*, 2003, 278, 19541–19548.

82. N. Krishnamurthy, K. Haraguchi, M. M. Greenberg and S. S. David, *Biochemistry*, 2008, 47, 1043-1050.

83. B. Dalhus, M. Forsbring, I. H. Helle, E. S. Vik, R. J. Forstrøm, P. H. Backe, I. Alseth and M. Bjørås, *Structure (Cambridge, MA, U. S.)*, 2011, 19, 117–127.

84. S. D. Bruner, D. P. G. Norman and G. L. Verdine, *Nature (London, U. K.)*, 2000, 403, 859–866.

85. S. M. Robey-Bond, R. Barrantes-Reynolds, J. P. Bond, S. S. Wallace and V. Bandaru, *Biochemistry*, 2008, 47, 7626–7636.

86. F. Faucher, S. S. Wallace and S. Doublié, *DNA Repair*, 2009, 8, 1283–1289.

87. M. Bjoras, L. Luna, B. Johnson, E. Hoff, T. Haug, T. Rognes and E. Seeberg, *EMBO J.*, 1997, 16, 6314–6322.

88. N. Guibourt, B. Castaing, P. A. Van Der Kemp and S. Boiteux, *Biochemistry*, 2000, 39, 1716–1724.

89. A. Banerjee, W. Yang, M. Karplus and G. L. Verdine, *Nature (London, U. K.)*, 2005, 434, 612–618.

90. P. A. van der Kemp, J. B. Charbonnier, M. Audebert and S. Boiteux, *Nucleic Acids Research*, 2004, 32, 570–578.

91. D. P. G. Norman, S. J. Chung and G. L. Verdine, *Biochemistry*, 2003, 42, 1564–1572.

92. T. A. Rosenquist, D. O. Zharkov and A. P. Grollman, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, 94, 7429–7434.

93. H. E. Krokan, R. Standal and G. Slupphaug, *J. Biochem.*, 1997, 325, 1–16.

94. S. Boiteux, T. R. O'Connor and J. Laval, *Embo J*, 1987, 6, 3177–3183.

95. D. Y. Jiang, Z. Hatahet, R. J. Melamede, Y. W. Kow and S. S. Wallace, *J. Biol. Chem.*, 1997, 272, 32230–32239.

96. Y. Saito, F. Uraki, S. Nakajima, A. Asaeda, K. Ono, K. Kubo and K. Yamamoto, *J. Bact.*, 1997, 179, 3783–3785.

97. I. Morland, V. Rolseth, L. Luna, T. Rognes, M. Bjoras and E. Seeberg, *Nucleic Acids Res*, 2002, 30, 4926-4936.

98. V. Bandaru, S. Sunkara, S. S. Wallace and J. P. Bond, *DNA Repair*, 2002, 1, 517–529.

99. E. Dogliotti, P. Fortini, B. Pascucci and E. Parlanti, *Prog Nucleic Acid Res Mol Biol*, 2001, 68, 3–27.

100. J. Tchou, V. Bodepudi, S. Shibutani, I. Antoshechkin, J. Miller, A. P. Grollman and F. Johnson, *J. Biol. Chem.*, 1994, 269, 15318–15324.

101. S. Boiteux, E. Gajewski, J. Laval and M. Dizdaroglu, *Biochemistry*, 1992, 31, 106–110.

102. L. H. Breimer, *Nucleic Acids Research*, 1984, 12, 6359–6367.

103. A. Karakaya, P. Jaruga, V. A. Bohr, A. P. Grollman and M. Dizdaroglu, *Nucleic Acids Research*, 1997, 25, 474–479.

104. G. E. Plum, A. P. Grollman, F. Johnson and K. J. Breslauer, *Biochemistry*, 1995, 34, 16148–16160.

105. J. C. Fromme and G. L. Verdine, *Nat Struct Mol Biol*, 2002, 9, 544–552.

106. E. I. Zaika, R. A. Perlow, E. Matz, S. Broyde, R. Gilboa, A. P. Grollman, D. O. Zharkov, *J. Biol. Chem.*, 2004, 279, 4849–4861.

107. D. O. Zharkov, T. A. Rosenquist, S. E. Gerchman and A. P. Grollman, *J. Biol. Chem.*, 2000, 275, 28607–28617.

108. C. T. Radom, A. Banerjee and G. L. Verdine, *J. Biol. Chem.*, 2007, 282, 9182–9194.

109. F. Faucher, S. M. Robey-Bond, S. S. Wallace and S. Doublié, *J. Mol. Biol.*, 2009, 387, 669–679.

110. J. C. Fromme and G. L. Verdine, *J. Biol. Chem.*, 2003, 278, 51543–51548.

111. F. Coste, M. Ober, T. Carell, S. Boiteux, C. Zelwer and B. Castaing, *J. Biol. Chem.*, 2004, 279, 44074–44083.

112. H. Li, A. V. Endutkin, C. Bergonzo, L. Fu, A. Grollman, D. O. Zharkov and C. Simmerling, *J. Am. Chem. Soc.*, 2017, 139, 2682–2692.

113. M. Pinak, *J. Comput. Chem.*, 2003, 24, 898–907.

114. C. M. Crenshaw, K. Nam, K. Oo, P. S. Kutchukian, B. R. Bowman, M. Karplus and G. L. Verdine, *J. Biol. Chem.*, 2012, 287, 24916–24928.

115. K. Song, V. Hornak, C. D. Santos, A. P. Grollman and C. Simmerling, *Biochemistry*, 2006, 45, 10886–10894.

116. S. J. Chung and G. L. Verdine, *Chem. Biol.*, 2004, 11, 1643–1649.

117. A. Banerjee and G. L. Verdline, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 15020–15025.

118. M. V. Rogacheva and S. A. Kuznetsova, *Russ. Chem. Rev.*, 2008, 77, 765–788.

119. M. Sugahara, T. Mikawa, T. Kumasaka, M. Yamamoto, R. Kato, K. Fukuyama, Y. Inoue and S. Kuramitsu, *The EMBO Journal*, 2000, 19, 3857–3869.

120. J. A. Gerlt and P. G. Gassman, *J. Am. Chem. Soc.*, 1993, 115, 11552–11568.

121. W. W. Cleland and M. M. Kreevoy, *Science (Washington, DC, U. S.)*, 1994, 264, 1887–1890.

122. J. C. Fromme, S. D. Bruner, W. Yang, M. Karplus and G. L. Verdine, *Nat Struct Mol Biol*, 2003, 10, 204–211.

123. I. Morland, L. Luna, E. Gustad, E. Seeberg and M. Bjørås, *DNA Repair*, 2005, 4, 381–387.

124. D. O. Zharkov, G. Golan, R. Gilboa, A. S. Fernandes, S. E. Gerchman, J. H. Kycia, R. A. Rieger, A. P. Grollman and G. Shoham, *EMBO J.*, 2002, 21, 789–800.

Chapter 2: Oxidation of Guanine

2.1 Introduction

Due to the prevalence and negative consequences of guanine oxidation, the mechanism of guanine oxidation, the characteristics of oxidative products, and the pathway for the repair of the damaged products have been the subject of numerous experimental studies.[1-6] In parallel, computational studies have investigated various aspects of damage to guanine,[7-18] and the associated repair processes[19-29] in bacteria[20-23] and human cells.[19, 24-31] Specifically, quantum mechanical calculations have been used to characterize the mechanisms for the formation of OG,[13, 32-33] FapyG,[11] Sp,[12] and Gh (Figure 2.1).[12] These works have used B3LYP, small Pople basis sets, and nucleobase models to map different reaction mechanisms for each damage product. These studies have identified the most favorable reaction pathways for the formation of a diverse set of guanine oxidative products. Furthermore, in agreement with experimental IR spectral data, the calculations revealed a range of products can be formed from a common intermediate generated upon guanine oxidation.[11] In the case of Sp and Gh, the preferred pathways predicted by DFT are consistent with the experimentally-observed pH dependency of their formation.[34-37]

As was mentioned in Chapter 1, in the previous computational studies,[11-13] the effects of the solvent and DNA environment on geometries have been neglected. Moreover, since the original mechanistic studies were published, more advanced DFT functionals that account for dispersion interactions that are critical for an accurate description of nucleic acids have been introduced, which may influence the previously reported structural and energetic results. Results of the present chapter reveal the importance of method on the energy barriers.

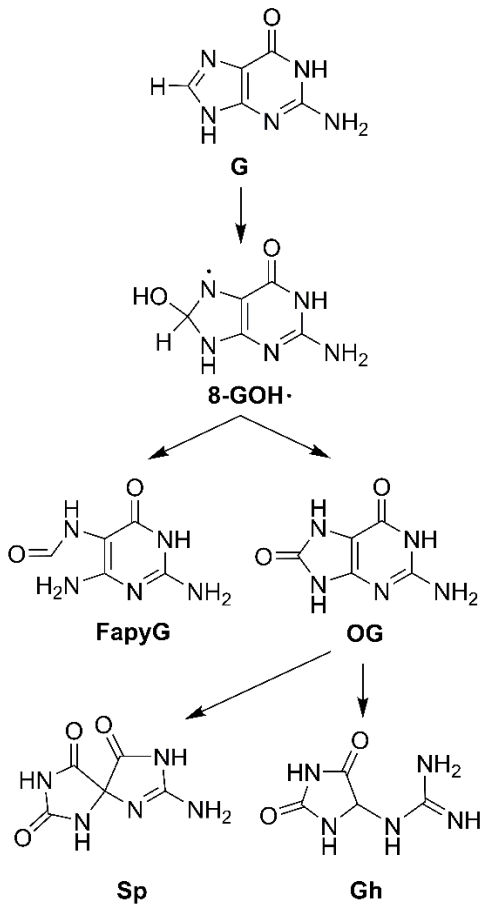

Figure 2.1. Structure of guanine (G), the 8-GOH• radical intermediate, and four abundant guanine oxidation products, namely 7,8-dihydro-8-oxoguanine or 8-oxoguanine (OG), 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG), 5-guanidinohydantoin (Gh), and 2-imino-5-5′-spirodihydantoin (Sp).

As a first step toward the systematic characterization of guanine oxidation pathways, this chapter assesses the effects of the computational approach on the structural parameters and energetics for the formation of FapyG as a representative DNA oxidative product. Four FapyG formation pathways initiated from the 8-hydroxy guanine radical (8-GOH•) have been previously characterized in the gas phase using B3LYP/6-31+G(d).[11] The most kinetically favorable mechanism (Figure 2.2) involves water-mediated hydrogen radical transfer from O8 of 8-GOH• to N7, followed by imidazole ring opening at C8–N9, and hydrogen radical addition to N9 by CH_3SH.[11] Starting from this previously

identified lowest energy pathway, the present chapter considers the effects of the environment (*i.e.*, the dielectric constant representing the surrounding environment), the level of theory (*i.e.*, the method and basis set), and the model size (*i.e.*, a nucleobase model versus a DNA helical model) on the geometrical parameters and relative energies of stationary points. Specifically, six different environments, four functionals combined with six basis sets, and six models are considered. Subsequently, a subset of representative computational approaches was chosen to further examine the effects of the computational model on the formation pathway for OG. Overall, while the geometrical parameters are not very sensitive to the level of theory, the results highlight the importance of including the effects of the reaction medium in the optimization routine, implementing advanced functionals and accounting for the surrounding DNA environment when evaluating the reaction energetics. As a result, the data presented in this chapter sets the stage for future investigations of DNA oxidation reactions that involve radical species.

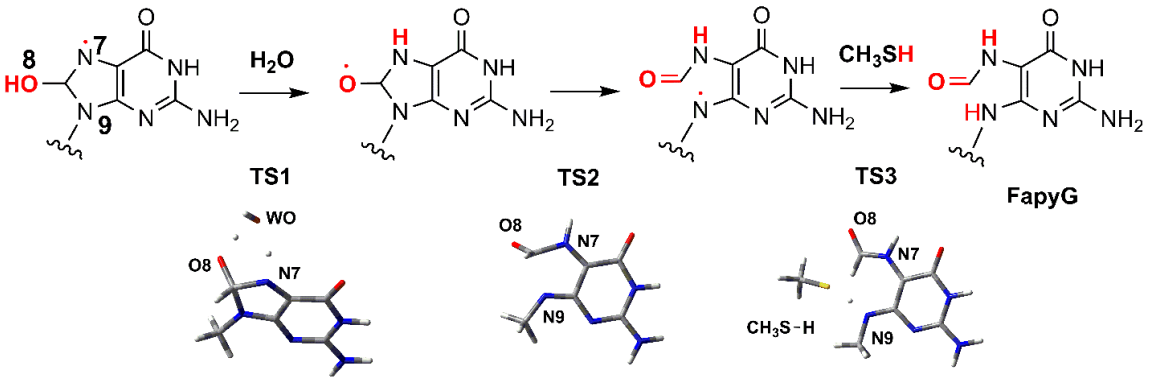

Figure 2.2. Previously proposed mechanism for the formation of FapyG from 8-GOH•[11] considered in the present work (top) and corresponding representative transition structures for the nucleobase model obtained at IEF-PCM-B3LYP/6-31G(d) level of theory.

2.2 Computational Details

Figure 2.2 outlines the reaction pathway considered in the present work for the formation of FapyG from 8-GOH•, which was previously identified to be the lowest energy pathway.[11] Since the formation of 8-GOH• from G via a hydroxyl radical has been determined to be barrierless,[13] this step is not considered herein. Initial guesses for all transition structures associated with the smallest nucleobase model were built based on previous work[11] and fully optimized. Subsequently, the corresponding intrinsic reaction coordinate (IRC) was followed in both directions. Optimizations were performed in six environments, namely the gas phase (ε = 1.0), diethyl ether (DEE, ε = 4.2), tetrahydrofuran (THF, ε = 7.4), 2-butanol (ε = 15.9), acetonitrile (ε = 35.7) and water (ε = 78.4), as described by the integral equation formalism-polarizable continuum model (IEF-PCM) method implemented in Gaussian 09 (revision D.01).[38] The structural, as well as energetic, effects of the reaction medium were investigated using the B3LYP/6-31+G(d) method previously implemented to characterize the lowest energy FapyG formation pathway.[11] Other basis sets considered were 6-31G(d), 6-31+G(d,p) and 6-311+G(2df,2p). Subsequently, B3LYP/6-311+G(2df,2p) solvent-phase single-point calculations were performed on the B3LYP gas-phase structures evaluated with 6-31G(d) and B3LYP/6-311+G(2df,2p) to assess the ability of this approach to recover the effects of the reaction medium.

Due to the potential role of dispersion interactions, the stationary points corresponding to all reaction steps for the nucleobase model were subsequently optimized with the B3LYP-D3, M06-2X and ωB97X-D functionals in conjunction with the 6-31G(d), cc-pVDZ, 6-31+G(d), 6-31+G(d,p), aug-cc-pVDZ and 6-311+G(2df,2p) basis sets in DEE (ε = 4.2). Calculations were also preformed using B3LYP in combination

with cc-pVDZ and aug-cc-pVDZ for completeness. Subsequently, B3LYP, B3LYP-D3, M06-2X and ωB97X-D single-point calculations were performed with the 6-311G(d,p), cc-pVTZ, and 6-311+G(2df,2p) basis sets on the B3LYP/6-31G(d) geometries to explore the ability to reproduce accurate reaction energetics at a reduced computational cost.

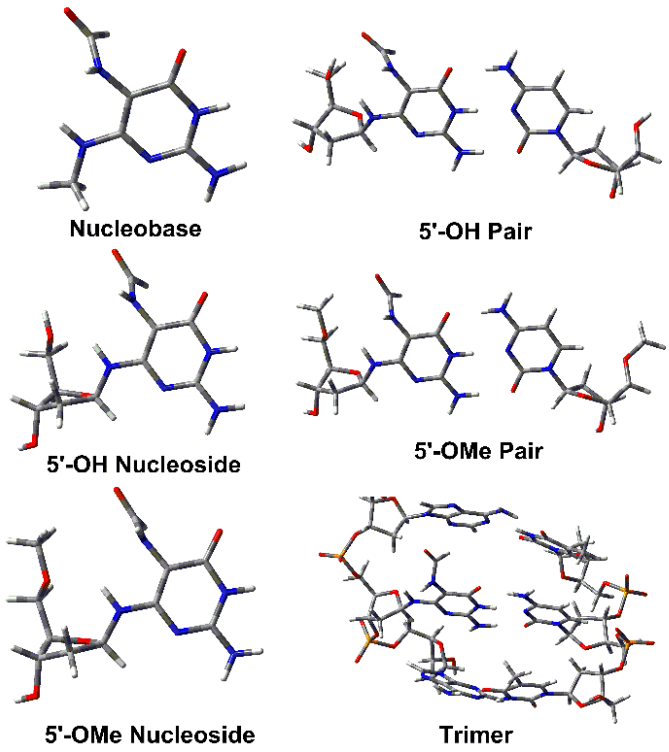

Figure 2.3. Models considered in the present work including the nucleobase, 5′-OH nucleoside, 5′-OMe nucleoside, 5′-OH nucleoside pair, 5′-OMe nucleoside pair, and trimer of nucleoside pairs. Models are provided for the final isolated product as representative examples.

Although the previous computational study of the FapyG formation mechanism considered increased nucleobase models that included $–CH_2OH$ or $–CH(CH_3)OCH_3$ at N9 to mimic 2′-deoxyribose,[11] the present chapter further explores the effects of the model size (Figure 2.3). Specifically, the smallest model considered replaces the hydrogen atom at N9 with a methyl group (denoted nucleobase model). The effects of 2′-deoxyribose were considered with –OH or –OMe linked to the 5′ carbon (denoted 5′-OH nucleoside

and 5′-OMe nucleoside, respectively). Subsequently, 8-GOH• was paired opposite C using the 5′-OH and 5′-OMe nucleoside models (designated 5′-OH pair and 5′-OMe pair, respectively). Finally, trimers of base pairs connected by the phosphate backbone were built using DNA templates implemented in HyperChem 8.0.8,[39] and capped with hydroxyl groups at the 3′- and 5′-ends. Since the base sequence can affect the DNA oxidation potential, with stacked guanines being most easily oxidized,[15-16] the trimer model consists of the damaged G stacked adenine (*i.e.*, the 5′-AGA trimer sequence). This sequence choice eliminates the complications associated with neighboring guanines. Future work should consider the effects of the lesion–site sequence context on the damage formation mechanism. Since all optimizations for the radical species were carried out using unrestricted approach, spin contaminations were closely monitored before and after annihilation. The largest spin contamination before the annihilation was ~0.86, while all the contaminations were eliminated by annihilation process. All calculations were conducted with Gaussian 09 (revision D.01).[38]

2.3 Results and Discussion

2.3.1 Formation of FapyG

2.3.1.1 Effects of the environment

To investigate the effects of the reaction medium on the structure and relative energies of the stationary points along the FapyG formation mechanism (Figure 2.2), the three reaction steps previously characterized in the gas phase with B3LYP/6-31+G(d)[11] were considered using B3LYP, the smallest nucleobase model (Figure 2.3), and four basis sets (*i.e.*, 6-31G(d), 6-31+G(d), 6-31+G(d,p), and 6-31+G(2df,2p)). Six environments were studied that cover a range of dielectric constants (ε), namely the gas phase (ε = 1.0), diethyl ether (DEE, ε = 4.2), tetrahydrofuran (THF, ε = 7.4), 2-butanol (ε = 15.9),

acetonitrile (ε = 35.7), and water (ε = 78.4). As the dielectric constant increases from 1 to 78.4, key reaction distances and angles in the transition state complexes (Figure 2.2, bottom) that correspond to the reaction coordinate change by ~ 0.14 Å (8%) and 3° (2%) at the B3LYP/6-31G(d) level, respectively (Table A2.1). Generally, the environmental effects on the bond distances are greater when larger basis sets are considered (up to 0.18 Å or ~ 13% for 6-311+G(2df,2p)), while the effects on the reaction angles are similar regardless of the basis set considered (Table A2.1). Similar trend is observed for the bond angle with largest change being 6° (4%) with 6-311+G(2df,2p)). Nevertheless, this data suggests that the B3LYP descriptions of the transition structures are influenced by the environment regardless of the basis set.

Compared to the transition structures, key geometrical parameters in the reactant complexes (representative structures for the employed models are shown in Figure A2.1) are less sensitive to the solvent for reaction steps 1 and 2 (less than ~ 3% in a distance and 4% in an angle change with environment). For step 3, the reactant geometries exhibit a larger dependence on the environment, with the largest difference being 0.167 Å (~ 7%) for N9···$HSCH_3$ distance in vacuum versus water when calculated with the largest 6-311+G(2df,2p) basis set (Table A2.2).

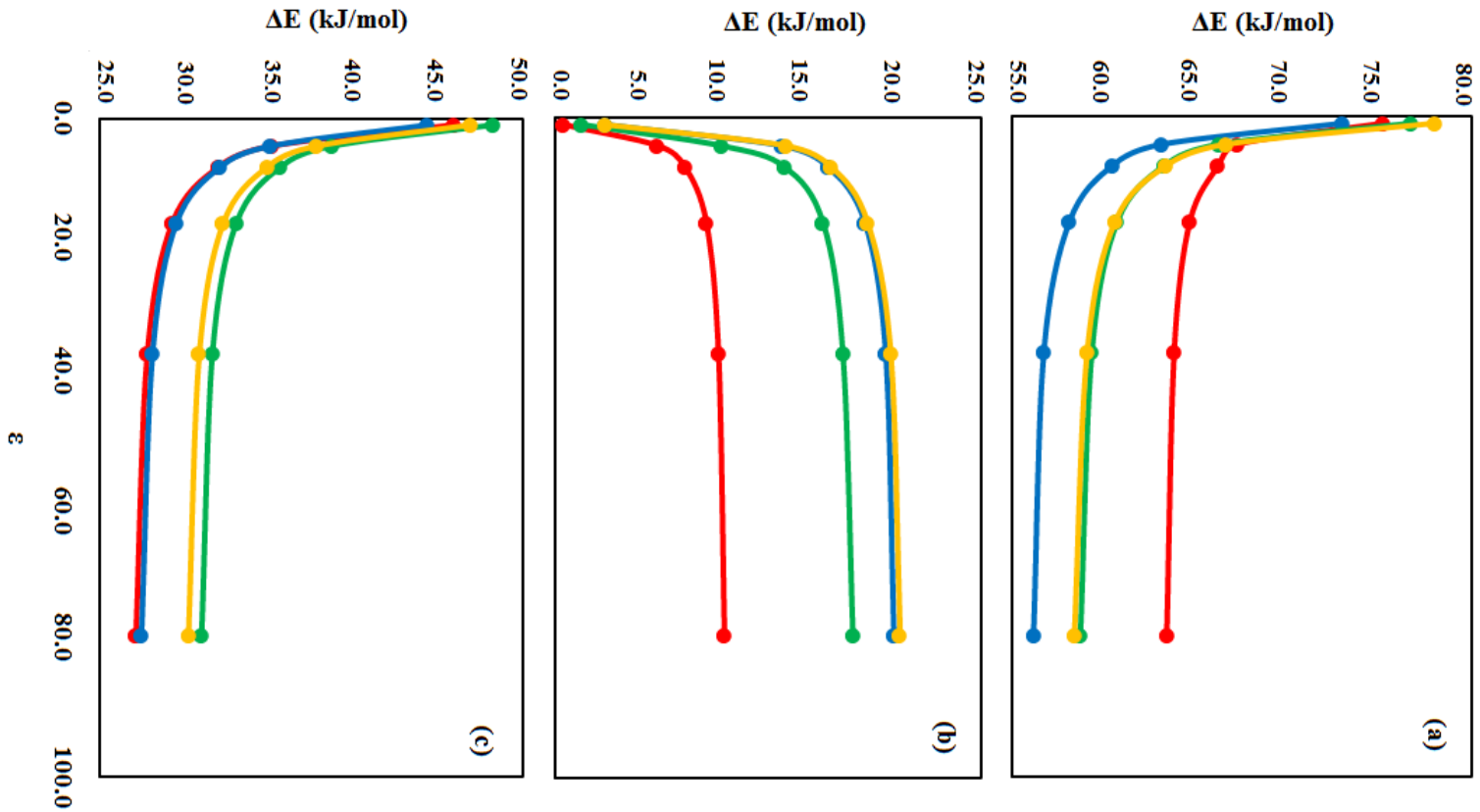

Figure 2.4. Correlation between the B3LYP energy barrier and the dielectric constant of the environment for (a) step 1, (b) step 2 and (c) step 3 along the FapyG formation pathway evaluated using 6-31(G) (red), 6-31+G(d) (yellow), 6-31+G(d,p) (blue), and 6-311+G(2df,2p) (green). Environments considered include the gas phase (ε = 1), diethyl ether (DEE, ε = 4.2), tetrahydrofuran (THF, ε = 7.4), 2-butanol (ε = 15.9), acetonitrile (ε = 35.7), and water (ε = 78.4).

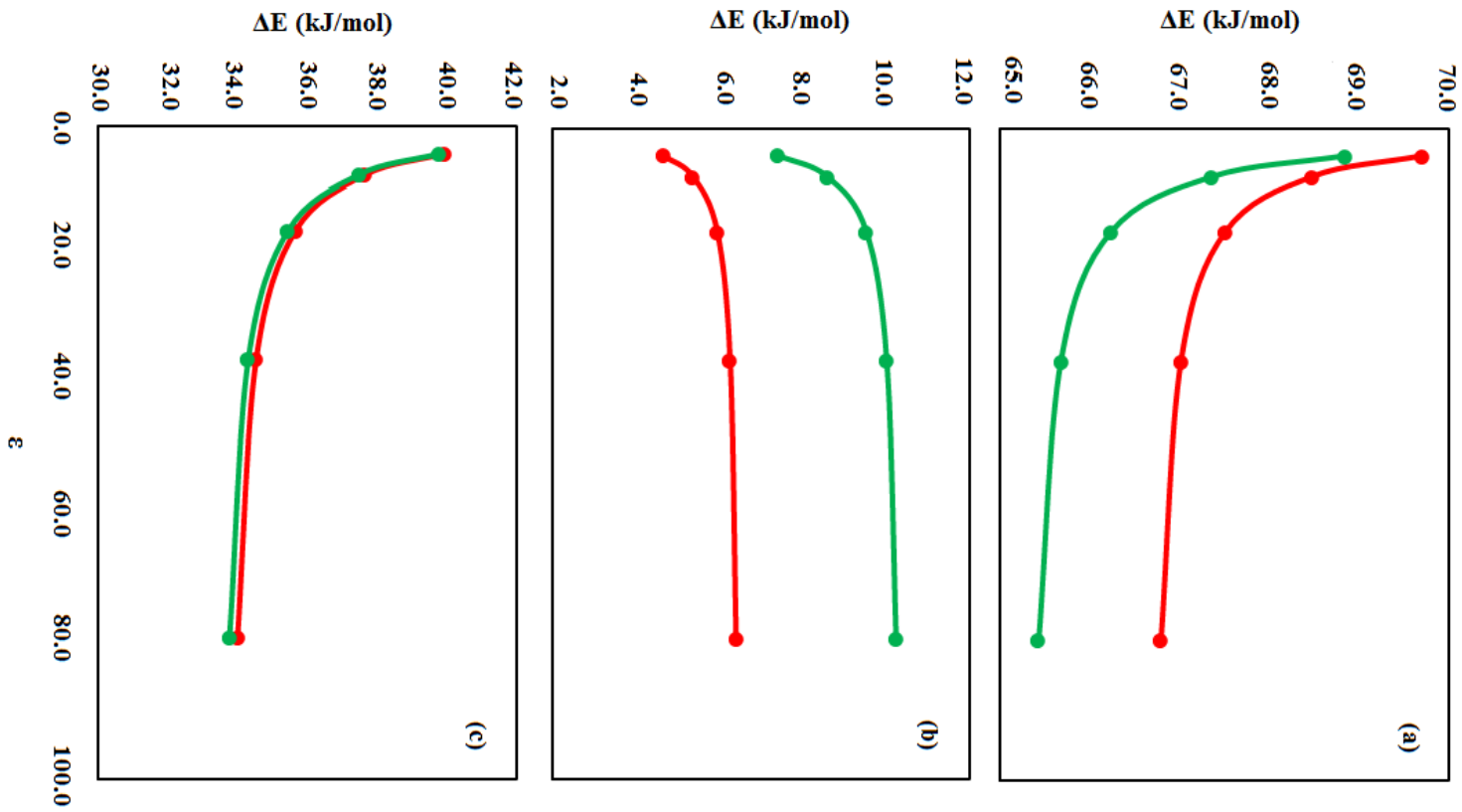

Figure 2.5. Correlation between B3LYP/6-311+G(2df,2p) single-point solvent-phase energy barrier and the dielectric constant evaluated using B3LYP/6-31G(d) gas-phase (red), B3LYP/6-311+G(2df,2p) gas-phase (blue), or solvent-phase B3LYP/6-311+G(2df,2p) geometries (green) for (a) step 1 (b) step 2, and (c) step 3 along the FapyG formation pathway. Environments considered include the gas phase (ε = 1), diethyl ether (DEE, ε = 4.2), tetrahydrofuran (THF, ε = 7.4), 2-butanol (ε = 15.9), acetonitrile (ε = 35.7), and water (ε = 78.4).

The predicted changes in key geometrical parameters along the reaction coordinate with an increase in the dielectric constant lead to a decrease in the barrier for steps 1 and 3, while the barrier for step 2 increases (Figure 2.4 and Table A2.3). Thus, the solvent effects on the energy barrier are related to the intrinsic nature of the reaction and the number of species involved. The effects are most pronounced (up to ~ 11 kJ/mol or 16% for 6-31+G(d)) when changing the environment from the gas phase to DEE (ε = 4.2), while incremental effects are seen for each subsequent increase in the dielectric constant. Indeed, the largest effect on the barrier between DEE (ε = 4.2) and THF (ε = 7.4) is only 3.7 kJ/mol (Table A2.3). Nevertheless, the effect of further increases in the dielectric constant of the reaction medium can be significant, with the difference between gas-phase and water (ε = 78.4) being up to 18 kJ/mol for 6-311+G(2df,2p). This trend holds true regardless of the basis set.

As commonly done in the literature, previous studies employed solvent-phase single-point calculations on gas-phase structures to characterize guanine oxidation pathways.[11-13] To determine whether the effect of the reaction medium can be accurately captured by this approach given the structural inferences of solvent highlighted above, this section compares the B3LYP/6-311+G(2df,2p) solvent-phase barriers calculated using the B3LYP/6-31G(d) or B3LYP/6-311+G(2df,2p) gas-phase geometries to those evaluated using the B3LYP/6-311+G(2df,2p) solvent-phase geometries (Figure 2.5 and Table A2.3 in parentheses). The trends in the reaction energies as a function of the dielectric constant predicted using solvent-phase geometries (Figure 2.5) prevail when solvent-phase single-point calculations are performed on gas-phase optimized geometries (Figure 2.5). However, in several cases, the reaction energies obtained using the gas-phase geometries deviate considerably from those obtained using the solvent-phase

geometries (Figure 2.5). Specifically, the largest change is ~ 7 kJ/mol for steps 1 and 2 or ~ 3 kJ/mol for step 3 with 6-311+G(2df,2p) in water. Although the difference is less significant for reaction steps 1 (11%) and 3 (7%), or when solvents other than water are considered, the difference for step 2 represents an ~ 40% change in the energy barrier. Therefore, depending on the reaction (*i.e.*, uni- or bi-molecular, polarity and polarizability of the reactants versus transition structure), the barriers obtained from solvent-phase single-point calculations on gas-phase geometries may not accurately reflect the reaction energetics. With advances in hardware and software over the past decade, solvent-phase optimizations are feasible. Therefore, although this approach may not have been practical for the large number of calculations and pathways examined in earlier studies,[11-13] solvent-phase calculations should be employed in future work mapping other guanine oxidative damage pathways.

Although previous computational investigations of the formation pathways of FapyG,[11] as well as other guanine oxidative products,[12-13, 32-33] considered the reaction energetics in aqueous solution, the reactions within duplex DNA is of primary interests. Since this work reveals an effect of the reaction environment on the structures of stationary points and calculated barriers, and previous works have estimated that the dielectric constant of DNA falls between ~ 2 and 8,[40-43] the remainder of the optimizations and single-point calculations for the damage formation pathways considered herein were performed in DEE (ε = 4.2). Although a range of dielectric constants has been previously proposed for the DNA duplex environment, this choice is justified due to negligible differences (less than 4 kJ/mol) for even the largest basis sets between the reactions performed in DEE (ε = 4.2) and THF (ε = 7.4). Interestingly, regardless of the basis set or environment implemented, the dependence of the energy

barriers on the dielectric constant of the reaction medium can be described with a logarithmic or power function (R^2 ~ 0.9, Table A2.4), which permits estimations of barriers in any environment with reasonable precision.

2.3.1.2 Effects of the level of theory

Previous computational investigations of the formation pathway of guanine oxidative products have consistently used the B3LYP functional combined with a range of basis sets for gas-phase optimizations.[11-13] The effects of the level of theory on the most favorable FapyG reaction pathway are investigated in the present work using the smallest nucleobase model (Figure 2.2). Specifically, four DFT functionals (*i.e.*, B3LYP, B3LYP-D3, M06-2X, and ωB97X-D) were combined with six basis sets (*i.e.*, 6-31G(d), cc-pVDZ, 6-31+G(d), 6-31+G(d,p), aug-cc-pVDZ, and 6-311+G(2df,2p)). Detailed analysis of the key geometrical parameters in the transition structures reveals that the overall trends in the changes in the distances and angles with increasing basis set size are consistent among the DFT functionals (Table A2.5). Specifically, as the basis set increases from 6-31G(d) to 6-311+G(2df,2p), the largest changes in key distances and angles for the transition state associated with step 1 are ~ 0.09 Å and 1°, respectively, regardless of the DFT functional. Furthermore, the geometrical changes with basis set size for the transition structures associated with steps 2 and 3 are typically even smaller (0.04 Å at step 2, and 0.01 Å and 2° at step 3, Table A2.5). On the other hand, the effects of basis set on the reactants are more significant. Specifically, the largest change in a distance in reactants with basis set (0.115 Å at step 1) is about 2 times greater than for the transition structures (~ 6%, Table A2.6). Furthermore, changes in the key angles with basis set along the reaction coordinate are larger for the reactants (with the largest being

9°, Table A2.6). Nevertheless, overall the basis set negligibly affects key reaction parameters for FapyG formation.

The key geometrical parameters of the transition structures predicted by the various DFT functionals are overall similar for a particular basis set. For example, for the largest 6-311+G(2df,2p) basis set, the deviations in the distances and angles calculated with B3LYP, B3LYP-D3, and ωB97X-D are smaller than 0.05 Å and 2° (Table A2.6). With the same basis set, the M06-2X distances deviate most significantly from those evaluated with the other functionals, with the largest change in a distance being 0.1 Å (step 1), nevertheless, the biggest alteration in an M06-2X angle is only 1° relative to other functionals (step 1, Table A2.6). These differences are small, representing a ~ 7 or 1% deviation in the distance or angle, respectively. Similarly, for the reactant complexes, the geometries predicted with different functionals are comparable for a given basis set. For the largest basis set, the greatest difference with a functional is 0.095 Å (5%) and 4.7° (3%) in a distance and an angle for steps 1 and 2, respectively (Table A2.6). For step 3, the deviations are more significant, with the largest changes being 0.302 Å (14%) and 13.3° (8%).

Overall, with only a few exceptions, all DFT functionals considered herein predict comparable geometrical parameters. Thus, the effects of both the functional and basis set on key reaction parameters for FapyG formation are negligible. In fact, for the largest 6-311+G(2df,2p) basis set, the effect of the functional on the structure (up to 7% deviation) is smaller than the effect of changing the environment from the gas phase to water (up to 13% deviation). This suggests that the least expensive B3LYP/6-31G(d) combination can be used to obtain structures of comparable accuracy to B3LYP-D3, M06-2X or ωB97X-D combined with larger basis sets.

In terms of the reaction energetics, for each DFT functional the largest changes in the energy barrier between the smallest 6-31G(d) and largest 6-311+G(2df,2p) basis sets for steps 1 and 3 are 2.0 kJ/mol (~ 3%) and 3.6 kJ/mol (~ 11%), respectively. However, the barrier for step 2 increases with the basis set by up to 7.3 kJ/mol or 35% (Table 2.1). The only exception is the barrier for step 1 predicted by ωB97X-D, which varies by 8.1 kJ/mol or 10% upon increasing the basis set size from 6-31G(d) to 6-311+G(2df,2p). Nevertheless, except for the smallest 6-31G(d), cc-pVDZ, and 6-31+G(d) basis sets, the difference between any two basis sets for a given functional is less than 4 kJ/mol (Table 2.1). Specifically, the largest deviation between two basis sets is 3.1, 2.9, 3.5, and 2.1 kJ/mol with B3LYP, B3LYP-D3, M06-2X, and ωB97X-D, respectively. Including the 6-31G(d), cc-pVDZ, and 6-31+G(d) basis sets in the comparison leads to the largest difference of 4.7, 10.5, 7.3, and 8.9 kJ/mol between two basis sets for the same functionals, respectively. Similarly, the largest difference between two basis sets in the energy barrier for step 2 for a given functional is less than 4 kJ/mol for the three largest basis sets (*i.e.*, 6-31+G(d,p), aug-cc-pVDZ, and 6-311+G(2df,2p)), while including the smaller basis sets leads to ~ 8 – 11 kJ/mol variations. Interestingly, at step 3, the largest variation for each functional occurs between the aug-cc-pVDZ and 6-311+G(2df,2p) basis sets (~ 7 – 8 kJ/mol, Table 2.1). Overall, with few exceptions, the barriers predicted for a given functional with the 6-31+G(d,p), aug-cc-pVDZ, and 6-311+G(2df,2p) basis sets are comparable while the agreement between 6-31+G(d,p) and 6-311+G(2df,2p) being better in most cases than that with aug-cc-pVDZ.

Due to the independence of the geometrical parameters on the basis set size, the energetics of the reactions were further evaluated using single-point calculations on the most computationally efficient B3LYP/6-31G(d) structures with different basis sets,

namely 6-31+G(d,p), 6-311G(d,p), cc-pVTZ, and 6-311+G(2df,2p) (Table 2.1, in parentheses). The results were compared to the 6-311+G(2df,2p) energies evaluated using geometries obtained with the largest basis set (the control). For a given functional, the single-point energies obtained with 6-31+G(d,p), cc-pVTZ and 6-311+G(2df,2p) significantly resemble the control values (on average within ~ 4.3 kJ/mol). Although deviations between the control and the 6-311G(d,p) single-point energies can be small, especially for steps 1 and 3, the differences can be larger for step 2 (up to ~ 8 kJ/mol with ωB97X-D). Overall, compared to 6-31+G(d,p) and 6-311G(d,p), the cc-pVTZ basis set best reproduces the control barriers, with the biggest deviation being 2.7 kJ/mol for the ωB97X-D functional at step 2 (Table 2.1).

For a given reaction step, the agreement in the barriers obtained with different functionals varies with the basis set regardless of the geometry used (Table 2.1). For steps 1 and 2, the B3LYP and B3LYP-D3 barriers are similar for a given basis set, while the corresponding M06-2X and ωB97X-D barriers are comparable to each other and larger than those obtained with B3LYP or B3LYP-D3. For step 1, the largest difference (~ 15 kJ/mol) occurs between B3LYP-D3 and M06-2X for 6-31+G(d). For step 2, the M06-2X and ωB97X-D barriers are ~ 2 times larger than the B3LYP and B3LYP-D3 energies evaluated using the same basis set, with the biggest difference (~ 11 kJ/mol) occurring between B3LYP and ωB97X-D for 6-311+G(2df,2p). For step 3, the B3LYP and ωB97X-D energies are similar, while the B3LYP-D3 and M06-2X barriers exhibit negligible differences. The biggest deviation for step 3 (~ 8 kJ/mol) occurs between M06-2X and ωB97X-D with 6-31+G(d).

Table 2.1. IEF-PCM (ε = 4.2) energy barriers (ΔE, kJ/mol) for the formation of FapyG calculated using the nucleobase model, and different functional and basis set combinations.[a]

	6-31G(d)	cc-pVDZ	6-31+G(d)	6-31+G(d,p)[b]	6-311G(d,p)[b]	aug-cc-pVDZ	cc-pVTZ[b]	6-311+G(2df,2p)[b]
					Step 1			
B3LYP	67.3	62.6	66.6	63.2 (62.6)	(69.0)	64.1	(66.7)	66.3 (65.9)
B3LYP-D3	67.7	65.0	57.2	62.8 (62.7)	(69.1)	63.7	(66.7)	65.7 (66.0)
M06-2X	74.4	68.0	75.3	69.6 (67.9)	(74.9)	71.7	(71.4)	73.1 (71.3)
ωB97X-D	77.6	70.8	72.6	68.7 (70.6)	(76.4)	70.8	(73.7)	69.5 (73.1)
					Step 2			
B3LYP	6.1	5.2	13.6	13.4 (11.6)	(5.3)	13.3	(7.5)	9.8 (10.0)
B3LYP-D3	6.6	5.8	13.9	13.8 (12.2)	(5.9)	13.7	(8.1)	11.5 (10.6)
M06-2X	14.1	12.7	21.8	21.9 (18.6)	(12.9)	21.9	(16.0)	20.1 (17.0)
ωB97X-D	13.6	12.6	23.2	23.0 (18.9)	(12.5)	23.0	(14.8)	20.9 (17.5)
					Step 3			
B3LYP	35.2	32.3	37.9	35.2 (34.7)	(34.1)	30.9	(36.9)	38.8 (38.1)
B3LYP-D3	29.6	26.5	31.0	28.6 (27.5)	(26.9)	24.5	(29.7)	31.5 (30.9)
M06-2X	28.9	27.8	30.9	28.8 (29.2)	(28.9)	24.8	(33.7)	32.5 (33.4)
ωB97X-D	36.4	33.7	38.7	35.9 (36.0)	(35.2)	30.6	(37.5)	39.0 (39.7)

[a] Representative transition structures for the 3 reaction steps are shown in Figure 2.2 (bottom). [b] Values in parentheses were obtained from single-point calculations on B3LYP/6-31G(d) geometries.

Overall, the above discussion supports investigations of the effects of increasing the size of the computational model on the FapyG formation pathway using single-point calculations on B3LYP/6-31G(d) geometries. Furthermore, when calculations with the largest 6-311+G(2df,2p) basis set are not practical due to model size, the cc-pVTZ basis set can be used in single-point calculations to obtain accurate energies. Since the barriers are dependent on the functional choice, the reaction energetics for larger models will be investigated using both B3LYP-D3 and M06-2X to differences in their performance for several reaction steps.

2.3.1.3 *Effects of the model size*

The effect of the computational model was examined by systematically increasing the model size from the nucleobase to the 5′-OH or 5′-OMe nucleoside to the 5′-OH or 5′-OMe pair and finally to the trimer as described in Computational Details (Figure 2.3). Based on the dependence of the reactant and transition state geometries on the environment and level of theory outlined in the previous sections, all structures were optimized with B3LYP/6-31G(d) in DEE ($\varepsilon = 4.2$). For the 5′-OH nucleoside and pair models, the 5′-OH interacts with O8 of the nucleobase along the reaction pathway (Figure 2.6), which likely influences the transition state geometry and reaction barrier. In fact, the largest deviations in the key distances or angles along the reaction pathway evaluated with the trimer model occurs for the 5′-OH nucleoside model (Table A2.7). Therefore, the 5′-OH nucleoside and pair models are not further discussed.

Table 2.2. IEF-PCM (ε = 4.2) energy barriers (ΔE, kJ/mol) for the formation of FapyG calculated using B3LYP/6-31G(d) geometries for various models.[a, b]

	Step 1		Step 2		Step 3	
Model	**B3LYP-D3**	**M06-2X**	**B3LYP-D3**	**M06-2X**	**B3LYP-D3**	**M06-2X**
Nucleobase	66.7 (66.0)	71.4 (71.3)	8.1 (10.5)	16.0 (17.2)	29.7 (27.5)	33.7 (33.4)
5'-OH Nucleoside	64.0 (65.1)	65.9 (68.4)	11.1 (12.9)	23.0 (23.4)	37.5 (37.1)	42.1 (42.7)
5'-OMe Nucleoside	67.6 (71.3)	75.9 (74.4)	0.6 (1.6)	1.5 (1.6)	31.7 (32.7)	36.6 (37.0)
5'-OH Pair	61.1 (61.8)	63.4 (73.8)	17.6 (19.7)	31.8 (32.5)	31.3 (33.3)	32.9 (33.7)
5'-OMe Pair	67.6 (67.3)	73.6 (75.6)	0.8 (1.3)	3.4 (3.5)	27.7 (29.4)	30.6 (28.3)
Trimer	46.8 –	55.8 –	8.1 (9.7)	16.2 –	26.3 –	25.9 –

[a] Representative transition structures for the 3 reaction steps are shown in Figures 2.6 and 2.7. [b] Barriers obtained using single-point calculations with cc-pVTZ and 6-311+G(2df,2p) in parentheses.

For step 1, increasing the model size from the nucleobase to the 5′-OMe nucleoside has negligible effects on key reaction distances in the TS (the largest change ~ 0.01 Å) and angles (0.1°, Table A2.7). Addition of the pairing base to the 5′-OMe nucleoside results in a 0.046 Å and 1.3° change in a distance or an angle, respectively. Further increase of the model size to a trimer results in a greater alteration in the distances (up to 0.110 Å), while the O···H···N7 angle changes by 1.2° (Table A2.7). The largest change in a distance or an angle at step 1 occurs between the 5′-OMe nucleoside and trimer models (0.145 Å or 10%, 1° or 2%, respectively, Table A2.7). The trimer model leads to the earliest transition state, with the largest O8H···OW and WO···HN7 distances, and smallest O8···HOW and WOH···N7 distances. At this step, the 5′-OMe nucleoside and pair models show the largest and smallest deviations in the geometrical parameters compared to the trimer model, respectively.

At step 2, addition of the sugar backbone decreases the C8···N9 distance by 0.134 Å, while the effect of the pairing base to the 5′-OMe nucleoside is negligible (0.006 Å, Table A2.7). Interestingly, for step 2, the nucleobase and trimer models predict similar geometries (with the largest difference being ~ 0.017 Å or < 1%) as increasing the model from the pair to the trimer leads to an increase in the reaction distance. The largest deviation occurs between the nucleobase and pair models (0.134 Å or 7%, Table A2.7). Thus, the effects of addition of the 2′-deoxyribose and the pairing base are compensated by the flanking bases.

Reaction distances in the transition state associated with step 3 only slightly change with the model from the nucleobase to the nucleoside (0.013 Å) or the nucleoside to the pair (0.016 Å), while the angles change by less than 1° (Table A2.7). The biggest change in a distance or an angle with further increase of the model size to the trimer is

0.020 Å and 3.7°, respectively. Overall, the geometrical parameter changes from the nucleoside to the pair and further to the trimer model are consistent. However, geometries predicted with the nucleobase model are closer to the trimer model compared to the two other models. The transition state with the trimer model is later, with the largest N9···HS and smallest S···HN9 distances among all models (Table A2.7). In comparison with other reaction steps, the reaction distances at step 3 are the least sensitive to the model size, with the greatest deviations being 0.036 Å and 4.5° (~ 3% between 5′-OMe nucleoside and trimer models, Table A2.7).

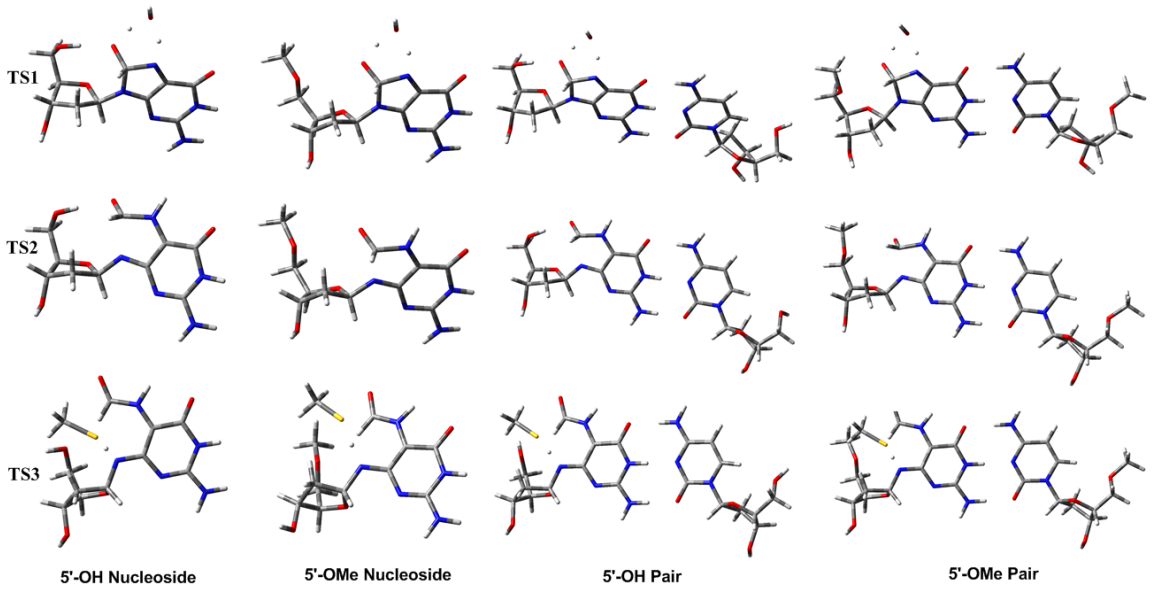

Figure 2.6. Transition state structures along the FapyG formation pathway for the 5′-OH and 5′-OMe nucleoside, and 5′-OH and 5′-OMe nucleoside pair models.

In comparison with the transition structures, the reactant geometries generally exhibit smaller changes with the model size (Table A2.8). Specifically, for steps 1 and 2, the largest change with the model size in a distance in the reactant complex (0.072 and 0.061 Å, Table A2.8) is about half that observed for the corresponding transition state structure (Table A2.7). However, at step 1, the angles in the reactants are more

significantly altered by the model size (~ 12° or 8%, Table A2.8). Moreover, at step 3, the model has a more significant impact on the N9···$HSCH_3$ distance in the reactant than the transition structure, with the largest change being 0.151 Å (~ 7%, Table A2.8).

In order to further assess the models, several structural parameters, including the χ dihedral angle (∠(O4′–C1′–N9–C4)) for models larger than nucleobase, and the pairing parameters (Figure A2.2) for models that include the opposing nucleobase were evaluated. Comparison of the transition structures for the models with the sugar backbone reveals that the average χ dihedral angle is ~ –100, –100, and –78° in TS1, TS2, and TS3, respectively (Table A2.9), and is comparable with the χ values for G in a G:C pair in DNA duplex (~ –95). In the reactant complexes, the χ values are generally similar to or larger than those for the corresponding transition states (Table A2.9). The pairing parameters (including C1′···C1′ distance, and the distances and angles between the atoms involved in the hydrogen bonds) along the reaction pathway were compared for the 5′-OH pair and trimer models. In general, the pairing parameters were maintained for all models and are comparable with experimental data (C1′···C1′ ~ 10.4 – 10.6 Å,[44, 45] Table A2.10). The differences of the pairing distances and angles for the reactant complexes and corresponding transition structures are negligible (Table A2.10). Overall, changes in geometrical parameters with the model size (~ 10 or 3% in a distance or angle, Tables A2.7 and A2.8) are comparable to those with the solvent (~ 13 or 2% in a distance and an angle, respectively, Tables A2.1 and A2.2), and are larger than those with the level of theory (functional/basis set) as discussed for the nucleobase model (Tables A2.5 and A2.6).

It is timely to emphasize that for the trimer model, the flanking bases (*i.e.*, 3′- and 5′- A and C) show distortions due to the introducing the OH• and CH_3SH moieties,

structural changes in the central G, and the lack of stacking interactions. As is seen in Figure 2.7, these distortions are less significant for step 2 where there is no other species is involved in the reaction.

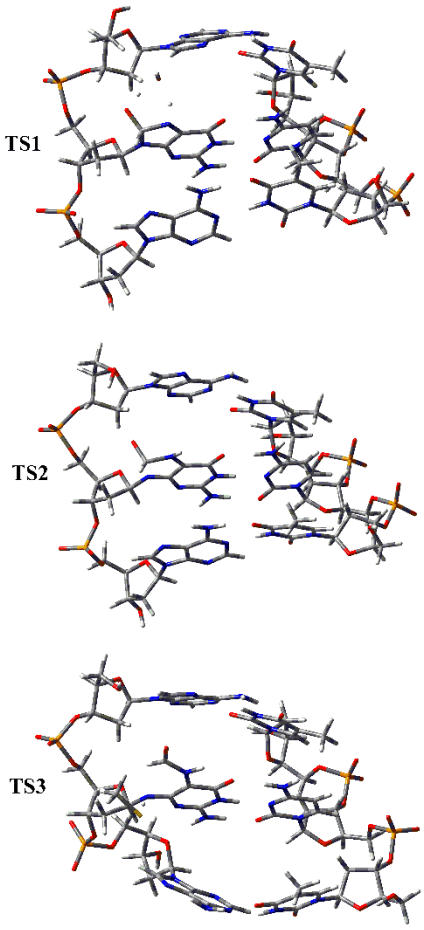

Figure 2.7. Transition structures along the FapyG formation pathway for the trimer model.

M06-2X and B3LYP-D3 single-point calculations with cc-pVTZ and 6-311+G(2df,2p) were carried out on the B3LYP/6-31G(d) optimized geometries for all reaction steps and models (Table 2.2). However, calculations with 6-311+G(2df,2p) were not practical for all steps for the trimer model. Thus, the following section compares the performance of the models based on the cc-pVTZ energies, while the discussion holds true for the 6-311+G(2df,2p) basis set.

The energy barrier of step 1 is not significantly affected as the model is increased from the nucleobase (66.7/71.4 kJ/mol) to the 5′-OMe nucleoside (67.6/75.9 kJ/mol), and further to the 5′-OMe pair (67.6/73.6 kJ/mol) (B3LYP-D3/M06-2X, Table 2.2). Thus, the largest change in energy with the model is only ~ 1 and 6% for B3LYP-D3 and M06-2X, respectively. In contrast, the barrier for the trimer model is 46.8 or 55.8 kJ/mol (~ 21 or 18 kJ/mol or ~ 44 and 32%) smaller than that of the 5′-OMe pair with B3LYP-D3 and M06-2X, respectively. At step 1, addition of 2′-deoxyribose or the pairing base to the nucleobase does not affect the barrier, while stabilization of the transition state with addition of the flanking bases leads to an earlier transition state with a smaller reaction barrier (Table 2.2). The similar distances for the reactant and transition structures at step 2 for the nucleobase and trimer models lead to comparable barriers (~ 8 kJ/mol with B3LYP-D3 and ~ 16 kJ/mol with M06-2X), while both the 5′-OMe nucleoside and pair models have significantly smaller barriers due to the early transition state (~ 1 – 3 kJ/mol, Table 2.2). Thus, the largest change in barrier with model for step 2 is ~ 7 or 14 kJ/mol with B3LYP-D3 or M06-2X. The relatively small barrier of this step leads to up to 93 and 91% (for B3LYP-D3 and M06-2X, respectively) model size effect on the energy barrier. At step 3, increasing the model to the nucleoside increases the barrier by ~ 2 or 3 kJ/mol (with B3LYP-D3 and M06-2X, Table 2.2), while addition of the pairing nucleoside leads to a barrier decrease (by 4 or 6 kJ/mol). Further increase of the model to the trimer reduces the barrier by an additional ~ 1 or 5 kJ/mol with both functionals. In agreement with the greatest variation in the geometries, the largest difference in the energy barriers occurs between the 5′-OMe nucleoside and trimer models, being ~ 5 and 11 kJ/mol (*i.e.,* ~ 17 and 29%) with B3LYP-D3 and M06-2X, respectively (Table 2.2).

Overall, while the 5′-OH nucleoside is commonly employed to study a variety of properties,[46-48] it might not be the best model for mechanistic investigations where the nucleobase and/or other species might directly interact with 5′-OH. Furthermore, changes in the geometries of stationary points and the reaction energetics upon further increasing the model size can be considerable depending on the intrinsic nature of the reaction. To further examine the effects of model size, the two reaction steps of the previously identified, preferred OG formation pathway are considered in the subsequent section using the same models and methods as discussed for FapyG.

2.3.2 Formation of OG

2.3.2.1 Effects of model size

To further evaluate the effect of model size on the geometry and energy barriers for the formation of guanine oxidative products, the formation of OG from the 8-GOH• radical is investigated. According to a previous study,[13] the lowest energy pathway for the formation of OG (Figure 2.8) involves the abstraction of a H• radical from O8 by an OH• radical, which results in a radical intermediate that contains two singly occupied orbitals at O8 and N7, and a water molecule. In the second step, the water molecule simultaneously abstracts a H• radical from C8 and delivers a H• radical to N7, which yields OG. In accordance with the previous section, stationary points were optimized with B3LYP/6-31G(d) in DEE, while the single-point calculations were completed using B3LYP-D3 and M06-2X combined with cc-pVTZ.

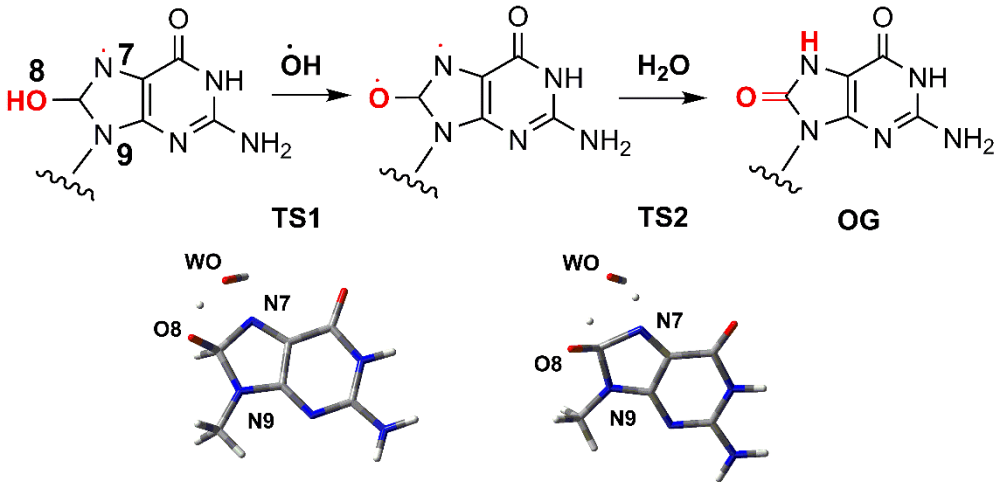

Figure 2.8. Previously proposed mechanism for the formation of OG from 8-GOH•[13] considered in the present work (top) and corresponding representative transition structures for the nucleobase model.

Since there are two singly occupied orbitals in both reaction steps, the system could be a singlet or a triplet. Although it might be assumed that the singlet is the dominant spin configuration, the second reaction proceeds without the aid of water when the singlet was considered for the nucleobase and 5′-OMe nucleoside models (*i.e.,* direct H• radical shift from C8 to the neighboring N7, Figure A2.3 and Table A2.11). Although this pathway is similar to the reaction reported in previous work,[13] a reaction with a direct H• radical shift between two neighboring atoms is less likely. Thus, this section focuses only on the reaction pathway involving a triplet intermediate that was not considered in the previous work.[13]

Transition structures associated with the two reaction steps for various models are shown in Figures 2.9 and 2.10 (geometries of reactants and transition structures, and energetics are provided in Tables A2.12 – A2.15 and 2.3). As shown for the FapyG formation pathway, the 5′-OH interacts with the reaction core during the formation of OG in the 5′-OH nucleoside and pair models (Figure 2.9), and therefore these models are not further discussed. At step 1, the largest change upon addition of the sugar backbone to the

nucleobase model in a distance (0.012 Å increase) and an angle (4.4° increase, Table A2.12). The largest change in a distance or an angle with addition of the pairing base to the 5′-OMe nucleoside is 0.018 Å and 7.2°, respectively in the opposite direction (decrease). Finally, further increase of the model to the trimer increases a distance by 0.027 Å and an angle by 12.4°. A comparison among all models at step 1 indicates overall small deviations in the transition distances, with the largest being 0.037 Å (~ 3%) between the 5′-OMe nucleoside and trimer models (Table A2.12). The key angle in the transition structure shows a large absolute variation with model size, being up to 12.4° or ~ 8%, which occurs between the 5′-OMe pair and trimer models. At this step, the effects of the model size on the geometries do not follow a clear trend and the opposing base (pair model) and the sugar backbone (5′-OMe model) have opposite effects on the geometrical parameters.

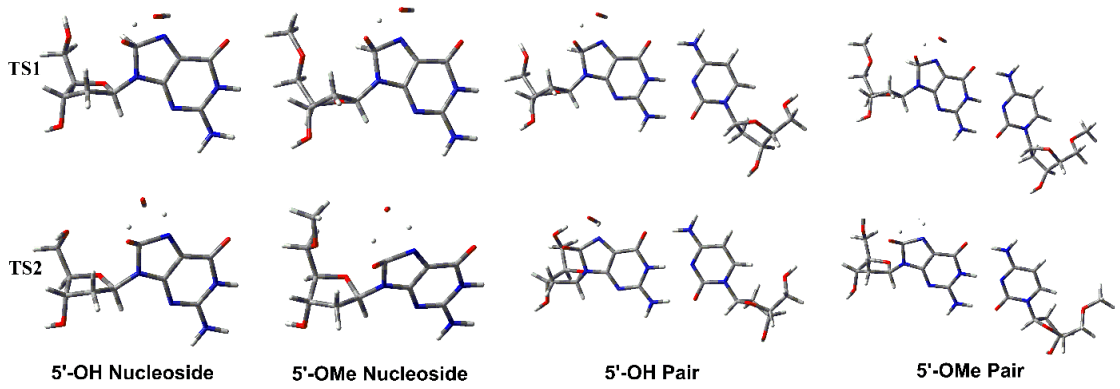

Figure 2.9. Transition structures along the OG formation pathway for the 5′-OH and 5′-OMe nucleoside, and 5′-OH and 5′-OMe nucleoside pair, models.

At step 2, increasing the model size almost consistently moves the geometries toward the values obtained with the trimer model (Table A2.12). The largest differences are seen between the nucleobase and trimer models. Most importantly, changes with the model size from the nucleobase to nucleoside, and to the pair model are smaller than

changes with further increase in the model from the pair to the trimer. Overall, the biggest deviation in a distance is 0.327 Å (~ 17%), which is dramatically larger compared to step 1. The largest change in an angle at this step is 9.2° between the nucleoside and trimer models (~ 8%, Table A2.12). At step 2, the flanking base has a considerable effect on the transition structure and leads to an earlier transition state.

In comparison with the transition structure, the reactant complexes (representative structures of the reactant complexes for the employed models are shown in Figure A2.4) show considerable sensitivity to the model size (Table A2.13). Specifically, the O8H···OW distance changes by up to 0.171 Å (~ 9%, Table A2.13) as the model increases from the nucleobase to the nucleoside. At the second step, the water molecule is located far from C8H in the pair model, resulting in an ~ 2 Å (76%) difference relative to the nucleobase model (Table A2.13). This also leads to a large change in the C8···H···OW angle for the pair model (47° or 72%) compared to the nucleobase. Thus, the triplet nature of the system impacts the intrinsic reaction coordinates and hence the reactants more profoundly with model size than the transition structures, where both single electrons are involved with strong interactions with other electrons in the reaction.

Table 2.3. Relative energy barriers (ΔE, kJ/mol) for the formation of OG calculated using IEF-PCM-B3LYP/6-31G(d) geometries for various models.[a, b]

	Step 1		Step 2	
Model	**B3LYP-D3**	**M06-2X**	**B3LYP-D3**	**M06-2X**
Nucleobase	13.5	53.5	78.0	107.0
5'-OMe Nucleoside	16.8	54.0	87.6	117.4
5'-OMe Pair	30.1	55.2	91.9	123.4
Trimer	24.1	71.0	71.9	96.8

[a] Representative transition structures for the 2 reaction steps are shown in Figures 2.9 and 2.10. [b] Barriers obtained using single-point calculations with cc-pVTZ.

The χ dihedral angle of the modified G throughout the reaction pathway is maintained in the reactant and transition structures among the models (on average ~ –90

and –87° at TS1 and TS2, respectively, Table A2.14). In fact, for OG formation, the χ is more consistent between the reactant and transition structures than those for the formation of FapyG (Table A2.6). The pairing parameters for the 5′-OMe pair and trimer models were also maintained along the reaction pathway and the differences between reactants and transition states are negligible (Table A2.15). Specifically, the C1′···C1′ distance at both steps for the pair and trimer models retain ~ 10.6 Å, and are comparable with experimental data and those during formation of FapyG (Table A2.10).

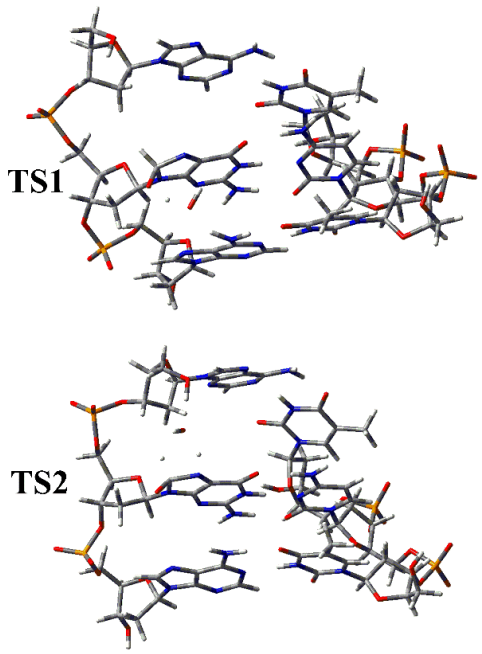

Figure 2.10. Transition structures along the OG formation pathway for the trimer model.

Similar to FapyG, the reaction barriers along the OG formation pathway depend on the model size (Table 2.3). Specifically, modeling the first step with B3LYP-D3, leads to larger barrier upon increasing the model from the nucleobase (13.5 kJ/mol) to the 5′-OMe nucleoside (16.8 kJ/mol) and further to the pair model (30.1 kJ/mol); however, the barrier for the trimer model (24.1 kJ/mol) is slightly smaller than for the pair model. Thus, the largest difference is ~ 55% between the nucleobase and 5′-OMe pair models

(Table 2.3). In comparison to the B3LYP-D3 barriers, the M06-2X barriers are considerably higher regardless of the model (by up to 47 kJ/mol). Moreover, the M06-2X barriers are significantly similar for the nucleobase, nucleoside and pair models (53.5 – 55.2 kJ/mol, Table 2.3), while the barrier for the trimer model is ~ 18 kJ/mol larger than the other models. At step 2, increasing the model from the nucleobase to the nucleoside (by ~ 10 kJ/mol) and further to the pair model (by ~ 5 kJ/mol) constantly increases the barrier regardless of the DFT functional. However, further increase of the model to the trimer results in the lowest barriers among all models for both B3LYP-D3 and M06-2X (Table 2.3). Similar to reaction step 1 during formation of FapyG, the flanking bases in the trimer model stabilize the transition state of step 2 along the OG formation. Despite deviations between the functionals, the barriers with the nucleobase and trimer models are closer together for a given functional, although the differences are not negligible (6.1 or 10.2 kJ/mol with B3LYP-D3 and M06-2X, respectively). The largest difference in the barrier for two models is 20.0 or 26.6 kJ/mol (~ 28% for the pair and trimer models) with B3LYP-D3 and M06-2X, accordingly (Table 2.3).

Comparing the effect of the model size on the geometry along the FapyG and OG formation pathways reveals that the geometries along the formation of OG are more sensitive to model size, with the largest differences in a distance or an angle between two models for OG are ~ 2 and ~ 3 times larger than those for FapyG, respectively. Overall, the OG formation pathway further underscores the significance of the model size on the geometries and energy barriers. Although the exact effect of a particular model size cannot be simply extrapolated, and is largely dependent on the reaction type and number of species involved in the reaction, the results presented in this Chapter highlight the effects of the DNA environment, particularly affecting the key geometrical parameters of

transition states by interacting with the reaction core. Since both the opposing and flanking bases influence the reaction, including these key elements surrounding the reaction core is necessary for the formation of guanine oxidative damage to obtain more accurate geometries and energetics.

2.4 Conclusion

The results discussed in the present chapter indicate that the geometrical parameters of the transition structures for the FapyG formation pathway obtained with the nucleobase model, and hence the energy barriers can be considerably influenced by the solvent, especially for dielectric constants close to 1. Furthermore, increasing the dielectric constant can increase or decrease the energy barriers depending on the reaction type. Nevertheless, for each step in the FapyG formation pathway, there is a non-linear correlation between the reaction energy barrier and the dielectric constant, which can be described with a logarithm or a power formula with a great accuracy ($R^2 \sim 0.9$).

The energy barriers obtained with solvent-phase single-point calculations on gas-phase geometries might not equal the barriers calculated with solvent-phase geometries. In comparison with the effect of the dielectric constant (ε = 1 to 78), the effects of the basis set size (6-31G(d) to 6-311+G(2df,2p)) for a given DFT functional are insignificant. Although the energy barriers are considerably altered by the functional/basis set combination, single-point calculations with the four functionals considered in this Chapter (*i.e.*, B3LYP, B3LYP-D3, M06-2X, and ωB97X-D) with the largest 6-311+G(2df,2p) basis set on B3LYP/6-31G(d) geometries accurately reproduce the barriers predicted using the geometries obtained with the corresponding functional and 6-311+G(2df,2p).

This chapter revealed that model size can considerably affect the geometries of the transition states associated with both FapyG and OG formation. However, the magnitude of the effect depends on the reaction type. Specifically, including the 2′-deoxyribose, the opposing nucleoside and the flanking bases can lead to earlier or later transition states. For example, the effect of the flanking base is particularly large for step 1 and step 2 during the formation of FapyG and OG, respectively. This suggests that inclusion of flanking bases for the mechanistic studies of DNA damage is important. However, increasing the model size has to be considered with caution to avoid artificial effects (*e.g.*, as was seen for 5′-OH models), while providing a better representation of the real system of interest and has a computational cost-accuracy balance. Moreover, energy barriers can be affected by the model size, although further analyses and considering similar reaction systems are required to unveil the exact effect of each model on the energy changes. Nevertheless, the results of the present chapter indicate that the molecular geometries and energetics of the transition states are generally more sensitive to the model size than the basis set for a given DFT functional, emphasizing the importance of balancing the computational model size and cost. Thus, future studies of guanine oxidation pathways aiming for accurate structural and energetic results should consider large models (trimer or larger), which requires QM/MM approach.

2.5 References

1. Breen, A. P.; Murphy, J. A., Reactions of Oxyl Radicals with DNA. *Free Radic. Biol. Med.* 1995, *18*, 1033–1077.

2. Muller, F. L.; Lustgarten, M. S.; Jang, Y.; Richardson, A.; Van Remmen, H., Trends in oxidative aging theories. *Free Radic. Biol. Med.* 2007, *43*, 477–503.

3. Neeley, W. L.; Essigmann, J. M., Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem Res Toxicol* 2006, *19*, 491–505.

4. Stover, J. S.; Ciobanu, M.; Cliffel, D. E.; Rizzo, C. J., Chemical and electrochemical oxidation of C8-arylamine adducts of 2′-deoxyguanosine. *J. Am. Chem. Soc.* 2007, *129*, 2074–2081.

5. Gimisis, T.; Cismaş, C., Isolation, Characterization, and Independent Synthesis of Guanine Oxidation Products. *Eur. J. Org. Chem.* 2006, *2006*, 1351–1378.

6. Burrows, C. J.; Muller, J. G., Oxidative nucleobase modifications leading to strand scission. *Chem. Rev. (Washington, DC, U. S.)* 1998, *98*, 1109–1151.

7. Lewis, K.; Copeland, K.; Hill, G., One-electron redox properties of DNA nucleobases and common tautomers. *Int. J. Quantum Chem.* 2014, *114*, 1678–1684.

8. Puiatti, M.; Vera, D. M.; Pierini, A. B., Species with negative electron affinity and standard DFT methods. Finding the valence anions. *Phys Chem Chem Phys* 2008, *10*, 1394–1399.

9. Gu, J.; Xie, Y.; Schaefer, H. F., Electron attachment to DNA single strands: gas phase and aqueous solution. *Nucleic Acids Res.* 2007, *35*, 5165–5172.

10. Psciuk, B. T.; Lord, R. L.; Munk, B. H.; Schlegel, H. B., Theoretical Determination of One-Electron Oxidation Potentials for Nucleic Acid Bases. *J. Chem. Theory Comput.* 2012, *8*, 5107–5123.

11. Munk, B. H.; Burrows, C. J.; Schlegel, H. B., Exploration of Mechanisms for the Transformation of 8-Hydroxy Guanine Radical to FAPyG by Density Functional Theory. *Chem. Res. Toxicol.* 2007, *20*, 432–444.

12. Munk, B. H.; Burrows, C. J.; Schlegel, H. B., An exploration of mechanisms for the transformation of 8-oxoguanine to guanidinohydantoin and spiroiminodihydantoin by density functional theory. *J. Am. Chem. Soc.* 2008, *130*, 5245–5256.

13. Jena, N. R.; Mishra, P. C., Mechanisms of formation of 8-oxoguanine due to reactions of one and two OH· radicals and the H_2O_2 molecule with guanine: A quantum computational study. *J. Phys. Chem. B* 2005, *109*, 14205–14218.

14. Psciuk, B. T.; Schlegel, H. B., Computational Prediction of One-Electron Reduction Potentials and Acid Dissociation Constants for Guanine Oxidation Intermediates and Products. *J. Phys. Chem. B* 2013, *117*, 9518–9531.

15. Saito, I.; Takayama, M.; Sugiyama, H.; Nakatani, K.; Tsuchida, A.; Yamamoto, M., Photoinduced DNA Cleavage via Electron Transfer: Demonstration That Guanine Residues Located 5′ to Guanine Are the Most Electron-Donating Sites. *J. Am. Chem. Soc.* 1995, *117*, 6406–6407.

16. Sugiyama, H.; Saito, I., Theoretical Studies of GG-Specific Photocleavage of DNA via Electron Transfer: Significant Lowering of Ionization Potential and 5‘-Localization of HOMO of Stacked GG Bases in B-Form DNA. *J. Am. Chem. Soc.* 1996, *118*, 7063–7068.

17. Verdolino, V.; Cammi, R.; Munk, B. H.; Schlegel, H. B., Calculation of pK(a) Values of Nucleobases and the Guanine Oxidation Products Guanidinohydantoin and Spiroiminodihydantoin using Density Functional Theory and a Polarizable Continuum Model. *J. Phys. Chem. B* 2008, *112*, 16860–16873.

18. Jena, N. R.; Mishra, P. C., Is FapyG mutagenic? Evidence from the DFT study. *ChemPhysChem* 2013, *14*, 3263–3270.

19. Sadeghian, K.; Ochsenfeld, C., Unraveling the Base Excision Repair Mechanism of Human DNA Glycosylase. *J. Am. Chem. Soc.* 2015, *137*, 9824–9831.

20. Sadeghian, K.; Flaig, D.; Blank, I. D.; Schneider, S.; Strasser, R.; Stathis, D.; Winnacker, M.; Carell, T.; Ochsenfeld, C., Ribose-Protonated DNA Base Excision Repair: A Combined Theoretical and Experimental Study. *Angew. Chem., Int. Ed. Engl.* 2014, *53*, 10044–10048.

21. Blank, I. D.; Sadeghian, K.; Ochsenfeld, C., A Base-Independent Repair Mechanism for DNA Glycosylase—No Discrimination Within the Active Site. *Scientific Reports* 2015, *5*, 10369.

22. Sowlati-Hashjin, S.; Wetmore, S. D., Computational Investigation of Glycosylase and β-Lyase Activity Facilitated by Proline: Applications to FPG and Comparisons to hOgg1. *J. Phys. Chem. B* 2014, *118*, 14566–14577.

23. Sowlati-Hashjin, S.; Wetmore, S. D., Quantum mechanical study of the β- and δ-lyase reactions during the base excision repair process: application to FPG. *Phys. Chem. Chem. Phys.* 2015, *17*, 24696–24706.

24. Kellie, J. L.; Wetmore, S. D., Mechanistic and Conformational Flexibility of the Covalent Linkage Formed during β-Lyase Activity on an AP-Site: Application to hOgg1. *J. Phys. Chem. B* 2012, *116*, 10786–10797.

25. Kellie, J. L.; Wilson, K. A.; Wetmore, S. D., An ONIOM and MD Investigation of Possible Monofunctional Activity of Human 8-Oxoguanine–DNA Glycosylase (hOgg1). *J. Phys. Chem. B* 2015, *119*, 8013–8023.

26. Sebera, J.; Trantírek, L.; Tanaka, Y.; Sychrovsky, V., Pyramidalization of the Glycosidic Nitrogen Provides the Way for Efficient Cleavage of the N-Glycosidic Bond of 8-OxoG with the hOGG1 DNA Repair Protein. *J. Phys. Chem. B* 2012, *116*, 12535–12544.

27. Sebera, J.; Trantirek, L.; Fukal, J.; Tanaka, Y.; Sychrovsky, V., Mechanism of Base Excision with the hOgg1 DNA Repair Enzyme. *Chemicke Listy* 2014, *108*, 364-367.

28. Sebera, J.; Trantirek, L.; Tanaka, Y.; Nencka, R.; Fukal, J.; Sychrovsky, V., The activation of N-glycosidic bond cleavage performed by base-excision repair enzyme hOGG1; theoretical study of the role of Lys 249 residue in activation of G, OxoG and FapyG. *RSC Advances* 2014, *4*, 44043–44051.

29. Sebera, J.; Hattori, Y.; Sato, D.; Reha, D.; Nencka, R.; Kohno, T.; Kojima, C.; Tanaka, Y.; Sychrovsky, V., The mechanism of the glycosylase reaction with hOGG1 base-excision repair enzyme: concerted effect of Lys249 and Asp268 during excision of 8-oxoguanine. *Nucleic Acids Res.* 2017, *45*, 5231–5242.

30. Schyman, P.; Danielsson, J.; Pinak, M.; Laaksonen, A., Theoretical Study of the Human DNA Repair Protein HOGG1 Activity. *J. Phys. Chem. A* 2005, *109*, 1713–1719.

31. Calvaresi, M.; Bottoni, A.; Garavelli, M., Computational Clues for a New Mechanism in the Glycosylase Activity of the Human DNA Repair Protein hOGG1. A Generalized Paradigm for Purine-Repairing Systems? *J. Phys. Chem. B* 2007, *111*, 6557–6570.

32. Llano, J.; Eriksson, L. A., Oxidation pathways of adenine and guanine in aqueous solution from first principles electrochemistry. *Phys. Chem. Chem. Phys.* 2004, *6*, 4707–4713.

33. Reynisson, H.; Steenken, S., DFT calculations on the electrophilic reaction with water of the guanine and adenine radical cations. A model for the situation in DNA. *Phys. Chem. Chem. Phys.* 2002, *4*, 527–532.

34. Luo, W.; Muller, J. G.; Rachlin, E. M.; Burrows, C. J., Characterization of Spiroiminodihydantoin as a Product of One-Electron Oxidation of 8-Oxo-7,8-dihydroguanosine. *Org. Lett.* 2000, *2*, 613–616.

35. Niles, J. C.; Wishnok, J. S.; Tannenbaum, S. R., Spiroiminodihydantoin and Guanidinohydantoin Are the Dominant Products of 8-Oxoguanosine Oxidation at Low Fluxes of Peroxynitrite: Mechanistic Studies with 18O. *Chem. Res. Toxicol.* 2004, *17*, 1510–1519.

36. Ye, Y.; Muller, J. G.; Luo, W.; Mayne, C. L.; Shallop, A. J.; Jones, R. A.; Burrows, C. J., Formation of ^{13}C-, ^{15}N-, and ^{18}O-Labeled Guanidinohydantoin from Guanosine Oxidation with Singlet Oxygen. Implications for Structure and Mechanism. *J. Am. Chem. Soc.* 2003, *125*, 13926–13927.

37. Suzuki, T.; Friesen, M. D.; Ohshima, H., Identification of Products Formed by Reaction of 3′,5′-Di-O-acetyl-2′-deoxyguanosine with Hypochlorous Acid or a Myeloperoxidase−H_2O_2−Cl^- System. *Chem. Res. Toxicol.* 2003, *16*, 382–389.

38. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., *et al. Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford CT, 2009.

39. *HyperChem*^{TM}, Professional 8.0.8; Hypercube, Inc.: Gainesville, Florida 32601, USA.

40. Jayaram, B.; Sharp, K. A.; Honig, B., The electrostatic potential of B-DNA. *Biopolymers* 1989, *28*, 975–993.

41. Stigter, D., An electrostatic model for the dielectric effects, the adsorption of multivalent ions, and the bending of B-DNA. *Biopolymers* 1998, *46*, 503–516.

42. Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability. Theory and applications. *Theor. Chem. Acc.* 2009, *124*, 11–28.

43. 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. U. S. A.* 2014, *111*, 3624–3630.

44. Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X. J.; Neidle, S.; Shakked, Z., *et al.*, A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* 2001, *313*, 229–237.

45. Zhou, H.; Hintze, B. J.; Kimsey, I. J.; Sathyamoorthy, B.; Yang, S.; Richardson, J. S.; Al-Hashimi, H. M., New insights into Hoogsteen base pairs in DNA duplexes from a structure-based survey. *Nucleic Acids Res* 2015, *43*, 3420–3433.

46. Kuska, M. S.; Majdi Yazdi, M.; Witham, A. A.; Dahlmann, H. A.; Sturla, S. J.; Wetmore, S. D.; Manderville, R. A., Influence of Chlorine Substitution on the Hydrolytic Stability of Biaryl Ether Nucleoside Adducts Produced by Phenolic Toxins. *The Journal of Organic Chemistry* 2013, *78*, 7176–7185.

47. Navarro-Whyte, L.; Kellie, J. L.; Lenz, S. A.; Wetmore, S. D., Hydrolysis of the damaged deoxythymidine glycol nucleoside and comparison to canonical DNA. *Phys Chem Chem Phys* 2013, *15*, 19343–19352.

48. Sharma, P.; Manderville, R. A.; Wetmore, S. D., Modeling the Conformational Preference of the Carbon-Bonded Covalent Adduct Formed upon Exposure of 2′-Deoxyguanosine to Ochratoxin A. *Chem. Res. Toxicol.* 2013, *26*, 803–816.

Chapter 3: Recognition of OG Glycosidic Conformers*

3.1 Introduction

As discussed in Chapter 1, both hOgg1 and FPG have been shown to exhibit an activity preference for OG repair based on the nucleobase opposite the lesion (Figure 1.9). Specifically, hOgg1 efficiently removes OG paired with C, but the excision rate for OG paired with A is negligible, being 1000- to 3000-fold slower[1, 2] than OG excision opposite C. Similarly, FPG more efficiently processes OG opposite C than A, but only at ~ 18- to 35-fold greater rate,[2, 3] leading to appreciable activity regardless of the opposing (C or A) nucleobase. Although several factors in the recognition step (*e.g.*, differences in base pair strengths,[3-4] direct contacts between the enzyme and the opposing base[5-7] or disturbance of the DNA–enzyme interface destabilizing the pre-catalytic complex[8]) have been proposed to play a role in dictating the observed hOgg1/FPG activity dependence on the opposing base, differences in the glycosidic orientation of OG paired opposite C and A may also influence how the substrate is bound in the active site and therefore affect the repair efficiency. Unfortunately, there is currently no direct comparison between the abilities of hOgg1 and FPG to accommodate multiple OG binding orientations in the literature.

In addition to the ambiguity surrounding the substrate binding orientation, the chemical reactions catalyzed by hOgg1 and FPG have not been completely explained. As discussed in Chapter 1, two different S_N2 mechanisms have been proposed for the initial

---

* Reproduced in part with permission from Sowlati-Hashjin S. and Wetmore S. D. Structural Insight into the Discrimination between 8−Oxoguanine Glycosidic Conformers by DNA Repair Enzymes: A Molecular Dynamics Study of Human Oxoguanine Glycosylase 1 and Formamidopyrimidine-DNA glycosylase *Biochemistry* **2018** DOI: 10.1021/acs.biochem.7b01292. Copyright 2018 American Chemical Society.

chemical steps of BER (Figure 1.12). Although quantum mechanical calculations have predicted that the ring-opening mechanism leads to a lower barrier for OG excision by both enzymes,[9-10] more evidence is required to unequivocally identify the favored pathway. Furthermore, each proposed mechanism assumes different roles for the active site residues. For example, Asp268 of hOgg1 has been proposed to initiate nucleoside hydrolysis[11] or stabilize the Schiff base intermediate.[12] In the second mechanism, Asp268of hOgg1[5, 12] and Glu3 of FPG[13-14] have been proposed to protonate 2′-deoxyribose of the damaged nucleotide to facilitate the ring-opening step (Figure 1.12b).

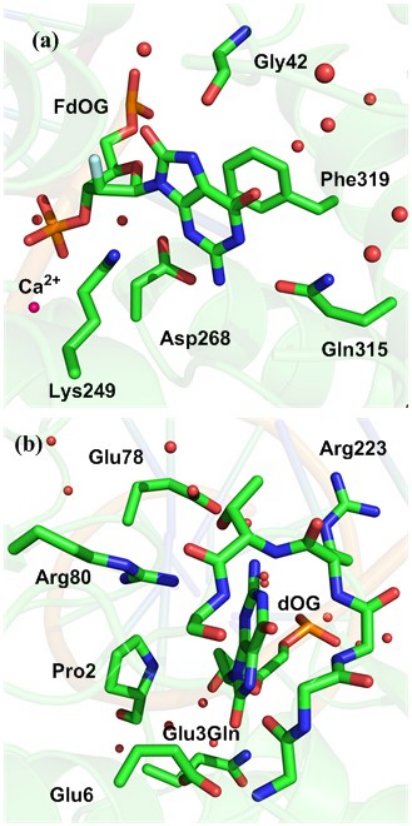

Figure 3.1. X-ray crystal structure of (a) 2′-fluoro-dOG (FdOG) bound in the hOgg1 active site (PDB ID: 3KTU) and (b) OG bound in the active site of the Glu3Gln FPG mutant (PDB ID: 1R2Y). Crystallographic water is shown as red spheres and Ca^{2+} ion as pink.

To shed light on how OG is bound and excised by hOgg1 and FPG, the present chapter uses molecular dynamics simulations to investigate changes in the active site

configuration and dynamics with the OG glycosidic conformation. By performing a detailed analysis of the interactions between the substrate, active site residues, and solvent, this work provides insight into how these enzymes accommodate different OG conformations and prevent further mutations. Importantly, my data suggests that the hOgg1 active site can more accurately discriminate against *syn*-OG than FPG, and thereby identifies the OG glycosidic bond orientation as one factor that contributes to the relative experimental excision rates. Furthermore, detailed analysis of active site interactions affords proposals regarding the roles of various hOgg1 and FPG residues in binding and/or discriminating against OG conformers. Finally, this new structural data sheds light on the relative importance of the previously proposed S_N2 catalytic mechanisms[13, 15] for the first chemical step catalyzed by both repair enzymes. Together, the present simulations rationalize the relative activities of these critical repair enzymes, and highlight similarities as well as key differences in how human and bacterial repair enzymes combat the major DNA oxidation product.

3.2 Computational Details

Starting structures for MD simulations were built using high resolution X-ray crystal structures. In the case of hOgg1, a crystal structure of the enzyme bound to DNA containing OG fluorinated at C2′ (2′-fluoro-dOG, FdOG) was used (PDB ID: 3KTU).[5] Missing residues (Gln80, Asp81, and Lys82) were added using PyMol[16] and GaussView,[17] and the C2′-fluorine atom of the OG nucleotide was replaced by a hydrogen atom. In the case of FPG, a crystal structure of the Glu3Gln mutant of *B. stearothermophilus* FPG bound to OG-containing DNA was used (PDB ID: 1R2Y),[18] with Gln3 substituted by the native Glu. For both crystal structures, missing heavy atoms and all hydrogen atoms were added using the LEaP module of Amber 12.[19] Protonation

states of ionizable residues were assigned using PROPKA 3.0,[21] with the exception of key active site residues. Specifically, for hOgg1, His270 was protonated since MD simulations suggest that this protonation state better maintains the active site orientation of the crystal structure.[21] Additionally, Lys249 and Cys253 were modelled as neutral since quantum mechanical calculations show that $Lys249^{+1}Cys254^{-1}$ is energetically unfavorable compared to Lys249Cys254 regardless of the Asp268 protonation state.[10, 21] Moreover, previous MD simulations revealed that the salt bridge increases the solvent density in the active site, which may prohibit catalysis.[21] Finally, although the dipole moment of $Lys249^{+1}Cys254^{-1}$ has been proposed to play an important role in the recognition step,[22] reversing the dipole direction with a Lys249Cys/Cys253Lys (KCCK) double mutant does not affect recognition, which may indicate that these residues are neutral.[11] Depending on the order of the deglycosylation and ring-opening steps (Figure 1.12), Asp268 may be anionic or neutral upon substrate binding. Specifically, if deglycosylation occurs first, Asp268 is likely anionic in order to stabilize the positive charge developing on 2′-deoxyribose.[25, 26] Alternatively, catalysis of an initial ring-opening step requires proton transfer to O4′ of 2′-deoxyribose, which could be facilitated by Asp268 that has been neutralized by Lys249 or solvent. Therefore, models were considered with anionic and neutral Asp268, and the solvent distribution in the active site was carefully monitored. Similarly, Glu3 in FPG was considered to be either anionic or neutral, while the Pro2 nucleophile was neutral. Hydrogen atoms were placed on the Oδ or Oε atoms of Asp268 or Glu3 automatically using the LEaP module of Amber 12, and were visually inspected to ensure the initial conformation adopted the most favorable proton orientation for an isolated amino acid. The OG substrate was bound in the active site in both the *anti* and *syn* conformations for each active site protonation state. Although

the base opposite the lesion affects enzymatic activity and may play a direct role in enzyme function, crystal structures of FPG reveal OG bound in the *syn* orientation despite the opposing C in the DNA duplex.[18] Therefore, to uncouple other contributing factors and thereby compare the effects of the glycosidic conformation on OG binding for hOgg1 and FPG, both OG conformers were considered and the opposing base maintained as C for both enzymes.

All systems were assigned AMBER parm99SB parameters[23] and parameters for non-standard OG taken from a previous work,[21] which employed parm99SB and GAFF parameters.[24, 25] In the case of FPG, parameters for the zinc ion were taken from ZAFF.[26] Each system was neutralized with Na^+ ions and solvated in an explicit TIP3P water box that was at least 8.0 Å from the edge of the DNA–enzyme complex. Constraints were imposed on covalent bonds involving hydrogen atoms using the SHAKE algorithm, while the particle mesh Ewald algorithm was used for long-range electrostatic interactions. The solvent molecules and ions were relaxed using 500 steps of steepest descent, and 500 steps of conjugate gradient minimization, while the protein and DNA were constrained using a 500.0 kcal mol^{-1} $Å^{-2}$ force constant. The entire system was then minimized using 1000 steps of unrestrained steepest descent, followed by 1500 steps of unrestrained conjugate gradient minimization. Subsequently, the system was heated from 0 to 300 K over 20 ps with restraints on the solute (10 kcal mol^{-1} $Å^{-2}$). Prior to the production phase, each system was equilibrated for 20 ps. The periodic boundary condition was employed for all MD simulations.

Once the preliminary models were established, each system was used for 20 ns of unconstrained MD pre-production simulations under NPT conditions (1 atm and 300 K). From each resulting trajectory, representative structures were chosen for subsequent 20 ns

pre-production simulations, which employed different initial velocities to enhance the sampling over the phase space. This iterative process was repeated until no large deviations in the active site were observed. At this point, two 20 ns MD production simulations were completed using different initial velocities. For all systems, the two replicas led to highly similar structures (rmsd with respect to the crystal structure of ~ 0.5 – 0.8 Å, Figures A3.1 and A3.2). Subsequently, one trajectory was chosen for a further 80 ns of production simulation under similar conditions. For all systems, the representative structures from the 20 and 100 ns simulations show significant resemblance (Figures A3.3 and A3.4). In the main text and the remainder of the Appendix A3, the 100 ns simulations results are presented.

Trajectory analysis was completed using the cpptraj module of Amber 12.[20] The root-mean-square deviation (rmsd) of the protein and DNA backbone over the production phase was analyzed for each simulation to ensure the system was stable (Table A3.1). Each trajectory was saved every 2 fs over the course of the production simulation. Clustering was completed for the trajectories using the hierarchical agglomerative algorithm (with $\varepsilon = 4$) based on the configurations of OG and key active site residues, namely Gly42, Lys249, Asp268, Gln315, and Phe319 for hOgg1, and Pro2, Glu3 for FPG. Although a single representative structure is shown for the cluster with the highest occupancy (see Table A3.1 for the corresponding rmsd and occupancy), the geometrical and energetic analyses were completed over all structures in the highest occupied cluster and the corresponding dynamical information is provided in Appendix A3.

Throughout the chapter, key average geometrical parameters are reported (standard deviations are provided in Appendix A3), including the nucleophilic distance (*i.e.*, the distance between Lys249/Pro2 and C1′ of 2′-deoxyribose or d(Nζ/N···C1′)), the

distance between Asp268 (Oδ) or Glu3 (Oε) and O4′ of 2′-deoxyribose (*i.e.*, d(Oδ/Oε···O4′)), and the relative orientation of the nucleophile (Lys249 or Pro2) and the OG nucleotide (*i.e.*, the ∠(Nζ/N–C1′–O4′) and ∠(Nζ/N–C1′–N9) nucleophilic attack angles). A 120° angle cut-off and a 3.4 Å distance cut-off were imposed between (donor and acceptor) heavy atoms for all hydrogen-bonding active site interactions. The linear interaction energies (LIE) are reported for key active site stacking and hydrogen-bonding interactions. The distribution of water in the active site was examined using a three-dimensional 20 Å × 20 Å × 20 Å grid centered on the OG nucleotide, with 0.5 Å spacing between grid points. In order to assess the possibility of proton transfer from bulk solvent to (anionic) Asp268 or Glu3, the average number of water molecules close to these residues was determined using a spherical solvation shell with a radius of 3.4 or 6.0 Å centered on one of three atoms (*i.e.*, O4′ of OG, and Oδ1 or Oδ2 of Asp268 for hOgg1, or Oε1 or Oε2 of Glu3 for FPG).

3.3 Results

3.3.1 hOgg1

3.3.1.1 anti-OG

The representative MD structure of *anti*-OG bound in the hOgg1 active site with anionic or neutral Asp268 (Figures 3.2 and A3.5) is similar to the starting crystal structure in terms of the relative orientation of key active site residues and OG (average active site backbone rmsd < 0.6 Å, with per residue rmsd provided in Table A3.2). Furthermore, OG maintains the *anti* conformation of the crystal structure (average $\chi \sim -104$ or $-113°$, Table A3.3). The hydrogen-bonding patterns between active site residues and OG are similar for anionic and neutral Asp268 (Table A3.4). Specifically, a hydrogen bond between N7–H of OG and the Gly42 backbone (Figure A3.6) is maintained over the trajectory (95 or

99% occupancy, Table A3.4), and the OG–Gly42 average interaction energy is –6.5 kcal/mol (Table A3.5). In both systems, N1–H and N2–H of OG form a hydrogen bond with Gln315, with similar hydrogen-bonding occupancies (91 – 99%) and hence average OG–Gln315 stabilization energies (~ –13 kcal/mol). Additionally, N2–H of OG forms a hydrogen bond with Oδ of Asp268, with an overall per residue interaction energy of –13.8 or –14.9 kcal/mol for anionic or neutral Asp268, respectively. The crystal structure position of anionic Asp268 relative to 2′-deoxyribose is also maintained through conserved hydrogen bonding to His270 and Met271 (~ 100% occupancy), although the hydrogen-bonding occupancies are reduced for neutral Asp268 (72 – 89%, Figure A3.6 and Table A3.6). Nevertheless, in both cases, Asp268 maintains the capping interactions for the α-helix (αM, Figure A3.6). Phe319 maintains the crystal structure stacked position (3.779 Å) with respect to OG for both systems (average of 3.613 and 3.596 Å for anionic and neutral Asp268, respectively, Figure A3.5 and Table A3.7), but Phe319 has a small overall contribution to substrate binding (~ –4 kcal/mol, Table A3.5).

Plots of the solvent distribution (Figure A3.7) show that water is distributed around the bound nucleobase, in particular within proximity to O6 of OG. In fact, a crystallographic water molecule hydrogen bonds with O6 for ~ 50% of the simulation (Figure A3.6), and stabilizes the DNA–protein complex by ~ –7 kcal/mol (Table A3.5). Furthermore, a crystallographic water repositions closer to anionic Asp268 during the simulation (Figure A3.6a). The presence of water near O4′ of 2′-deoxyribose, or an Oδ atom of anionic or neutral Asp268 (Table A3.8) is particularly important when Asp268 is anionic in order to promote ring opening through protonation of O4′, either directly or through Asp268. In fact, OG(O4′)···H_2O and Asp268(Oδ)···H_2O hydrogen bonds exist for 27 and 94% of the simulation time, respectively (Figure A3.6, and Tables A3.4 and

A3.6). Although a significant amount of water also exists in this region when Asp268 is neutral (Figure A3.7 and Table A3.6), an OG(O4′)···Asp268(Oδ–H) hydrogen bond is maintained for 24% of the simulation (Figure A3.6 and Table A3.4).

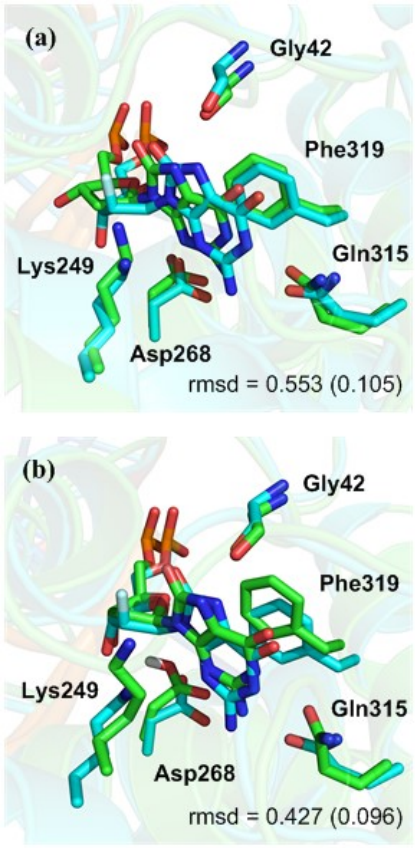

Figure 3.2. Overlay of the crystal structure (PDB ID: 3KTU; cyan) and MD representative structure (green) of *anti*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity only the hydrogen atom of neutral Asp268 (Oδ–H) is shown. The rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

In order to obtain structural insight into the feasibility and relative importance of the previously examined S_N2 mechanisms for the first BER step (Figure 1.12), key reaction parameters for each mechanism were evaluated, namely the nucleophilic distance, the nucleophilic attack angle, and the distance between Asp268 and O4′ of 2′-deoxyribose. The d(Nζ···C1′) nucleophilic distance is 4.374 or 4.127 Å for hOgg1 with anionic or neutral Asp268, respectively (Table A3.7). This distance is in agreement with

the nucleophilic distance in the reactant complex predicted by previous QM calculations (3.109 Å).[10] Although the ∠(Nζ–C1′–N9) angle is 62.7° or 59.2° for anionic or neutral Asp268, respectively (Table A3.9), which deviates significantly from the 180° angle anticipated for an initial S_N2 deglycosylation step, the average ∠(Nζ–C1′–O4′) angle is 153.9° or 149.9° for anionic or neutral Asp, respectively (Table A3.9). The predicted ∠(Nζ–C1′–O4′) angle is similar to that previously reported from QM calculations (136.5° or 157.6° in the reactant or transition state complex, respectively).[10] Furthermore, Asp268 is in close proximity to O4′ of 2′-deoxyribose regardless of the protonation state (*i.e.*, d(Oδ1···O4′) < 3.708 Å, Table A3.7).

3.3.1.2 syn-OG

OG maintains the *syn* orientation throughout the simulation when bound in the hOgg1 active site (average χ = 61.8° or 76.2° for hOgg1 with anionic or neutral Asp268, respectively; Table A3.3). In comparison with the crystal structure and MD data for *anti*-OG bound to hOgg1, the active site configuration changes when *syn*-OG is considered (average active site backbone rmsd ~ 0.5 Å; Figure 3.3, with per residue rmsd provided in Table A3.2). Specifically, anionic Asp268 is significantly displaced due to repulsion with O8 of *syn*-OG (Figures 3.3a and A3.8a), with a resulting average interaction energy of ~ –2 kcal/mol (Table A3.5). The new location of Asp268 disrupts the helix cap such that Asp268 hydrogen bonds with Met271 (95%) and Val269 (99%), but not His270 (Figure A3.9 and Table A3.6), which changes the first turn of the αM helix type from 3.6_{13} to 3_{10}. Additionally, the position of a crystallographic Ca^{2+} ion changes to afford coordination with Asp268 (Figure A3.9a). This partially solvated ion is used by the enzyme to stabilize the negative charge of the phosphate groups proximal to the lesion site. The new Ca^{2+}

position induces a conformational change in the DNA backbone on the 3′-side of OG, which alters the position of the nucleotide relative to Lys249 (Figures 3.3 and A3.8a). Other changes in the active site include interactions between Gly42 and the N1 and N2 amino hydrogens of *syn*-OG (20 and 77%, Figure A3.9 and Table A3.4), which collectively stabilize the DNA–enzyme complex by ~ –7 kcal/mol (Table A3.5). On the other hand, Gln315 interacts with OG through N7–H (41%) and O6 (76%), and stabilizes the system by ~ –4.9 kcal/mol. The average distance between Phe319 and *syn*-OG (~ 3.9 Å) is slightly larger than *anti*-OG (~ 3.6 Å, Table A3.7), and hence the associated stacking interaction for *syn*-OG is slightly reduced by ~ 0.5 kcal/mol (Table A3.5).

Active site conformational changes are also induced by *syn*-OG when Asp268 is neutral (Figures 3.3b and A3.8b). In this case, Asp268 shifts relative to 2′-deoxyribose and forms an OG(O8)···Asp268(Oδ–H) hydrogen bond (89% occupancy and average overall OG–Asp268 interaction energy of –10.2 kcal/mol; Figure A3.9, Tables A3.4 and A3.5). In this orientation, Asp268 maintains helix-capping interactions with His270 (98%); however, the Asp268(Oδ)···Met271(N–H) interaction is reduced to 18% and replaced by Asp268(O)···Met271(N–H) (72%), which induces the 3.6_{13} to 3_{10} change in the first turn of the αM helix (Figure A3.9b and Table A3.6). Gly42 forms hydrogen bonds with OG at N1–H and N2–H (93 and 88%, Figure A3.9b and Table A3.4) and stabilizes the DNA–enzyme complex by –9.0 kcal/mol (Table A3.5). However, repulsive interactions between OG (O6 and O8) and Gln315 (Oε) reduce the associated hydrogen-bonding occupancy (17%, Table A3.4) and interaction energy (~ –1 kcal/mol, Table A3.5).

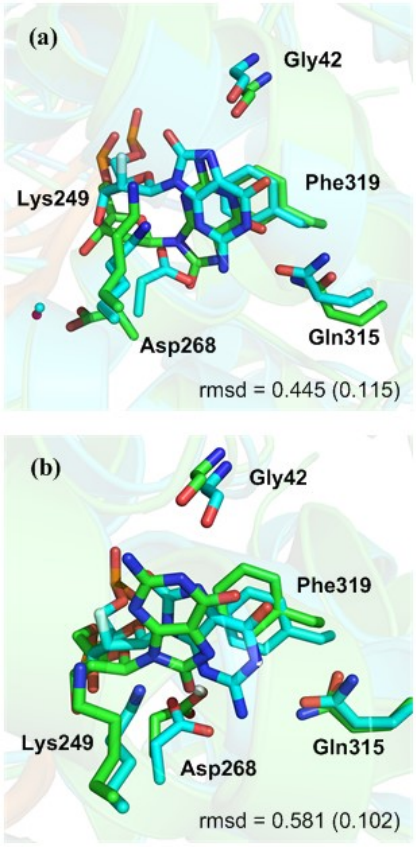

Figure 3.3. Overlay of the crystal structure (PDB ID: 3KTU; cyan) and MD representative structure (green) of *syn*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity only the hydrogen atom of neutral Asp268 (Oδ–H) is shown. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

The overall distribution of solvent in the active site when *syn*-OG is bound is similar to that of *anti*-OG (Figure A3.10). OG (O6) hydrogen bonds to a crystallographic water molecule for ~ 28 – 36% of the simulation (Table A3.4), with an associated contribution to binding of ~ –3 to –4 kcal/mol (Table A3.5). The average number of water molecules centered on anionic Asp268 for *syn*-OG is greater than for *anti*-OG (Table A3.8), which is due to the displacement of this residue toward the solvent exposed backbone, and hence does not imply a higher probability of catalysis. In fact, although the average number of solvent molecules near O4′ is comparable for *syn* and *anti*-OG (Table

A3.8), O4′ of *syn*-OG forms a hydrogen bond with solvent for only 13% or 14% of the simulation for anionic or neutral Asp268, respectively (Table A3.4).

The d(Nζ···C1′) nucleophilic distance is 5.505 or 4.605 Å for *syn*-OG bound to hOgg1 containing anionic or neutral Asp268, respectively, which is larger than for *anti*-OG (Table A3.7). Furthermore, compared to *anti*-OG, the ∠(Nζ–C1′–N9) angle increases slightly for *syn*-OG with anionic (78.3°) or neutral (74.4°) Asp268 (Table A3.9), but is still not ideal for S_N2 deglycosylation upon nucleophilic attack. The ∠(Nζ–C1′–O4′) angle is significantly smaller (109.6°, Table A3.9) and the d(Oδ1···O4′) distance is significantly larger (6.587 Å, Table A3.7) for *syn*-OG than *anti*-OG with anionic Asp268 due to the displacement of the Asp268 sidechain. Nevertheless, this distance for neutral Asp268 (4.246 Å, Table A3.7) is comparable to that for *anti*-OG. Although the ∠(Nζ–C1′–O4′) nucleophilic attack angle for *syn*-OG with neutral Asp268 (136.0°, Table A3.9) is comparable to *anti*-OG (~ 150°, Table S9), the displaced Asp268 sidechain makes the attack implausible from this starting point.

3.3.2 FPG

3.3.2.1 anti-OG

When *anti*-OG is bound in the FPG active site with anionic or neutral Glu3 (Figure 3.4), the crystal structure positions of the backbones of active site residues are maintained (average rmsd ~ 0.6 Å), although some conformational changes occur in the sidechains (per residue rmsd provided in Table A3.10). *anti*-OG (average χ = –113.7 and –101.6° with anionic or neutral Glu3, respectively; Table A3.3) extensively interacts with the loop located on the O6–N7 side of the nucleobase (Figures A3.11 and A3.12, and Table A3.11). The major hydrogen bonds between the loop and the OG lesion involve Ser220, Val222, Arg223, Thr224, and Tyr225 (Figures A3.12a and A3.13a).

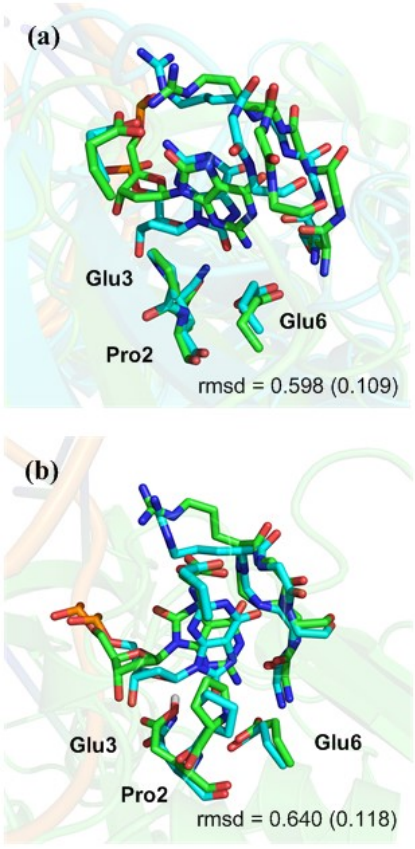

Figure 3.4. Overlay of the crystal structure (PDB ID: 1R2Y; cyan) and the MD representative structure (green) of *anti*-OG bound in the FPG active site with (a) anionic or (b) neutral Glu3. For clarity only the hydrogen atom of neutral Glu3 (Oε–H) is shown. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are shown.

The strongest interaction between the nucleobase and the loop region occurs with Arg223 (up to ~ –31 kcal/mol, Table A3.12). Beyond the loop, few residues interact with the nucleobase. For both anionic and neutral Glu3, the strongest interaction outside the loop occurs between OG(N2–H) and Glu6 (Figures A3.12b and A3.13b), which contributes to the average interaction energy of ~ –16 kcal/mol (Table A3.13). The position of the anionic Glu3 sidechain relative to 2′-deoxyribose is primarily maintained through hydrogen bonding to Gly173, Asn174, Ile175, and Tyr176 (Table A3.13). The hydrogen bonds between Glu3, and Gly173 and Asn174 vanish for neutral Glu3. Instead, neutral Glu3 interacts with OG at O3′ and O4′ for 25% and 22% of the simulation,

respectively (Table A3.13). On the other hand, the OG(O8)···Glu78(Oε) contact is the most repulsive interaction regardless of the Glu3 protonation state (~ 10 – 12 kcal/mol, Table A3.12).

Similar to hOgg1 (Figure A3.7), the active site of FPG is solvent exposed (Figure A3.14), with large solvent distribution around OG. The average number of solvent molecules in the vicinity of Glu3 and the sugar ring is larger for anionic than neutral Glu3 for both solvent shells (Table A3.14). In fact, O4′ of the sugar and anionic Glu3 form hydrogen bonds with solvent for 39 and 55% of the simulation time, respectively (Figure A3.12b, and Tables A3.11 and A3.13). On the other hand, neutral Glu3 directly hydrogen bonds to O4′ of 2′-deoxyribose (22% occupancy, Table A3.11). The d(N···C1′) nucleophilic distance is 4.219 or 5.316 Å for anionic or neutral Glu3, respectively (Table A3.7). Regardless of the Glu3 protonation state, the ∠(N–C1′–O4′) angle is better aligned for a substitution reaction that involves an initial ring-opening step (~ 160°, Tables A3.9), while the ∠(N–C1′–N9) angle is < 81° with either Glu3 protonation state. Moreover, although the distance between anionic Glu3 (Oε) and O4′ is large (up to ~ 7.5 Å), neutral Glu3 lies close to O4′ (< 4.6 Å, Table A3.7) due to hydrogen bonds between Glu3(Oε–H) and O3′, O4′, and N3 of OG.

3.3.2.2 syn-OG

The overall structure of the FPG active site bound to *syn*-OG with either anionic or neutral Glu3 resembles the experimental crystal structure (average active site backbone rmsd ~ 0.5 Å with anionic or neutral Glu3, respectively, Figures 3.5 and A3.15, with per residue rmsd provided in Table A3.10). Regardless of the Glu3 protonation state, OG maintains a *syn* orientation throughout the simulation, with χ adopting values (~ 57° – 67°) smaller than the crystal structure (χ ~ 101°, Table A3.3). Similar to *anti*-OG, *syn*-OG

extensively interacts with the FPG active site loop regardless of the Glu3 protonation state (Figures A3.16a and A3.17a). Most hydrogen-bond occupancies (Table A3.11) and interaction energies (Table S12) between OG and the loop residues (Ser220, Arg223, Thr224, and Tyr225) are comparable for *anti*- and *syn*-OG. However, unlike *anti*-OG, the strongest interaction with *syn*-OG occurs for Glu78 (–26.1 and –24.8 kcal/mol with anionic or neutral Glu3, respectively) rather than Arg223 (Table A3.12). In fact, the repulsive interaction between the Arg223 sidechain and *syn*-OG significantly affects the interaction energy (~ –3 kcal/mol). In addition to Arg223, differences in *anti*- and *syn*-OG binding occur in the interactions involving residues outside the loop, namely Glu3, Glu6, and Glu78 (Figures A3.16b and A3.17b). Specifically, unlike *anti*-OG, *syn*-OG interactions with Glu3, Glu6, and Arg80 are destabilizing (by ~ 1 – 12 kcal/mol), while the Glu78 interaction is stabilizing (by ~ –20 to –23 kcal/mol, Table A3.12).

Despite destabilizing interactions between *syn*-OG (O8 and O4′) and Glu3 (Table A3.12), the Glu3 sidechain maintains the crystal structure position relative to O4′ for *syn*-OG binding in the FPG active site (Figure 3.5). Furthermore, similar to *anti*-OG in FPG, hydrogen-bond interactions between anionic Glu3 and Gly173, Asn174, Ile175, and Tyr176 preserve the position of the Glu3 sidechain relative to 2′-deoxyribose ring (Table A3.13). With neutral Glu3, hydrogen bonds between Glu3 and Gly173 or Asn174 completely vanish, which is only partly compensated by hydrogen bonding to OG (O3′ and O4′, Table A3.13). On the other hand, neutral Glu3 forms an additional hydrogen bond with Glu6 (55% occupancy, Figure A3.17b and Table A3.13).

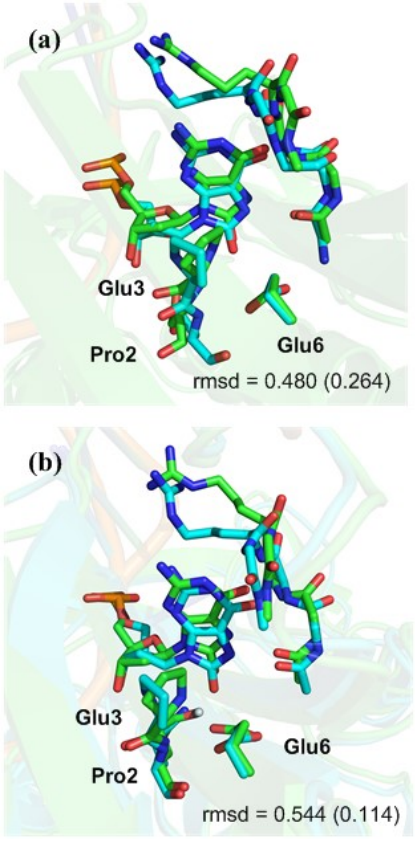

Figure 3.5. Overlay of the crystal structure (PDB ID: 1R2Y; cyan) and the MD representative structure (green) of *syn*-OG bound in the FPG active site with (a) anionic or (b) neutral Glu3. For clarity only the hydrogen atom of neutral Glu3 (Oε–H) is shown. The rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are shown.

Similar to *anti*-OG, there is significant water density in the vicinity of anionic Glu3 and the sugar ring when *syn*-OG is bound in the FPG active site (Figure A3.18). Indeed, the average number of solvent molecules centered on O4′ of the sugar or Oε of anionic Glu3 is comparable for *anti*- and *syn*-OG (Table A3.14). Furthermore, O4′ of *syn*-OG forms hydrogen bonds with solvent for 27 or 51% of the simulation trajectory when anionic or neutral Glu3 is considered, respectively (Table A3.11). While anionic Glu3 hydrogen bonds with solvent for 69% of the simulation (Table A3.13), neutral Glu3 forms a hydrogen bond with O4′ (21% occupancy, Table A3.11).

The position of the Pro2 nucleophile relative to OG is comparable for *anti*- and *syn*-OG regardless of the Glu3 protonation state (Figure A3.15), with an average d(N···C1′) distance of 3.998 or 4.542 Å for *syn*-OG and anionic or neutral Glu3, respectively (Table A3.7). Similarly, the ∠(N–C1′–N9) and ∠(N–C1′–O4′) angles are ~ 70 and 160° for *syn*-OG (Table A3.9), and are comparable to the angles discussed for *anti*-OG. Both d(Oε···O4′) distances are larger for anionic than neutral Glu3 (Table A3.7), presumably due to the OG(O4′)···Glu3(Oε–H) and OG(O3′)···Glu3(Oε–H) hydrogen bonds for neutral Glu3 that are absent in the case of anionic Glu3.

3.4 Discussion

The present chapter data provides a structural basis for understanding how OG binds in the active sites of hOgg1 and FPG, and rationalizes experimental data highlighting the differential ability of these critical repair enzymes to excise OG. Although *anti*-OG forms a Watson-Crick hydrogen-bonded pair with C, *syn*-OG forms a Hoogsteen pair with A, and experimental studies have shown that the activities of hOgg1 and FPG are uniquely dependent on the opposing base. Specifically, hOgg1 removes OG paired with C, while the excision rate is negligible when OG is paired with A. In contrast, FPG exhibits appreciable activity towards OG regardless of the pairing base, albeit with reduced efficiency when the opposing base is A. Although factors early in the recognition step, such as disruption of the DNA–enzyme interface destabilizing the pre-catalytic complex, weaker base pairing or direct contacts between the enzyme and the opposing base, have been proposed to play a role in dictating the activity dependence on the opposing base, the present study uses MD simulations to investigate the impact of the OG glycosidic conformation on the structure and dynamics of the hOgg1 and FPG active sites.

Representative MD structures of the hOgg1 active site with *anti*-OG bound are similar to the crystal structure starting point (Figure 3.2). Specifically, regardless of the protonation state of Asp268, Gly42 forms a strong hydrogen bond with N7–H of OG, and Gln315, Phe319 and an active site water form additional stabilizing interactions with the substrate, which collectively maintain the active site position of *anti*-OG. As a result, both the Lys249 nucleophile and Asp268 are properly positioned within vicinity of the lesion to initiate base excision. Since *anti*-OG that was originally paired with C is bound in the active site of hOgg1 in experimental crystal structures, the catalytically competent DNA–enzyme complex predicted by MD upon *anti*-OG binding is consistent with the experimentally-observed activity of hOgg1 toward OG paired with C.

In contrast to *anti*-OG, the hOgg1 active site hydrogen-bonding pattern changes considerably when *syn*-OG is bound regardless of the Asp268 protonation state (Figure 3.3). Specifically, Gly42 forms hydrogen bonds with N1–H and N2–H of *syn*-OG rather than N7–H of *anti*-OG. Furthermore, neutral Asp268 hydrogen bonds with O8 of OG, while anionic Asp268 is solvent exposed and coordinates with a crystallographic Ca^{2+}. In both cases, Asp268 is substantially displaced with respect to OG. As a consequence, Asp268 at least partially loses the capacity to form the αM helix cap, which likely destabilizes the protein as previously reported for the Asp268Glu and Asp268Gln mutants that similarly perturb the α-helix. Furthermore, although Lys249 retains a position conducive for nucleophilic attack when Asp268 is neutral, the active site position of Lys249 is completely disrupted for anionic Asp268. Since Asp268 has been shown to be a key catalytic residue based on mutation to Ala, Asn, Glu or Gln negatively influencing protein stability or catalytic activity, the predicted disruption of Asp268 likely at least in part contributes to the experimentally-observed up to 3000-fold slower excision rate of

OG paired with A than C. In fact, the present structural data highlights that Asp268 can play a key discriminatory role, which complements previous proposals for the function of this residue (*i.e.*, deprotonating the nucleophile, protonating O4′, stabilizing the Schiff base, or acting as the general base during the β-elimination reaction) and other hypotheses regarding how hOgg1 discriminates against an OG:A pair (*i.e.*, direct inspection of the complementary base and thermal instability of base pairs). Interestingly, although *C. ac.* Ogg has additional active site residues compared to hOgg1 that may influence the substrate binding orientation (namely Gln279 and Pro239), *C. ac*. Ogg contains Asp241 in the same position as Asp268 in hOgg1. By analogy to hOgg1, my data suggests Asp241 in *C. ac.* Ogg will likely be displaced upon *syn*-OG binding, and thereby provides a possible rationalization for the observed *anti*-OG binding orientation and base excision for *C. ac.* Ogg regardless of the (C or A) opposing base.

Unlike hOgg1, key reaction parameters, including the position of Pro2, Glu3, and active site water, are conducive for catalysis when either *anti*- or *syn*-OG is bound to FPG (Figures 3.3 and 3.4). However, the lesion–active site interactions vary with the OG glycosidic conformation, which may impede *syn*-OG excision. Specifically, *anti*-OG binding is primarily facilitated by stabilizing interactions between N2–H of the lesion and Glu6, and contacts with the recognition loop (notably Arg223). In contrast, these interactions are repulsive when *syn*-OG is bound to FPG. Instead, *syn*-OG binding is primarily facilitated through strong contacts with Glu78, as well as interactions with the recognition loop. This finding correlates with the proposal that a stable interaction between OG and the Lys217 sidechain of *E. coli* FPG (the homologue residue of Arg223 in *B. st.* FPG considered in this study) may be required for the excision, and may distinguish between OG and unmodified G. However, the present results emphasize that

the OG–Arg223 interaction only persists for the *anti* conformer, suggesting that Arg223 may be important for discriminating between OG glycosidic conformers, which complements the potential role of differentiating between G and OG. Furthermore, in addition to Arg223, my calculations suggest Glu3, Glu6, and Glu78 may help distinguish between *anti*- and *syn*-OG.

Most importantly, the present simulations show that FPG can accommodate both the *anti* and *syn* OG conformers, and both glycosidic orientations lead to a catalytically conducive active site configuration, suggesting that FPG can potentially excise OG regardless of the glycosidic orientation. Indeed, FPG benefits from a flexible active site architecture and makes extensive interaction with both OG conformers. Thus, the differentiation between the two OG orientations depends on subtle changes in the interaction between the lesion and active site residues. This contrasts the substantial alterations in the locations of key catalytic residues (Lys249 and Asp268) seen for hOgg1. Nevertheless, these observations are in direct correlation with the ability of FPG to excise a wider range of damaged DNA bases compared to hOgg1, which has been proposed to arise due to the flexibility of the FPG active site.

In addition to broadening our understanding of how different OG glycosidic conformers are accommodated in the hOgg1 and FPG active sites, my simulations provide structural information about the DNA–enzyme pre-reaction complexes. Although QM/MM mechanistic studies are required to fully understand the reaction mechanism, the present structural data sheds light on an existing controversy in the literature regarding the order of the first repair steps catalyzed by bifunctional DNA glycosylases.[9, 10, 21, 34-40] Specifically, two S_N2 mechanisms have been previously proposed (Figure 1.12), which differ in whether deglycosylation or sugar ring opening occurs upon nucleophilic attack

(Lys249 in hOgg1 or Pro2 in FPG). Regardless of the OG conformation, the protonation state of Asp268 (hOgg1) or Glu3 (FPG), or the enzyme considered, the ∠(Nζ/N–C1′–O4′) nucleophilic angle corresponding to ring opening is closer to the ideal 180° expected for an S_N2 reaction than the ∠(Nζ/N–C1′–N9) angle corresponding to deglycosylation in the DNA–enzyme reactant complex. Furthermore, my simulations provide direct evidence of active site water positioned between anionic Asp268/Glu3 and O4′ or direct hydrogen bonding between neutral Asp268/Glu3 and O4′, which would allow Asp268/Glu3 to facilitate an initial ring-opening step, and correlates with experimental data highlighting the critical nature of these residues for the function of hOgg1[5, 10-12] and FPG.[9, 13, 40, 41] Thus, although my simulations provide only a glimpse of the active site alignment in the pre-catalytic complex, my data adds to the mounting evidence that the most probable catalytic mechanism employed by bifunctional glycosylases involves initial nucleophilic attack in the Nuc→C1′→O4′ direction and ring opening as the first repair step. Importantly, the present data provides quality starting points for future QM/MM mechanistic studies that further explore the two proposed reaction mechanisms for both enzymes and different bound OG conformations.

3.5 Conclusion

Overall, this chapter predicts that hOgg1 and FPG employ different strategies to differentiate between OG glycosidic conformers, which rationalizes experimentally observed differences in their activities.[1-3] Specifically, *syn*-OG induces significant changes in the position of key catalytic residues in the hOgg1 active site, whereas subtle conformational changes occur upon binding *syn*-OG in FPG. Nevertheless, both hOgg1 and FPG are proposed to invoke a similar catalytic mechanism following OG binding to initiate base excision repair. Future work must continue to probe the activity of these

crucial repair enzymes by investigating the active site dynamics and catalytic excision mechanism for a diverse set of substrates, which will enhance our understanding of how human and bacterial cells remove major DNA oxidation products from the genome. With this goal in mind, the next chapter considers binding of FapyG, another major oxidative damage product of G, in the active sites of hOgg1 and FPG.

3.6 References

1. Leipold, M. D.; Workman, H.; Muller, J. G.; Burrows, C. J.; David, S. S., Recognition and Removal of Oxidized Guanines in Duplex DNA by the Base Excision Repair Enzymes hOGG1, yOGG1, and yOGG2. *Biochemistry* 2003, *42*, 11373–11381.

2. Krishnamurthy, N.; Haraguchi, K.; Greenberg, M. M.; David, S. S., Efficient Removal of Formamidopyrimidines by 8-Oxoguanine Glycosylases. *Biochemistry* 2008, *47*, 1043–1050.

3. Tchou, J.; Bodepudi, V.; Shibutani, S.; Antoshechkin, I.; Miller, J.; Grollman, A. P.; Johnson, F., Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA. *J. Biol. Chem.* 1994, *269*, 15318–15324.

4. Plum, G. E.; Grollman, A. P.; Johnson, F.; Breslauer, K. J., Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex. *Biochemistry* 1995, *34*, 16148–16160.

5. Bruner, S. D.; Norman, D. P. G.; Verdine, G. L., Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature (London, U. K.)* 2000, *403*, 859–866.

6. Faucher, F.; Wallace, S. S.; Doublié, S., Structural basis for the lack of opposite base specificity of Clostridium acetobutylicum 8-oxoguanine DNA glycosylase. *DNA Repair* 2009, *8*, 1283–1289.

7. Zharkov, D. O.; Rosenquist, T. A.; Gerchman, S. E.; Grollman, A. P., Substrate Specificity and Reaction Mechanism of Murine 8-Oxoguanine-DNA Glycosylase. *J. Biol. Chem.* 2000, *275*, 28607–28617.

8. Popov, A. V.; Endutkin, A. V.; Vorobjev, Y. N.; Zharkov, D. O., Molecular dynamics simulation of the opposite-base preference and interactions in the active site of formamidopyrimidine-DNA glycosylase. *BMC Struct. Biol.* 2017, *17*, 5.

9. Sadeghian, K.; Flaig, D.; Blank, I. D.; Schneider, S.; Strasser, R.; Stathis, D.; Winnacker, M.; Carell, T.; Ochsenfeld, C., Ribose-Protonated DNA Base Excision Repair: A Combined Theoretical and Experimental Study. *Angew. Chem., Int. Ed. Engl.* 2014, *53*, 10044–10048.

10. Sadeghian, K.; Ochsenfeld, C., Unraveling the Base Excision Repair Mechanism of Human DNA Glycosylase. *J. Am. Chem. Soc.* 2015, *137*, 9824–9831.

11. Dalhus, B.; Forsbring, M.; Helle, I. H.; Vik, E. S.; Forstrøm, R. J.; Backe, P. H.; Alseth, I.; Bjørås, M., Separation-of-Function Mutants Unravel the Dual-Reaction Mode of Human 8-Oxoguanine DNA Glycosylase. *Structure (Cambridge, MA, U. S.)* 2011, *19*, 117–127.

12. Norman, D. P. G.; Chung, S. J.; Verdine, G. L., Structural and Biochemical Exploration of a Critical Amino Acid in Human 8-Oxoguanine Glycosylase. *Biochemistry* 2003, *42*, 1564–1572.

13. Zharkov, D. O.; Shoham, G.; Grollman, A. P., Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair* 2003, *2*, 839–862.

14. Sugahara, M.; Mikawa, T.; Kumasaka, T.; Yamamoto, M.; Kato, R.; Fukuyama, K.; Inoue, Y.; Kuramitsu, S., Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, Thermus thermophilus HB8. *The EMBO Journal* 2000, *19*, 3857–3869.

15. Nash, H. M.; Bruner, S. D.; Scharer, O. D.; Kawate, T.; Addona, T. A.; Spooner, E.; Lane, W. S.; Verdine, G. L., Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr Biol* 1996, *6*, 968–980.

16. *The PyMOL Molecular Graphics System*, 1.3; Schrödinger, LLC.

17. Dennington, R. K., T.; Millam, J. *GaussView*, 5; Semichem Inc.: Shawnee Mission KS, 2009.

18. Fromme, J. C.; Verdine, G. L., DNA Lesion Recognition by the Bacterial Repair Enzyme MutM. *J. Biol. Chem.* 2003, *278*, 51543–51548.

19. Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M., et al., *Amber 12; University of California: San Francisco, CA, 2012*.

20. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* 2011, *7*, 525–537.

21. Kellie, J. L.; Wilson, K. A.; Wetmore, S. D., An ONIOM and MD Investigation of Possible Monofunctional Activity of Human 8-Oxoguanine–DNA Glycosylase (hOgg1). *J. Phys. Chem. B* 2015, *119*, 8013–8023.

22. Banerjee, A.; Yang, W.; Karplus, M.; Verdine, G. L., Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature (London, U. K.)* 2005, *434*, 612–618.

23. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* 2006, *65*, 712–725.

24. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation

Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules *J. Am. Chem. Soc.* 1996, *118*, 2309–2309.

25. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* 2004, *25*, 1157–1174.

26. Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz, K. M., Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF). *J. Chem. Theory Comput.* 2010, *6*, 2935–2947.

27. Radom, C. T.; Banerjee, A.; Verdine, G. L., Structural characterization of human 8-oxoguanine DNA glycosylase variants bearing active site mutations. *J. Biol. Chem.* 2007, *282*, 9182–9194.

28. Sidorenko, S. V.; Grollman, A. P.; Jaruga, P.; Dizdaroglu, M.; Zharkov, D. O. Substrate specificity and excision kinetics of natural polymorphic variants and phosphomimetic mutants of human 8-oxoguanine-DNA glycosylase. *FEBS J* 2009, *276*, 5149–5162.

29. Bjørås, M.; Seeberg, E.; Luna, L.; Pearl, L. H.; Barrett, T. E., Reciprocal "flipping" underlies substrate recognition and catalytic activation by the human 8-oxo-guanine DNA glycosylase11Edited by J. Karn. *J. Mol. Biol.* 2002, *317*, 171–177.

30. Robey-Bond, S. M.; Barrantes-Reynolds, R.; Bond, J. P.; Wallace, S. S.; Bandaru, V., Clostridium acetobutylicum 8-Oxoguanine DNA Glycosylase (Ogg) Differs from Eukaryotic Oggs with Respect to Opposite Base Discrimination. *Biochemistry* 2008, *47*, 7626–7636.

31. Zaika, E. I.; Perlow, R. A.; Matz, E.; Broyde, S.; Gilboa, R.; Grollman, A. P.; Zharkov, D. O., Substrate Discrimination by Formamidopyrimidine-DNA Glycosylase: A mutational analysis. *J. Biol. Chem.* 2004, *279*, 4849–4861.

32. Boiteux, S.; Gajewski, E.; Laval, J.; Dizdaroglu, M., Substrate specificity of the Escherichia coli Fpg protein formamidopyrimidine-DNA glycosylase: excision of purine lesions in DNA produced by ionizing radiation or photosensitization. *Biochemistry* 1992, *31*, 106–110.

33. David, S. S.; Williams, S. D., Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev. (Washington, DC, U. S.)* 1998, *98*, 1221–1261.

34. Schyman, P.; Danielsson, J.; Pinak, M.; Laaksonen, A., Theoretical Study of the Human DNA Repair Protein HOGG1 Activity. *J. Phys. Chem. A* 2005, *109*, 1713–1719.

35. Calvaresi, M.; Bottoni, A.; Garavelli, M., Computational Clues for a New Mechanism in the Glycosylase Activity of the Human DNA Repair Protein hOGG1. A

Generalized Paradigm for Purine-Repairing Systems? *J. Phys. Chem. B* 2007, *111*, 6557–6570.

36. Shim, E. J.; Przybylski, J. L.; Wetmore, S. D., Effects of Nucleophile, Oxidative Damage, and Nucleobase Orientation on the Glycosidic Bond Cleavage in Deoxyguanosine. *J. Phys. Chem. B* 2010, *114*, 2319–2326.

37. Sebera, J.; Trantírek, L.; Tanaka, Y.; Sychrovský, V., Pyramidalization of the Glycosidic Nitrogen Provides the Way for Efficient Cleavage of the N-Glycosidic Bond of 8-OxoG with the hOGG1 DNA Repair Protein. *J. Phys. Chem. B* 2012, *116*, 12535–12544.

38. Kellie, J. L.; Wetmore, S. D., Mechanistic and Conformational Flexibility of the Covalent Linkage Formed during β-Lyase Activity on an AP-Site: Application to hOgg1. *J. Phys. Chem. B* 2012, *116*, 10786–10797.

39. Sebera, J.; Trantirek, L.; Tanaka, Y.; Nencka, R.; Fukal, J.; Sychrovsky, V., The activation of N-glycosidic bond cleavage performed by base-excision repair enzyme hOGG1; theoretical study of the role of Lys 249 residue in activation of G, OxoG and FapyG. *RSC Advances* 2014, *4*, 44043–44051.

40. Blank, I. D.; Sadeghian, K.; Ochsenfeld, C., A Base-Independent Repair Mechanism for DNA Glycosylase—No Discrimination Within the Active Site. *Scientific Reports* 2015, *5*, 10369.

41. Lavrukhin, O. V.; Lloyd, R. S., Involvement of Phylogenetically Conserved Acidic Amino Acid Residues in Catalysis by an Oxidative DNA Damage Enzyme Formamidopyrimidine Glycosylase. *Biochemistry* 2000, *39*, 15266–15271.

Chapter 4: Recognition of FapyG Glycosidic Conformers

4.1 Introduction

FapyG (Figure 4.1) is known to be highly mutagenic in mammalian cells, inducing G:C to T:A mutations.[1] Similar to OG, *anti*-FapyG can pair with C, but FapyG can also form equally stable *anti*-FapyG:A and *syn*-FapyG:A pairs (Figure 4.1).[2] As was discussed in Chapter 1, both hOgg1 and FPG show lower discrimination between FapyG:C and FapyG:A;[3, 4] however, this reduced discrimination is not as profound as hOgg1 (Figure 1.9). In Chapter 3, several factors that potentially play a role in the observed differential excision rates for OG opposite different nucleobases, such as the different base pairing strength in DNA duplex,[4, 5] the opposing base,[6-8] and destabilization of DNA–enzyme interface in the pre-catalytic complex[9] were highlighted. Currently, there is no crystal structure of FapyG bound in the hOgg1 active site; however, available crystal structures show that G[10] or OG[6, 11, 12] initially paired with C in the DNA duplex bind in the hOgg1 active site in the *anti* conformation. This suggests that the conformation of the nucleobase is retained upon flipping into the active site pocket. Thus, the glycosidic conformation of the lesion may also affect the recognition and/or activity of the enzyme.

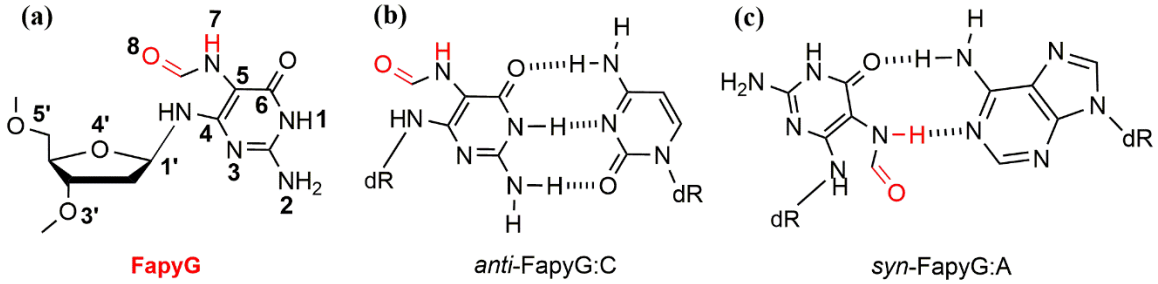

Figure 4.1. (a) 2D representative structure of 6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG). (b) *anti*-FapyG:C pair, and (c) *syn*-FapyG:A pair.

In comparison with OG, FapyG has been poorly studied. A previous MD study[13] on FapyG suggested that the *anti* conformer is recognized in *E. coli* FPG through interactions between FapyG, and Met73 and Lys217 (Met77 and Arg223 in *B. st.* FPG), which were not observed for *syn*-FapyG. Another MD study on *anti*-FapyG bound in *B. st.* FPG led to similar results.[14] The authors highlighted that replacing Glu78 with a serine (to represent the *E. coli* FPG sequence) leads to a more planar structure for FapyG as the Glu78–FapyG interaction vanishes.[14] Nevertheless, the potential effects of the FapyG glycosidic conformation on the reaction parameters required for enzymatic activity were not addressed.

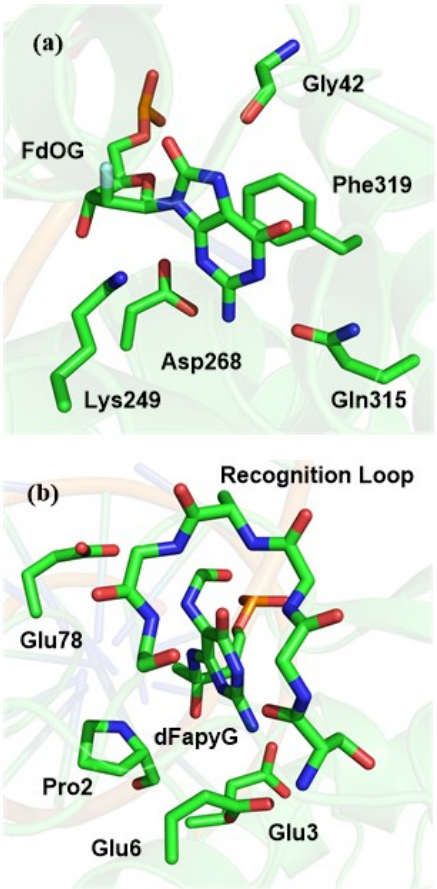

Figure 4.2. X-ray crystal structure of (a) 2′-fluoro-dOG (FdOG) bound in the hOgg1 active site (PDB ID: 3KTU) and (b) dFapyG bound in the FPG active site (PDB ID: 1XC8).

To complement Chapter 3, and further investigate the potential role of the glycosidic conformation on hOgg1 and FPG activity from a structural perspective, the current chapter employs MD simulations to examine the structure and dynamics of the hOgg1 and FPG active sites upon binding different FapyG conformers. Comparing the data from the present chapter with those for different OG glycosidic conformers (Chapter 3) sheds light on differences in the preferred binding modes of these two major oxidative lesions in the hOgg1 and FPG active sites, and provides a structural basis to rationalize the observed relative enzymatic activities (Figure 1.9) based on a lesion-recognition complex. As in Chapter 3, key nucleophilic distances and angles were monitored in order to further evaluate the probability of the proposed catalytic mechanism for the initial chemical step as discussed in Chapters 1 and 3.

4.2 Computational Details

The computational procedure used in this chapter is similar to that employed in Chapter 3. In the case of hOgg1, no crystal structure of the enzyme bound to FapyG is available; thus, a crystal structure of the enzyme bound to DNA containing OG fluorinated at C2′ (2′-fluoro-dOG, FdOG) was used as a starting point (PDB ID: 3KTU, Figure 4.2a).[6] This choice is reasonable since both OG and FapyG are guanine oxidation products, and both lesions are excised efficiently by hOgg1. Similarly, in the case of FPG, a crystal structure of the Glu3Gln mutant of *B. st.* FPG bound to OG-containing DNA was used (PDB ID: 1R2Y, Figure 4.2.b),[15] with Gln3 substituted by the native Glu. Although a crystal structure of (*anti*) carba-FapyG (4′-oxygen of FapyG replaced by a CH_2 group; PDB ID: 1XC8)[16] bound in the *Lactococcus lactis* (*L. lac.*) FPG active site is available, the 1R2Y starting structure was used to ensure an accurate basis for comparing

the dynamic behavior of active site–FapyG interactions to those for OG discussed in the previous chapter. The carba-FapyG taken from the crystal structure (1XC8) was converted to FapyG using GaussView. The obtained FapyG replaced the OG in both crystal structures of the enzymes with no steric clashes between the lesion and the active site residues. As done in Chapter 3, His270 in the hOgg1 active site was protonated since MD simulations suggest this protonation state better maintains the active site orientation of the crystal structure.[17] Additionally, nucleophilic residues (*i.e.,* Lys249 in hOgg1 and Pro2 in FPG) were modelled as neutral. Moreover, models were considered with anionic and neutral Asp268 or Glu3 in hOgg1 and FPG, respectively. The FapyG substrate was bound in the active site in both the *anti* and *syn* conformations for each active site protonation state, and C was maintained as the pairing base in the complementary strand.

For all models, AMBER parm99SB parameters[18] were used. The non-standard FapyG parameters were taken from previous work,[14] which was shown to reproduce the nonplanar structure observed in the crystal structure (PDB ID: 1XC8).[16] The same protocol for the pre-production phase, the unconstrained MD, and trajectory analysis was used as described in Chapter 3 using the cpptraj module of Amber 12.[19] Most importantly, 20 ns unconstrained MD simulations under NPT conditions (1 atm and 300 K) were conducted as pre-production simulations. The representative structures from each trajectory were chosen for subsequent pre-production simulations with different initial velocities to improve the phase space sampling. Once the iteration led to active site configuration without any significant difference, two 20 ns MD simulations were conducted with different initial velocities. The obtained structures from the replicas for each system were highly similar (Figures A4.1 and A4.2). Thus, for each system a further 80 ns MD simulation was performed on one trajectory. Although a single representative

structure is shown for the cluster with the highest occupancy (see Table A4.1 for the corresponding rmsd and occupancy), the geometrical and energetic analyses were completed over all structures of the highest occupied cluster and the corresponding dynamical information (standard deviations) is provided in the Appendix A4.

4.3 Results

4.3.1 hOgg1

4.3.1.1 anti-FapyG

Regardless of the Asp268 protonation state, the representative MD structures of *anti*-FapyG bound in the hOgg1 active site and the starting crystal structure are overall very similar (Figure 4.3, average active site backbone rmsd < 0.7 Å, with rmsd of clusters and per residue rmsd provided in Tables A4.1 and A4.2, respectively).

FapyG maintains an *anti* conformation (average $\chi \sim -110°$, Table A3). Additionally, the ∠(C4–C5–C8–O8) dihedral angle is ~ 44 or 35° with anionic and neutral Asp268, respectively, while the ∠(C5–N7–C8–O8) torsional angle assumes larger average value (~ 13°) with anionic Asp268 than for neutral Asp268 (~ 3°, Table A4.4). The hydrogen-bonding patterns between active site residues and FapyG are similar for anionic and neutral Asp268, with slight differences (Table A4.5). Specifically, the hydrogen-bond occupancy between N7–H of FapyG and the Gly42 backbone (Figure 4.4) is 22 or 57% for anionic or neutral Asp268, respectively (Table A4.5), with an average FapyG–Gly42 interaction energy of ~ –2.0 or –2.8 kcal/mol (Table A4.6). In both systems, N1–H and N2–H of FapyG interact extensively with Gln315 (74 – 86%, and average stabilization energy of ~ –12.5 kcal/mol). Another noticeable lesion–active site interaction is a hydrogen bond between Oδ of Asp268 and N2–H of FapyG (74 or 85% occupancy), with an overall per residue interaction energy of –10.5 or –13.7 kcal/mol for

anionic or neutral Asp268, respectively. Importantly, anionic Asp268 preserves the crystal structure position relative to 2′-deoxyribose. Phe319 retains the crystal structure stacked position (3.779 Å) with respect to FapyG, although it is tilted and located slightly farther away from the nucleobase for anionic Asp268 (average of 4.028 and 4.009 Å for anionic and neutral Asp268, respectively, Figure 4.3 and Table A4.7). Regardless, Phe319 has a small overall contribution to substrate binding (~ –4.0 kcal/mol, Table A4.6).

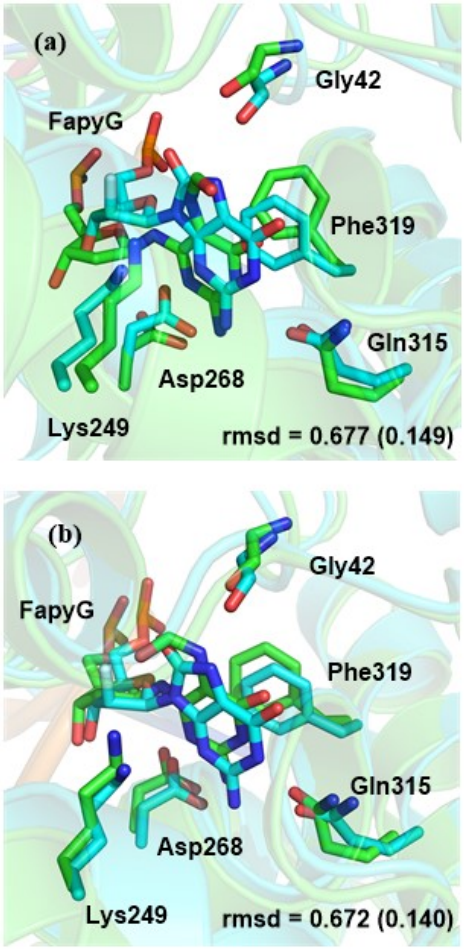

Figure 4.3. Overlay of the crystal structure (PDB ID: 3KTU; cyan) and MD representative structure (green) of *anti*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. Average rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided. For clarity, hydrogen atoms are not shown.

Plots of the solvent distribution (Figure A4.7) show that there is a large water density above (on the 5′-side) and below (3′-side) the 2′-deoxyribose of the damaged nucleotide. The average number of solvent molecules in close proximity of the lesion (Table A4.8) reflects the solvent accessibility of the hOgg1 active site when *anti*-FapyG is bound. Indeed, for both Asp268 protonation states, O6 of FapyG hydrogen bonds with solvent for ~ 60% of the simulation (stabilization energy ~ –3.5 kcal/mol, Table A4.6). Another important solvent–lesion interaction is the FapyG(O4′)···H_2O contact, which exists for 27 or 31% of the simulation for anionic or neutral Asp268, respectively (Table A4.5). Furthermore, for neutral Asp268, a FapyG(O4′)···Asp268(Oδ–H) hydrogen bond is maintained for 39% of the simulation time (Figure A4.6 and Table A4.5).

The d(Nζ···C1′) nucleophilic distance is ~ 4.2 Å regardless of the Asp268 protonation state (Table A4.7). The ∠(Nζ–C1′–N9) attack angle is ~ 65° (Table A4.9), while the average ∠(Nζ–C1′–O4′) angle is ~ 157° (Table A4.9). Furthermore, the proximity of Asp268 to O4′ of 2′-deoxyribose (with d(Oδ···O4′) < 4.123 Å, Table A4.7) indicates the possibility of a proton transfer to O4′ during the first chemical step.

4.3.1.2 syn-FapyG

The hOgg1 active site configuration when *syn*-FapyG is bound is comparable (Figures 4.6 and A4.2) to the crystal structure or MD data for *anti*-FapyG bound to hOgg1 (average active site backbone rmsd < 0.8 Å, with respect to the crystal structure with per residue rmsd provided in Table A4.2). The average χ dihedral angle for *syn*-FapyG bound in the hOgg1 active site is 64.7° or 88.6° with anionic or neutral Asp268, respectively (Table A4.3). The average ∠(C4–C5–C8–O8) angle falls between ~ –50 and –60°, and ∠(C5–N7–C8–O8) is ~ 4 or 5° (Table A4.4). Comparing the values of these

angles for *anti*- and *syn*-FapyG indicates the flexibility of the opened-imidazole ring, which can adopt an orientation with favorable interactions (Table A4.4). Some hydrogen-bonding interactions inevitably change with the glycosidic conformer of the bound nucleotide (Figure A4.9). The main lesion–active site interactions include the Gly42 backbone hydrogen bonding with the N1 and N2 hydrogens of *syn*-FapyG (~ 30 – 75%, Figure A4.9 and Table A4.5), which leads to a more stabilizing DNA–enzyme interaction (by ~ –3 kcal/mol, Table A4.6) compared to that for *anti*-FapyG. Moreover, the *syn*-FapyG(N7–H)···(Oε)Gln315 and *syn*-FapyG(O6)···H–Nε(Gln315) interactions (87 – 69%) are comparable to *anti*-FapyG(N1/2–H)···Oε(Gln315) occupancies (Table A4.5). The interactions between *syn*-FapyG and Gln315 stabilizes the binding by ~ –11.5 kcal/mol regardless of the Asp268 protonation state. Unlike reported for *syn*-OG (Chapter 3), the Asp268 sidechain is not displaced either due to repulsion (anionic Asp) or hydrogen bonding (neutral Asp) with O8 of FapyG. In fact, the flexibility of the opened-imidazole ring of FapyG permits N7–C8=O to avoid repulsive interactions with Asp268 (Figure A4.9). In fact, with both anionic and neutral Asp268, the O8 of FapyG forms a hydrogen bond (with low occupancy) with the Asp268 backbone (for clarity it is not shown in the figure). Attractive interactions between anionic Asp268 and N9–H, or neutral Asp268 and N9–H, O3′, and O4′ collectively stabilize the binding with an average interaction energy –5.0 or –5.6 kcal/mol for anionic and neutral Asp, respectively (Table A4.6). Thus, regardless of the protonation state, Asp268 maintains its crystal structure position relative to the 2′-deoxyribose ring (d(Oδ···O4′) ~ 3.6 or 4.8 Å, Table A4.7), with O3′···H–Oδ(Asp268) and O4′···H–Oδ(Asp268) hydrogen bonds being formed for 80 and 34% of the simulation time, respectively, in the case of neutral Asp268. Finally, the average *syn*-FapyG···Phe319 distance (~ 4.1 Å, Table A4.7), and the associated stacking

interaction energy (up to ~ –4 kcal/mol, Table A4.6) are similar to that for *anti*-FapyG for both Asp268 protonation states.

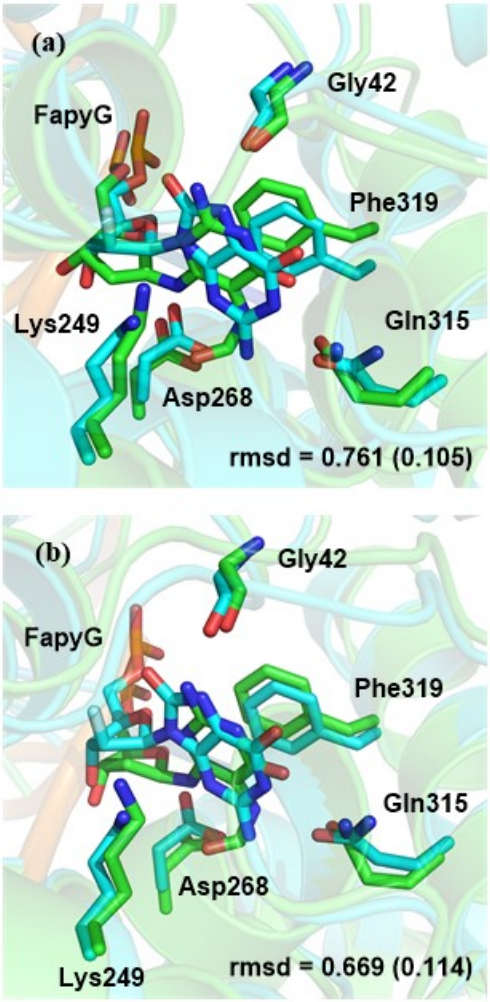

Figure 4.4. Overlay of the crystal structure (PDB ID: 3KTU; cyan) and MD representative structure (green) of *syn*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. Average rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

Despite the significant solvent density around the 2′-deoxyribose of *syn*-FapyG (Figure A4.10), solvent molecules mainly occupy the larger (6.0 Å) shell (Table A4.8) and FapyG–solvent hydrogen-bonding occupancies are reduced compared to the corresponding interactions for *anti*-FapyG. Specifically, O6 of *syn*-FapyG hydrogen bonds to water for ~ 15 – 18% of the simulation (Table A4.5), with an associated

contribution to binding of ~ –1.8 or –1.1 kcal/mol with anionic or neutral Asp268, respectively (Table A4.6). However, O4′ of *syn*-FapyG when bound in the acitve site with anionic Asp268, forms a hydrogen bond with solvent for 37% of the simulation. This hydrogen bond is maintained only 8% of the trajectory with neutral Asp268. Regardless of the Asp268 protonation state, the O4′ of *syn*-FapyG forms a hydrogen bond (with water or neutral Asp268) which may increase the probability of ring-opening reaction involved in the first chemical step in comparison with *syn*-OG (Chapter 3).

The d(Nζ···C1′) nucleophilic distance is 4.958 or 4.419 Å for *syn*-FapyG bound to hOgg1 containing anionic or neutral Asp268, respectively, which is slightly larger than for *anti*-FapyG (Table A4.7). The ∠(Nζ–C1′–N9) angle is ~ 70° regardless of the Asp268 protonation state. Interestingly, with anionic Asp268, the ∠(Nζ–C1′–O4′) angle is significantly smaller (109.1°, Table A4.9) than for neutral Asp268 (146.6°) or for *anti*-FapyG (~ 157°) regardless of the Asp268 protonation state. Finally, as was mentioned above, both anionic and neutral Asp268 residues remain in close proximity of the sugar ring of *syn*-FapyG, which makes the potential proton transfer to O4′ plausible (Table A4.7).

4.3.2 FPG

4.3.2.1 anti-FapyG

The overall structure of *anti*-FapyG bound in the FPG active site with anionic or neutral Glu3 resembles the crystal structure (Figures 4.5 and A4.11), with slight conformational changes in the sidechains (Figure 4.5, average active site backbone rmsd < 0.8 Å, with per residue rmsd provided in Table A4.10). FapyG adopts the *high anti* conformation (average χ = –114.6 and –99.4° with anionic or neutral Glu3, respectively;

Table A4.3) and extensively interacts with the FPG recognition loop (Ser220, Val222, Arg223, Thr224, and Tyr225, Figure A4.12 and Table A4.11). Despite the difference in the χ values from the crystal structure (–64.3°, 1XC8) and those obtained from the MD simulations and the crystal structure (Table A4.3), the average χ values for FapyG are very similar to those for OG when bound in the FPG active site. The average ∠(C4–C5–C8–O8) angle is ~ 118 or –102° for anionic and neutral Glu3, respectively (Table A4.4). This difference indicates the opened-imidazole ring can rotate to form more favorable interactions with the loop residues. Specifically, in the case of anionic Glu3, O8 of FapyG forms hydrogen bonds with the Thr224 sidechain, and N7–H of FapyG interacts with solvent, while, for neutral Glu3, O8 hydrogen bonds with Arg223, and FapyG(N7–H) hydrogen bonds with Thr224(Oγ) and Glu78(Oε) (Table A4.11 and Figure 4.9). The average ∠(C5–N7–C8–O8) dihedral angle is ~ –3 or 5° with anionic or neutral Glu3, respectively (Table A4.4).

The major hydrogen bonds between the loop and the FapyG lesion are similar to those previously observed for *anti*-OG (Chapter 3), and involve Ser220, Val222, Arg223, Thr224, and Tyr225 (Figure A4.12). However, the flexibility of the opened-imidazole ring allows FapyG to avoid repulsive contacts and form attractive interactions by rotating around the N7–C8 bond. For example, the *anti*-FapyG···Glu78 interaction is attractive due to (albeit low-occupancy, 12 – 18%) hydrogen bonding between N7–H and the Glu78 sidechain (average energy ~ –1 kcal/mol). This is unlike the case of OG where the interaction between Glu78 and *anti*-OG is repulsive (~ 10 kcal/mol, Chapter 3). Overall, the interaction energies between the nucleobase and the recognition loop are consistent when anionic or neutral Glu3 are considered (Table A4.12). Outside the loop, interactions between FapyG (N2 hydrogens) and the Glu6 sidechain are the strongest, with an average

strength of ~ –16 kcal/mol for FPG with anionic or neutral Glu3, respectively (Table A4.12). The most noticeable dependence of the active site interactions on the Glu3 protonation state for *anti*-FapyG is the ability of neutral Glu3 to hydrogen bond with O3′ and O4′ of FapyG (60% and 26%, respectively, Table A4.11), which provides greater stability to the active site (by ~ –8 kcal/mol, Table A4.12).

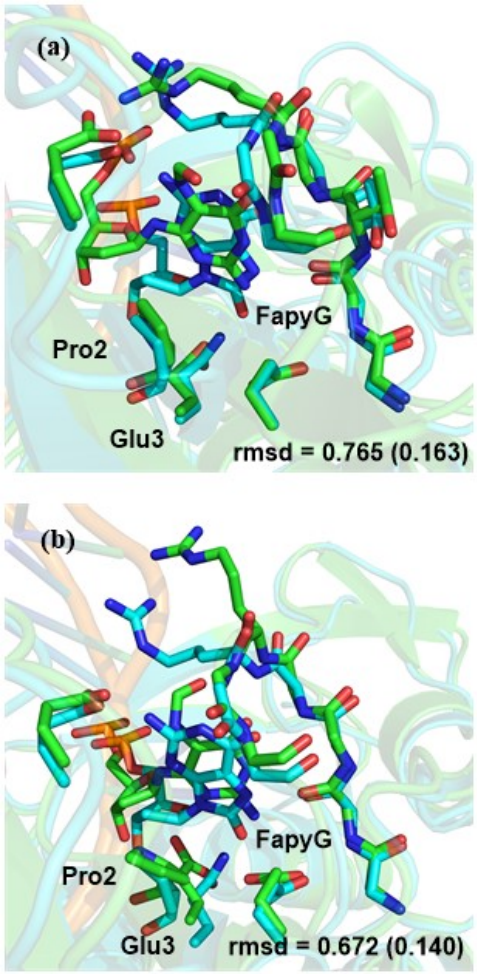

Figure 4.5. Overlay of the crystal structure (PDB ID: 1R2Y, cyan) and MD representative structure (green) of *anti*-FapyG bound in the FPG active site with (a) anionic or (b) neutral Glu3. Average rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure (1R2Y) are provided.

In the FPG active site, the solvent density is larger on the 5′-side than the 3′-side of the FapyG sugar plane (Figures A4.14). The average number of solvent molecules in the vicinity of Glu3 and the sugar ring is larger for anionic than neutral Glu3 in both

solvent shells (Table A4.13). In fact, O4′ of the sugar and the anionic Glu3 sidechain form hydrogen bonds with solvent for 39 and 64% of the simulation time, respectively (Figure A4.12 and Table A4.11). On the other hand, neutral Glu3 directly hydrogen bonds to O4′ of 2′-deoxyribose (26% occupancy, Figure A4.13 and Table A4.11).

The d(N···C1′) nucleophilic distance is 4.221 or 4.326 Å for anionic or neutral Glu3, respectively (Table A4.7). Regardless of the Glu3 protonation state, the ∠(N–C1′–N9) and ∠(N–C1′–O4′) attack angles are ~ 75° and ~ 158°, respectively (Table A4.9). Moreover, the distance between neutral Glu3 (Oε) and O4′ is shorter (~ 3.8 Å) compared to that for anionic Glu3 (~ 4.8 Å, Table A4.7).

4.3.2.2 syn-FapyG

The MD representative structure of the FPG active site bound to *syn*-FapyG with either anionic or neutral Glu3 (Figure A4.15) resembles the crystal structures (Figure 4.6, average active site backbone rmsd < 0.9 Å, with per residue rmsd provided in Table A4.10). A *syn* orientation is maintained throughout the simulation regardless of the Glu3 protonation state (average χ ~ 70° and 94°, Table A4.3) due to extensive interactions with the FPG active site (Figure A4.16). Moreover, regardless of the Glu3 protonation state, the ∠(C4–C5–C8–O8) average angle is ~ –48°, which locates FapyG(O8) away from Ser220(O), and thus prevents electrostatic repulsion (Figure A4.15), while allowing Ser220(O) and Glu6(N–H) to form hydrogen bonds with FapyG at N7–H and O8, respectively (Figure 4.15).

As observed for *anti*-FapyG, the average ∠(C5–N7–C8–O8) angle are similar for both Glu3 protonation states (~ 3°, Table A4.4). The major hydrogen-bonding interactions occur between O6, N1–H, and N2–H of FapyG and the active site residues in

the recognition loop, namely Thr224(Oγ), Arg223(N–H), and Tyr225(N–H) (Table A4.11). Beyond the loop, the strongest lesion–active site interactions take place between N2 hydrogens of *syn*-FapyG and the Glu78 sidechain (–17.2 and –19.0 kcal/mol for FPG with anionic or neutral Glu3, respectively, Table A4.12). *syn*-FapyG (N7–H/N9–H or O8) forms attractive interactions with the Glu6 sidechain (Figures A4.16 and A4.17), although the average stabilization energy for *syn*-FapyG is small (~ –2 kcal/mol, Table A4.12) compared to *anti*-FapyG. The Glu3 sidechain maintains the crystal structure position relative to O4′ for *syn*-FapyG bound in the FPG active site (Figures 4.6 and A4.15), with the Glu3(Oε)···(O4′)FapyG distance equal to or less than those for *anti*-FapyG (Table A4.7). Similar to *anti*-FapyG, the interaction between Glu3 and *syn*-FapyG stabilizes the binding. Specifically, the average interaction energy between the lesion and anionic Glu3 is –4 kcal/mol. Neutral Glu3 forms hydrogen bonds with O4′ for 31% of the simulation (Figure A4.7), which together with Glu3(Oε–H)···(O3′)FapyG hydrogen bond stabilize the binding by ~ –5 kcal/mol (Table A4.11).

For the *syn*-FapyG bound in the FPG active site with anionic Glu3, the active site solvent is mainly distributed on the 5′-side of the FapyG sugar plane (Figure A4.18). Although there are on average 1 or 2 water molecules close to the Glu3 sidechain and O4′, respectively, these molecules occupy the larger solvent shell (Table A4.13). This leads to a low O4′···H–O(H_2O) hydrogen-bond occupancy (*i.e.,* 10%, ~ 4 times less than for *anti*-FapyG). Similarly, for neutral Glu3, water density in the vicinity of anionic Glu3 and the sugar ring is low (Figure A4.18); however, neutral Glu3 forms a hydrogen bond with O4′ (31% occupancy, Table A4.11).

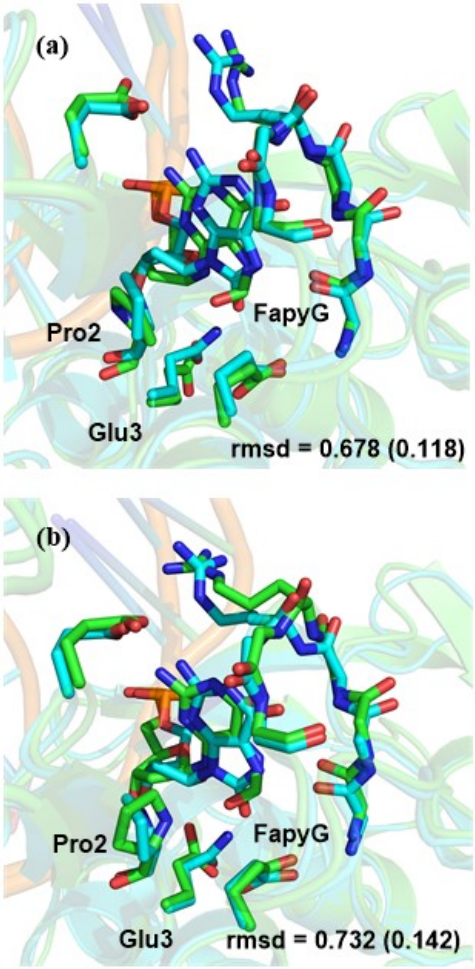

Figure 4.6. Overlay of the crystal structure (PDB ID: 1R2Y, cyan) and MD representative structure (green) of *syn*-FapyG bound in the FPG active site with (a) anionic or (b) neutral Glu3. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure (1R2Y) are provided.

Regardless of the Glu3 protonation state, the average d(N···C1′) nucleophilic distance is slightly larger for *syn*-FapyG (~ 4.8 Å) relative to *anti*-FapyG (~ 4.3 Å, Table A4.7). The ∠(N–C1′–N9) angle is ~ 69 or 86° and the ∠(N–C1′–O4′) angle is ~ 150° (Table A4.9) for the active site with anionic and neutral Glu3. Both d(Oε···O4′) distances are larger for anionic than neutral Glu3 (Table A4.7), presumably due to the OG(O4′)···Glu3(Oε–H) and OG(O3′)···Glu3(Oε–H) hydrogen bonds for neutral Glu3 that are absent for anionic Glu3.

4.4 Discussion

To complement MD simulations investigating the binding of OG in the hOgg1 and FPG active sites (Chapter 3), the present chapter examined the dynamic behaviour of the hOgg1 and FPG active site when bound to different FapyG glycosidic conformers. hOgg1 has been shown to remove OG and FapyG when paired opposite C at similar rates;[3] however, while hOgg1 remains relatively inert towards the OG:A pair,[3] the discrimination between FapyG:C and FapyG:A is significantly reduced.[3] In contrast, although FPG removes FapyG and OG more efficiently opposite C, FPG also shows appreciable activity toward OG:A and FapyG:A.[3, 4] In addition to factors that affect the lesion recognition step (*i.e.*, factors in the DNA duplex or prior to full insertion in the active site),[4-9] the lesion glycosidic conformation may affect the activity of the enzyme toward certain damage products by forming different interactions, inducting active site configurational changes, or destabilizing the DNA–enzyme complex (active site pocket). In this study, MD simulations have provided a structural basis for the previously observed activity of both enzymes.

Overlays of the MD representative structure of the hOgg1 active site with *anti*-FapyG bound and the crystal of *anti*-OG bound are similar (Figure 4.3). The main difference arises due to the flexibility of the opened-imidazole ring of FapyG, which reduces the hydrogen-bond occupancy between N7–H and Gly42. In fact, the decrease in this interaction occupancy is compensated by the interaction with solvent (Figure A4.6). Although the Phe319 sidechain is tilted with respect to the crystal structure position in the representative structure with anionic Asp268 (Figure 4.3), this does not significantly affect the average interaction energy or the position of the lesion relative to other catalytic

residues. Thus, *anti*-FapyG is well accommodated and stabilized in the hOgg1 active site, and the conducive positions of catalytic residues (*i.e.,* Lys249 and Asp268) necessary to initiate the catalysis are maintained. Overall, *anti*-FapyG binding resembles *anti*-OG binding previously reported for and is consistent with experimental data suggesting similar OG and FapyG excision rates.[3]

The MD representative structures for *syn*-FapyG bound in the hOgg1 active site reveal inevitable differences in the lesion–active site hydrogen-bonding patterns compared to *anti*-FapyG (Figure 4.4). Specifically, the FapyG···Gly42 and FapyG···Gln315 interactions alter with the FapyG glycosidic conformation. However, unlike what was reported for *syn*-OG (Chapter 3), rotation about the glycosidic bond to yield *syn*-FapyG does not significantly alter the positions of hOgg1 active site residues relative to the bound lesion. In fact, active site residues form considerable stabilizing interactions with *syn*-FapyG and preserve the reaction parameters required to facilitate the catalysis. This finding is consistent with the appreciable activity towards FapyG paired opposite A.

Overall, regardless of the Asp268 protonation state, the hydrogen-bonding interactions and per residue substrate–active site stabilization energies are comparable for *anti*- and *syn*-FapyG in the hOgg1 active site. Moreover, the nucleophilic distance and angle of attack are comparable and conducive for the catalytic reaction according to an initial ring-opening mechanism.[20-23] Thus, the similar excision rates of OG and FapyG opposite C by hOgg1 could be related to the overall similarities between the *anti* conformers of the bound OG and FapyG in the active site. Therefore, the observed large difference between hOgg1 catalytic rate for OG:C/OG:A (3000-fold difference)[3] and FapyG:C/FapyG:A (47-fold difference)[3] might be related to the different effects of the

syn-OG and *syn*-FapyG conformers on the overall configuration of active site residues in the pre-catalytic complex.

Comparable to OG bound in the FPG active site (Chapter 3), both *anti*- and *syn*-FapyG extensively interact with the FPG active site residues. Nevertheless, some specific lesion–active site interactions change with the lesion conformation. Indeed, the flexibility of the opened-imidazole ring enables *anti*-FapyG to form more hydrogen bonds with the FPG recognition loop. For example, O8 and/or N7–H of *anti*-FapyG interacts with the Thr224 backbone and sidechain, and O8 with the Arg223 backbone. Moreover, N9 of FapyG hydrogen bonds with the active site solvent. In contrast, the major interactions between *syn*-FapyG and the active site occur between the N1 and N2 hydrogens, and O6 and the FPG recognition loop. The main interactions at N7 and O8 occur with the Ser220 or Glu3 backbone, respectively. Thus, in comparison with OG, both the *anti* and *syn*-FapyG conformations form stabilizing interactions with the active site residues. Specifically, although the interaction between *anti*-OG(O8) and Glu78 is repulsive (Chapter 3), *anti*-FapyG (at N7–H and C8–H) forms attractive interactions with Glu78 (Table A4.12). Moreover, the *syn*-OG and *syn*-FapyG interactions with Glu3 and Glu6 are different (*i.e.*, OG(O8) and the Glu3 or Glu6 sidechain repel each other and destabilize the DNA–enzyme complex, while FapyG(O8) forms attractive interactions with the Glu3 and Glu6 backbones and sidechains; Figure 4.12). Finally, although Arg223 has similar stabilizing effects on the FapyG glycosidic conformers, this residue only stabilizes *anti*-OG (Chapter 3). Regardless, as reported for OG, neither FapyG glycosidic conformer alter the catalytic parameters (*i.e.,* nucleophilic distance and angle) or the position of lesion-stabilizing residues with respect to the bound substrate. Thus, the simulated structures suggest the FPG active site alignment is conducive for the initial

chemical step for both FapyG glycosidic conformers, which is in agreement with experimental data indicating the reduced effect of the opposing base for FapyG over OG.[3, 4]

The present structural data on the FapyG glycosidic conformers bound in the hOgg1 and FPG active sites show that the ∠(Nζ/N–C1′–O4′) angle that corresponds to the nucleophilic attack leading to ring opening is closer to 180°, the expected angle for an S_N2 reaction. This is in agreement with previous data for OG (Chapter 3) bound in the active site and data reported from QM/MM calculations[22] (136.5° or 157.6° in the reactant or transition state complex for hOgg1, respectively). The increase of the supporting data for the ring-opening reaction over the deglycosylation as the first chemical step for hOgg1 and FPG suggest that this mechanism may be applied by other bifunctional glycosylases or at least hOgg1 and FPG homologue glycosylases such as Nth and Nei.

4.5 Conclusion

Overall, MD simulation results presented in this chapter reveal that the effects of the FapyG glycosidic conformation on the active site of both hOgg1 and FPG are subtle, and the key catalytic parameters are generally highly conserved upon changing the damaged nucleotide conformation. The present data suggests that the level of discrimination against the FapyG glycosidic conformers when bound in the active site is reduced compared to OG for both enzymes, which is in direct agreement with experimental data suggesting similar excision rates for OG and FapyG.[3] Notably, this chapter provides the first evidence that the hOgg1 active site is conducive for the initial chemical step when either *anti*- or *syn*-FapyG is bound. In addition to other factors that have been previously identified to affect the hOgg1 recognition and removal process of

OG and FapyG opposite C or A in the duplex, this chapter has shown that the glycosidic conformation plays an important role in the processing of OG in the active site, but this factor is less crucial for FapyG than OG and compared to FPG, providing strong structural basis for the different hOgg1 enzymatic excision rates for OG and FapyG paired with C or A.

4.6 References

1. M. A. Kalam, K. Haraguchi, S. Chandani, E. L. Loechler, M. Moriya, M. M. Greenberg and A. K. Basu, *Nucleic Acids Research*, 2006, 34, 2305–2315.

2. N. R. Jena and P. C. Mishra, *ChemPhysChem*, 2013, 14, 3263–3270.

3. N. Krishnamurthy, K. Haraguchi, M. M. Greenberg and S. S. David, *Biochemistry*, 2008, 47, 1043–1050.

4. J. Tchou, V. Bodepudi, S. Shibutani, I. Antoshechkin, J. Miller, A. P. Grollman and F. Johnson, *J. Biol. Chem.*, 1994, 269, 15318–15324.

5. G. E. Plum, A. P. Grollman, F. Johnson and K. J. Breslauer, *Biochemistry*, 1995, 34, 16148–16160.

6. S. D. Bruner, D. P. G. Norman and G. L. Verdine, *Nature (London, U. K.)*, 2000, 403, 859–866.

7. F. Faucher, S. S. Wallace and S. Doublié, *DNA Repair*, 2009, 8, 1283–1289.

8. D. O. Zharkov, T. A. Rosenquist, S. E. Gerchman and A. P. Grollman, *J. Biol. Chem.*, 2000, 275, 28607–28617.

9. A. V. Popov, A. V. Endutkin, Y. N. Vorobjev and D. O. Zharkov, *BMC Struct. Biol.*, 2017, 17, 5.

10. C. M. Crenshaw, K. Nam, K. Oo, P. S. Kutchukian, B. R. Bowman, M. Karplus and G. L. Verdine, *J. Biol. Chem.*, 2012, 287, 24916–24928.

11. A. Banerjee, W. Yang, M. Karplus and G. L. Verdine, *Nature (London, U. K.)*, 2005, 434, 612–618.

12. C. T. Radom, A. Banerjee and G. L. Verdine, *J. Biol. Chem.*, 2007, 282, 9182–9194.

13. R. A. Perlow-Poehnelt, D. O. Zharkov, A. P. Grollman and S. Broyde, *Biochemistry*, 2004, 43, 16092–16105.

14. K. Song, V. Hornak, C. de los Santos, A. P. Grollman and C. Simmerling, *J. Comput. Chem.*, 2008, 29, 17–23.

15. J. C. Fromme and G. L. Verdine, *J. Biol. Chem.*, 2003, 278, 51543–51548.

16. F. Coste, M. Ober, T. Carell, S. Boiteux, C. Zelwer and B. Castaing, *J. Biol. Chem.*, 2004, 279, 44074–44083.

17. J. L. Kellie, K. A. Wilson and S. D. Wetmore, *J. Phys. Chem. B*, 2015, 119, 8013–8023.

18. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, 65, 712–725.

19. D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz. et al., *Amber 12; University of California: San Francisco, CA, 2012*.

20. K. Sadeghian, D. Flaig, I. D. Blank, S. Schneider, R. Strasser, D. Stathis, M. Winnacker, T. Carell and C. Ochsenfeld, *Angew. Chem., Int. Ed. Engl.*, 2014, 53, 10044–10048.

21. I. D. Blank, K. Sadeghian and C. Ochsenfeld, *Scientific Reports*, 2015, 5, 10369.

22. K. Sadeghian and C. Ochsenfeld, *J. Am. Chem. Soc.*, 2015, 137, 9824–9831.

23. D. O. Zharkov, G. Shoham and A. P. Grollman, *DNA Repair*, 2003, 2, 839–862.

Chapter 5: Deglycosylation and β-elimination Reactions in FPG*

5.1 Introduction

The unique features of FPG compared to hOgg1 have attracted attention towards this bacterial enzyme. Despite an abundance of experimental studies on different aspects of the BER process facilitated by FPG,[1-6] and structural information gained from available DNA–enzyme crystal structures,[6-8] several mechanistic features of the reactions catalyzed by FPG remain unclear and require further detailed investigation. The first step in the generally accepted mechanism utilized in the bifunctional activity of hOgg1 and FPG (Figure 1.11) involves nucleophilic attack (by lysine in hOgg1 and proline in FPG), which displaces OG. This is followed by the ring-opening step, which is initiated by (direct or general base–assisted) proton transfer from the amine nucleophile to O4′ in deoxyribose, leading to a Schiff base intermediate. Although recent studies suggest that ring-opening may occur prior to deglycosylation,[9-10] ring-opening occurs after deglycosylation in the most broadly accepted mechanism.[7, 11-12] Next, β-lyase activity occurs through proton abstraction from C2′ by a general base, resulting in 3′-phosphate elimination. The final step recovers the enzyme through hydrolysis of the Schiff base and leaves the nicked DNA strand for further processing.

Previous computational studies of the chemical step catalyzed by hOgg1 have primarily focused on the deglycosylation phase,[13-18] with only select studies considering the ring-opening step.[13, 18] Specifically, Kellie and Wetmore have mapped several

---

* Reproduced in part with permission from Sowlati-Hashjin S. and Wetmore S. D. Computational Investigation of Glycosylase and β−Lyase Activity Facilitated by Proline: Applications to FPG and Comparisons to hOgg1 *J. Phys. Chem. B* **2014** (118) 14566−14577. Copyright 2014 American Chemical Society.

possible pathways for the hOgg1 bifunctional mode using the smallest chemically-relevant model (a modified (*anti*) nucleoside-3′-monophosphate and a truncated lysine nucleophile).[18] In fact, this was the first computational study to consider the lyase reaction facilitated by a bifunctional glycosylase. Interestingly, this work also supports the proposed monofunctional activity of hOgg1 by identifying the low energy intermediate as the DNA–protein crosslink that occurs before ring-opening, but after proton transfer from the amine, which would lead to monofunctional activity if hydrolyzed. Unfortunately, much less is known about the chemical steps employed in the mechanism of action of FPG compared to hOgg1. To date, most experimental studies on FPG have focused on the structure of the DNA–enzyme crosslink and the lesion recognition step,[6-8, 19] with only a few investigations proposing mechanisms for the β-lyase[7-8, 20] and δ-lyase steps.[20] Similarly, quantum mechanical studies on FPG are limited to the deglycosylation reaction step.[9-10, 21]

To further understand the mechanism employed by bifunctional glycosylases, and FPG in particular, this chapter investigates potential deglycosylation and β-elimination mechanisms employed by FPG. To allow the greatest number of pathways to be considered, the smallest chemically-relevant model is employed, which includes a modified nucleoside-3′-monophosphate and truncated proline nucleophile. Although this model will not fully recover all details of the chemical reactions occurring in the FPG active site, this approach will reveal the intrinsically preferred mechanisms of the main phases in the overall process. This will identify pathways that must be considered in future large-scale modeling of the enzymatic reaction, where investigating all mechanistic options is computationally prohibited. Furthermore, the current model is equivalent to the model employed to investigate the bifunctional activity of hOgg1 (lysine as the

nucleophile).[18] Therefore, this will permit accurate comparison between the two enzymes in order to identify any differences in the inherent chemistry originating due to the nucleophile.

5.2 Computational Details

5.2.1 Model

Although many crystal structures are available for various FPG mutants and modified substrates, only one high-resolution structure is available with the enzyme bound to DNA containing an intact 8-oxoguanine nucleoside (PDB ID: 1R2Y).[11] Therefore, the starting models were based on this crystal structure. Although the N-terminal formylmethionine is cleaved post-translationally in *E. coli*,[22] the N-terminus is not cleaved in some mutant FPG enzymes,[23-24] which leads to the proline nucleophile being designated as residue 2. This chapter adopts this convention to match the crystal structure numbering.

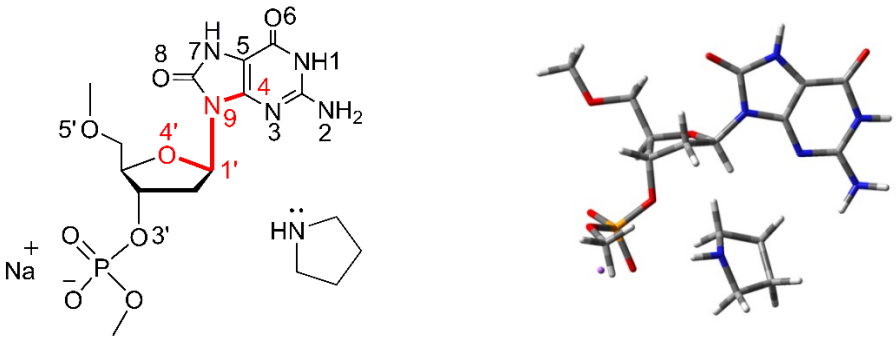

Figure 5.1. 2D (left) and 3D (right) representations of the computational model employed in this chapter, which includes 3′-monophosphate dOG (capped with methyl group at O5′- and 3′-phosphate) and proline (modeled as pyrrolidine). A sodium counterion is included to charge neutralize the phosphate moiety. The χ dihedral angle (∠(O4′–C1′–N9–C4)) is shown in bold.

The initial geometry of *syn* OG and the proline nucleophile (Pro2) was obtained from the crystal structure 1R2Y.[11] This choice for OG ensures that structural changes to the DNA backbone inflicted when the glycosylase flips the base out of the DNA helix are

taken into account. The *anti* conformer was generated by rotating OG about the glycosidic bond and followed by a geometry optimization. In order to investigate the β-elimination step catalyzed by FPG, a modified nucleoside-3′-monophosphate is used (Figure 5.1) in which the 3′- and 5′-terminal oxygen atoms are capped by methyl (Me) groups rather than hydrogens to avoid artificial hydrogen bonds between a terminal hydrogen and the nucleobase, sugar or nucleophile that does not occur in DNA. Neglect of the C5′-phosphate is justified by a previous study showing that, although geometries may change slightly, the calculated energy barriers remain similar upon inclusion of this group.[25]

To neutralize the model, a sodium ion was added equidistant to the oxygen atoms in the 3′-phosphate group, which has been shown to result in better geometries than protonated or anionic models.[26-27] It is timely to note that the effect of replacing the sodium ion with a potassium ion was considered for the elimination step since the metal ion likely plays the most important role in this step of the overall reaction. However, only slight changes to the calculated geometries were observed (< 0.18 Å change in key bond lengths and ~ 3.5° change in key angles; Figure A5.1). As a result, the calculated energetics of this step were unaltered by changing the counterion, and a sodium ion was used to model the remaining reaction steps. The proline nucleophile is modelled as pyrrolidine (Figure 5.1). This model choice is consistent with the previous study that used a (truncated) lysine nucleophile.

5.2.2 Mechanisms

Although there is evidence that OG is not tightly bound in the FPG active site after excision,[7] OG has been proposed to be involved in subsequent (C2′–H abstraction)

reaction steps.[12, 28] Therefore, to bypass existing confusion in the literature regarding the identity of key catalytic residues, the excised OG plays the role of the general base and acid as required throughout the reaction pathway in the present work. Although other residues in the FPG active site likely fulfill these roles during the enzyme-catalyzed reaction, this choice keeps the model small and consistent throughout the reaction and allows considering all possible reaction pathways. Nevertheless, this function requires inevitable motion of OG at different mechanistic steps, which may deviate from accessible OG orientations in the FPG active site. Most importantly, the model implemented does not affect the main conclusions regarding the necessity of an active site base and acid to catalyze the various chemical steps. As a result, the small model provides important geometrical information that can be used in future large-scale modeling of the DNA–protein complex, which will accurately identify the roles of active site residues.

The pathways considered in the current study for the first three mechanistic steps in the bifunctional activity of FPG are schematically shown in Figure 5.2. The first step (a) is dOG deglycosylation, producing a covalently bound DNA–protein crosslink and the OG anion. This step is considered for both the *syn* and *anti* conformations of dOG. In the second step (b), the abasic intermediate undergoes a Schiff base rearrangement through proton transfer from N of the original proline nucleophile to O4′ of deoxyribose, which leads to a ring-opened intermediate. This step can be accompanied by either direct proton transfer or proton transfer assisted by a general base and acid. In the direct pathway, the proton is directly transferred from N of proline to O4′ of the sugar, causing the ring to open. This pathway is considered for both the *anti* and *syn* conformations of dOG. In the assisted pathway, a base (OG in the present model) first abstracts the proton from N, and subsequently delivers the proton to O4′. Proton abstraction is considered to be performed

by O8 or N9 of OG in two separate pathways. The final step (c) is 3′-phosphate elimination through β-lyase activity. This phase begins with C2′–H abstraction, which is performed by O8 or N9 of OG, and requires translational motion of OG from the proximity of O4′ to C2′. This step is followed by the final 3′-phosphate elimination.

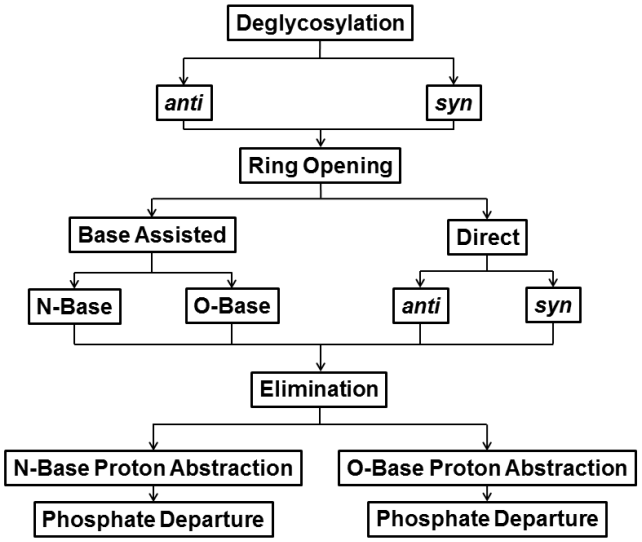

Figure 5.2. Reaction pathways considered in the present chapter including a) deglycosylation initiated from either the *anti* or *syn* conformation of OG, b) ring-opening with either direct (from the *anti* or *syn* conformer) or base-assisted proton transfer from N to O4′ (by N9 or O8 of the excised OG), and c) elimination coupled with proton abstraction by either N9 or O8 of the excised OG followed by 3′-phosphate departure.

The nomenclature employed in previous work involving a lysine nucleophile (hOgg1)[18] is used in the present work to facilitate meaningful comparisons between the studies. Specifically, consistent with Scheme 1, the structures corresponding to the deglycosylation step end in (a), while stationary points referring to the ring-opening and elimination steps end in (b) and (c), respectively. For example, the transition state for the deglycosylation step is labelled TS(a), while the associated intermediate is denoted IC(a). When a mechanistic step involves more than one intermediate, a prime is used to distinguish between the structures. For example, the carbanion intermediate formed in the third reaction phase is denoted IC(c').

5.2.3 Methodology

Reaction potential energy surfaces (PESs) were searched for each mechanistic pathway (Figure 5.2) by optimizing the corresponding transition states (TS). Subsequently, the intrinsic reaction coordinate (IRC) was carefully followed as far as possible and the resulting structures were fully optimized to generate relaxed reactant, intermediate and product complexes. Within this model, the product of each chemical step may not exactly correspond to the reactant for the subsequent reaction. This is mainly because of inevitable translational motion of OG due to its multiple roles along the reaction pathway. Unfortunately, however, efforts to characterize transitions between the intermediates were not successful. Therefore, structures obtained from reverse IRCs were considered to be the intermediate that links the previous reaction step. Nevertheless, deviations in the structures primarily correspond to motion of the excised nucleobase, necessary for this residue to act as general base and acid, and therefore the major conclusions of this chapter will not be affected. Furthermore, the deglycosylation products are energetically and structurally the same as the reactants of the direct ring-opening reaction, with the energy difference falling within chemical accuracy (4 kJ/mol) for both the *anti* and *syn* OG conformations.

All stationary points were fully optimized at the B3LYP/6-31G(d) level of theory in the presence of bulk solvent as described by the IEF-PCM method implemented in Gaussian 09.[29] Previous work shows that structural and energetic changes upon addition of diffuse functions are negligible for a similar reaction.[18] Frequency calculations were performed at the same level of theory to discern the nature of stationary points, and obtain zero-point vibrational (ZPVE) and Gibbs free energy corrections. All single-point

calculations were completed using SMD-M06-2X/6-311+G(2df,2p), which has been shown to perform well for S_N2 and elimination reactions.[30] To investigate the effects of basis set expansion on the calculated structures and energetics, the lowest energy pathway was re-calculated with B3LYP/6-31G(d,p) (Table A5.1 and Figure A5.2). The basis set expansion generally leads to a relatively small change in the barriers height, with the largest effect occurring in the deglycosylation transition state (~ 14 kJ/mol or less than 10%). Additionally, important geometrical features do not deviate significantly between the two basis sets, with the largest difference in bond length being 0.027 Å and the largest difference in bond or dihedral angle being 0.7°.

A wide range of dielectric responses have been predicted for different proteins by molecular simulation methods, which range from $\varepsilon \approx 3$ in the center of globular proteins to a gradual increase to 10 at protein boundaries.[31-32] Based on the shape of FPG and the location of the active site, the dielectric constant of diethyl ether (4.24) was used to estimate the effects of the surrounding protein in all calculations. Although this estimation may not reflect the exact reaction environment, this approach is more accurate than gas-phase calculations. Unless otherwise mentioned, all reported energies correspond to (unscaled) Gibbs free energies obtained from the single-point calculations. All calculations were carried out using Gaussian 09 (Revision C.01).[29]

5.3 Results

The relative SMD-M06-2X/6-311+G(2df,2p) Gibbs free energies of the multiple pathways considered in the present study for the dOG glycosylase and β-lyase activity of FPG (Figure 5.2) are provided in Table 5.1 and the corresponding Gibbs reaction surfaces are shown in Figure 5.3. Each step in the overall mechanism is separately considered directly below, while a comparison to hOgg1 and the biological implications will be

addressed in the Discussion section. Due to the neglect of discrete active site residues in the computational model, the calculated energy barriers reported in the present study will vary from experimental barriers, with the magnitude of the effect potentially varying for each reaction step. Nevertheless, this model recovers several main geometrical features throughout the reaction as determined by comparison to experimental crystal structures. Therefore, the model implemented will fulfil the primary goal of obtaining information about the conformational flexibility of a large number of pathways, which can be used in future large-scale modeling of the entire DNA–protein complex.

5.3.1 Deglycosylation

The present chapter considers the deglycosylation reaction for both starting dOG conformations (Figure 5.4). Interestingly, there is good agreement between the calculated *syn* reactant complex and crystal structures corresponding to dOG (PDB ID: 1R2Y)[7] or an abasic site (PDB ID: 3TWM)[33] bound in the enzyme active site (Figure A5.3), which lends confidence to the computational approach.

Both calculated reactants (RC, Figure 5.4) possess C3′-*exo* sugar puckering and the model proline is positioned on the opposite side of the sugar with respect to the nucleobase and O5'. The reaction proceeds via an S_N2 mechanism, while a parallel S_N1-type reaction could not be identified. The transition states for this step (TS(a), Figure 5.4) have a (breaking) glycosidic bond length of 2.333 or 2.305 Å and a C1′···N distance of 2.399 or 2.444 Å for the *anti* or *syn* conformer, respectively. Therefore, the TS occurs slightly earlier for the *syn* conformer. Nevertheless, the ∠(N···C1′···N9) angle is similar in the two TSs, being 151.2° for the *anti* and 149.3° for the *syn* conformer. The structural differences lead to energetic differences (Table 5.1), where the deglycosylation barrier for the *anti* conformer (148.6 kJ/mol) is slightly larger than that for *syn* dOG (140.2 kJ/mol).

Furthermore, the *syn* reactant is 4.0 kJ/mol higher in energy than the *anti* reactant, and the *syn* intermediate for the deglycosylation reaction (IC(a), Figure 5.4) is 11.5 kJ/mol more stable.

Table 5.1. Relative Gibbs free energies (kJ/mol) for stationary points characterized along the deglycosylation and β-elimination pathways characterized in the present chapter.[a,b]

	Deglycosylation			
	anti		***syn***	
RC	0		4	
TS(a)	148.6		140.2	
	Ring-Opening			
	Direct		**Assisted**	
	anti	***syn***	**O-base**	**N-base**
IC(a)	36.9	25.4	–44.9	–11.5
TS(b)	180.8	178.7	–55.2	–13.7
IC(b')		–	11.7	–58.9
TS(b')		–	35.6	39.9
	Elimination			
	O-base Proton Abstraction		**N-base Proton Abstraction**	
IC(b)	51.4		38.6	
TS(c)	72.9		74.1	
	3′-PO_4 Departure		**3′-PO_4 Departure**	
IC(c')	47.6		5.1	
TS(c')	128.3		100.5	
IC(c)	94.7		65.4	

[a] Energies reported relative to a common (*anti*) reactant complex (Figure 5.3). [b] SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) values including unscaled correction to Gibbs energy.

It has been suggested that enzymatic deglycosylation of purine nucleosides can be acid-catalyzed.[34] Therefore, in order to investigate the effect of nucleobase activation prior to deglycosylation, OG was protonated at available donor sites, namely N3, O6 or O8.

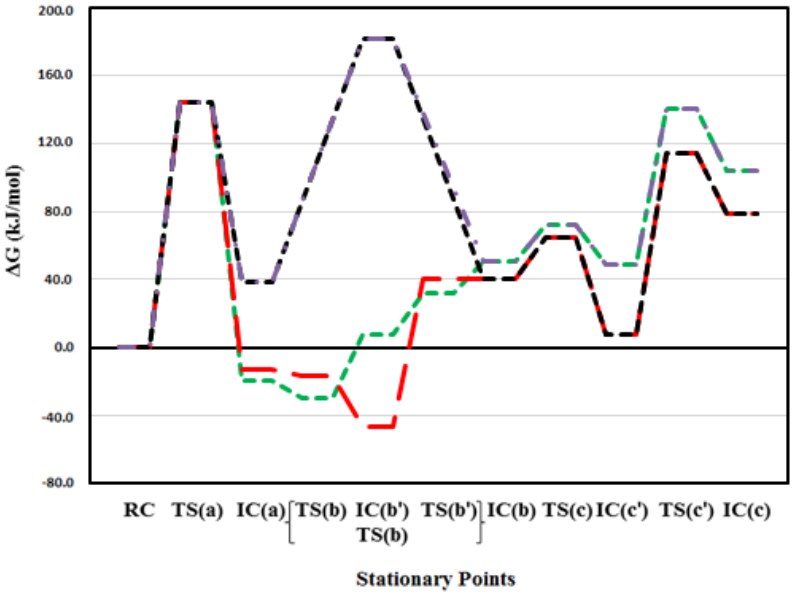

Figure 5.3. SMD-M06-2X/6-311+G(2df,2p)//PCM-B3LYP/6-31G(d) relative Gibbs energies (kJ/mol) for the deglycosylation and β-lyase pathways considered with respect to the *anti* dOG reactant, including N-base assisted ring-opening followed by N-base 3'-phosphate elimination (red, long dashed), O-base assisted ring-opening followed by 3'-phosphate elimination facilitated by O8 of OG (green, short dashed), direct ring-opening followed by O-base 3'-phosphate elimination (purple, dot and dashed), and direct ring-opening followed by N-base 3'-phoshphate elimination (black, short dashed).

Protonation of the nucleobase at N3, O6 or O8 lowers the deglycosylation barrier of the *anti* conformer to 104.6, 102.1 or 69.8 kJ/mol, respectively. Moreover, the corresponding intermediate of this step falls at –27.6, –42.2 or –73.0 kJ/mol, and these intermediates are more stable than the corresponding neutral forms (Table 5.1). The barrier for the *syn* conformer protonated at O6 or O8 (100.3 and 59.0 kJ/mol, respectively) is even smaller. However, efforts to locate the deglycosylation TS associated with protonated N3 of the *syn* conformer were not successful. Regardless, in agreement with previous studies,[34] activation of the nucleobase through protonation prior to deglycosylation is one way the enzyme can significantly catalyze the reaction. However, since the excised OG must play the role of the general base and acid in several

subsequent reaction steps in the present work, neutral OG is implemented throughout to characterize the remainder of the mechanistic pathway.

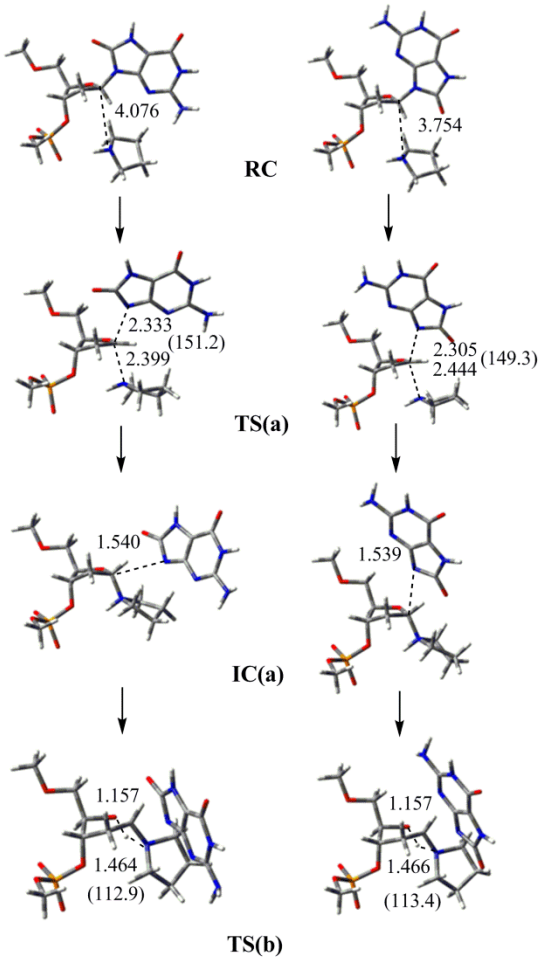

Figure 5.4. Structures along the deglycosylation (RC, TS(a) and IC(a)) and direct ring-opening (TS(b)) steps initiated from the *anti* (left) and *syn* (right) dOG conformers. Important distances (Å) and angles (deg., in parentheses) obtained at the PCM-B3LYP/6-31G(d) level of theory are provided. The sodium counterion is omitted for clarity.

5.3.2 Ring-Opening

5.3.2.1 Direct Proton Transfer

The ring-opening pathway can occur via direct proton transfer from N of the model proline to O4′ of deoxyribose, which can be initiated from the *anti* or *syn* conformer. The intermediates corresponding to the direct pathway (IC(a)) are directly connected to the deglycosylation transition states (Figure 5.4). Regardless of the starting

dOG conformation, this reaction requires significant bending of the proline linkage with respect to the sugar moiety (IC(a) → TS(b), Figure 5.4). As a result, the geometrical parameters for the reaction core of the associated transition states are nearly identical for both (*anti*/*syn*) dOG conformers. Specifically, the distance between the proton and O4′ is the same (1.157 Å), and d(N···H) deviates by only 0.002 Å for the *anti* and *syn* pathways. TS(b) falls 180.8 or 178.7 kJ/mol above the overall reactants and $\Delta G^{\ddagger}$ is 143.9 or 153.3 kJ/mol for the *anti* or *syn* orientation, respectively (Table 5.1 and Figure 5.3). These significant barriers suggest that there may be an alternative mechanism for the ring-opening step that yields a lower barrier.

5.3.2.2 Base-Assisted Proton Transfer

Previous work on the dOG glycosylase and β-lyase reaction catalyzed by hOgg1 indicates that a base-assisted proton transfer mechanism can lower the ring-opening barrier.[18] In this pathway (Figure 5.5), a two-step mechanism is considered in which OG abstracts the proton from N in the first step and the proton is delivered from OG to O4′ in the second step (Figure 5.2). Consistent with the previous study on hOgg1 bifunctional activity,[18] either O8 or N9 of the excised OG can act as the general base, and the corresponding pathways are denoted as O-base and N-base, respectively (Figure 5.5). However, since proline has only one proton at N, either OG must move from the O5′ to the O3′ side of the sugar or the proline linkage must flip to direct the proton towards OG in order for proton abstraction to occur. When OG moves to the O3′ side of the sugar, only O8 can capture the proton from the linkage since steric constraints prevent close contact between N9 of OG and the N–H of proline. Alternatively, when the proline flips, a TS could not be identified for proton abstraction by O8.

In addition to the intermediates (IC(a), Figure 5.4) associated with direct proton transfer, the deglycosylation transition states (TS(a), Figure 5.4) are connected to two intermediates corresponding to the base-assisted ring-opening pathway (IC(a), Figure 5.5). The intermediates along the assisted pathways are considerably more stable than those for the direct mechanism (–44.9 kJ/mol for O-base and –11.5 kJ/mol for N-base assisted (Figure 5.5) compared to 36.9 and 25.4 kJ/mol for *anti* and *syn* direct (Figure 5.4) ring-opening pathways, respectively). From these intermediates, TS(b) for the assisted mechanism has different geometrical features for the two OG base pathways. Specifically, the C1′···H distance is 1.356 Å for O-base, but only 1.171 Å for N-base. On the other hand, the N9···H distance (1.444 Å) is larger than O8···H (1.168 Å). The proton transfer from N to OG is barrierless with the given model and the present level of theory. In the case of the N-base pathway, the IC(b') product is the most stable intermediate over the entire reaction pathway (Figure 5.5). However, this structure is artificially stabilized through interactions between the excised OG and the sodium ion complexed to the C3′-phosphate, which would not occur in the enzymatic environment. The corresponding O-base intermediate lies ~12 kJ/mol above the reactants.

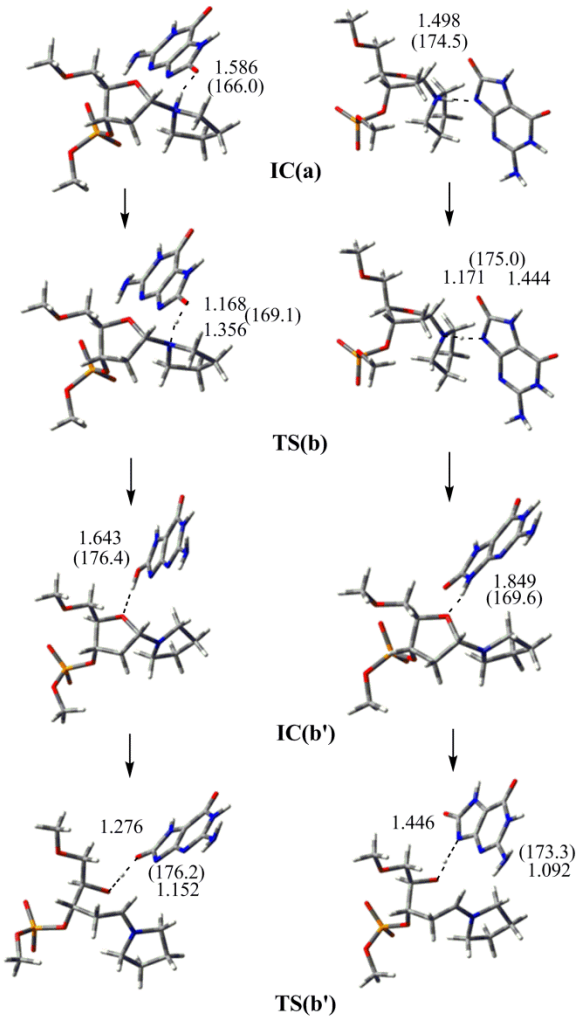

Figure 5.5. Structures along the assisted ring-opening step (IC(a), TS(b), IC(b') and TS(b')) with protonation facilitated by O8 (O-base, left) and N9 (N-base, right) of dOG. Important distances (Å) and angles (deg, in parentheses) obtained at the PCM-B3LYP/6-31G(d) level of theory are provided. The sodium counterion is omitted for clarity

The second step in the assisted ring-opening pathway is delivery of the proton captured by OG to O4′. In the TSs (TS(b'), Figure 5.5), d(N9···H) (1.446 Å) is longer than d(O8···H) (1.276 Å), and d(O4′···H) is shorter for the N-base (1.092 Å) than the O-base (1.152 Å) pathway, which follows the trend for the proton abstraction step. In the N-base pathway, there is a weak (2.415 Å) C1′–H···N3 hydrogen-bonding interaction, while there is a slightly stronger C1′–H···N9 interaction (2.228 Å) in the O-base pathway. Interestingly, N9 abstracts (∠(N···H···N9)) and delivers (∠(O4′···H···N9)) the proton at

a similar angle (deviation less than 2°). On the other hand, the abstraction and the delivery angles for O8 differ by ~ 7°. Although ΔG of the proton delivery is similar for the O-base and N-base mechanisms (35.6 and 39.9 kJ/mol respectively), the stability of IC(b') for the N-base pathway leads to a considerable $\Delta G^{\ddagger}$ (98.8 kJ/mol) compared to O-base pathway ($\Delta G^{\ddagger}$ = 23.9 kJ/mol, Table 5.1). Since the imine linkage produced in this reaction step is flexible,[18] the IC(b) intermediates (Figure 5.6) will be taken as the products of both the direct and assisted pathways in order to consider the next (β-elimination) reaction step.

Overlays of the calculated ring-opened Schiff base intermediate (IC(b) for the O-base or N-base pathway, Figure 5.5) and crystal structures corresponding to a borohydride-trapped abasic site complex (PDB ID: 1L1Z)[7] or a Schiff base intermediate (PDB ID: 1K82)[8] are provided in Figure A5.4. Differences in the calculated structures and 1K82 mainly arise in the ∠(Cδ–N–C1′) angle, which is smaller in the constrained crystal (100.1°) than fully optimized structures (123.3° and 126.1° for O-base and N-base intermediates, respectively). Moreover, the calculated d(N–C1′) bond length is shorter (1.286 and 1.290 Å for O-base and N-base intermediates, respectively) than that in the crystal structure (1.392 Å), which is indicative of greater sp^2 character of the Schiff base in the fully optimized geometry. On the other hand, the sp^3 hybridization of C1′ in the reduced abasic site intermediate (PDB ID: 1L1Z) leads to larger deviations compared to the calculated Schiff base intermediate (Figure A5.4). Indeed, the ∠(N–C1′–C2′) angle suggests the experimental structure is nearly tetrahedral at C1′ (111.6°) compared to the sp^2 character in the optimized structures (125.0° and 124.3° for O-base and N-base intermediates, respectively). The observed N–C1′ bond length (1.474 Å) further supports the sp^3 nature of C1' in 1L1Z. Despite these deviations, the present small model fits

within crystal structures that mimic different stages of the reaction and thereby recovers important structural features of the crosslink that will be useful in large-scale DNA–protein modeling.

5.3.3 β-Elimination

The third phase of the overall process facilitated by FPG is β-elimination (Figure 5.2), which consists of proton abstraction from C2′ followed by O3′ elimination. Similar to the proton abstraction in the assisted ring-opening step, the β-elimination step can be facilitated by either O8 or N9 of the excised OG, denoted as the O-base and N-base pathway, respectively. In this step, either the *pro*-R or *pro*-S proton may be abstracted; however, since *pro*-R proton abstraction requires overcoming a larger barrier due to necessary torsion in the C1′–C2′–C3′ angle,[18, 20] only *pro*-S proton abstraction is considered in this study.

The reaction occurs via an E1cB mechanism within the constraints of the model implemented. The E1cB elimination proceeds through a carbanion intermediate (IC(c'), Figure 5.6) since N of the proline–sugar linkage stabilizes the negative charge developing on C2′ upon proton abstraction. This partial stabilization likely prevents the E2 mechanism. Despite the fact that the potential carbocation intermediate formed along an E1 pathway may be partially stabilized by the imine linkage, the E1 elimination reaction is improbable due to the absence of a strong acid and good leaving group.

In the transition states corresponding to proton abstraction (TS(c), Figure 5.6), d(H···N9) (1.421 Å) is longer than d(H···O8) (1.254 Å), which is similar to the proton abstraction in the assisted ring-opening step (TS(b), Figure 5.5). However, d(C2′···H) is shorter for N-base than O-base, which is opposite to the situation for TS(b). $\Delta G^{\ddagger}$ for TS(c) is 15.0 kJ/mol larger for N-base than O-base with respect to the corresponding reactants

IC(b). However, since the intermediate IC(b) is more stable for N-base, the corresponding $\Delta G^{\ddagger}$ is 14.0 kJ/mol larger.

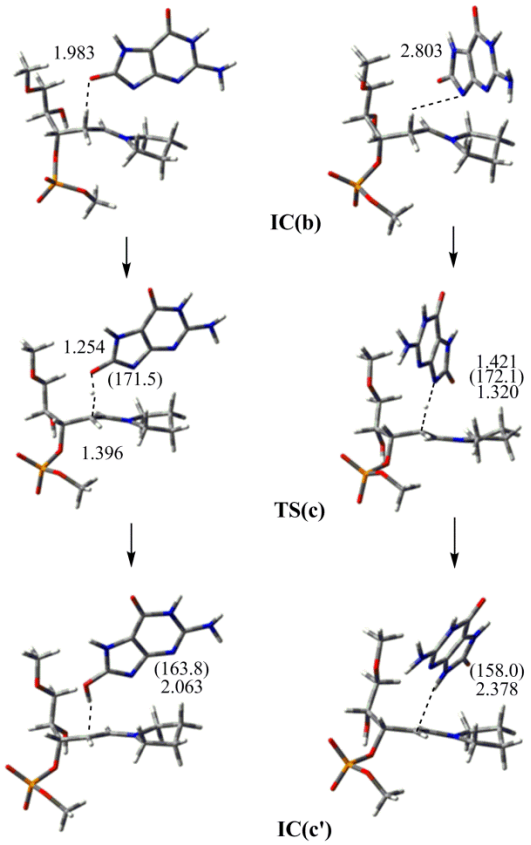

Figure 5.6. Structures along the proton abstraction step prior to β-elimination (IC(b), TS(c) and IC(c')) facilitated by O8 (O-base, left) and N9 (N-base, right) of dOG. Important distances (Å) and angles (deg, in parentheses) obtained at PCM-B3LYP/6-31G(d) level of theory are provided. Sodium counterion is omitted for clarity.

Proton abstraction is followed by O3′ elimination. In the transition states (TS(c'), Figure 5.7), the d(C3′···O3′) distance is 2.529 Å or 2.576 Å for the O-base or N-base pathway, respectively. In the final elimination products (IC(c), Figure 5.7), d(C2′···H) is 2.503 Å to O8 and 2.464 Å to N9, which are longer and shorter than in the corresponding TSs, respectively. Hydrogen-bonding interactions between the oxygen of the excised phosphate group and O4′–H stabilize the elimination products. However, the mobility of

the phosphate group towards O4'–H arises from the truncated model used in the present work, and the final products of the enzymatic reaction may be significantly different.

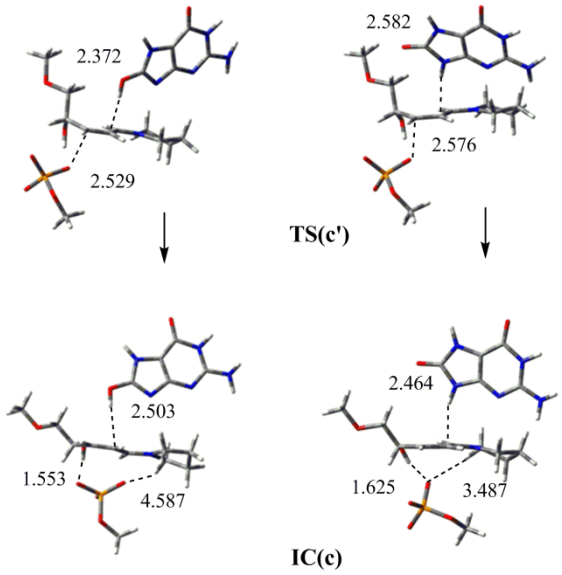

Figure 5.7. Structures along the elimination step (TS(c') and IC(c)) associated with the O-base (left) and N-base (right) pathways. Important distances (Å) and angles (deg, in parentheses) obtained at the PCM-B3LYP/6-31G(d) level of theory are provided. The sodium counterion is omitted for clarity.

5.4 Discussion

Although an abundance of experimental and computational work has shed light on many aspects of the mechanism of action of hOgg1, only a few computational studies have considered the chemical steps involved in the bifunctional activity of FPG.[9-10, 21] Furthermore, little experimental evidence is available for the details of each phase of the FPG reaction. Since hOgg1 and FPG differ in the nucleophile,[20, 35] active site residues,[20, 35] lesion specificity,[20, 35-37] possible activity modes,[38-40] and efficiency,[5] extension of previous computational studies on hOgg1 to FPG may not be possible. Therefore, the present work completed the first computational study of the chemical steps associated with the glycosylase and β-lyase activity of FPG (Figure 5.2). By using the same approaches as implemented in a previous study on hOgg1,[18] an accurate comparison

between the intrinsic chemistry facilitated by FPG and hOgg1 is obtained that provides information about the preferred reaction pathways and the associated energetic barriers, as well as potential roles for active sites residues. This work thereby provides clues about the influence of the nucleophile on the relative efficiencies and specificities of these enzymes. To this end, each phase of the FPG catalyzed reaction will be separately discussed below, while drawing important comparisons to hOgg1.

5.4.1 Deglycosylation

Excision of the damaged nucleobase is the first step catalyzed by FPG during the BER pathway. Experimental data supports different mechanisms for this step.[7, 11] Specifically, Pro2 has been proposed to be perfectly positioned for (S_N2) attack on C1′,[7] while another study indicates that Pro2 lies close to C1′, but is not well aligned for an S_N2 reaction.[11] In agreement with a previous computational study on the FPG[21] and hOgg1[18] glycosylation activity, the model implemented in the current work only resulted in an S_N2 deglycosylation pathway. Interestingly, the deglycosylation transition state occurs later when proline acts as the nucleophile rather than lysine,[18] where the d(N···C1′) distances in the transition state for *anti* dOG deglycosylation are similar (only 0.018 Å longer for proline), while the d(C1′···N9) distance is (0.108 Å) shorter for proline.

Previous studies suggest that OG can adopt either the *anti* or *syn* conformation in the FPG recognition pocket, with the *anti* conformer being more favorable.[41-42] Further support for this statement is provided in Chapter 3. Therefore, deglycosylation initiated from either OG conformer was considered. The *anti*-OG transition state occurs later than the *syn* variant, with d(N···C1′) being 0.065 Å longer and d(C1′···N9) being 0.028 Å shorter for *anti*-OG. Furthermore, although the *syn* conformer lies 4.0 kJ/mol above the *anti* conformer, the *syn* deglycosylation barrier is 8.4 kJ/mol lower in energy than for the

anti orientation. This suggests that with the given model and level of theory, *syn*-OG is the intrinsically most favored conformer for excision. Nevertheless, interactions between OG and the enzyme may affect these relatively small calculated energy differences.

The deglycosylation barriers for the *anti* (148.6 kJ/mol) and *syn* (140.2 kJ/mol) OG conformers using proline are similar to those for lysine calculated with the same level of theory (147.4 kJ/mol),[18] as well as other computational investigations.[14-15, 17] Thus, calculations predict that deglycosylation is the rate-determining step (for the most favourable, assisted ring-opening pathway). However, experimental studies have determined that the lyase step is the overall rate-limiting step in the BER chemical process.[43-44] Thus, consistent with other glycosylases,[45-46] including hOgg1,[18] FPG must significantly reduce the deglycosylation barrier. The hOgg1 catalyzed deglycosylation has been previously proposed to be facilitated by OG hydrogen bonding with (or being protonated via) an active site water or amino acids.[12, 28, 38, 47-49] Indeed, this model indicates that *anti*-OG protonation significantly reduces the deglycosylation barrier (by 50 – 80 kJ/mol depending on the OG site used for abstraction) and produces a more stable intermediate (by 55 – 105 kJ/mol). Unfortunately, however, no residue capable of protonating OG at O6 can be found in the experimental X-ray structure (1R2Y) of the enzyme–substrate complex (Figure A5.5). Nevertheless, at least partial activation of OG may occur through hydrogen-bonding interactions between O6 of OG and Thr221, Val222, Arg223 and Thr224 (Figure A5.5).[11] Alternatively, it has been proposed that Glu3 could protonate OG at the O8 position (Figure A5.5),[11, 34] which may be possible depending on the protonation state upon binding.

This chapter anticipates that including any of these residues in a larger computational model will lower the calculated deglycosylation barrier through, for

example, partial protonation of the departing base via hydrogen-bonding interactions. In fact, it has been previously shown that hydrogen-bonding interactions can significantly affect the acidity of the canonical and damaged nucleobases,[50-51] including OG, and therefore decrease the deglycosylation barrier.[52] Hence, partial protonation presumably will lead to a barrier height between those calculated in the present study for the neutral and fully protonated OG deglycosylation pathways.

In addition to nucleobase activation, the catalytically important Asp268 has been proposed to stabilize the positive charge accumulating on the DNA-protein crosslink in the deglycosylation step catalyzed by hOgg1.[38, 53] If Glu3 is anionic upon binding or transfers a proton to O6 of OG upon base departure, then Glu3 may stabilize the positive charge forming at this point along the FPG catalyzed reaction. This proposal is supported by the fact that there is a short distance between Glu3 (or the mutant Gln3) and O4′ (3.1 to 3.4 Å) in many crystal structures.[6-7, 11] Furthermore, electrostatic stabilization is a role commonly assigned to Asp/Glu residues positioned near the sugar in the active site of most DNA glycosylases.

5.4.2 Ring-Opening

The ring-opening step (Figure 5.2) differs from the previously considered lysine-catalyzed reaction since there is only one proton transfer possibility from N of proline (compared to two protons on Nζ of lysine). Transfer of the N proton either directly or via a general base was considered in the present work. Direct proton transfer from N to O4′ coupled with ring-opening was initiated from either the *anti* (Figure 5.3, right) or *syn* (Figure 5.3, left) conformer of excised OG. Although key geometrical features (*i.e.*, d(O4′···H) and d(N···H)) are nearly identical for the two reaction pathways (deviate by less than 0.002 Å), the corresponding distances are shorter for the (*anti* dOG) lysine

reaction (by 0.039 Å and 0.048 Å, respectively).[18] Nevertheless, as found for the lysine nucleophile, the barriers for direct proton transfer from cross-linked proline (~ 140 to 150 kJ/mol) are high due to the significant torsion required on the N–C1′ bond. Therefore, proton transfer facilitated by the excised OG was considered, which significantly reduces the ring-opening barrier (by ~ 115 kJ/mol for abstraction facilitated by O8 of OG). In fact, in agreement with the previous study on lysine, the base-assisted proton transfer is barrierless (by up to approximately –10 kJ/mol; Table 5.1 and Figure 5.3). Although proton delivery from OG to O4′ necessitates overcoming a barrier of ~ 25 kJ/mol (for the more favourable O-base pathway), this is much less than the favoured direct proton transfer barrier.

The barrierless proton abstraction from N indicates that the crosslink easily delivers the proton to moieties in the proximity of the linkage. Because the N proton and departed OG are on opposite sides of the sugar moiety, proton transfer to OG is unlikely in the enzyme. Nevertheless, proton delivery to O4′ may be facilitated by a different active site residue in FPG. In fact, the predicted ease of the proton delivery supports proposed proton transfer from the cross-linked N to Glu3 through a tight hydrogen-bonding network of crystallographic water molecules,[20] which would then allow Glu3 to protonate O4′. The proposal that Glu3 acts as the general acid is further supported by evidence that O4′ of the sugar moiety forms a strong hydrogen bond to Oε2 of Glu3,[20] and that Glu3 is in close proximity of O4′ in most crystal structures.

The previous computational study on hOgg1 glycosylase and β-lyase activity[18] identified the lowest energy point on the PES as the DNA–protein crosslink after lysine is deprotonated, but before the ring-opening step. This novel finding was the first computational support for the observed dissociation of hOgg1 from the abasic site by the

human AP Endonuclease (APE1, EC#: 4.2.99.18) prior to the β-elimination step,[38-40, 44, 54] and monofunctional activity of hOgg1 *in vivo*.[39-40, 54-55] In the present work, the equivalent point on the PES (without artificial stabilization due to the computational model) for FPG glycosylase and β-lyase activity lies ~ 12 kJ/mol above the reactants. Therefore, unlike the case of hOgg1, the current work does not strongly support a monofunctional activity mode for FPG, which may explain the lack of experimental evidence for FPG monofunctional activity. Furthermore, the potential absence of analogous monofunctional activity may rationalize the greater efficiency, as well as unique β-lyase activity of FPG over hOgg1.

5.4.3 β-Elimination

Although one experimental study suggests that OG diffuses out of the active site after glycosidic bond cleavage,[7] OG has been proposed to be involved in the C2′–H abstraction step.[12, 28] Alternatively, although a crystallographic water molecule has been implicated in this step,[7] the water does not hydrogen bond with the protein, and is not observed in the related *E. coli* FPG and Nei enzymes.[20] Thus, the identity of the general base for the C2′–H abstraction step is currently unclear. As a result, two pathways were characterized for the elimination step in the present work, which begin with C2′–H abstraction by OG as the general base. The proton abstraction occurs with a relatively small barrier (approximately 20 – 35 kJ/mol depending on the OG site involved, Table 5.1).

The elimination step is complete when the 3′-phosphate dissociates. As for lysine,[18] this step occurs through an E1cB mechanism, which involves a carbanion intermediate that is stabilized by the positive charge on the crosslink. In the transition state, the key O3′···C3′ distance is similar for the proline and lysine nucleophiles, and

therefore changes in the Gibbs free energy for the elimination step ranges from approximately 100 – 135 kJ/mol regardless of the (O-base or N-base) pathway or nucleophile (proline or lysine). Although the elimination barrier is slightly lower than the deglycosylation and direct ring-opening barriers, the presence of a sodium cation in this model artificially reduces the core barrier compared to an isolated phosphate group. Therefore, the phosphate group must be activated for departure in the enzyme. In FPG, this may be accomplished through interactions between the 3′-phosphate and conserved Lys56, Asn168, and Arg258 residues.[8] Interestingly, this role was proposed to be filled by a calcium dication in the active site and/or the solvent exposure of the phosphate moiety in hOgg1.[18] The importance of this stabilization predicted by calculations coupled with the greater anticipated stabilization provided by FPG provides additional clues regarding the greater efficiency, and lack of observed monofunctional activity, for FPG.

5.5 Conclusion

For the first time, this chapter uses the smallest computational model feasible (modified nucleoside-3′-monophosphate with a truncated proline nucleophile) to fully explore possible pathways for the bifunctional activity of bacterial FPG. The overall process was divided into three main phases, namely deglycosylation, (deoxyribose) ring-opening, and β-elimination, and unique findings were obtained about each step that help explain how this important enzyme can facilitate this seemingly difficult chemistry. Specifically, this work indicates that the deglycosylation barrier must be significantly reduced by the protein. Although a direct proton transfer from N of the cross-linked proline to O4′ of the sugar moiety requires overcoming a significant barrier in the second (ring-opening) phase, proton transfer assisted by a general base and acid is more favourable, which may be achieved in the enzyme through a hydrogen-bonding network

of water. The key conclusions from this chapter on FPG glycosylase and β-lyase activity parallel those from a previous study on the corresponding human enzyme (hOgg1) using a similar computational model. This suggests that the majority of differences between hOgg1 and FPG bifunctional activity do not arise due to differences in the intrinsic chemistry of the nucleophile, but rather variety in other active site residues or different catalytic mechanisms. The potential roles of active site residues during the catalytic steps, such as activation of OG during deglycosylation or protonation of O4′ to facilitate the ring-opening step, can be considered with the use of larger models.

5.6 References

1. Tchou, J.; Kasai, H.; Shibutani, S.; Chung, M. H.; Laval, J.; Grollman, A. P.; Nishimura, S., 8-oxoguanine (8-hydroxyguanine) DNA Glycosylase and its Substrate-Specificity. *Proc. Natl. Acad. Sci. U. S. A.* 1991, *88*, 4690–4694.

2. Tchou, J.; Bodepudi, V.; Shibutani, S.; Antoshechkin, I.; Miller, J.; Grollman, A. P.; Johnson, F., Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA. *J. Biol. Chem.* 1994, *269*, 15318–15324.

3. Tchou, J.; Grollman, A. P., The Catalytic Mechanism of FPG Protein - Evidence for a Schiff-Base Intermediate and Amino-Terminus Localization of the Catalytic Site. *J. Biol. Chem.* 1995, *270*, 11671–11677.

4. Zharkov, D. O.; Rieger, R. A.; Iden, C. R.; Grollman, A. P., NH2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. *J. Biol. Chem.* 1997, *272*, 5335–5341.

5. Krishnamurthy, N.; Haraguchi, K.; Greenberg, M. M.; David, S. S., Efficient Removal of Formamidopyrimidines by 8-Oxoguanine Glycosylases. *Biochemistry* 2008, *47*, 1043–1050.

6. Coste, F.; Ober, M.; Carell, T.; Boiteux, S.; Zelwer, C.; Castaing, B., Structural Basis for the Recognition of the FapydG Lesion (2,6-Diamino-4-hydroxy-5-formamidopyrimidine) by Formamidopyrimidine-DNA Glycosylase. *J. Biol. Chem.* 2004, *279*, 44074–44083.

7. Fromme, J. C.; Verdine, G. L., Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nat Struct Mol Biol* 2002, *9*, 544–552.

8. Gilboa, R.; Zharkov, D. O.; Golan, G.; Fernandes, A. S.; Gerchman, S. E.; Matz, E.; Kycia, J. H.; Grollman, A. P.; Shoham, G., Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *J. Biol. Chem.* 2002, *277*, 19811–19816.

9. Sadeghian, K.; Flaig, D.; Blank, I. D.; Schneider, S.; Strasser, R.; Stathis, D.; Winnacker, M.; Carell, T.; Ochsenfeld, C., Ribose-Protonated DNA Base Excision Repair: A Combined Theoretical and Experimental Study. *Angew. Chem., Int. Ed. Engl.* 2014, *53*, 10044–-10048.

10. Blank, I. D.; Sadeghian, K.; Ochsenfeld, C., A Base-Independent Repair Mechanism for DNA Glycosylase—No Discrimination Within the Active Site. *Scientific Reports* 2015, *5*, 10369.

11. Fromme, J. C.; Verdine, G. L., DNA Lesion Recognition by the Bacterial Repair Enzyme MutM. *J. Biol. Chem.* 2003, *278*, 51543–51548.

12. Fromme, J. C.; Bruner, S. D.; Yang, W.; Karplus, M.; Verdine, G. L., Product-assisted catalysis in base-excision DNA repair. *Nat Struct Mol Biol* 2003, *10*, 204–211.

13. Osakabe, T.; Fujii, Y.; Hata, M.; Tsuda, M.; Neya, S.; Hoshino, T., Quantum chemical study on base excision mechanism of 8-oxoguanine DNA glycosylase: Substrate-assisted catalysis of the N-glycosidic linkage cleavage reaction. *Chem-Bio Inf. J.* 2004, *4*, 73–92.

14. Schyman, P.; Danielsson, J.; Pinak, M.; Laaksonen, A., Theoretical Study of the Human DNA Repair Protein HOGG1 Activity. *J. Phys. Chem. A* 2005, *109*, 1713–1719.

15. Calvaresi, M.; Bottoni, A.; Garavelli, M., Computational Clues for a New Mechanism in the Glycosylase Activity of the Human DNA Repair Protein hOGG1. A Generalized Paradigm for Purine-Repairing Systems? *J. Phys. Chem. B* 2007, *111*, 6557–6570.

16. Zheng, Y.; Xue, Y.; Yan, S. G., The effects of oxidation and protonation on the N-glycosidic bond stability of 8-oxo-2'-deoxyguanosine: DFT study. *THEOCHEM* 2008, *860*, 52–57.

17. Shim, E. J.; Przybylski, J. L.; Wetmore, S. D., Effects of Nucleophile, Oxidative Damage, and Nucleobase Orientation on the Glycosidic Bond Cleavage in Deoxyguanosine. *J. Phys. Chem. B* 2010, *114*, 2319–2326.

18. Kellie, J. L.; Wetmore, S. D., Mechanistic and Conformational Flexibility of the Covalent Linkage Formed during β-Lyase Activity on an AP-Site: Application to hOgg1. *J. Phys. Chem. B* 2012, *116*, 10786–10797.

19. de Jesus, K. P.; Serre, L.; Zelwer, C.; Castaing, B., Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA. *Nucleic Acids Research* 2005, *33*, 5936–5944.

20. Zharkov, D. O.; Shoham, G.; Grollman, A. P., Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair* 2003, *2*, 839–862.

21. Zheng, J.-H.; Tan, H.-W.; Chen, G.-J., Theoretical study on the mechanism of the DNA repair protein Fpg. *Int. J. Quantum Chem.* 2011, *111*, 2454–2463.

22. Boiteux, S.; O'Connor, T. R.; Laval, J., Formamidopyrimidine-DNA glycosylase of Escherichia coli: cloning and sequencing of the fpg structural gene and overproduction of the protein. *Embo J* 1987, *6*, 3177–3183.

23. Rabow, L. E.; Kow, Y. W., Mechanism of action of base release by *Escherichia coli* Fpg protein: Role of lysine 155 in catalysis. *Biochemistry* 1997, *36*, 5084–5096.

24. Laval, J.; Jurado, J.; Saparbaev, M.; Sidorkina, O., Antimutagenic role of base-excision repair enzymes upon free radical-induced DNA damage. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* 1998, *402*, 93–102.

25. Lenz, S. A. P.; Kellie, J. L.; Wetmore, S. D., Glycosidic Bond Cleavage in Deoxynucleotides: Effects of Solvent and the DNA Phosphate Backbone in the Computational Model. *J. Phys. Chem. B* 2012, *116*, 14275–14284.

26. Millen, A. L.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxyl-2'-deoxyguanosine Nucleotide Adducts. *J. Phys. Chem. B* 2010, *114*, 4373–4382.

27. Churchill, C. D. M.; Wetmore, S. D., Developing a computational model that accurately reproduces the structural features of a dinucleoside monophosphate unit within B-DNA. *Phys. Chem. Chem. Phys.* 2011, *13*, 16373–16383.

28. Bruner, S. D.; Norman, D. P. G.; Verdine, G. L., Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature (London, U. K.)* 2000, *403*, 859–866.

29. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., et al. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford CT, 2009.

30. Zhao, Y.; Truhlar, D. G., Density Functional Calculations of E2 and S(N)2 Reactions: Effects of the Choice of Density Functional, Basis Set, and Self-Consistent Iterations. *J. Chem. Theory Comput.* 2010, *6*, 1104–1108.

31. Gilson, M. K.; Honig, B. H., The Dielectric-Constant of a Folded Protein. *Biopolymers* 1986, *25*, 2097–2119.

32. Simonson, T.; Perahia, D.; Bricogne, G., Intramolecular Dielectric Screening in Proteins. *J. Mol. Biol.* 1991, *218*, 859–886.

33. Duclos, S.; Aller, P.; Jaruga, P.; Dizdaroglu, M.; Wallace, S. S.; Doublié, S., Structural and biochemical studies of a plant formamidopyrimidine-DNA glycosylase reveal why eukaryotic Fpg glycosylases do not excise 8-oxoguanine. *DNA Repair* 2012, *11*, 714–725.

34. Berti, P. J.; McCann, J. A. B., Toward a Detailed Understanding of Base Excision Repair Enzymes: Transition State and Mechanistic Analyses of N-Glycoside Hydrolysis and N-Glycoside Transfer. *Chem. Rev. (Washington, DC, U. S.)* 2006, *106*, 506–555.

35. Brooks, S. C.; Adhikary, S.; Rubinson, E. H.; Eichman, B. F., Recent Advances in the Structural Mechanisms of DNA Glycosylases. *Biochim. Biophys. Acta* 2013, *1834*, 247–271.

36. Gasparutto, D.; Muller, E.; Boiteux, S.; Cadet, J., Excision of the oxidatively formed 5-hydroxyhydantoin and 5-hydroxy-5-methylhydantoin pyrimidine lesions by *Escherichia coli* and Saccharomyces cerevisiae DNA N-glycosylases. *Biochim. Biophys. Acta, Gen. Subj.* 2009, *1790*, 16–24.

37. Wiederholt, C. J.; Delaney, M. O.; Pope, M. A.; David, S. S.; Greenberg, M. M., Repair of DNA containing Fapy·dG and its beta-C-nucleoside analogue by formamidopyrimidine DNA glycosylase and MutY. *Biochemistry* 2003, *42*, 9755–9760.

38. Dalhus, B.; Forsbring, M.; Helle, I. H.; Vik, E. S.; Forstrøm, R. J.; Backe, P. H.; Alseth, I.; Bjørås, M., Separation-of-Function Mutants Unravel the Dual-Reaction Mode of Human 8-Oxoguanine DNA Glycosylase. *Structure (Cambridge, MA, U. S.)* 2011, *19*, 117–127.

39. Hill, J. W.; Hazra, T. K.; Izumi, T.; Mitra, S., Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic Acids Research* 2001, *29*, 430–438.

40. Morland, I.; Luna, L.; Gustad, E.; Seeberg, E.; Bjørås, M., Product inhibition and magnesium modulate the dual reaction mode of hOgg1. *DNA Repair* 2005, *4*, 381–387.

41. Perlow-Poehnelt, R. A.; Zharkov, D. O.; Grollman, A. P.; Broyde, S., Substrate Discrimination by Formamidopyrimidine-DNA Glycosylase: Distinguishing Interactions within the Active Site. *Biochemistry* 2004, *43*, 16092–16105.

42. El, Z.; Perlow Ra Fau - Matz, E.; Matz E Fau - Broyde, S.; Broyde S Fau - Gilboa, R.; Gilboa R Fau - Grollman, A. P.; Grollman Ap Fau - Zharkov, D. O.; DO, Z., Substrate discrimination by formamidopyrimidine-DNA glycosylase: a mutational analysis. *J. Biol. Chem.* 2004, *279*, 4849–4861.

43. Bjoras, M.; Luna, L.; Johnson, B.; Hoff, E.; Haug, T.; Rognes, T.; Seeberg, E., Opposite base-dependent reactions of a human base excision repair enzyme on DNA containing 7,8-dihydro-8-oxoguanine and abasic sites. *EMBO J.* 1997, *16*, 6314–6322.

44. Zharkov, D. O.; Rosenquist, T. A.; Gerchman, S. E.; Grollman, A. P., Substrate Specificity and Reaction Mechanism of Murine 8-Oxoguanine-DNA Glycosylase. *J. Biol. Chem.* 2000, *275*, 28607–28617.

45. Drohat, A. C.; Stivers, J. T., *Escherichia coli* uracil DNA glycosylase: NMR characterization of the short hydrogen bond from His187 to uracil O2. *Biochemistry* 2000, *39*, 11865–11875.

46. Brinkmeyer, M. K.; Pope, M. A.; David, S. S., Catalytic Contributions of Key Residues in the Adenine Glycosylase MutY Revealed by pH-dependent Kinetics and Cellular Repair Assays. *Chem. Biol.* 2012, *19*, 276–286.

47. Norman, D. P. G.; Bruner, S. D.; Verdine, G. L., Coupling of Substrate Recognition and Catalysis by a Human Base-Excision DNA Repair Protein. *J. Am. Chem. Soc.* 2001, *123*, 359–360.

48. Bjoras, M.; Seeberg, E.; Luna, L.; Pearl, L. H.; Barrett, T. E., Reciprocal "flipping" underlies substrate recognition and catalytic activation by the human 8-oxo-guanine DNA glycosylase. *J. Mol. Biol.* 2002, *317*, 171–177.

49. Radom, C. T.; Banerjee, A.; Verdine, G. L., Structural characterization of human 8-oxoguanine DNA glycosylase variants bearing active site mutations. *J. Biol. Chem.* 2007, *282*, 9182–9194.

50. Hunter, K. C.; Wetmore, S. D., Environmental Effects on the Enhancement in Natural and Damaged DNA Nucleobase Acidity Because of Discrete Hydrogen-Bonding Interactions. *J. Phys. Chem. A* 2007, *111*, 1933–1942.

51. 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*, 6351–6362.

52. Millen, A. L.; Archibald, L. A. B.; Hunter, K. C.; Wetmore, S. D., A kinetic and thermodynamic study of the glycosidic bond cleavage in deoxyuridine. *J. Phys. Chem. B* 2007, *111*, 3800–3812.

53. Norman, D. P. G.; Chung, S. J.; Verdine, G. L., Structural and Biochemical Exploration of a Critical Amino Acid in Human 8-Oxoguanine Glycosylase. *Biochemistry* 2003, *42*, 1564–1572.

54. Vidal, A. E.; Hickson, I. D.; Boiteux, S.; Radicella, J. P., Mechanism of stimulation of the DNA glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step. *Nucleic Acids Research* 2001, *29*, 1285–1292.

55. Kuznetsov, N. A.; Koval, V. V.; Zharkov, D. O.; Nevinsky, G. A.; Douglas, K. T.; Fedorova, O. S., Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic Acids Research* 2005, *33*, 3919–3931.

Chapter 6: β- and δ-elimination Reactions Facilitated by FPG*

6.1 Introduction

In addition to the deglycosylation and the β-elimination facilitated by FPG discussed in Chapter 5, very little is known about the mechanism of action of the β- and δ-lyase reactions catalyzed by FPG. Specifically, the general base that abstracts the proton in the β-elimination reaction is currently unidentified.[1] Although there has been some speculation about the general base that abstracts the proton, even fewer details are available about the phosphate elimination reaction. As discussed in Chapter 1, crystal structures reveal that the 3′- and 5′-phosphates with respect to the damaged nucleoside are in close proximity to conserved residues in FPG.[2-3] Specifically, the 3′-phosphate interacts with Lys57 and Arg259, while the 5′-phosphate interacts with Asn169, Arg259, and Tyr242.[2] Indeed, it has been proposed that Lys57 and Arg259 protonate the 3′- and 5′-phosphate moieties, respectively, while the other residues further stabilize the increased negative charge formed upon phosphate departure through hydrogen-bonding contacts.[2, 4] Although this evidence suggests that acidic residues in the proximity of the phosphate leaving groups are important for elimination, the reaction mechanisms for the excision steps have yet to be uncovered.

In light of the above experimental unknowns, several computational studies have investigated different aspects of the BER pathway facilitated by FPG.[5-8] In addition, quantum mechanical studies have investigated the catalytic reaction mechanism of the

---

* Reproduced in part from Sowlati-Hashjin S. and Wetmore S. D. Quantum mechanical study of the β- and δ-lyase reactions during the base excision repair process: application to FPG *Phys. Chem. Chem. Phys.* **2015** (17) 24696−24706 with permission from PCCP Owners Societies.

first BER step under the assumption that deglycosylation occurs prior to sugar-ring opening.[5-6]

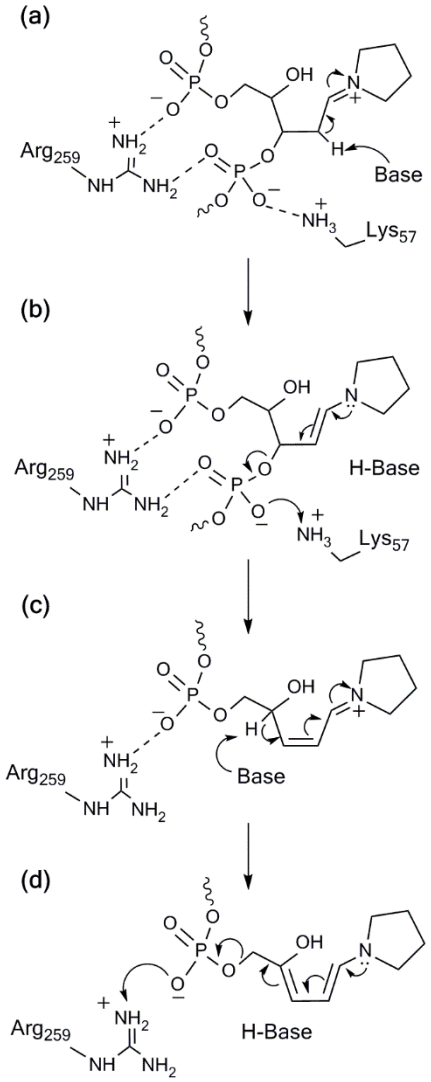

Figure 6.1. Proposed chemical steps during β- and δ-lyase activity of FPG. Specifically, (a) an active-site base abstracts the *pro*-S proton, which is followed by (b) the elimination of the 3′-phosphate. Subsequently, (c) C4′–H is removed by a general base and (d) δ-elimination occurs. In FPG, the 3′-phosphate has been proposed to be protonated by Lys57 and further stabilized by Arg259, while the 5′-phosphate has been proposed to be protonated by Arg259 and further stabilized by Tyr237 and Asn169. Asn169 and Tyr237 are not shown for clarity.

Recent computational studies have characterized a deglycosylation mechanism involving ring opening preceding glycosidic bond cleavage.[7-8] Beyond the initial

deglycosylation step, several unique pathways for the β-elimination reaction were previously characterized using modified nucleoside-3′-monophosphate models and OG^- as the general base.[9] However, the nucleoside-3′-monophosphate models employed in the previous study were neutralized by a sodium ion,[9] which prevents the phosphate protonation that has been proposed to facilitate the reaction.[2, 4] Furthermore, no computational study to date has considered the mechanism of action of the δ-elimination reaction.

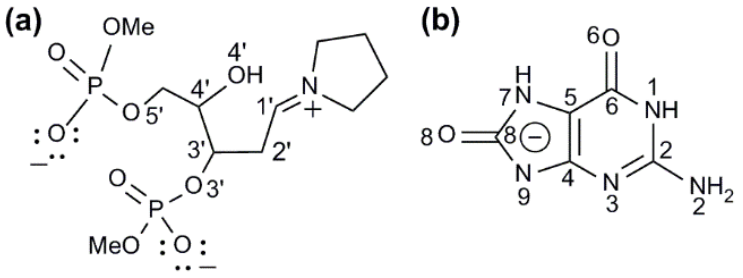

Figure 6.2. 2D representation and atomic numbering of a) 1-(4-hydroxy-3,5-diyl dimethyl bis(phosphate)pentylidene) pyrrolidinium, and b) the negatively charged OG employed in the initial computational model.

The present chapter uses quantum mechanical methods and a small, yet chemically-relevant, model as a first step towards gaining currently missing information about the mechanistic details of the β- and δ-elimination reactions catalyzed by a bifunctional glycosylase during BER. In contrast to Chapter 5and previous computational studies on the β-elimination step,[9] anionic phosphate models are used to investigate the role of phosphate activation in the elimination reactions. Furthermore, the current model includes both 3′- and 5′-phosphate residues to allow the first comparison between the β- and δ-elimination reactions in attempts to reveal key mechanistic and energetic similarities and/or differences between these BER steps. This approach will unveil whether differences in the inherent chemistry helps explain why some bifunctional

glycosylases only facilitate the β-elimination reaction. By using a simplistic model to characterize different potential pathways for the β- and δ-lyase reactions, the most relevant routes will be identified that can subsequently be considered using large-scale DNA–protein models in the future.

6.2 Computational Details

To simultaneously investigate the β- and δ-elimination reactions catalyzed by FPG, 1-(4-hydroxy-3,5-diyl dimethyl bis(phosphate)pentylidene) pyrrolidinium was used as the core component of the present model (Figure 6.2a). This component was built from the ring-opened deoxyribose intermediate optimized in the previous chapter by replacing the 5′-methoxy group with a 5′-phosphate group. The 3′- and 5′-phosphates are capped with methyl groups in the present model to avoid hydrogen bonding between the DNA backbone and nucleobase (OG) that cannot occur in the DNA–protein system. Support for the chosen starting model comes from a comparison between the associated relaxed structure and crystal structures of FPG bound to a borohydride-trapped abasic site (PDB ID: 1K82)[4] or a Schiff base intermediate (PDB ID: 1L1Z;[2] Figure A6.1). Specifically, deviations between the present model and the crystal structures mainly arise due to crystallization techniques used to trap the DNA–protein crosslink (such as mutating key residue(s) or employing inhibitors), different C1′ hybridizations (sp^3 in the crystal structures versus sp^2 in the optimized model), and the flexibility of the terminal phosphates in the present model that are bound within a DNA oligomer in the crystal structures.

As in Chapter 5, OG^- (or H–OG) is added to the model to act as the general base (or acid) as required for the reaction to proceed (Figure 6.2b). Furthermore, OG^- can accept (or H–OG can deliver) a proton via O8 or N9, and the associated pathways are

denoted as O-base and N-base throughout the present chapter, respectively. Nevertheless, OG^- or H–OG is removed from the model for the reactions in which a general acid or base is not required. Similarly, the cleaved 3′-phosphate is removed from the present model following the β-elimination step to prevent artificial interactions between the eliminated group and the remaining sugar–proline crosslink that are unlikely to occur in the DNA–protein system.

The reaction potential energy surfaces (PESs) were searched by characterizing transition states (TSs), and following the corresponding intrinsic reaction coordinate (IRC) in both (forward and backward) directions. All resulting stationary points were fully optimized at the B3LYP/6-31G(d) level of theory in the presence of bulk solvent as described by the IEF-PCM method implemented in Gaussian 09.[10] The dielectric was approximated as diethyl ether ($\varepsilon = 4.24$) in all calculations based on the shape of FPG and the active site location. Within the model, the product of some chemical steps (denoted as "primed" intermediates) do not exactly correspond to the reactant for the subsequent reaction due to inevitable translational motion of OG (the general base/acid) during the reaction. In these cases, the structure obtained from the reverse IRC was considered to be the intermediate for the previous reaction step. Unless otherwise mentioned, all reported energies correspond to relative Gibbs energies (including scaled (0.9806) zero-point vibrational energy and unscaled thermal corrections) obtained from single-point calculations at the SMD-M06-2X/6-311+G(2df,2p) level of theory. All calculations were carried out using Gaussian 09 (Revision C.01).[10]

6.3 Results and Discussion

As outlined in the Introduction, the chemical steps facilitated by FPG include deglycosylation, ring opening, β-elimination and δ-elimination. In this study, possible

pathways for the β- and δ-elimination phases of BER facilitated by FPG are characterized (Figure 1.12). The relative SMD-M06-2X/6-311+G(2df,2p) Gibbs energies for all reaction steps are provided in Tables 6.1 and 6.2, and the corresponding Gibbs reaction surfaces are shown in Figure 6.4. The IEF-PCM-B3LYP/6-31G(d) relative energies (with and without (scaled) zero-point vibrational energy corrections) are provided in the Appendices (Tables A6.1 – A6.4). In the sections below, important structural features of the stationary points and the energy barriers for all reaction pathways characterized are compared to each other and the literature when possible, and the biochemical implications of the present findings that are important for future large-scale modeling on complete DNA–enzyme models are discussed.

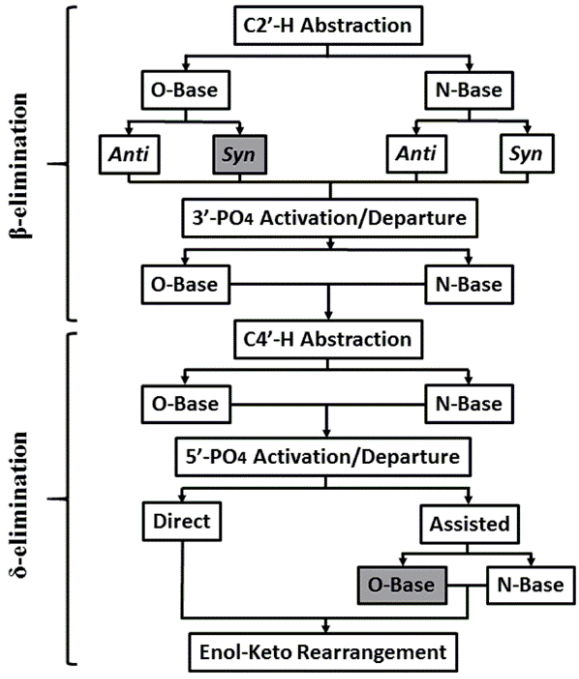

Figure 6.3. Reaction pathways for β- and δ-elimination facilitated by FPG considered in the present study.

6.3.1 β-Lyase Step

As described in Section 6.1, the β-elimination reaction catalyzed by FPG occurs in conjunction with proton abstraction from Cβ (C2′ of the deoxyribose) by a general base (Figure 6.1). Although the *pro*-S or *pro*-R proton can be abstracted in this step, structural information suggests that the elimination reaction proceeds by *pro*-S proton removal in *E. coli* FPG. Specifically, the average (ν2) torsion angle around C1′–C2′–C3′–C4′ in *E. coli* FPG (ν2 = –81.6°) is close to the value required for *pro*-S proton removal through an anti stereochemical elimination (ν2 ≈ –60°).[1] Moreover, a previous computational study on the β-elimination reaction determined that *pro*-S proton abstraction has a smaller associated barrier than *pro*-R proton.[9] Therefore, only *pro*-S proton removal was considered for the reaction facilitated by FPG in this chapter.

In agreement with a previous computational study[9] and as was shown in Chapter 5, the elimination reaction occurs in two successive steps with proton abstraction preceding phosphate cleavage (*i.e.*, an E1cB mechanism). In the present model, four different routes were characterized for the first *pro*-S proton abstraction step (Figure 6.3). Specifically, the proton abstraction from the reactant (RC1) was considered to be facilitated by O8 (denoted as O-base) or N9 (denoted as N-base) of OG^-. Furthermore, since proton abstraction immediately follows deglycosylation, OG^- was placed in an orientation with respect to the rest of the model that corresponds to an *anti* or *syn* nucleoside conformation.

Table 6.1. Relative Gibbs free energies (ΔG, kJ/mol) for stationary points characterized along the *anti* O-base and N-base pathways for the β-elimination reaction.[a]

Reaction Step	Stationary Point	O-base	N-base
C2′–H Abstraction	RC1	0.0	0.0
	TS1	22.0	51.5
	IC1	–0.6	–2.6
3′-PO$_4$ Elimination	TS2	19.3	–
	IC2	–18.7	–

[a] Relative energies were obtained with SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) and include (scaled) zero-point vibrational energy and (unscaled) thermal corrections.

For O-base *pro*-S proton abstraction, the initial *syn* OG^- pathway led to a transition state with a similar structure (Figures A6.2), but (~ 40 kJ/mol) higher barrier (Table A6.1), than isolated from an initial *anti* OG^- configuration. Thus, this pathway is not further discussed. Along the *anti* O-base *pro*-S proton abstraction pathway (Figure 6.5, left), the transition state (TS1, Figure 6.5, left) occurs with d(O8···H) = 1.246 Å, d(C2′···H) = 1.388 Å, and ∠(C2′···H···O8) = 174.3°. Both d(O8···H) and d(C2′···H) are slightly shorter (by 0.045 and 0.008 Å, respectively) than previously estimated using a nucleoside-3′-monophosphate model (Chapter 5) since interactions between the 5′-phosphate and OG^- (mainly N7–H) in the current model delay the TS. In the associated intermediate (IC1′; Figure 6.5, left), the proton is completely transferred to OG (d(H–O8) = 0.992 Å) and a planar C3′–C2′–C1′–N arrangement is formed. H–OG maintains an N7–H···O interaction with the 5′-phosphate, which pulls H–OG away from C2′ (*i.e.*, d(O8–H···C2′) = 2.114 Å compared to 2.063 Å in the absence of the 5′-phosphate from the previous chapter).

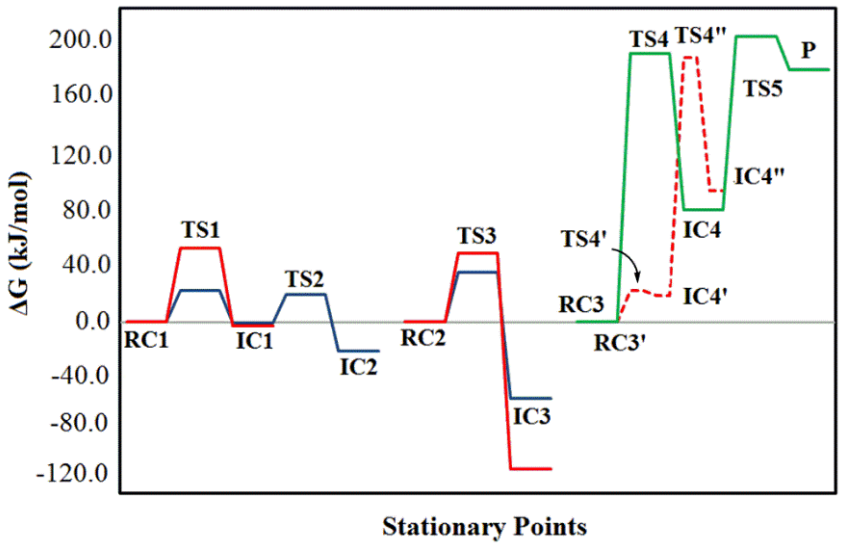

Figure 6.4. The SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) Gibbs energies (kJ/mol) relative to the corresponding reactant for each pathway considered in the present study, including the N-base (red, solid), O-base (blue, solid), the N-base assisted 5′-PO_4 elimination (red, dashed), and the direct 5′-PO_4 elimination followed by enol-keto rearrangement (green, solid).

When C2′–H abstraction is facilitated by N9 (N-base pathway, Figure 6.3), the proton abstraction barrier is significantly (~ 32 kJ/mol) lower when initiated with an *anti* rather than a *syn* OG^- orientation (Table A6.2). This difference mainly arises because the optimized reactant from the reverse IRC for the *syn* TS contains an intact glycosidic bond and an opened ring (Figure A6.3). Thus, a larger calculated barrier arises for the syn than *anti* OG^- pathway since the barrier also accounts for the deglycosylation reaction. Although the deglycosylation BER step is not the focus of this chapter, a controversy in the literature surrounding the order of the deglycosylation and ring opening steps is acknowledged.[1, 7] Using a similar model, data from the previous chapter suggests that (*syn*-OG) deglycosylation requires ~ 140 kJ/mol when occurring prior to ring opening. However, the current work predicts the calculated deglycosylation barrier to be 83.4 kJ/mol when occurring after ring opening in conjunction with C2′–H abstraction. Thus, the present data suggests that the deglycosylation reaction can occur with a much lower

energy barrier once the deoxyribose ring is opened. This finding is supported by the most recent experimental and computational studies on the deglycosylation of OG[7] and FapyG[8] facilitated by FPG, which proposed that an initial acid-catalyzed ring-opening step is followed by barrierless deglycosylation. Regardless, since the present investigation is concerned with the (β- and δ-) lyase activity of FPG, focus is placed on the *anti* OG^- N-base pathway for β-lyase activity below.

In the *pro*-S proton abstraction transition state associated with the anti OG^- N-base pathway (TS1; Figure 6.5, right), d(N9···H) = 1.400 Å and d(C2′···H) = 1.334 Å, which are comparable to the distances found using a model that lacks the C5′-phosphate moiety (d(N9···H) = 1.421 Å and d(C2′···H) = 1.320 Å). Furthermore, the ∠(C2′···H···N9) angle is 169.2°, which is only slightly (~ 2°) larger than previously obtained using the smaller model (Chapter 5). TS1 leads to an intermediate (IC1′; Figure 6.5, right) with a planar C3′–C2′–C1′–N arrangement and H–OG in the proximity of the crosslink (d(N9–H··· C2′) = 2.288 Å). However, H–OG is further removed from the crosslink than in the intermediate associated with the O-base pathway (by 0.174 Å), likely due to repulsive contacts between O8 and the 5′-phosphate.

Comparison of the (*anti* OG^-) transition states for C2′–H abstraction facilitated by O8 and N9 reveals that d(N9···H) is significantly (0.154 Å) longer than d(O8···H), while the C2′···H distance is 0.058 Å shorter for the N-base than O-base pathway (Figure 6.5). Moreover, the proton transfer occurs with a smaller angle for the N-base (∠(C2′···H···N9) = 169.4°) than the O-base (∠(C2′···H···O8) = 174.3°) pathway. These structural differences result in a larger barrier for the N-base (51.5 kJ/mol) than O-base (22.0 kJ/mol) reaction. Although the O-base abstraction barrier is in good agreement with that calculated using a nucleoside-3′-monophosphate model and the same level of theory

(21.5 kJ/mol), the N-base abstraction barrier is larger than in the absence of the 5′-phosphate (26.0 kJ/mol; Chapter 5), which may be due to destabilizing interactions between O8 and O5′ in the transition structure associated with the current model. In fact, in the N-base pathway, OG retains the initial *anti* orientation in the current study (Figure 6.5), but rearranges to a *syn* orientation in the nucleoside-3′-monophosphate model, which leads to a stabilizing interaction between O5′ and N2–H.

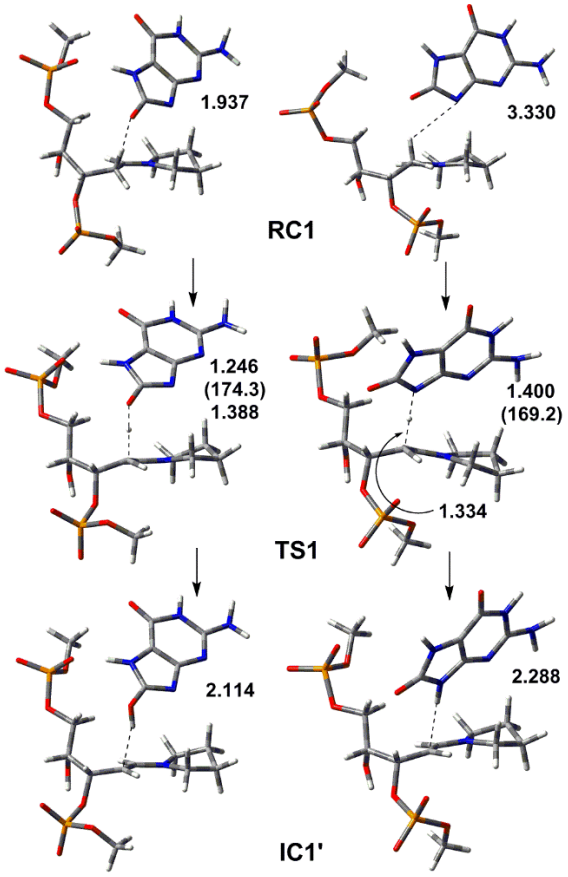

Figure 6.5. Structures characterized along the *anti* O-base (left) and *anti* N-base (right) pathways for the C2′–H abstraction reaction. Important distances (Å) and angles (deg, in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) are provided.

Following the proton abstraction step, the C3′–PO_4 bond cleavage completes the β-elimination reaction (Figure 6.1b). Since a negatively charged phosphate is a poor leaving group, efforts to locate a transition state for the elimination reaction without

protonating the phosphate group were unsuccessful. Indeed, the 3′-phosphate with respect to the damaged nucleoside is exposed to conserved Lys57 and Arg259 active site amino acids in crystal structures of FPG bound to DNA (Figure 1.13b).[2, 4] The proposed mechanism for the C3′–PO_4 bond cleavage involves protonation of a terminal oxygen in the phosphate moiety by Lys57, with additional stabilization provided by Arg259.[4] In the present model, free H–OG (*i.e.*, OG^- formed following deglycosylation and C2′ proton abstraction by O8 (O-base) or N9 (N-base)) acts as the general acid to deliver a proton to the phosphate group (Figure 6.6). To model this reaction step, H–OG was moved from the O5′-side in the intermediate from the C2′ abstraction step (IC1′; Figure 6.5, left) to the O3′-side of the original deoxyribose moiety to generate IC1 (Figure 6.6), the reactant for the C3′-phosphate activation/departure. Although activation was considered from both O8 and N9, the transition state for 3′-phosphate protonation by N9 could not be located, likely due to steric or electrostatic clashes between O8 of H–OG and O4′ of the sugar-phosphate backbone.

The 3′-phosphate protonation and elimination reactions are concerted (TS2, Figure 6.6). In the transition state, d(O3′···C3′) = 1.529 Å and H–OG (O8) delivers the proton to the phosphate at an ∠(O8···H···O3′) angle equal to 164.4°, with d(O8···H) = 1.361 Å and d(H···O3′) = 1.103 Å. The d(O3′···C3′) distance in the transition state is significantly shorter than when the phosphate group is neutralized by a sodium ion (2.529 and 2.576 Å for the O-base and N-base pathways, respectively; Chapter 5). The calculated relative Gibbs free energy for the 3′-phosphate cleavage step is 19.9 kJ/mol (Table 6.1). The β-elimination reaction results in an intermediate in which 3′-PO_4 is protonated at O3′ and the O3′–C3′ bond is completely cleaved (IC2, Figure 6.6).

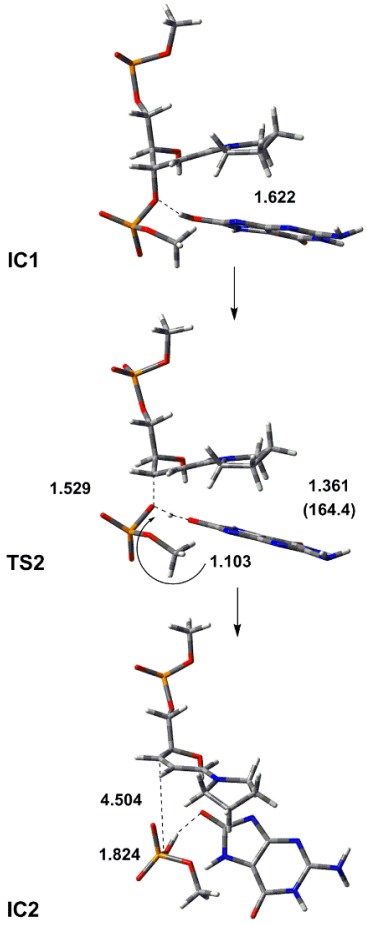

Figure 6.6. Structures characterized along the O-base pathway for the 3′-PO_4 activation/elimination reaction. Important distances (Å) and angles (deg, in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) are provided.

Note that the 3′-phosphate elimination could only be characterized with the present model once the leaving group was protonated, and the newly calculated barrier (19.9 kJ/mol) is significantly less than that previously reported when the phosphate moiety is solely stabilized by a sodium cation (80.7 – 95.4 kJ/mol; Chapter 5). Therefore, protonation of the phosphate moiety seems to be required prior to expulsion, which is in-line with the elimination mechanism conjectured from experimental crystal structures.[4] Furthermore, this mechanism is supported by longer distances (weaker interactions) between the 3′-phosphate with respect to the damaged nucleoside and active site residues

in hOgg1 (Figure 1.13a),[11] which shows only weak β-lyase activity.[12] Interestingly, the contact distances between FPG active site residues (Lys57 and Arg259) and the 3′-phosphate are shorter to the terminal oxygen atoms (~ 2.5 – 3.1 Å) than O3′ (~ 4.9 – 5.1 Å) in the lesion recognition complex (PDB ID: 1R2Y),[13] as well as Schiff base (PDB ID: 1K82)[4] and abasic site (PDB ID: 1L1Z23[2] and 3TWM[14]) intermediates (Figure 1.13); however, all efforts to protonate a terminal oxygen atom of the phosphate group in the present model were unsuccessful, and instead only O3′ could be protonated. Although this discrepancy could be an artifact of the computational model, this finding may suggest that active site residues adopt new positions as the reaction proceeds that alter active site–phosphate interactions compared with the static pictures provided by crystal structures or other species are involved in the reaction. For example, water molecules can be found in the proximity of the 3′-phosphate,[13] which may protonate O3′. Regardless of the acid that protonates the phosphate group, additional stabilizing phosphate–enzyme interactions (Figure 1.13) may further lower the barrier height determined with the model considered. Therefore, the relative importance, dynamics and interplay of the multiple hydrogen bonds between FPG, active site water and the 3′-phosphate found in experimental crystal structures must be more carefully considered using a complete DNA–enzyme model.

6.3.2 δ-Lyase Step

Current knowledge about the mechanism of the δ-lyase activity facilitated by FPG is even less than that for the β-lyase step. In parallel to the β-lyase reaction, it has been proposed that a (currently unidentified) base abstracts a proton from C4′ and, subsequently, the 5′-phosphate is eliminated (Figure 6.1c – d).[1] Unlike for the β-elimination step, there is only one proton abstraction possibility (Figure 1.13). In this model, C4′–H is initially abstracted by either O8 or N9 of OG^- (Figure 6.7), which

produces a conjugated π-system over C4′–C3′–C2′–C1′–N (Figure 6.7). The associated transition structure is earlier for the N-base than O-base pathway (TS3, Figure 6.8), with d(C4′···H) being 0.023 Å shorter and d(X···H) being 0.101 Å longer (X = N9 for N-base and O8 for O-base pathway). The earlier TS for the N-base pathway may be due to hydrogen bonding between O4′–H and O8 of OG, which is not feasible along the O-base pathway since O4′–H is instead oriented towards the 5′-phosphate (TS3, Figure 6.8). For the O-base pathway, OG^- abstracts the C2′ and C4′ protons with a similar angle (∠(C2′···H···O8) = 174.3° and ∠(C4′···H···O8) = 171.9°) in the β- and δ-elimination steps. In contrast, the proton abstraction angles differ by ~ 11° for the N-base pathway (∠(C2′···H···N9) = 169.2° and ∠(C4′···H···N9) = 158.2°), although both angles are smaller than in the corresponding O-base pathway. In the intermediate complex (IC3, Figure 6.8), O8 hydrogen bonds with O4′–H and N9–H interacts with 5′-PO_4 along the N-base pathway, while such specific interactions are absent in the O-base intermediate.

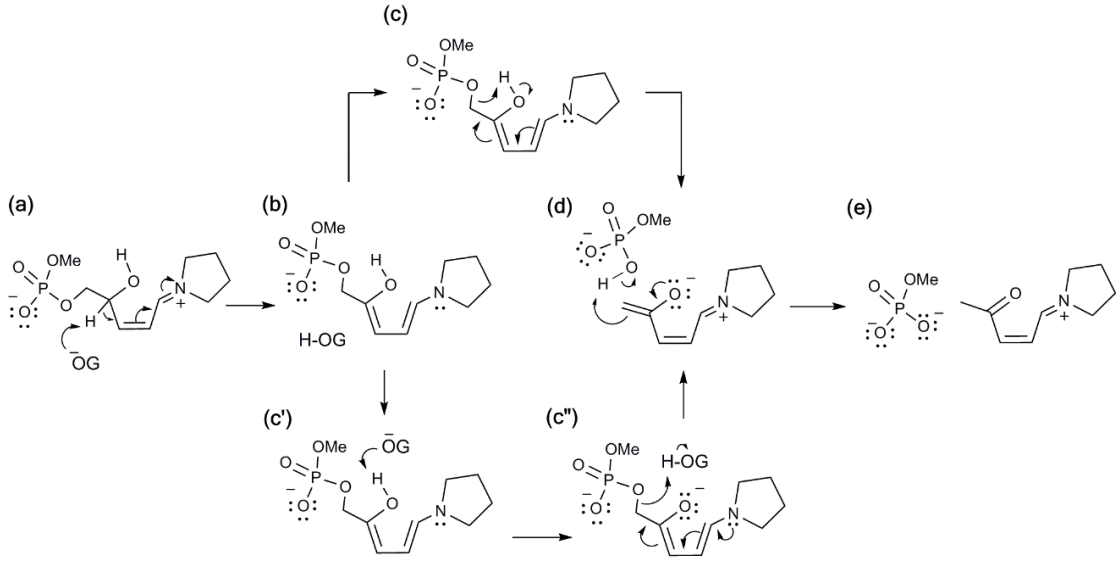

Figure 6.7. Pathways characterized in the present chapter for the δ-elimination reaction facilitated by FPG, including direct (b – c) and assisted (b – c' – c") phosphate protonation/elimination routes.

The energy barriers associated with proton abstraction in the δ-elimination step are 34.7 or 48.0 kJ/mol for the O-base or N-base pathways, respectively (Table 6.2). The higher barrier for the N-base pathway stems at least in part from hydrogen bonding between O4′–H and O8 in the associated reactant (RC2, Figure 6.8). Both transition structures (TS3, Figure 6.8) show bending of the sugar–phosphate moiety relative to the corresponding reactant (RC2). This bending occurs since it is energetically more favorable for the molecule to form an extended conjugated π-system after proton abstraction by aligning the p-orbital of C4′ in a parallel fashion with respect to the p-orbitals of C3′, C2′, C1′, and N (Figure A6.4), which results in a highly stable intermediate (IC3, Figure 6.8 and Table 2). The intermediate associated with the N-base pathway is more stable with respect to the corresponding reactant (–102.8 kJ/mol) than the O-base intermediate (–53.8 kJ/mol) due to the discrete hydrogen bonds between H–OG and the sugar-phosphate moiety. Although the C4′ and C2′ hydrogen abstraction barriers are similar for the N-base pathways (within 3 kJ/mol), the C4′–H abstraction pathway requires an additional 12 kJ/mol than C2′–H abstraction, likely due to the aforementioned bending in the transition state.

As in the β-elimination step, 5′-phosphate departure follows proton abstraction from the neighbouring (C4′) carbon. Furthermore, a transition state for the 5′-phosphate elimination could not be located in the absence of leaving group protonation. However, in the case of δ-elimination, efforts to protonate the phosphate group by (either O8 or N9) H–OG (the intermediate (IC3) following the C4′ abstraction step) were also unsuccessful. Instead, two unique pathways were characterized for 5′-PO_4 protonation/departure that involve the O4′ proton.

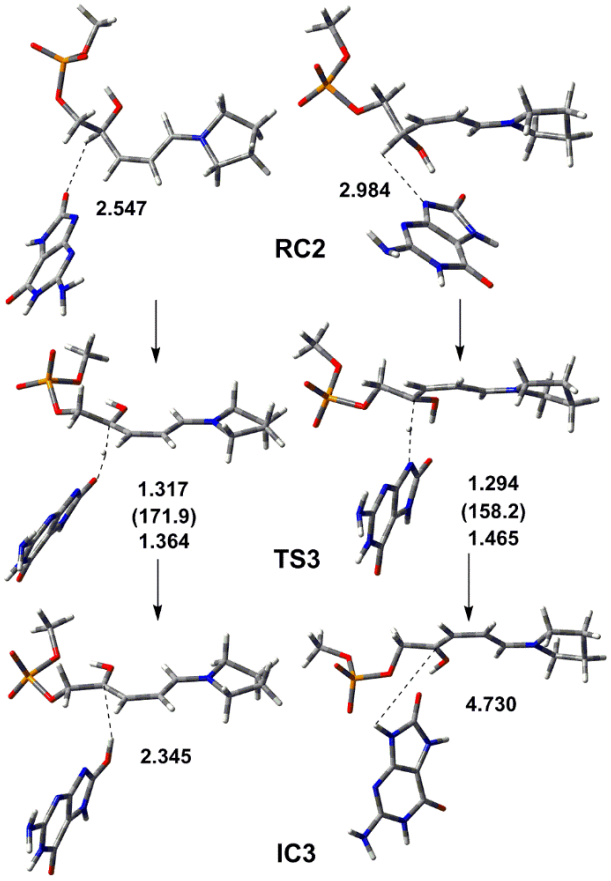

Figure 6.8. Structures characterized along the O-base (left) and N-base (right) pathways for the C4′–H abstraction reaction. Important distances (Å) and angles (deg, in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) are provided.

In the first (direct) pathway, 5′-PO_4 is activated through direct proton transfer from O4′ to O5′ (Figure 6.7c). In the second (assisted) pathway, the O4′ proton is initially transferred to OG^- and 5′-PO_4 is then activated for departure (at O5′) by H–OG (Figure 6.7c′ – c″). Although technically either O8 or N9 could abstract the O4′ proton, only the N-base pathway could be characterized. Furthermore, in order for O4′–H abstraction to take place, H–OG in the previous intermediate (IC3) is replaced by OG^- in the reactant for the assisted pathway (RC3′).

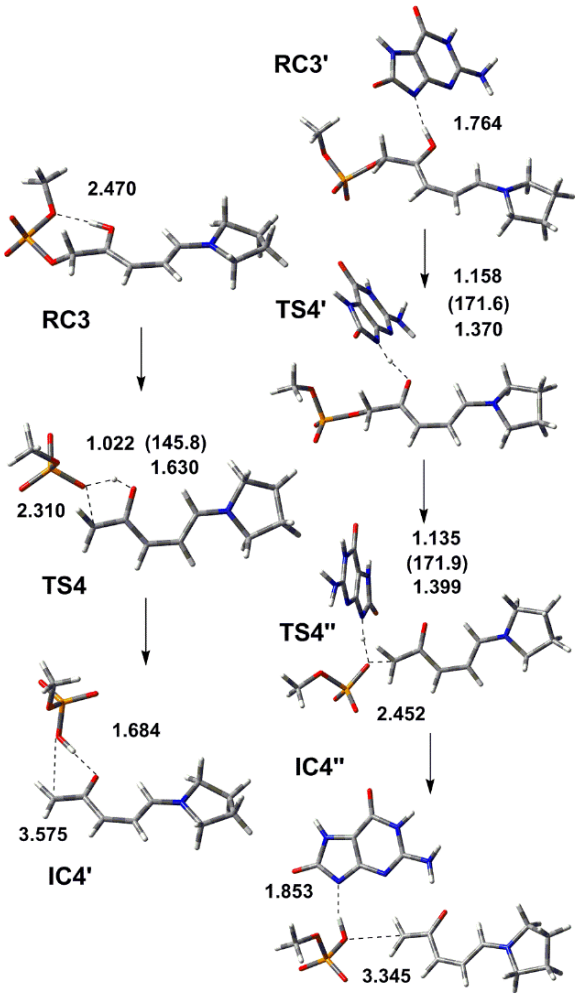

Figure 6.9. Structures characterized along the direct (left) and acid-assisted (right) pathways for the 5′-PO_4 activation/elimination reaction. Important distances (Å) and angles (degrees in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) are provided.

In the transition state for O4′ proton abstraction along the assisted pathway (TS4′, Figure 6.9), d(O4′···H) and d(H···N9) are 1.370 and 1.158 Å, respectively, and ∠(O4′···H···N9) = 171.6°. As in the β-elimination reaction, the phosphate protonation and cleavage steps are simultaneous in both the assisted (TS4″) and direct (TS4) pathways (Figure 6.9). In the assisted pathway, d(O5′···C5′) = 2.452 Å, which is significantly longer than in the transition state for the β-elimination (d(O3′···C3′) = 1.529 Å) and slightly longer than in the direct pathway (d(O5′···C5′) = 2.310 Å). Furthermore, proton delivery occurs at a narrower angle (∠(O4′···H···O5′) = 145.8°) in the direct

pathway compared to the assisted 5′-PO_4 cleavage (∠(N9···H···O5′) = 171.9°) and the β-elimination step (∠(O8···H···O3′) = 164.4°). The energy barrier for the 5′-PO_4 elimination is 188.6 kJ/mol for the direct route and 166.5 kJ/mol for the assisted pathway, which involves an additional 21.9 kJ/mol barrier associated with O4′ proton abstraction. This suggests that there is not an energetic preference for one pathway over the other within the present model. The immediate product of the δ-elimination reaction is an enolate moiety (IC4′/IC4″), which rearranges to the final ketone product through a proton transfer from the expelled phosphate to C5′ (Figure 6.7d – e). In the transition state (TS5, Figure 6.10), the phosphate group delivers the proton with ∠(O5′···H···C5′) = 171.6°, d(O5′···H) = 1.512 Å and (C5′···H) = 1.200 Å. The associated calculated barrier is large (121.7 kJ/mol, Table 6.2) due to perturbation of the conjugated π-system upon protonation of C5′ (Figure 6.7d – e), which completes the δ-lyase reaction (P, Figure 6.10).

Table 6.2. Relative Gibbs free energies (ΔG, kJ/mol) for stationary points characterized along the O-base and N-base pathways for the δ-elimination reaction, as well as the enol-keto rearrangement.[a]

Reaction Step	Stationary Point	O-base	N-base
C4′–H Abstraction	RC2	0.0	0.0
	TS3	34.7	48.0
	IC3	–53.8	–102.8
	Stationary Point	**Direct**	**Assisted**
5′-PO_4 Elimination	RC3/RC3′	0.0	0.0
	TS4′	–	21.9
	TS4/IC4′	188.6	18.8
	TS4″	–	185.3
	IC4/IC4″	78.9	92.0
Enol-keto Rearrangement	TS5	200.6	–
	P	176.8	–

[a] Relative energies were obtained with SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) and include (scaled) zero-point vibrational energy and (unscaled) thermal corrections.

As discussed above, a transition state for the 5′-phosphate elimination reaction could not be isolated without phosphate protonation. Analysis of experimental crystal structures suggests that the 5′-phosphate interacts with Asn169, Arg259 and Tyr237 in the FPG active site (Figure 6.4a), with Arg259 being proposed to protonate a terminal oxygen for phosphate departure.[2, 4] However, as discussed for 3′-phosphate elimination, efforts to locate a transition state corresponding to protonation of a terminal oxygen atom were unsuccessful and instead protonation only occurred at O5′ in the present model. It is interesting to note that FPG active site residues are generally closer to O5′ (< 4.6 Å) than O3′ (< 5.3 Å) in a range of crystal structures,[2, 4, 13-14] which may suggest that protonation of the 5′-phosphate at O5′ is more feasible. Nevertheless, the distances between the active site residues and terminal oxygen atoms in the 5′-phosphate group are shorter (up to 3.6 Å) than to O5′ (up to 4.6 Å) as discussed for the 3′-phosphate.

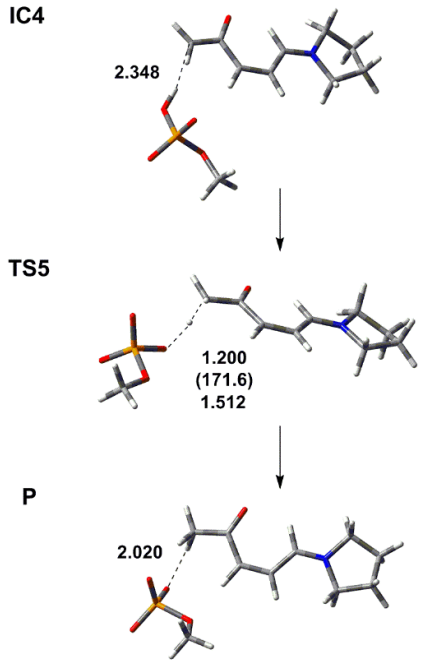

Figure 6.10. Structures characterized for the enol-keto rearrangement step. Important distances (Å) and angles (deg, in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) are provided.

According to the present model, the calculated barrier for the 5′-PO_4 elimination (166.5 – 188.6 kJ/mol) is significantly larger than anticipated for an enzyme-catalyzed reaction. Although this is likely in part due to the small size of the present model, which neglects additional interactions with the 5′-phosphate occurring in the FPG active site, the present results highlight the intrinsic lower reactivity of the C5′–O5′ bond compared to the C3′–O3′ bond. As discussed for the 3′-PO_4 elimination step, larger computational models are required to elucidate the reaction mechanism and the role of active site amino acids and/or solvent in this important process. Regardless, the calculated barrier obtained for the δ-elimination step is much larger than determined for the β-elimination step (19.9 kJ/mol). This predicted difference between the β- and δ-elimination reactions is in agreement with experimental evidence suggesting that δ-elimination is the rate-determining step for OG removal by FPG.[15] Furthermore, this observation may explain why bifunctional glycosylases that lack intensive interactions with 5′-PO_4 (such as hOgg1; Figure 1.12a) do not show δ-lyase activity.[12]

6.4 Conclusion

In the current chapter, the β- and δ-elimination reactions facilitated by FPG were examined. The model predicts that both the β- and δ-elimination reactions proceed through a step-wise (E1cB) mechanism. Attempts to characterize the transition state corresponding to phosphate elimination without protonation were unsuccessful, and the model predicts phosphate protonation and elimination to be simultaneous events. In the case of 5′-side cleavage, two nearly isoenergetic pathways for phosphate protonation/elimination were characterized, which involved either (intramolecular) proton transfer from O4′ or protonation assisted by a general acid (OG^- in the current model abstracts the O4′ proton and delivers the proton to the 5′-phosphate group). Regardless,

these calculations suggest that the additional charge developing on the phosphate moiety must be highly stabilized in the glycosylase active site in order for the reaction to occur. It is anticipated that the multiple active site interactions with the phosphate groups seen in experimental crystal structures of FPG bound to DNA will further lower the barriers calculated with this model. Interestingly, the current model indicates that O3′ and O5′ are intrinsically more likely to accept a proton than the terminal oxygen atoms of the phosphate group, which contradicts conclusions drawn from interaction distances measured in experimental crystal structures. Furthermore, the difference in energy barriers for the β- and δ-elimination reactions indicates that the chemistry of these two reactions are inherently different. Although the focus is placed on FPG, the model choice allows the results to be related to other bifunctional DNA glycosylases that use a proline residue to cleave the glycosidic bond, and that nick the DNA strand at the 3′- and 5′-sides of the damaged nucleoside (such as endonuclease VIII (Nei) and Neil). Moreover, since the nature of the amine nucleophile has been shown to have negligible effects on the mechanism and energetics of the β-elimination step in the BER pathway, current results for the 3′-phosphate elimination are valid for enzymes that employ a lysine nucleophile and exhibit only β-lyase activity (such as hOgg1, *E. coli* endonuclease III (Nth) and (human) hNTH1).

6.5 References

1. Zharkov, D. O.; Shoham, G.; Grollman, A. P., Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair* 2003, *2*, 839–862.

2. Fromme, J. C.; Verdine, G. L., Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nat Struct Mol Biol* 2002, *9*, 544–552.

3. Zharkov, D. O.; Golan, G.; Gilboa, R.; Fernandes, A. S.; Gerchman, S. E.; Kycia, J. H.; Rieger, R. A.; Grollman, A. P.; Shoham, G., Structural analysis of an Escherichia coli endonuclease VIII covalent reaction intermediate. *EMBO J.* 2002, *21*, 789–800.

4. Gilboa, R.; Zharkov, D. O.; Golan, G.; Fernandes, A. S.; Gerchman, S. E.; Matz, E.; Kycia, J. H.; Grollman, A. P.; Shoham, G., Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *J. Biol. Chem.* 2002, *277*, 19811–19816.

5. Shim, E. J.; Przybylski, J. L.; Wetmore, S. D., Effects of Nucleophile, Oxidative Damage, and Nucleobase Orientation on the Glycosidic Bond Cleavage in Deoxyguanosine. *J. Phys. Chem. B* 2010, *114*, 2319–2326.

6. Zheng, J.-H.; Tan, H.-W.; Chen, G.-J., Theoretical study on the mechanism of the DNA repair protein Fpg. *Int. J. Quantum Chem.* 2011, *111*, 2454–2463.

7. Sadeghian, K.; Flaig, D.; Blank, I. D.; Schneider, S.; Strasser, R.; Stathis, D.; Winnacker, M.; Carell, T.; Ochsenfeld, C., Ribose-Protonated DNA Base Excision Repair: A Combined Theoretical and Experimental Study. *Angew. Chem., Int. Ed. Engl.* 2014, *53*, 10044–10048.

8. Blank, I. D.; Sadeghian, K.; Ochsenfeld, C., A Base-Independent Repair Mechanism for DNA Glycosylase—No Discrimination Within the Active Site. *Scientific Reports* 2015, *5*, 10369.

9. Kellie, J. L.; Wetmore, S. D., Mechanistic and Conformational Flexibility of the Covalent Linkage Formed during β-Lyase Activity on an AP-Site: Application to hOgg1. *J. Phys. Chem. B* 2012, *116*, 10786–10797.

10. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., *et al. Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford CT, 2009.

11. Bruner, S. D.; Norman, D. P. G.; Verdine, G. L., Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature (London, U. K.)* 2000, *403*, 859–866.

12. Krishnamurthy, N.; Haraguchi, K.; Greenberg, M. M.; David, S. S., Efficient Removal of Formamidopyrimidines by 8-Oxoguanine Glycosylases. *Biochemistry* 2008, *47*, 1043–1050.

13. Fromme, J. C.; Verdine, G. L., DNA Lesion Recognition by the Bacterial Repair Enzyme MutM. *J. Biol. Chem.* 2003, *278*, 51543–51548.

14. Duclos, S.; Aller, P.; Jaruga, P.; Dizdaroglu, M.; Wallace, S. S.; Doublié, S., Structural and biochemical studies of a plant formamidopyrimidine-DNA glycosylase reveal why eukaryotic Fpg glycosylases do not excise 8-oxoguanine. *DNA Repair* 2012, *11*, 714–725.

15. Fedorova, O. S.; Nevinsky, G. A.; Koval, V. V.; Ishchenko, A. A.; Vasilenko, N. L.; Douglas, K. T., Stopped-Flow Kinetic Studies of the Interaction between Escherichia coli Fpg Protein and DNA Substrates. *Biochemistry* 2002, *41*, 1520–1528.

Chapter 7: Phosphate Elimination and Crosslink Hydrolysis Reactions

7.1 Introduction

As was mentioned in previous chapters, after the deglycosylation step, the backbone on the 3′- and 5′-sides of the lesion must be cleaved, either by a bifunctional glycosylase or an AP-endonuclease (Figure 1.7). While hOgg1 only facilitates the β-elimination reaction,[1] FPG catalyzes both the β- and δ-elimination reactions.[2] Moreover, deglycosylation is catalyzed by hOgg1 about 10-fold faster than β-elimination, with a "pause" between the two steps that results in "uncoupled" glycosylase and lyase activities.[3] In fact, studies have suggested that hOgg1 only facilitates the deglycosylation step and the observed lyase activity is a "secondary effect" of the presence of Lys249 in the active site.[4] In contrast, FPG catalyzes the deglycosylation and both elimination reactions with appreciable rates, with δ-elimination being considered to be the rate-limiting step.[5] Nevertheless, both enzymes process an abasic site (AP-site, Figure 7.1c – d), with the DNA backbone being nicked at the 3′- or both the 3′- and 5′-sides by hOgg1 and FPG, respectively.[6, 7]

Despite some proposals for the mechanism of the elimination reactions as discussed in Chapters 1, 5, and 6, several details regarding the lyase activities of hOgg1 and FPG are currently unknown. For example, the identity of the base that initiates the reaction by abstracting a proton (C2′–H or C4′–H during β- or δ-elimination catalysis, respectively) is still unidentified for both enzymes. Previous proposals for the general base (*e.g.,* cleaved OG for hOgg1[8] or a water molecule for FPG)[9, 10] cannot be applied for all bifunctional glycosylases (Chapter 1) or do not explain the lyase activities of hOgg1 and FPG on abasic sites. Moreover, OG has been suggested to inhibit elimination catalysis for hOgg1.[11] In Chapter 6, it was shown that the sequential β- and δ-elimination

reactions exhibit intrinsic differences for the employed model. Moreover, the level of stabilization of the leaving group (the 3′- and 5′-phosphates) may explain the differential backbone cleavage rates in hOgg1 and FPG (Figure 1.13).

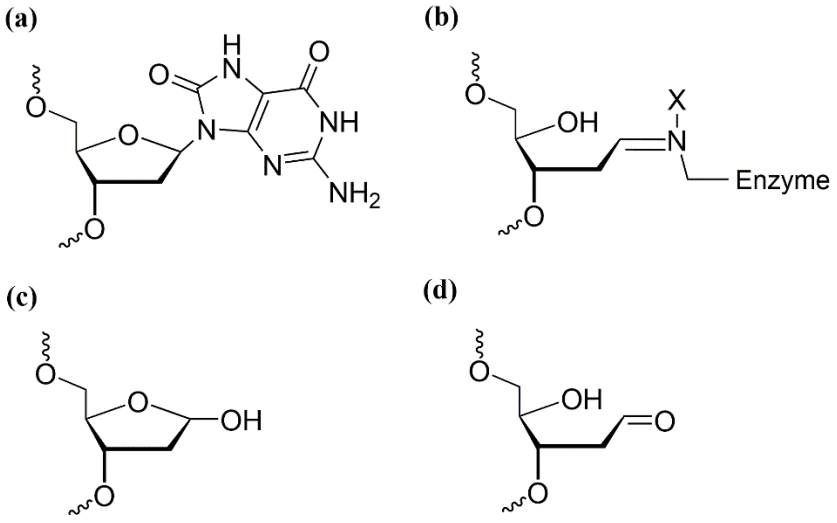

Figure 7.1. Schematic representation of (a) 2′-deoxy-8-oxoguanosine (dOG). (b) Schiff base crosslink between DNA and bifunctional glycosylase, with X being either an electron lone-pair or a hydrogen atom for hOgg1, and CH_2 of Pro2 for FPG. (c) Ring-closed AP-site (product of a monofunctional glycosylase). d) Ring-opened AP-site.

To address the aforementioned uncertainties, this chapter employs various truncated models of the Schiff base crosslink, as well as an abasic site (Figure 7.1d), relevant to the hOgg1 and FPG active sites, and has examined several pathways for the elimination reactions and the hydrolysis of the DNA–enzyme crosslinks. Specifically, the order and mechanism of the elimination and hydrolysis reactions on crosslinks with different protonation states relevant to hOgg1 and FPG are examined. The results of the present chapter suggest that the protonation state of the crosslink, order of the elimination and hydrolysis steps, and the mechanism of the elimination are different for hOgg1 and FPG. Moreover, this chapter suggests a general base for the elimination reaction in hOgg1, explains the lower efficiency of hOgg1 for the β-lyase step, and rationalizes the lack of δ-lyase activity. Although the truncated models used in this study do not fully

reflect the substrate–active site interactions, they allow for investigation of a large number of possible reactions. MD simulations were also used to evaluate the reactant complex relevant to the newly proposed β-elimination reaction for hOgg1.

7.2 Computational Details

Previous computational studies have considered the lysine-crosslink associated with hOgg1 catalysis to be cationic after the deglycosylation and ring-opening steps.[12-14] However, two recent studies have proposed new mechanisms for the deglycosylation reaction catalyzed by hOgg1,[15, 16] and predict the lysine-crosslink intermediate to be neutral once the deoxyribose ring is opened and OG is removed. On the other hand, the proline-crosslink associated with FPG catalysis is necessarily cationic after deglycosylation and ring-opening steps due to the number of covalent bonds the nitrogen of the proline ring forms in the Schiff base (Figure 7.1b).[17, 18]

In the present chapter, both cationic and neutral lysine-crosslinks, and the cationic proline-crosslink are considered to investigate the elimination and hydrolysis reactions catalyzed by hOgg1 and FPG, respectively. For the hydrolysis and β-elimination reactions, the crosslink models were capped by 5′-OCH_3 and 3′-$CH_3OPO_3^-Na^+$ group (Figure 7.2), while for the δ-elimination reaction 5′-$CH_3OPO_3^-Na^+$ models were used. Since the δ-elimination takes place after the β-elimination step, the models used for the δ-elimination reaction were not capped at the 3′-side. The general acid/bases modelled as OG, or the sidechain of Asp, Glu, or Lys in different protonation states are introduced in Figure A7.1. Based on the protonation state of the crosslink, pathways are divided into three main groups, namely cationic lysine-crosslink (group I, 3 routes, Figure 7.2), cationic proline-crosslink (group II, 2 routes, Figure 7.3), and neutral lysine-crosslink (group III, 3 routes, Figure 7.4). Pathways within a route are denoted as P followed by the

number of the pathway. For each group and route, reaction mechanisms (pathways) are shown separately in Figures 7.5 – 7.12. Note that some reactions take place in more than one pathway, and thus some steps in different pathways have the same barriers.

The potential energy surfaces (PESs) for the reactions were mapped with a similar procedure described in Chapters 5 and 6. All geometries were obtained at the B3LYP/6-31G(d) level of theory in the presence of bulk solvent as described by the IEF-PCM (ε = 4.2) method implemented in Gaussian 09 (Revision D.01).[19] Unless otherwise mentioned, all reported energies correspond to relative Gibbs free energies (including unscaled thermal corrections) obtained from single-point calculations at the SMD-M06-2X/6-311+G(2df,2p) level of theory. Cartesian coordinates of the transition state structures and the corresponding reactant complexes are provided in Appendix A7.

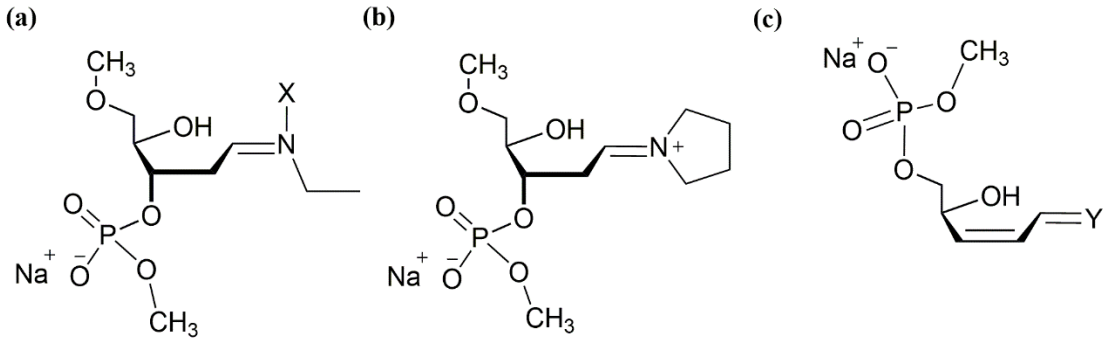

Figure 7.2. Schematic representation of the initial models employed in the present chapter for the (a) lysine crosslink, with X being either an electron lone-pair or a hydrogen atom, (b) proline crosslink, and (c) model used for the δ-elimination step, with Y corresponding to lysine or proline crosslink, or oxygen atom.

A similar protocol as described in Chapter 3 was employed for the MD simulations. Initial structures for the MD simulations were built using a high resolution X-ray crystal structure of hOgg1 bound to an abasic site (abasic furan containing DNA, PDB ID: 1N3A).[20] The abasic furan was replaced with a ring-opened AP-site. Protonation states of ionizable residues were assigned using PROPKA 3.0,[21] with the exception of key

active site residues. Specifically, His270 was considered to be protonated, while Lys249 and Cys253 were modelled as neutral. Asp268 was considered to be either anionic or neutral. All systems were assigned AMBER ff14SB[22] parameters, while the non-standard ring-opened AP-site was assigned GAFF parameters[23, 24] using the ANTECHAMBER module of Amber 12.[25] Restrained electrostatic potential (RESP) charges for the non-standard residue was assigned using the R.E.D.v.III.4 program,[26] with the geometry optimized at the B3LYP/6-31G(d) level of theory using Gaussian 09 (revision D.01).[19] The production phase included a 20 ns MD simulation without constraints under NPT conditions (1 atm and 300 K). The periodic boundary condition was employed for all steps. The root-mean-square deviations (rmsds) for the protein and DNA backbone over the production phase were analyzed for each simulation to ensure the systems were stable (Table A7.1).

7.3 Results

7.3.1 Group I. Cationic Lysine-crosslink (hOgg1)

The β-elimination reaction for cationic lysine-crosslink begins with C2′–H abstraction and may proceed through an intermediate when catalyzed by OG or Asp, or a simultaneous 3′-phosphate protonation/departure when water initiates the catalysis (Figure 7.3). Two routes were considered following the β-elimination reaction: 1) hydrolysis of the crosslink or 2) δ-elimination (Figure 7.3, R1 and R2, respectively). The hydrolysis reaction (R1) occurs in two steps. In the first step, a base (modelled as Asp^{-1}) is required to abstract a proton from a nucleophilic water, and the resulting hydroxide ion attacks the crosslink at C1′, which breaks the π bond between C1′ and Nζ of lysine. In the second step, C1′–OH transfers a proton to Nζ of lysine via a water molecule, which completes the hydrolysis reaction (Figure 7.3, R1).

The δ-elimination reaction (Figure 7.3, R2) is initiated by a water molecule, and includes simultaneous proton abstraction (C4′–H) and 5′-phosphate protonation/departure similar to that discussed for the β-elimination. Once δ-elimination is complete, the crosslink is hydrolyzed through two steps that are similar to those discussed for R1.

The C2′–H abstraction step facilitated by Asp^{-1} is barrierless (Table A7.2), while the activation energy of this step catalyzed by OG^{-1} is ~ 17.7 – 54.3 kJ/mol depending on the identity of the OG^{-1} atom (N9 or O8) that receives the proton (Table A7.3).[12] Protonation of 3′-PO_4 leads to a lower barrier for the β-elimination step (76.8 kJ/mol) compared to the elimination of an anionic phosphate group (77.2 – 123.0 kJ/mol, Table A7.3).[12] The Gibbs energies associated with the two hydrolysis steps are 61.5 and 55.9 kJ/mol, respectively (Table A7.2). The energy barriers for the β- and δ-elimination reactions catalyzed by water (Figure 7.2, R2) are 105.2 and 107.7 kJ/mol, respectively (Table A7.2). Both hydrolysis steps after the δ-elimination have larger energy barriers (75.6 and 61.6 kJ/mol) compared to the hydrolysis after the β- but prior to the δ-elimination step. Efforts to characterize a transition state for hydrolysis of the crosslink facilitated by Asp^{-1} and H_2O were unsuccessful. In fact, when two water molecules were used to catalyze the hydrolysis, an addition reaction (Figure A7.2) with a low barrier (ΔG = 48.4 kJ/mol, Table A7.4) took place. The resulting intermediate from this addition reaction does not undergo hydrolysis, likely because the tetra-coordinated nitrogen (Nζ of Lys) cannot undertake concurrent bond breaking and forming. Moreover, a water catalyzed β-elimination from the intermediate has a barrier of 230.6 kJ/mol (Table A7.4). Even after the elimination reaction, hydrolysis of the crosslink could not be characterized.

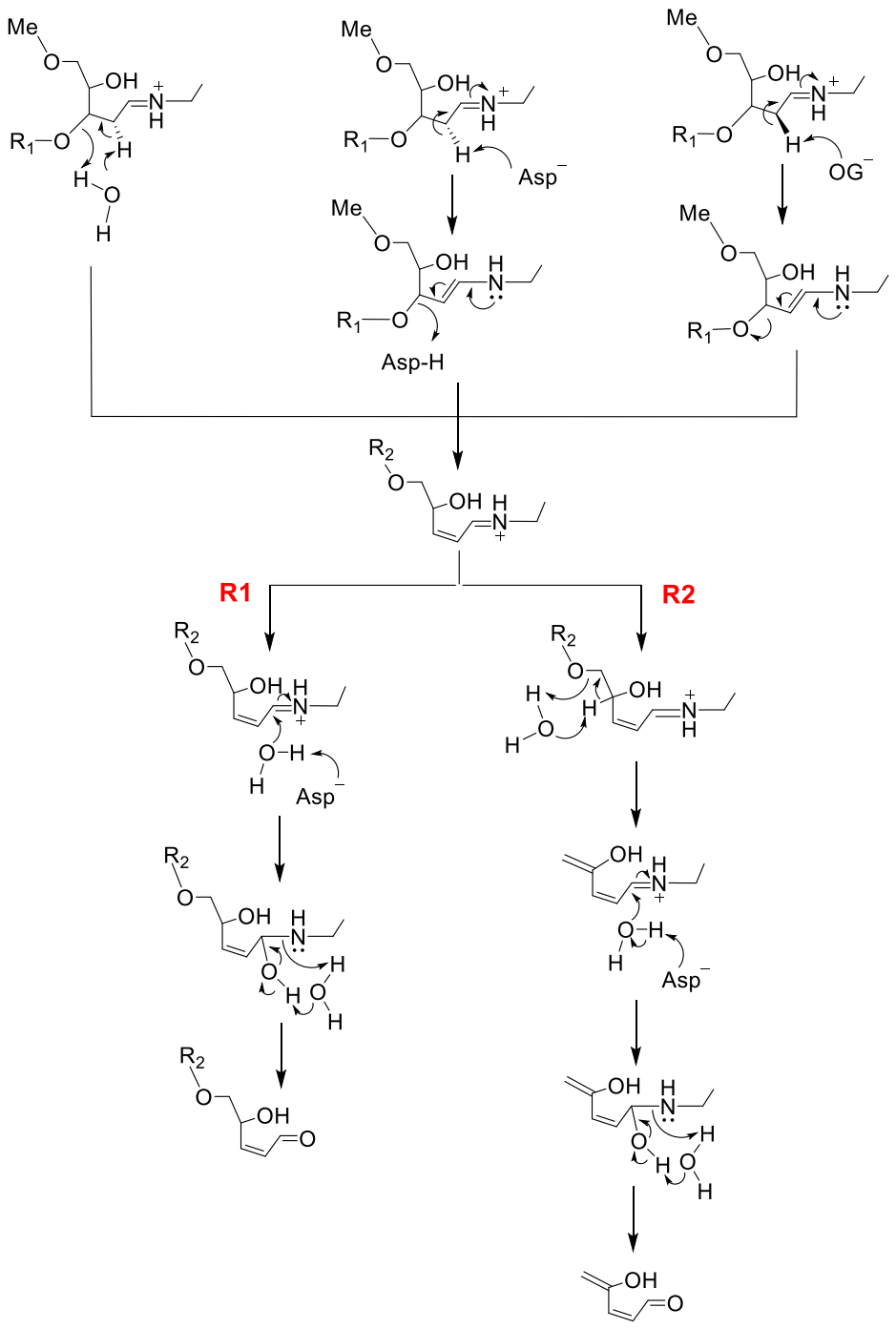

Figure 7.3. Reaction mechanisms considered for the cationic lysine-crosslink (group I), with R_1 and R_2 = $Na^+CH_3PO_4^-$. The Gibbs energies are provided in Table A7.2. Modelled acids and bases are defined in Figure A7.1.

7.3.2 Group II. Cationic proline-crosslink (FPG)

Since the cationic crosslink associated with lysine was shown to be a poor substrate for hydrolysis (group I), only elimination reactions were considered for the proline-crosslink (Figure 7.4). Previous works have predicted the barrier for C2′–H abstraction by OG^{-1} to be 21.5 – 51.5 kJ/mol, depending on the identity of the OG^{-1} atom (N9 or O8) that accepts the proton (Table A7.3).[27, 28]

As in the cationic lysine-crosslink, the β-elimination catalyzed by water for the proline-crosslink occurs in one step (Figure 7.4). This step can be followed by a two-step hydrolysis of the crosslink, and subsequent water-catalyzed δ-elimination (Figure 7.4, R1). Alternatively, the δ-elimination may take place prior to the hydrolysis of the crosslink (Figure 7.4, R2). Similar to the hydrolysis of the cationic lysine-crosslink, a base (modelled as Glu^{-1}) must abstract a proton from a water molecule to initiate the first step of the hydrolysis (Figure 7.4). Enol-keto rearrangement may take place after the hydrolysis (Figure 7.4, R1 and R2-P1) or after the δ-elimination but before the hydrolysis (Figure 7.4, R2-P2). In fact, after the hydrolysis is complete, the rearrangement is enzymatically irrelevant since the substrate has been processed.

The Gibbs energy of the β-elimination reaction (105.7 kJ/mol, Table A7.5) is almost identical to that for the cationic-lysine crosslink (105.2 kJ/mol, Table A7.2). However, the δ-elimination has a higher barrier (121.9 kJ/mol, Table A7.5) compared to that for the lysine counterpart (107.7 kJ/mol, Table A7.2). The δ-elimination step after the hydrolysis of the crosslink has an even higher barrier (175.7 kJ/mol, Table A7.5). The first and second steps of the hydrolysis before the δ-elimination have barriers of 56.5 and 44.8 kJ/mol, respectively. The barriers for these reactions after the δ-elimination step are 48.1 and 61.0 kJ/mol, respectively (Table A7.5). The enol-keto step prior to the

hydrolysis lowers the barrier for the first step of hydrolysis (20.1 kJ/mol) while the barrier for the second step lies between the barriers for the other two pathways (55.8 kJ/mol, Table A7.5). Nevertheless, since the barrier for the rearrangement is relatively high (148.2 kJ/mol, Table A7.5), it is unlikely to take place prior to the hydrolysis.

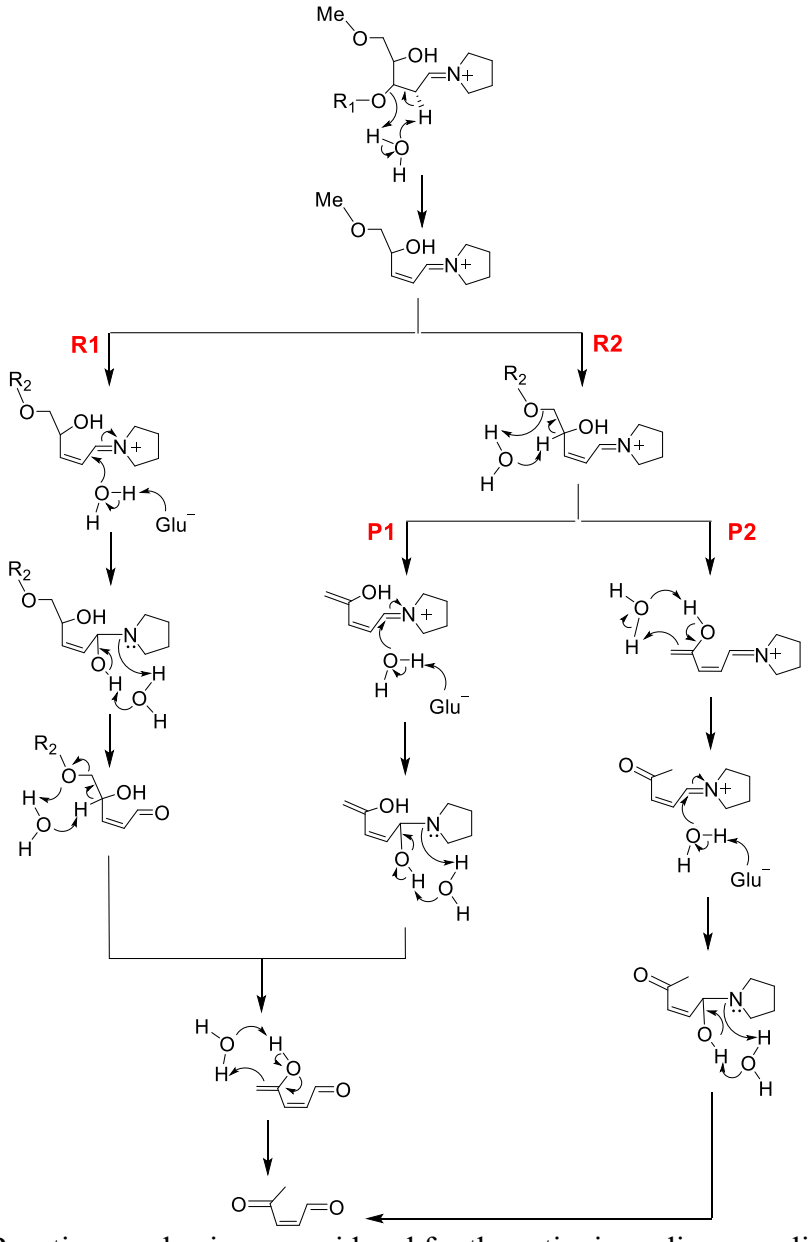

Figure 7.4. Reaction mechanisms considered for the cationic proline-crosslink (group II) with R_1 and $R_2 = Na^+CH_3PO_4^-$. The Gibbs energies are provided in Table A7.5. Modelled acids and bases are defined in Figure A7.1.

7.3.3 Group III. Neutral lysine-crosslink (hOgg1)

The neutral lysine-crosslink can undergo various reactions, which are categorized in three main routes summarized in Figure A7.3.

7.3.3.1 Route 1. P1 and P2 (Figure 7.5): The first route includes a β-elimination step catalyzed by water or neutral Asp, which simultaneously abstracts the C2′–H proton (*Pro*-R) and protonates O3′ of the 3′-PO_4, facilitating phosphate departure. Alternatively, OG^{-1} abstracts the *Pro*-S proton, which leads to simultaneous phosphate departure (Figure 7.5). The resulting intermediate is hydrolyzed (Figure 7.5, R1-P1) or undergoes water-catalyzed δ-elimination prior to a two-step hydrolysis reaction (Figure 7.5, R1-P2).

The Gibbs energy barriers for the β-elimination catalyzed by water or Asp^{-1} are 258.9 or 233.8 kJ/mol, respectively (Table A7.6). Using two bridging water molecules for the C2′–H abstraction and phosphate protonation steps decreased the barrier slightly (241.4 kJ/mol, Table A7.6). The δ-elimination reaction catalyzed by water has a comparable barrier (235.2 kJ/mol, Table A7.6). Interestingly, the barrier for the β-elimination catalyzed by OG^{-1} is only 115.8 kJ/mol, which indicates that OG could act as a cofactor for the β-elimination step if deglycosylation results in a nucleobase anion. However, the barrier when neutral OG is the base is 207.4 kJ/mol (Table A7.3), which makes this reaction unlikely to occur. The first step of the hydrolysis has a similar barrier regardless of whether it occurs before (122.2 kJ/mol) or after (121.3 kJ/mol, Table A7.6) the δ-elimination reaction, while the barrier for the second step is slightly smaller before (55.9 kJ/mol) rather than after (61.6 kJ/mol, Table A7.6) the δ-elimination.

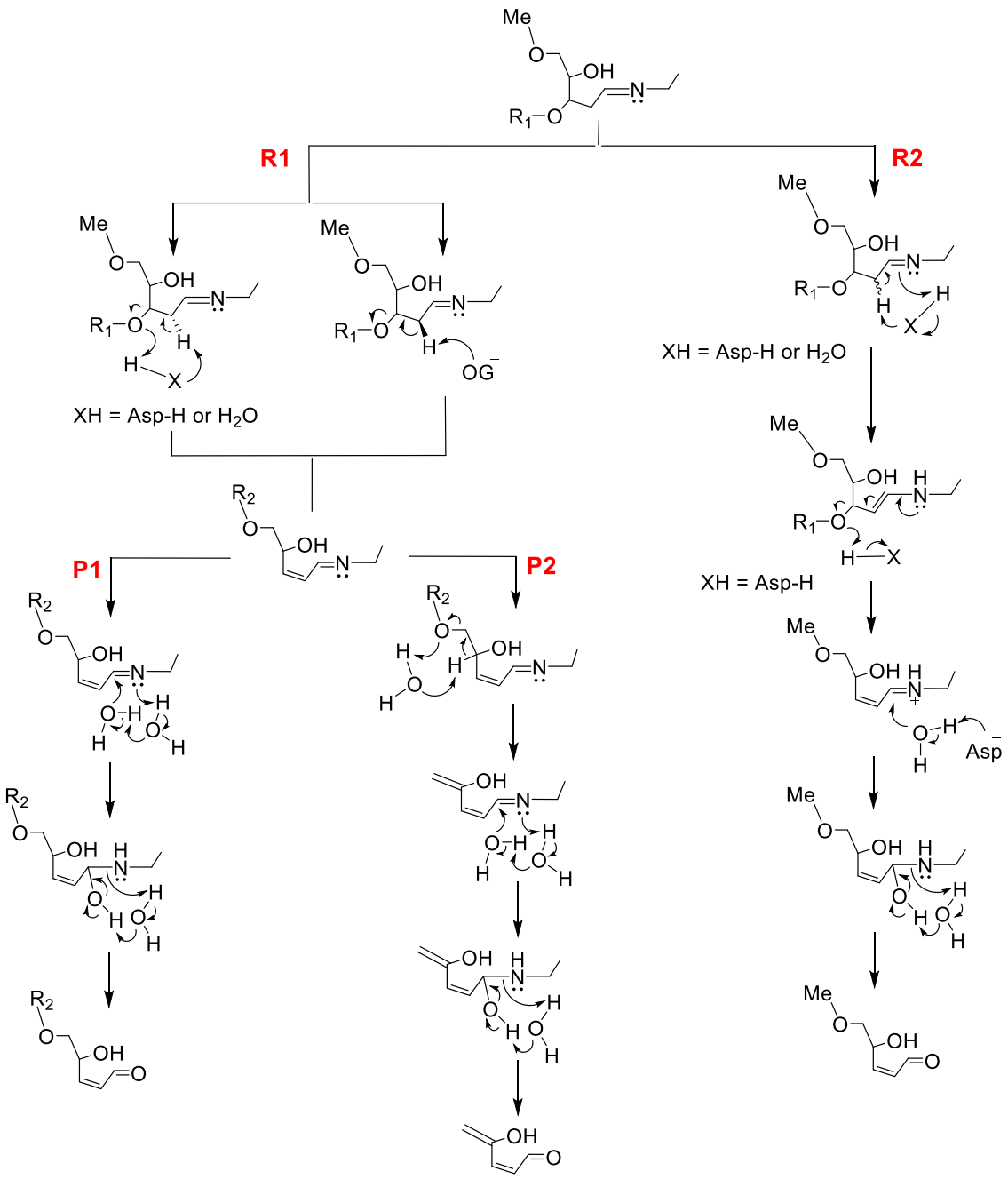

Figure 7.5. Reaction mechanisms considered for the neutral lysine-crosslink (group III: left R1-P1 and right R1-P2) with R_1 and R_2 = $Na^+CH_3PO_4^-$. The Gibbs energies are provided in Table A7.6. Modelled acids and bases are defined in Figure A7.1.

7.3.3.2 Route 2 (Figure 7.5): This route begins with simultaneous C2′–H abstraction and protonation of the crosslink by Asp or water. Subsequent phosphate departure takes place with or without receiving a proton (from Asp–H, with only the former shown in Figure

7.5). The resulting cationic crosslink undergoes a two-step hydrolysis. The first step requires water to lose a proton to a base (Asp^{-1}), which is followed by subsequent OH^- attack at C1′ of the crosslink. The final step includes breaking the bond between Nζ of lysine and C1′ of the ring-opened deoxyribose, which is catalyzed by solvent.

The barrier for the abstraction/protonation is almost two times higher when catalyzed by water (143.6 kJ/mol, Table A7.7) compared to Asp (75.8 kJ/mol, Table A7.7). Moreover, in comparison with direct elimination (122.4 kJ/mol), the 3′-PO_4 departure has a lower barrier when activated by Asp–H (76.8 kJ/mol). The first and second steps of the hydrolysis have energy barriers of 61.5 and 55.9 kJ/mol, respectively (Table A7.7).

7.3.3.3 Route 3. P1 (Figure 7.6): R3-P1 begins with hydrolysis. Subsequently, C2′–H abstraction takes place by Asp^{-1} or water, facilitating the 3′-PO_4 departure. Removing C2′–H by Asp^{-1} leads to phosphate departure without protonation/activation, while the C2′–H abstraction and 3′-PO_4 protonation take place simultaneously when catalyzed by water. Hydrolysis of the crosslink follows the elimination reaction and completes the pathway.

The barrier for the first step of hydrolysis step is 117.5 kJ/mol (Table A7.8). Interestingly, the barrier for elimination without protonation (Asp^{-1} catalyzed) is 187.1 kJ/mol, while the water-catalyzed process has a barrier of 261.0 kJ/mol (Table A7.8). The elimination reaction is followed by the second step of hydrolysis (55.9 kJ/mol, Table A7.8), which completes the pathway.

7.3.3.4 Route 3. P2 (Figure 7.6): In R3-P2, after the first hydrolysis step, the crosslink receives a proton from water upon C2′–H abstraction. This reaction completes the crosslink hydrolysis in a different way compared to previous pathways, leaving an enol

moiety to undergo the elimination reaction (Figure 7.6). In the final step, O3′ of the phosphate receives a proton from C1′–OH and departs. The simultaneous C2′–H abstraction and crosslink protonation reaction appears to be the rate-limiting step (206.2 kJ/mol, Table A7.9), while the phosphate departure has a relatively small barrier (84.1 kJ/mol, Table A7.9).

7.3.3.5 Route 3. P3 (Figure 7.6): In this pathway, the crosslink is fully hydrolyzed by two water molecules in a two-step process (Figure A7.4). The resulting intermediate is a ring-opened keto-abasic site, while lysine dissociated in a neutral protonation state (Figure 7.6). The next step is the β-elimination catalyzed by water (1 step), Asp^{-1} (2 steps), or Lys (2 steps). The δ-elimination reaction is also facilitated by a water molecule after the β-elimination is completed.

The C2′–H abstraction (*pro*-R) by Asp^{-1} has a barrier of 13.1 kJ/mol, while Lys abstracts either the *pro*-S (C2′–H in *exo*-ring-opened) or *pro*-R (C2′–H in *endo*-ring-opened) with associated barriers of 93.3 or 61.2 kJ/mol, respectively (Figure 7.6 and Table A7.10). The Gibbs energies associated with subsequent protonation of 3′-PO_4 is 32.1 for Asp and 91.0 or 49.0 kJ/mol for Lys^{+} following the *pro*-S or *pro*-R proton abstraction, respectively (Figure A7.5 and Table A7.10). Not only are the catalytic reactions (deglycosylation and the 3′-side backbone cleavage) complete at this step, but the enzyme is also no longer cross-linked to the processed substrate and therefore is ready to dissociate (Figure A7.6, IC4). Both the water-catalyzed β- (233.5 kJ/mol) and δ-elimination (175.7 kJ/mol, Table A7.10) reactions have large barriers, which are only decreased to 191.4 kJ/mol using two water molecules. These barriers indicate that one-step elimination reactions on a ring-opened abasic site are very unlikely.

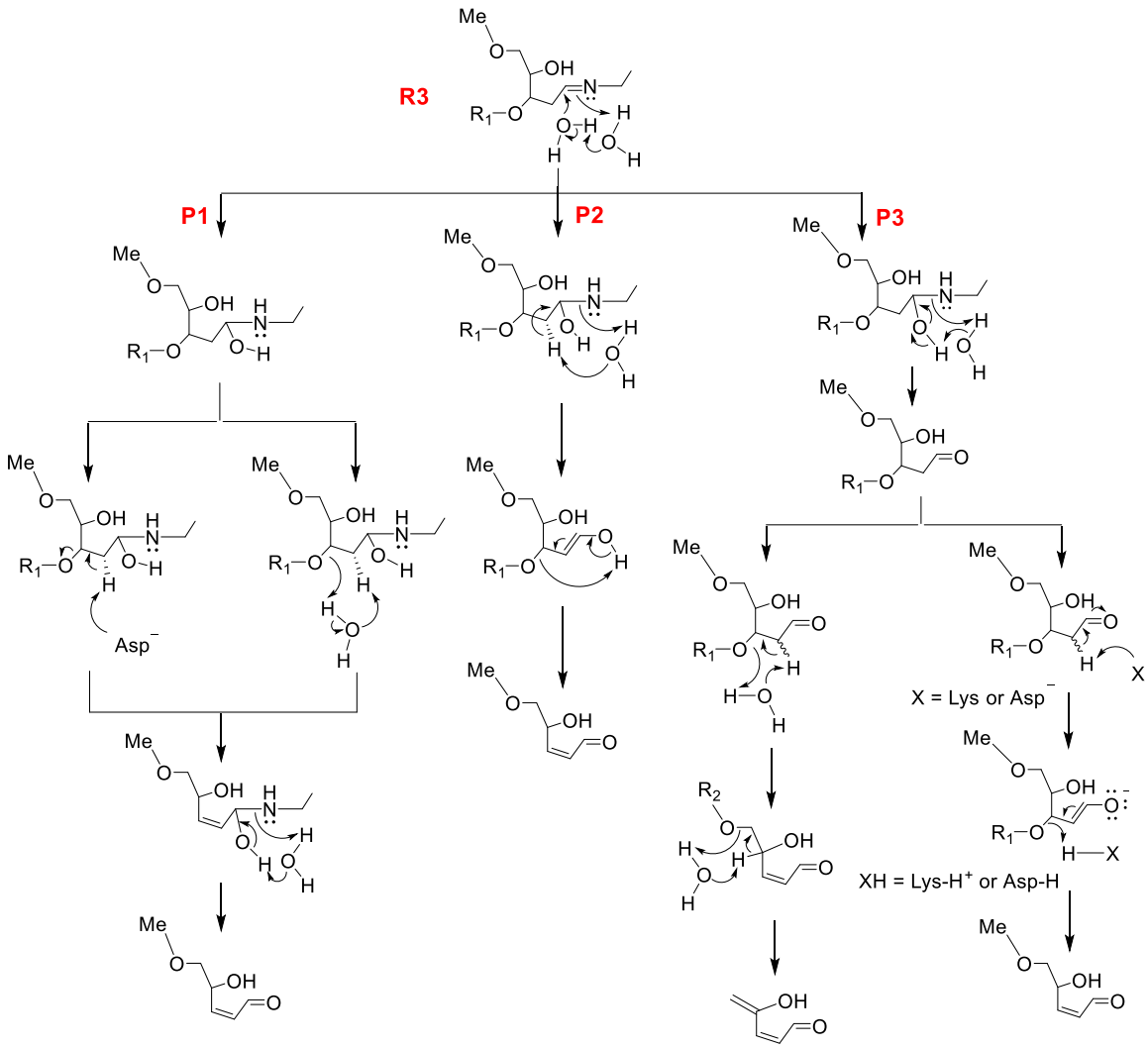

Figure 7.6. Reaction mechanisms considered for the neutral lysine-crosslink (group III: R3) with R_1 and $R_2 = Na^+CH_3PO_4^-$. The Gibbs energies are provided in Table A7.7. Modelled acids and bases are defined in Figure A7.1.

To examine the feasibility of the R3-P3 mechanism (Figure 7.6), MD simulations were carried out on the abasic site bound in the hOgg1 active site with anionic or neutral Asp268, and neutral Lys249. The MD representative structures show the positions of Lys249 and Asp268 relative to the substrate (Figure 7.7). With anionic Asp268, the average distance between Lys(Nζ) and the *pro*-S or *pro*-R (C2′–H) is 3.489 (±0.641) or 3.320 (±0.469) Å, respectively. When neutral Asp268 was considered, the Lys249(Nζ)···*pro*-R (C2′–H) distance is 3.780 (±0.678) Å, while the orientation of the

pro-S (C2′–H) relative to Lys249(Nζ) makes the abstraction unlikely. In both cases, Asp268 is too far from the C2′ hydrogens (Figure 7.7). The DFT reactant complex structures for the elimination reaction catalyzed by Lys (Figure A7.5, R3) resemble the MD representative structures (Figure 7.7). The δ-elimination reaction is considered to be catalyzed by a water molecule since neither Lys nor Asp are aligned to initiate the elimination by abstracting the C4′–H. The Gibbs energy of the one-step water-catalyzed δ-elimination reaction is large (175.7 kJ/mol, Table A7.10). When coupled with the deficient stabilizing interactions for the 5′-PO_4 in hOgg1, this provides further explanation for the lack of δ-lyase activity of hOgg1.

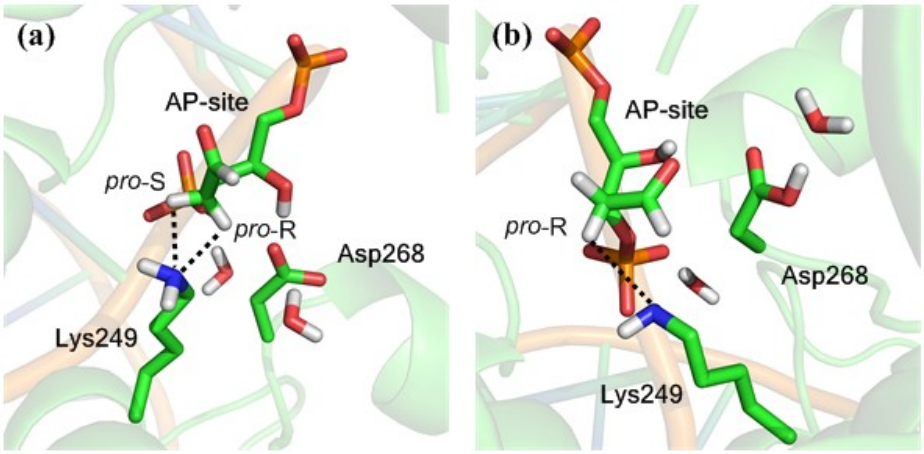

Figure 7.7. Representative structure of hOgg1 active site bound to the ring-opened AP-site with (a) anionic Asp268 and (b) neutral Asp268. In structure (a) the average C2′–H···Lys249(Nζ) distance is 3.489 (±0.641) and 3.320 (±0.469) Å for *pro*-S or *Pro*-R proton, respectively. In structure (b) the average *pro*-R···Lys249(Nζ) distance is 3.780 (±0.678) Å.

7.4 Discussion

The neutral and cationic lysine-crosslink, and the cationic proline-crosslink were employed to examine the hydrolysis and elimination reactions in an attempt to unveil the reaction mechanisms relevant to the elimination reactions facilitated by hOgg1 and FPG. By comparing the energy barriers and kinetic data, as well as active site structural data,

some pathways can be ruled out, leaving only one catalytic reaction mechanism that is more likely to take place for each enzyme.

7.4.1 Cationic Lysine- and Proline-crosslinks

All attempts to hydrolyze the cationic crosslinks prior to the β-elimination step were unsuccessful. Instead, an addition reaction occurs, which was only considered for lysine-crosslink (Figure A7.2) since the β-elimination reaction from the resulting intermediate occurs with a large barrier (230.6 kJ/mol, Table A7.4). On the other hand, both cationic lysine- and proline-crosslinks easily undergo the β-elimination reaction (Tables A7.2 and A7.5). The Gibbs energy of the one-step water-catalyzed β-elimination is almost identical for both crosslinks (~ 105 kJ/mol). When catalyzed by OG^{-1}, the barrier for C2′–H abstraction is 17.7 – 54.3 or 21.5 – 51.5 kJ/mol for cationic lysine- or proline-crosslink, respectively.[12, 27, 28] Although this step is barrierless with Asp^{-1}, the position of Asp^{-1} relative to the substrate in the active site makes this reaction unlikely. The Gibbs energy of the δ-elimination reaction catalyzed by water indicates the feasibility of this step for both the lysine- and proline-crosslinks (107.7 and 121.9 kJ/mol, respectively). This is in line with the δ-lyase activity for FPG, but cannot explain the experimentally-observed lack of δ-lyase activity for hOgg1, and thus this pathway can be ruled out for hOgg1.

7.4.2 Neutral lysine-crosslink

The present results reveal that an elimination reaction from a neutral lysine-crosslink (Figure 7.5, R1) or the intermediate resulting from a first hydrolysis step for the neutral crosslink hydrolysis (Figure 7.6, R3-P1) is highly restricted due to the large energy barrier (> 200 kJ/mol, Tables A7.6 and A7.8). This could arise for two reasons:

First, unlike the cationic crosslink, neither the neutral Schiff base nor the intermediate can accommodate the developing negative charge upon C2′ proton abstraction, and hence the C3′–O3′ bond necessarily breaks in chorus with the proton abstraction (even if the 3′-PO_4 is not activated/protonated). Second, concurrent C2′–H abstraction and phosphate departure requires a general base to abstract the *pro*-S (C2′–H) to fulfill the stereochemistry of an E2 elimination reaction. Thus, the s*yn* elimination reactions in the R1 or R3-P1 mechanisms have remarkably large barriers. In R1 (Figure 7.5), the cleaved OG could be a candidate to abstract the *pro*-S proton of C2′. Although this mechanism may satisfy the proposed role of OG as a cofactor in the elimination reaction catalyzed by hOgg1,[9, 10] OG must be anionic in order to be an efficient base. However, the present results indicate that neutral OG cannot efficiently catalyze this step due to the large barrier (> 200 kJ/mol, Table A7.3). Moreover, according to recent computational studies, the cleaved OG is either cationic[15] or neutral.[16] Furthermore, it has also been suggested that OG inhibits the catalysis of the elimination reaction by hOgg1.[9] This may be related to the fact that the cleaved OG lingers in the hOgg1 active site,[9] (unlike the case of FPG)[8] likely making the active site less accessible to the solvent molecules required for the hydrolysis. Thus, a catalytic process promoted by OG can be doubted. Additionally, the δ-elimination reaction is highly unlikely to take place through a R3-P3 mechanism due to the large energy barrier (235.2 kJ/mol, Table A7.6).

Another pathway considered for the neutral lysine-crosslink begins with C2′–H abstraction and protonation of the crosslink by either a water molecule or Asp (Figure 7.6). While the barrier for this step is large when involving water (143.6 kJ/mol, Table A7.7), the barrier is only 75.8 kJ/mol when Asp is the general base/acid. However,

catalysis by Asp is unlikely due to the required significant dislocation of Asp relative to the crystallographic position in the anticipated reactant complex.

The simultaneous C2′–H abstraction and protonation of the crosslink after the first hydrolysis step (Figure 7.6, R3-P2) is also unlikely due to the large associated energy barriers (> 200 kJ/mol, Table A7.9). In contrast, complete hydrolysis of the crosslink (Figure 7.6, R3-P3) results in an intermediate that undergoes a β-elimination reaction catalyzed by Asp^{-1} or Lys. The Gibbs energy of the elimination facilitated by water is large (~ 294 or 191 kJ/mol with one or two water molecules, respectively), mainly due to the unfavorable *syn* E2 elimination. Although in comparison with water- and Lys-catalyzed elimination both the abstraction (~ 13 kJ/mol) and elimination (~ 32 kJ/mol) barriers are lower when catalyzed by Asp, the position of Asp with respect to the ring-opened abasic site and particularly relative to the C2′–H has to change significantly relative to the crystal structure position of the reactant complex. This considerable movement is necessary for both steps, which makes this mechanism less likely to occur. Moreover, C2′–H abstraction by Asp^{-1} that is mediated by a water molecule could not be characterized. Alternatively, the dissociated Lys249, which is in close proximity to both C2′–H and O3′, is able to promote the β-elimination reaction. Once the C2′–H is abstracted, the proton can be shuttled to O3′, facilitating the 3′-side cleavage (Figure A7.5, IC3). This finding is in line with the results of Chapter 6, where it was shown the protonation of the 3′-phosphate group takes place at O3′ rather than a terminal oxygen. Therefore, considering the energy barriers and structural data, initial hydrolysis of a neutral lysine-crosslink is the most likely mechanism (R3-P3) among all pathways considered.

7.4.3 Enzymatic Implications

As discussed in the previous section, the neutral lysine-crosslink is a better substrate for the hydrolysis reaction, while the cationic lysine- and proline-crosslink undergo the β-elimination reaction and cannot be hydrolyzed prior to the β-elimination step. It is timely to emphasize that including the enzyme active site residues could affect the barriers and possibly the catalytic mechanism, and the present discussion is mainly focused on the inherent chemical properties and feasibility of reaction mechanisms based on the protonation states of the crosslinks relevant to hOgg1 and FPG. However, the similar energy barriers for the cationic lysine- and proline-crosslinks might be indicative that – in agreement with recent computational studies[15, 16] – the lysine-crosslink is neutral rather than cationic.

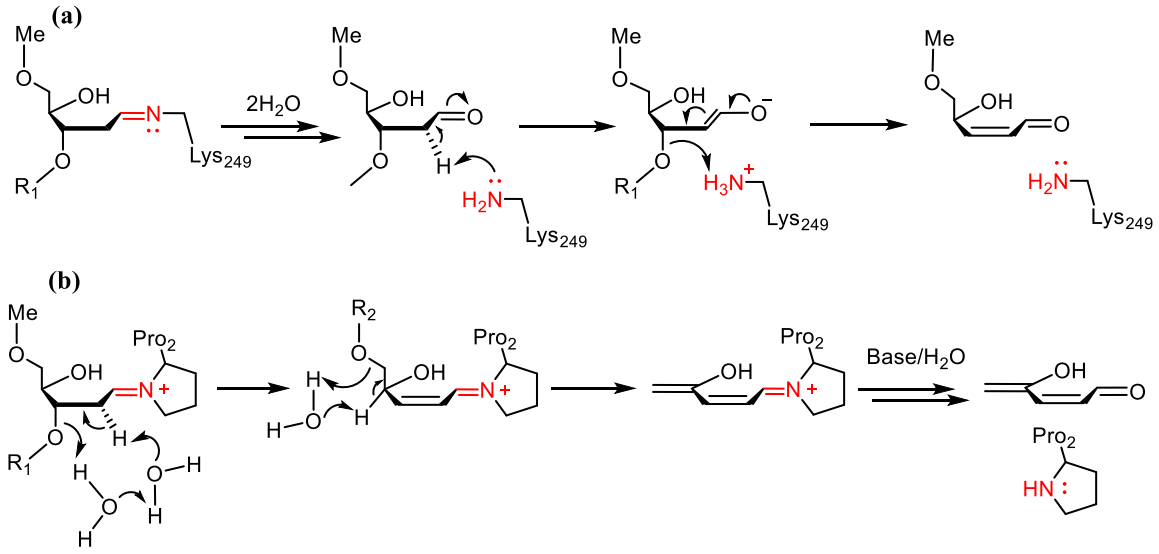

Figure 7.8. Reaction mechanism proposed in the present chapter for (a) the β-lyase catalyzed by hOgg1 and (b) the β- and δ-lyase catalyzed by FPG, with R_1 and R_2 = $Na^+CH_3PO_4^-$.

Figure 7.8 compares the reaction mechanisms proposed in this chapter for (a) a neutral and (b) a cationic crosslink. The reactions from a neutral crosslink undergoes: (1) a 2-step hydrolysis, (2) C2′–H abstraction by Lys249, and (3) protonation and elimination

of the 3′-PO_4 by Lys–H^+. On the other hand, the cationic crosslink undergoes: (1) a 1-step water-catalyzed β-elimination, (2) a 1-step water-catalyzed δ-elimination, and (3) a 2-step hydrolysis of the crosslink. Thus, not only the orders of the elimination and hydrolysis reactions are reversed for neutral and cationic crosslinks, but the β-elimination reaction mechanisms differ (*i.e.,* 1-step *syn* E2 for cationic crosslink versus a sequential proton abstraction and the C3′–O3′ bond cleavage (E1cB) for the neutral crosslink). The proposed β-elimination mechanism facilitated by Lys249 as the general base (hOgg1) could be experimentally tested using deuterium tagged 2′-dOG (at C2′) and Lys249 (at Nζ). Specifically, once the elimination reaction is completed, the cleaved 3′-phosphate must have a deuterium atom.

7.5 Conclusion

In this chapter, several pathways were studied to evaluate the order and mechanism of the hydrolysis and elimination reactions initiated from different Schiff base intermediates that are relevant to the reactions catalyzed by FPG and hOgg1 during BER pathway. By comparing the energy barriers of the reactions, and the structural and kinetic experimental data, some reaction pathways were determined to be kinetically unfeasible, and new mechanisms are proposed for the lyase activity of hOgg1 and FPG. Specifically, the different protonation states of the Schiff base crosslinks (cationic relevant to FPG and neutral for hOgg1) result in different reaction pathways, with the cationic crosslink undergoing elimination reaction and the neutral crosslink favoring hydrolysis. Hydrolysis of the neutral crosslink (hOgg1) dissociates Lys249, which assumes the general base role and initiates the β-elimination reaction step. In fact, this proposal can be used to explain the observed similar rates for strand cleavage (β-elimination) by hOgg1 on OG and abasic site substrates. Moreover, the lack of δ-elimination catalysis by hOgg1 could be related to

the fact that a water-catalyzed δ-elimination reaction on the processed neutral abasic site has a large energy barrier and the general base at the β-elimination step (Lys) is not suitably located to initiate δ-elimination. Thus, in addition to the difference in the inherent chemistry of the β- and δ-elimination reactions discussed in Chapter 6, the present chapter provides additional explanations for the lack of δ-lyase activity of hOgg1. Finally, both the β- and δ-elimination reactions are energetically feasible from a cationic crosslink (either Lys or Pro) at comparable rates, with δ-elimination having a slightly larger barrier. This is in agreement with the experimental evidence for the rate-limiting step observed for repair catalyzed by FPG. Future large-model studies of the hOgg1 active site with a neutral lysine-crosslink and comparison with the cationic proline-crosslink in FPG will provide more insight, and further explain the observed differences in the lyase activities of hOgg1 and FPG, as well as the associated catalytic rates.

7.6 References

1. P. J. Berti and J. A. B. McCann, *Chem. Rev. (Washington, DC, U. S.)*, 2006, 106, 506–555.

2. D. O. Zharkov, G. Shoham and A. P. Grollman, *DNA Repair*, 2003, 2, 839–862.

3. N. Krishnamurthy, K. Haraguchi, M. M. Greenberg and S. S. David, *Biochemistry*, 2008, 47, 1043–1050.

4. B. Dalhus, M. Forsbring, I. H. Helle, E. S. Vik, R. J. Forstrøm, P. H. Backe, I. Alseth and M. Bjørås, *Structure (Cambridge, MA, U. S.)*, 2011, 19, 117–127.

5. O. S. Fedorova, G. A. Nevinsky, V. V. Koval, A. A. Ishchenko, N. L. Vasilenko and K. T. Douglas, *Biochemistry*, 2002, 41, 1520–1528.

6. J. Tchou, V. Bodepudi, S. Shibutani, I. Antoshechkin, J. Miller, A. P. Grollman and F. Johnson, *J. Biol. Chem.*, 1994, 269, 15318–15324.

7. D. O. Zharkov, T. A. Rosenquist, S. E. Gerchman and A. P. Grollman, *J. Biol. Chem.*, 2000, 275, 28607–28617.

8. J. C. Fromme and G. L. Verdine, *Nat Struct Mol Biol*, 2002, 9, 544–552.

9. S. D. Bruner, D. P. G. Norman and G. L. Verdine, *Nature (London, U. K.)*, 2000, 403, 859–866.

10. J. C. Fromme, S. D. Bruner, W. Yang, M. Karplus and G. L. Verdine, *Nat Struct Mol Biol*, 2003, 10, 204–211.

11. I. Morland, L. Luna, E. Gustad, E. Seeberg and M. Bjørås, *DNA Repair*, 2005, 4, 381–387.

12. J. L. Kellie and S. D. Wetmore, *J. Phys. Chem. B*, 2012, 116, 10786–10797.

13. M. Calvaresi, A. Bottoni and M. Garavelli, *J. Phys. Chem. B*, 2007, 111, 6557–6570.

14. P. Schyman, J. Danielsson, M. Pinak and A. Laaksonen, *J. Phys. Chem. A*, 2005, 109, 1713–1719.

15. K. Sadeghian and C. Ochsenfeld, *J. Am. Chem. Soc.*, 2015, 137, 9824–9831.

16. J. Sebera, Y. Hattori, D. Sato, D. Reha, R. Nencka, T. Kohno, C. Kojima, Y. Tanaka and V. Sychrovsky, *Nucleic Acids Res*, 2017, 45, 5231–5242.

17. I. D. Blank, K. Sadeghian and C. Ochsenfeld, *Scientific Reports*, 2015, 5, 10369.

18. K. Sadeghian, D. Flaig, I. D. Blank, S. Schneider, R. Strasser, D. Stathis, M. Winnacker, T. Carell and C. Ochsenfeld, *Angew. Chem., Int. Ed. Engl.*, 2014, 53, 10044–10048.

19. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. M. Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, Revision D.01 edn., 2009.

20. D. P. G. Norman, S. J. Chung and G. L. Verdine, *Biochemistry*, 2003, 42, 1564–1572.

21. M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, 7, 525–537.

22. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, 65, 712–725.

23. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1996, 118, 2309–2309.

24. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, 25, 1157–1174.

25. D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz and a. et., *Amber 12; University of California: San Francisco, CA, 2012*.

26. F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski and P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, 12, 7821–7839.

27. S. Sowlati-Hashjin and S. D. Wetmore, *J. Phys. Chem. B*, 2014, 118, 14566–14577.

28. S. Sowlati-Hashjin and S. D. Wetmore, *Phys. Chem. Chem. Phys.*, 2015, 17, 24696–24706.

Chapter 8: Global Conclusions and Future Work

Computational chemistry methods have been shown to be a useful means for obtaining atomic-level details about a variety of chemical and biochemical systems. Depending on the nature of the system and the properties of interest, various computational approaches can be employed to gain in-depth insight at a reasonable cost. In this thesis, QM methods and different model sizes have been used to study chemical reactions related to DNA damage and repair processes. Moreover, MD simulations were employed to investigate the dynamics of several lesion–active site complexes within the context of the lesion recognition step associated with the hOgg1 and FPG repair enzymes prior to catalytic removal of the damage. Specifically, the reaction pathways of guanine oxidation to form the OG and FapyG, the effect of OG and FapyG glycosidic conformations on the active site conformations of hOgg1 and FPG, and several chemical steps during the BER for OG removal, including deglycosylation, ring opening, β-elimination, δ-elimination, and Schiff base (crosslink) hydrolysis, have been investigated. In this chapter, the major concluding remarks for each individual step, namely damage, recognition, and repair are highlighted, and future directions for computational studies of DNA damage and repair are outlined.

8.1 Damage Formation: Mechanism of Oxidation of Guanine

In Chapter 2, the effects of solvent, the level of theory (combination of several DFT functionals and basis sets), and the model size on the structures and energy barriers along the formation of FapyG and OG were examined. The results show that the geometrical parameters of the transition structures for the nucleobase model, and hence the energy barriers can be considerably altered by the solvent. Moreover, solvent-phase

single-point calculations on gas-phase geometries may not necessarily reproduce the solvent-phase energies. However, the correlation between the energy barrier and the dielectric constant can be accurately described by logarithm or power formulae. Hence, for a given reaction, the energy barrier can be predicted by the corresponding formula. In comparison to the effects of solvent, the level of theory has a smaller effect on the geometries, which justifies the choice of B3LYP/6-31G(d) for studying the reactions with larger models. Despite negligible effects on the geometries, energetics are considerably affected by the level of theory; however, the data set of four DFT functionals combined with different basis sets reveals that single-point calculations on B3LYP/6-31G(d) geometries reproduce the predicted barriers using structures optimized with the corresponding functional/basis set. Chapter 2 shows that the geometrical parameters vary with the model size, which in turn affect the barriers. Unfortunately, there is not a simple correlation between the model size and the energetics. Nevertheless, certain models, such as the 5′-OH nucleoside and 5′-OH nucleoside pair models, were shown to typically introduce large deviations in geometries and energetics in comparison with other models, and should not be used in similar mechanistic studies. Despite a somewhat unclear trend in the effects of the model size on energy barriers discussed in Chapter 2, it is possible to derive some general *rules of thumb* based on the effects of the model on the transition geometries.

In order to evaluate the significance of the model size effects reported in Chapter 2, complementary studies on other reactions are required. In fact, the effects of the pairing base could be more significant for the formation of other guanine oxidative products, such as Gh and Sp, due to more complicated reaction pathways (*i.e.,* the number of reaction steps and species involved in the pathway).[1] Furthermore, both Gh and Sp have two

stereoisomers, formation of which could be affected by the flanking bases. Additionally, in the DNA duplex, OG could be paired with C or adopt the *syn* conformation and pair with A. Thus, not only the hydrogen-bonding face of OG changes, but the pairing C and A bases have different electronic properties that could influence the reaction mechanism by playing roles such as the general acid or base during the reaction. QM/MM approaches may also be applied to provide geometries and energetics that are more relevant to the DNA environment.

8.2 Lesion Recognition: OG and FapyG

Chapter 3 employed MD simulations to examine how the *anti* and *syn* conformers of OG are accommodated in the hOgg1 and FPG active sites as a first step toward understanding the difference in the experimentally observed activity of these enzymes on the OG:C and OG:A pairs (Figure 1.9).[2] When *anti*-OG is bound, the structural data indicates that hOgg1 active site residues are properly aligned to initiate catalytic base departure, while the geometrical parameters required for the catalytic reaction are not conserved for *syn*-OG. On the other hand, the FPG catalytic residues are suitably aligned for both OG conformers, with *anti*-OG being more favorably bound. Thus, the data suggests that the differential ability of hOgg1 and FPG to accommodate the *anti* and *syn*-OG glycosidic conformations is an important factor that contributes to the relative experimental excision rates.[2-3]

To complement Chapter 3, Chapter 4 employs MD simulations to examine the lesion–active site dynamics when *anti*- or *syn*-FapyG is bound in the active sites of hOgg1 and FPG. The structural data for the lesion and active site residues reveals that hOgg1 and FPG can accommodate both FapyG conformers, and key catalytic residues

and active site solvent are properly aligned to initiate catalysis. This contrasts the OG glycosidic conformers bound in hOgg1, where *syn*-OG induced significant alterations in the active site, and thereby diminished the probability for catalysis. These data suggest that the differential ability of hOgg1 to accommodate both *anti* and *syn*-FapyG could be one factor that determines the enzyme activity toward FapyG:A relative to OG:A. On the other hand, similar to the case of OG, FPG catalytic residues are suitably aligned for both FapyG conformers, with only subtle changes in the interaction patterns for the *anti* and *syn* conformers. Moreover, the greater flexibility allows FapyG to form more stabilizing interactions within the FPG active site compared to OG, which provides a plausible explanation for the slightly higher activity of FPG toward FapyG in comparison with OG regardless of the opposing base.

8.3 Repair Mechanisms: Glycosylase and β-elimination Reactions

In Chapter 5, the glycosylase and β-elimination reactions were considered using a truncated model relevant to FPG, and the results were compared to a previous study relevant to hOgg1.[4] The large calculated barriers indicate that FPG must actively catalyze the three main phases of the overall reaction, namely deglycosylation, 2′-deoxyribose ring-opening and β-elimination, and provide clues about how this is achieved through comparison to accurate crystal structures. The barriers for the base departure step facilitated by proline (Chapter 5) versus lysine[4] indicate that the nucleophile nature has a small effect on the barrier within the considered reaction mechanism. However, this step is further facilitated by several other residues in hOgg1 and FPG due to the unique composition of the active sites, which explains the distinct substrate specificity of each enzyme. Although the barrier for the final (elimination) step is less than those of the

deglycosylation and direct ring-opening phases, considerable energy is still required and therefore the enzyme must also catalyze this step. Indeed, several FPG active site residues (including Lys57, Arg259 and Asn169 sidechains) interact with the 3′-phosphate group and may provide this critical stabilization. This suggests that the differences between hOgg1 and FPG bifunctional activity could be partly due to differences in stabilization of the leaving group by the active site residues. Specifically, the β-elimination step is facilitated by several tight hydrogen-bonding interactions in FPG, while the 3′-phosphate is only interacting with a distant asparagine in the hOgg1 active site (Asn149, Figure 1.12a), which in part justifies the greater efficiency of FPG over hOgg1. Nevertheless, the nucleophile implemented may rationalize why monofunctional activity has been reported for hOgg1, but not FPG, *in vivo*. Specifically, only the lysine nucleophile stabilizes a low energy intermediate along the reaction pathway that can be hydrolyzed to afford monofunctional activity.

8.4 Repair Mechanisms: β- and δ-elimination Reactions

In Chapter 6, the inherent chemistry of the β- and δ-elimination reactions were examined. The results obtained with the employed model predict that both these reactions proceed through a step-wise (E1cB) mechanism in which (C2′ or C4′) proton abstraction is followed by cleavage of the phosphate backbone. Attempts to characterize the transition state corresponding to phosphate elimination without protonation were unsuccessful, and the model predicts phosphate protonation and elimination to be simultaneous events. Most importantly, the calculated barrier for 5′-phosphate elimination is larger than for 3′-phosphate elimination. This intrinsic difference between the β- and δ-elimination reactions at least in part rationalizes why the δ-elimination is the rate-

determining step for some bifunctional glycosylases (such as bacterial FPG) and explains why some bifunctional glycosylases do not exhibit δ-lyase activity (such as human hOgg1).

8.5 Repair Mechanisms: β- and δ-elimination and DNA-enzyme Crosslink Hydrolysis

In Chapter 7, the mechanism and the order of the β- and δ-elimination, and the hydrolysis of the crosslink in different Schiff base intermediates relevant to the reactions catalyzed by FPG and hOgg1 during the BER pathway were characterized. According to the results discussed in this chapter, the different protonation states of the Schiff base crosslinks (cationic relevant to FPG and possibly neutral in hOgg1) lead to different reaction pathways. Specifically, a cationic crosslink undergoes successive β- and δ-elimination reactions, whereas hydrolysis of the crosslink is more likely to be the first step for a neutral Schiff base. Hydrolysis of the neutral crosslink for hOgg1 dissociates Lys249, which is then free to assume the role of the general base and initiate the β-elimination reaction. In fact, this proposal can be used to explain the observed similar rates for strand cleavage (β-elimination) by hOgg1 on OG and on abasic sites. Moreover, the lack of δ-lyase activity for hOgg1 could be at least partly explained by the lack of stabilizing interactions with the 5′-phosphate in the active site (Figure 1.12a), the large barrier associate with water catalyzed δ-elimination reaction on the processed neutral abasic site, and Lys249 (general base in the β-elimination step) is positioned far from C4′–H to initiate the δ-elimination reaction. On the other hand, both the β- and δ-elimination reactions are energetically feasible on a cationic Schiff base (proline and lysine crosslinks) and the rates are comparable, with δ-elimination having a slightly larger

intrinsic barrier. This is in agreement with experimental evidence suggesting that the δ-elimination reaction is the rate-limiting step catalyzed by FPG. The results from Chapter 7 are also applicable to other enzymes with similar active site nucleophiles. Specifically, the observed lyase activities of Nei and Nth, which belong to the same structural families as FPG and hOgg1, respectively, could be explained by the mechanisms proposed in Chapter7.

8.6 Future Work

Truncated models such as those employed in the present thesis are particularly useful for investigating potentially important reaction mechanisms where the use of larger models may not be feasible. However, due to excluding potentially important moieties in the reaction pathway, truncated models may not fully reproduce the potential energy surface for a given enzymatic pathway. Once the energetically most relevant reactions have been mapped using truncated models, the reactions should be examined with larger models (using QM cluster and QM/MM approaches) to confirm the mechanistic conclusions. In this respect, the reactions studied in this thesis (*i.e.*, Chapters 2, 5, 6, and 7) can be further explored with larger models. In fact, employing larger models could not only improve the energetics, but could in some cases disclose the potential roles of individual moieties that were neglected in the truncated models.

Employing larger models is of great importance for mechanistic studies, such as the formation of Sp and Gh, which involve multiple steps and various moieties.[1] More specifically, both Sp and Gh have two stereoisomers, formation of which could be affected by the flanking bases in the DNA duplex. Additionally, in the DNA duplex, OG could be paired with C or adopt the syn conformation and pair with A. Thus, not only are

the hydrogen-bonding faces of OG different in OG:C and OG:A pairs, which can affect the electron density of OG and hence reaction barriers, but also C and A have different chemical properties and could change the reaction mechanism by playing roles such as the general acid or base during the reaction. Increasing the model size for the reactions considered in Chapters 5, 6, and 7 would involve including the active site residues that can potentially affect the reaction mechanism or the reaction barrier. Thus, future mechanistic studies need to employ other approached, such as QM/MM, which can provide further details and improve energetics at reasonable computational costs.

MD simulations presented in this thesis investigated the effects of the lesion glycosidic conformations on the active site pocket and the dynamic behaviour of catalytic residues in a pre-catalytic complex as a part of the lesion recognition process carried out by hOgg1 and FPG. The lesion recognition step, however, is not limited to binding of the lesion to the active site. In fact, several steps need to be completed prior to insertion of the damaged nucleoside into the active site. Searching for the lesion in the DNA duplex, eversion of the damaged nucleoside, filling the consequent void in the DNA helix and stabilizing the estranged nucleobase in the complementary strand are all parts of the recognition phase. Although scanning the DNA duplex by the enzyme cannot be modeled by conventional MD methods, eversion of the lesion, and interactions between the enzyme and the complementary strand during and after the lesion excursion can be investigated to supplement data provided in this thesis on the recognition step.

In summary, this thesis highlights several important differences in the recognition step and chemical reactions catalyzed by hOgg1 and FPG during the BER pathway. Future studies should strive to further explain the structural basis for the observed differences in the activities of these enzymes and examine other glycosylases, such as

Nei, Neil1, and Nth to determine whether these characteristics are general among a broader spectrum of enzymes.

8.7 References

1. Munk, B. H.; Burrows, C. J.; Schlegel, H. B., J. Am. Chem. Soc. 2008, 130, 5245–5256.

2. Krishnamurthy, N.; Haraguchi, K.; Greenberg, M. M.; David, S. S., Biochemistry 2008, 47, 1043–1050.

3. Tchou, J.; Bodepudi, V.; Shibutani, S.; Antoshechkin, I.; Miller, J.; Grollman, A. P.; Johnson, F., J. Biol. Chem. 1994, 269, 15318–15324.

4. Kellie, J. L.; Wetmore, S. D., J. Phys. Chem. B 2012, 116, 10786–10797.

Appendices

Appendix A2: Supplementary Information for Chapter 2

Table of Contents for Appendix A2

Table A2.1. Key B3LYP geometrical parameters for the transition structures associated with FapyG formation calculated with different basis sets and environments.

	Step 1						Step 2		Step 3		
	Distance				Angle		Distance	Angle	Distance		Angle
6-31G(d)	O8H···OW	O8···HOW	WO···HN7	WOH···N7	O8···H···OW	O···H···N7	C8···N9	N9···C8···N7	N9···HS	S···HN9	S···H···N9
Vacuum	1.215	1.231	1.326	1.199	159.8	148.0	1.812	90.9	1.291	1.638	170.6
DEE	1.274	1.176	1.384	1.163	161.7	148.4	1.919	90.0	1.299	1.635	169.9
THF	1.289	1.163	1.400	1.154	162.2	148.6	1.932	89.8	1.306	1.641	170.8
2-Butanol	1.301	1.154	1.413	1.147	162.5	148.7	1.942	89.8	1.317	1.634	170.9
Acetonitrile	1.307	1.149	1.421	1.143	162.7	148.7	1.947	89.7	1.325	1.627	171.0
Water	1.310	1.147	1.424	1.142	162.7	148.8	1.949	89.7	1.329	1.624	171.0
6-31+G(d)											
Vacuum	1.260	1.192	1.361	1.176	158.8	146.9	1.884	90.4	1.289	1.634	170.9
DEE	1.384	1.108	1.489	1.115	160.9	146.3	1.977	89.6	1.291	1.645	170.9
THF	1.415	1.092	1.514	1.106	161.2	146.4	1.987	89.5	1.298	1.642	171.4
2-Butanol	1.442	1.080	1.530	1.100	161.4	146.6	1.999	89.4	1.309	1.635	171.5
Acetonitrile	1.458	1.074	1.536	1.099	161.5	146.9	2.001	89.4	1.317	1.630	171.5
Water	1.466	1.071	1.538	1.098	161.5	147.0	2.004	89.4	1.321	1.627	171.6
6-31+G(d,p)											
Vacuum	1.227	1.210	1.351	1.175	159.3	147.2	1.878	90.5	1.283	1.630	170.0
DEE	1.321	1.134	1.449	1.125	161.6	147.0	1.975	89.6	1.285	1.642	170.4
THF	1.348	1.116	1.477	1.113	162.1	146.9	1.986	89.6	1.291	1.639	170.9
2-Butanol	1.370	1.104	1.495	1.107	162.4	147.0	1.994	89.5	1.303	1.632	171.0
Acetonitrile	1.381	1.097	1.511	1.102	162.8	146.9	1.999	89.5	1.310	1.626	171.2
Water	1.386	1.095	1.507	1.104	162.7	147.2	2.003	89.5	1.315	1.623	171.2
6-311+G(2df,2p)											
Vacuum	1.237	1.198	1.354	1.172	160.4	148.1	1.843	90.8	1.282	1.630	170.0
DEE	1.346	1.116	1.466	1.116	162.5	147.5	1.958	89.6	1.287	1.640	170.6
THF	1.375	1.099	1.493	1.106	162.9	147.5	1.974	89.4	1.293	1.637	171.0
2-Butanol	1.400	1.086	1.512	1.100	163.2	147.6	1.984	89.3	1.305	1.630	171.2
Acetonitrile	1.414	1.079	1.520	1.096	163.3	147.7	1.989	89.3	1.312	1.625	171.3
Water	1.420	1.076	1.524	1.095	163.3	147.8	1.991	89.3	1.316	1.622	171.4

Table A2.2. Key B3LYP geometrical parameters for the reactants associated with FapyG formation calculated with different basis sets and environments.

	Step 1						Step 2		Step 3		
	Distance				Angle		Distance	Angle	Distance		Angle
6-31G(d)	O8H···OW	O8···HOW	WO···HN7	WOH···N7	O8···H···OW	O···H···N7	C8···N9	N9···C8···N7	N9···HS	S···HN9	S···H···N9
Vacuum	1.828	0.991	0.988	1.933	157.0	142.6	1.672	92.6	2.412	1.354	171.8
DEE	1.820	0.989	0.988	1.909	162.3	143.1	1.613	94.0	2.292	1.357	174.5
THF	1.853	0.989	0.989	1.897	156.4	146.4	1.603	94.3	2.267	1.358	174.9
2-Butanol	1.855	0.989	0.989	1.898	156.3	146.5	1.596	94.5	2.241	1.359	175.2
Acetonitrile	1.855	0.989	0.989	1.898	156.3	146.5	1.593	94.6	2.236	1.359	175.3
Water	1.855	0.989	0.990	1.896	156.4	146.7	1.593	94.6	2.233	1.360	175.4
6-31+G(d)											
Vacuum	1.895	0.986	0.986	1.951	153.5	140.9	1.630	93.9	2.428	1.355	167.5
DEE	1.911	0.986	0.988	1.935	153.8	143.6	1.580	95.5	2.349	1.356	171.7
THF	1.921	0.985	0.988	1.935	153.5	144.2	1.566	96.1	2.327	1.357	171.9
2-Butanol	1.920	0.986	0.988	1.925	153.0	144.7	1.570	96.0	2.319	1.357	172.1
Acetonitrile	1.926	0.985	0.988	1.925	152.8	145.2	1.564	96.2	2.299	1.357	174.0
Water	1.928	0.985	0.988	1.924	152.7	145.4	1.564	96.2	2.294	1.357	174.0
6-31+G(d,p)											
Vacuum	1.895	0.986	0.986	1.951	153.5	140.9	1.630	93.9	2.416	1.353	168.0
DEE	1.934	0.981	0.985	1.913	153.6	145.0	1.580	95.6	2.327	1.355	172.2
THF	1.944	0.980	0.986	1.902	153.2	146.0	1.572	95.8	2.309	1.355	171.7
2-Butanol	1.951	0.981	0.986	1.897	152.9	146.8	1.569	96.0	2.294	1.356	171.8
Acetonitrile	1.978	0.979	0.985	1.901	153.5	148.5	1.564	96.2	2..277	1.356	173.7
Water	1.939	0.981	0.985	1.914	153.6	145.4	1.564	96.1	2.263	1.357	173.2
6-311+G(2df,2p)											
Vacuum	1.936	0.976	0.979	1.931	154.2	144.0	1.642	93.6	2.561	1.346	168.3
DEE	1.977	0.975	0.981	1.904	153.5	148.5	1.581	95.3	2.452	1.347	173.5
THF	1.997	0.974	0.982	1.891	152.9	149.9	1.574	95.6	2.428	1.348	173.2
2-Butanol	1.965	0.975	0.982	1.906	153.0	147.6	1.565	95.9	2.413	1.348	172.6
Acetonitrile	1.980	0.975	0.982	1.903	153.5	148.7	1.566	95.9	2.401	1.348	174.3
Water	1.998	0.974	0.982	1.891	153.0	150.0	1.563	96.0	2.394	1.348	174.5

Table A2.3. B3LYP energy barriers (ΔE, kJ/mol) for the FapyG formation pathway calculated with the nucleobase model, and different basis sets and environments.[a, b]

		6-31G(d)[c]	*6-31+G(d)*	*6-31+G(d,p)*	*6-311+G(2df,2p)*[d]
Step 1	***Vacuum***	75.2 (75.8)	78.0	73.0	76.7
	DEE	67.3 (69.7)	66.6	63.2	66.3 (68.8)
	THF	66.2 (68.5)	63.4	60.5	63.3 (67.4)
	2-Butanol	64.7 (67.5)	60.6	58.1	60.7 (66.2)
	Acetonitrile	63.8 (67.0)	59.1	56.7	59.3 (65.7)
	Water	63.4 (66.8)	58.4	56.2	58.7 (65.4)
Step 2	***Vacuum***	0.6 (1.3)	3.0	3.0	1.6
	DEE	6.1 (4.7)	13.6	13.4	9.8 (7.4)
	THF	7.7 (5.4)	16.2	16.0	13.5 (8.6)
	2-Butanol	8.9 (6.0)	18.4	18.2	15.7 (9.5)
	Acetonitrile	9.6 (6.3)	19.7	19.4	16.9 (10.0)
	Water	10.0 (6.4)	20.3	19.9	17.5 (10.3)
Step 3	***Vacuum***	45.9 (48.0)	46.9	44.4	48.3
	DEE	35.2 (39.9)	37.9	35.2	38.8 (39.8)
	THF	32.0 (37.7)	34.9	32.2	35.7 (37.5)
	2-Butanol	29.3 (35.7)	32.3	29.6	33.1 (35.4)
	Acetonitrile	27.9 (34.5)	30.9	28.2	31.7 (34.3)
	Water	27.2 (34.0)	30.3	27.5	31.1 (33.8)

[a] Representative transition structures for the 3 reaction steps are shown in Figure 2.2. [b] Environments considered include the gas phase (ε = 1), diethyl ether (DEE, ε = 4.2), tetrahydrofuran (THF, ε = 7.4), 2-butanol (ε = 15.9), acetonitrile (ε = 35.7), and water (ε = 78.4). [c] Values in parentheses correspond to solvent-phase 6-311+G(2df,2p) single-point calculations on 6-31G(d) gas-phase optimized geometries. [d] Values in parentheses correspond to solvent-phase 6-311+G(2df,2p) single-point calculations on 6-311+G(2df,2p) gas-phase optimized geometries.

Table A2.4. Correlations between the B3LYP/6-311+G(2df,2p) energy barrier and the dielectric constant calculated using different geometries.

		Dielectric Constant (ε)							
		1.0	**4.2**	**7.4**	**15.9**	**35.7**	**78.4**	**Log description *(R^2)***	**Power description *(R^2)***
Optimization	**Step**	**ΔE (kJ/mol)**						**ΔE =**	**ΔE =**
solvent-phase 6-311+G(2df,2p)	**1**	76.7	66.3	63.3	60.7	59.3	58.7	$-4.0\ \ln(\varepsilon) + 73.7$ *(0.879)*	$73.7\ (\varepsilon)^{-0.06}$ *(0.900)*
	2	1.6	9.8	13.5	15.7	16.9	17.5	$3.7\ \ln(\varepsilon) + 3.9$ *(0.894)*	$3.0\ (\varepsilon)^{0.5}$ *(0.738)*
	3	48.3	38.8	35.7	33.1	31.7	31.1	$-3.9\ \ln(\varepsilon) + 45.6$ *(0.894)*	$45.7\ (\varepsilon)^{-0.1}$ *(0.926)*
gas-phase 6-31G(d)	**1**	75.8	69.7	68.5	67.5	67.0	66.8	$-1.0\ \ln(\varepsilon) + 70.6$ *(0.890)*	$70.7\ (\varepsilon)^{-0.01}$ *(0.893)*
	2	1.3	4.7	5.4	6.0	6.3	6.4	$0.6\ \ln(\varepsilon) + 4.1$ *(0.898)*	$4.2\ (\varepsilon)^{0.1}$ *(0.874)*
	3	48.0	39.9	37.7	35.7	34.5	34.0	$-2.0\ \ln(\varepsilon) + 42.0$ *(0.915)*	$42.3\ (\varepsilon)^{-0.05}$ *(0.926)*
gas-phase 6-311+G(2df,2p)	**1**	–	68.8	67.4	66.2	65.7	65.4	$-1.1\ \ln(\varepsilon) + 69.9$ *(0.881)*	$69.9\ (\varepsilon)^{-0.02}$ *(0.886)*
	2	–	7.4	8.6	9.5	10.0	10.3	$0.9\ \ln(\varepsilon) + 6.5$ *(0.895)*	$6.8\ (\varepsilon)^{0.1}$ *(0.870)*
	3	–	39.8	37.5	35.4	34.3	33.8	$-2.0\ \ln(\varepsilon) + 41.9$ *(0.914)*	$42.2\ (\varepsilon)^{-0.06}$ *(0.925)*

Table A2.5. Key geometrical parameters for the transition structures along the FapyG formation pathway calculated using different functional and basis set combinations.

	Step 1						Step 2		Step 3		
	Distance				Angle		Distance	Angle	Distance		Angle
B3LYP	O8H···OW	O8···HOW	WO···HN7	WOH···N7	O8···H···OW	O···H···N7	C8···N9	N9···C8···N7	N9···HS	S···HN9	S···H···N9
6-31G(d)	1.274	1.176	1.384	1.163	161.7	148.4	1.919	90.0	1.296	1.647	170.7
cc-pVDZ	1.228	1.209	1.361	1.173	162.8	149.5	1.899	90.3	1.284	1.651	170.9
6-31+G(d)	1.384	1.108	1.489	1.115	160.9	146.3	1.977	89.6	1.291	1.645	170.9
6-31+G(d,p)	1.321	1.134	1.449	1.125	161.6	147.0	1.975	89.6	1.285	1.642	170.4
aug-cc-pVDZ	1.332	1.159	1.476	1.116	162.6	147.1	1.974	89.6	1.296	1.634	171.0
6-311+G(2df,2p)	1.346	1.116	1.466	1.116	162.5	147.5	1.958	89.6	1.287	1.640	170.6
B3LYP-D3											
6-31G(d)	1.282	1.170	1.394	1.156	161.9	148.2	1.917	90.0	1.296	1.634	172.1
cc-pVDZ	1.232	1.207	1.362	1.172	162.9	149.3	1.901	90.3	1.285	1.639	172.3
6-31+G(d)	1.388	1.107	1.499	1.111	161.0	146.1	1.967	89.8	1.296	1.629	172.8
6-31+G(d,p)	1.329	1.130	1.465	1.119	161.7	146.6	1.964	89.8	1.289	1.627	172.4
aug-cc-pVDZ	1.336	1.127	1.491	1.110	162.8	146.8	1.960	89.8	1.300	1.619	172.6
6-311+G(2df,2p)	1.352	1.113	1.480	1.111	162.7	147.2	1.948	89.8	1.290	1.626	172.4
M06-2X											
6-31G(d)	1.211	1.218	1.347	1.180	161.0	148.0	1.955	89.2	1.267	1.653	170.7
cc-pVDZ	1.170	1.258	1.334	1.186	162.0	149.1	1.941	89.5	1.211	1.711	170.5
6-31+G(d)	1.273	1.167	1.405	1.149	160.7	146.8	1.992	89.0	1.260	1.654	171.1
6-31+G(d,p)	1.230	1.193	1.385	1.154	160.9	147.1	1.989	89.1	1.250	1.660	170.9
aug-cc-pVDZ	1.243	1.184	1.407	1.144	162.4	147.7	1.991	89.0	1.270	1.636	170.6
6-311+G(2df,2p)	1.253	1.170	1.400	1.143	162.3	147.8	1.977	89.1	1.253	1.653	171.1
ωB97X-D											
6-31G(d)	1.275	1.163	1.383	1.154	161.9	148.2	1.953	89.5	1.249	1.671	171.0
cc-pVDZ	1.223	1.202	1.354	1.169	162.9	149.3	1.938	89.8	1.217	1.704	171.2
6-31+G(d)	1.396	1.093	1.502	1.104	160.8	145.5	1.998	89.2	1.242	1.671	171.6
6-31+G(d,p)	1.327	1.120	1.459	1.115	161.6	146.3	1.994	89.3	1.232	1.676	171.3
aug-cc-pVDZ	1.339	1.115	1.489	1.106	162.8	146.3	1.997	89.2	1.266	1.641	171.1
6-311+G(2df,2p)	1.353	1.102	1.479	1.106	162.6	146.7	1.988	89.1	1.240	1.667	171.3

Table A2.6. Key geometrical parameters for the reactants along the FapyG formation pathway calculated using different functional and basis set combinations.

	Step 1						Step 2		Step 3		
	Distance				Angle		Distance	Angle	Distance		Angle
B3LYP	O8H···OW	O8···HOW	WO···HN7	WOH···N7	O8···H···OW	O···H···N7	C8···N9	N9···C8···N7	N9···HS	S···HN9	S···H···N9
6-31G(d)	1.820	0.989	0.988	1.910	162.3	143.1	1.612	94.1	2.292	1.357	174.5
cc-pVDZ	1.804	0.990	0.990	1.874	159.7	146.2	1.606	94.3	2.297	1.366	174.6
6-31+G(d)	1.911	0.986	0.988	1.935	153.8	143.6	1.580	95.5	2.349	1.356	171.7
6-31+G(d,p)	1.934	0.981	0.985	1.913	153.6	145.0	1.580	95.6	2.327	1.355	172.2
aug-cc-pVDZ	1.946	0.979	0.985	1.893	153.5	147.4	1.579	95.6	2.368	1.363	173.6
6-311+G(2df,2p)	1.977	0.975	0.981	1.904	153.5	148.5	1.581	95.3	2.452	1.347	173.5
B3LYP-D3											
6-31G(d)	1.846	0.989	0.990	1.881	156.0	145.8	1.608	94.2	2.077	1.362	176.2
cc-pVDZ	1.812	0.988	0.987	1.908	156.4	141.8	1.606	94.3	2.067	1.371	176.0
6-31+G(d)	1.930	0.984	0.989	1.891	152.6	145.5	1.579	95.6	2.098	1.361	178.2
6-31+G(d,p)	1.932	0.981	0.986	1.887	152.6	145.6	1.577	95.6	2.085	1.360	177.9
aug-cc-pVDZ	1.936	0.979	0.987	1.862	152.8	148.4	1.578	95.6	2.086	1.368	177.1
6-311+G(2df,2p)	1.967	0.975	0.983	1.843	152.9	148.8	1.574	95.6	2.150	1.352	178.8
M06-2X											
6-31G(d)	1.844	0.985	0.983	1.931	154.3	142.8	1.581	94.4	2.127	1.353	169.6
cc-pVDZ	1.823	0.984	0.983	1.904	157.5	144.5	1.578	94.7	2.173	1.360	177.3
6-31+G(d)	1.888	0.982	0.982	1.962	151.9	139.8	1.567	95.5	2.145	1.351	170.0
6-31+G(d,p)	1.919	0.977	0.978	1.960	150.6	139.8	1.558	95.8	2.146	1.350	171.1
aug-cc-pVDZ	1.947	0.975	0.978	1.926	150.1	144.2	1.555	95.8	2.213	1.357	158.9
6-311+G(2df,2p)	1.955	0.972	0.975	1.938	150.0	144.1	1.557	95.6	2.270	1.343	165.5
ωB97X-D											
6-31G(d)	1.845	0.982	0.979	1.991	155.3	138.1	1.580	94.9	2.123	1.354	175.5
cc-pVDZ	1.823	0.982	0.980	1.956	157.5	139.7	1.580	95.0	2.124	1.363	176.7
6-31+G(d)	1.889	0.980	0.978	2.030	154.7	137.9	1.555	96.1	2.143	1.353	174.6
6-31+G(d,p)	1.893	0.977	0.979	1.913	159.6	143.5	1.555	96.1	2.128	1.353	175.5
aug-cc-pVDZ	1.919	0.975	0.980	1.903	154.4	146.3	1.553	96.1	2.152	1.361	174.4
6-311+G(2df,2p)	1.960	0.971	0.976	1.906	152.4	147.1	1.554	95.9	2.205	1.346	175.7

Table A2.7. Key B3LYP/6-31G(d) geometrical parameters for the transition structures along the FapyG formation pathway calculated using different models.[a]

	Step 1						Step 2		Step 3		
	Distance				Angle		Distance	Angle	Distance		Angle
Model	**O8H···OW**	**O8···HOW**	**WO···HN7**	**WOH···N7**	**O8···H···OW**	**O···H···N7**	**C8···N9**	**N9···C8···N7**	**N9···HS**	**S···HN9**	**S···H···N9**
Nucleobase	1.269	1.179	1.378	1.167	161.8	148.6	1.919	90.0	1.296	1.647	170.7
5'-OH Nucleoside	1.226	1.215	1.290	1.225	160.2	149.6	1.973	90.0	1.294	1.661	169.6
5'-OMe Nucleoside	1.259	1.188	1.372	1.169	161.7	148.6	1.785	91.7	1.295	1.660	169.9
5'-OH Pair	1.254	1.190	1.313	1.208	160.7	149.9	1.996	89.8	1.317	1.648	171.3
5'-OMe Pair	1.294	1.160	1.418	1.145	163.0	148.7	1.779	91.7	1.311	1.652	170.7
Trimer	1.404	1.087	1.500	1.112	163.3	147.5	1.902	90.4	1.331	1.636	174.4

[a] Representative transition structures for the 3 reaction steps are shown in Figures 2.6 and 2.7.

Table A2.8. Key B3LYP/6-31G(d) geometrical parameters for the reactants along the FapyG formation pathway calculated using different models.[a]

	Step 1						Step 2		Step 3		
	Distance				Angle		Distance	Angle	Distance		Angle
Model	**O8H···OW**	**O8···HOW**	**WO···HN7**	**WOH···N7**	**O8···H···OW**	**O···H···N7**	**C8···N9**	**N9···C8···N7**	**N9···HS**	**S···HN9**	**S···H···N9**
Nucleobase	1.820	0.989	0.988	1.910	162.3	143.1	1.612	94.1	2.292	1.357	174.5
5'-OH Nucleoside	1.786	0.994	0.987	1.940	158.5	142.2	1.598	95.5	3.032	1.350	167.0
5'-OMe Nucleoside	1.839	0.989	0.987	1.929	157.9	144.2	1.669	92.9	2.443	1.354	174.2
5'-OH Pair	1.801	0.993	0.989	1.913	158.1	143.9	1.587	95.8	2.914	1.350	169.5
5'-OMe Pair	1.892	0.985	0.987	1.931	152.2	145.2	1.643	93.5	2.384	1.356	174.7
Trimer	1.844	0.985	0.994	1.855	164.0	146.7	1.608	93.9	2.315	1.358	176.0

[a] Representative reactant structures for the 3 reaction steps are shown in Figure A2.1.

Table A2.9. B3LYP/6-31G(d) χ dihedral angle in the transition structures and the reactant complexes (in parentheses) along the FapyG formation pathway calculated using models that include 2′-deoxyribose.[a]

	Step 1	Step 2	Step 3
5′-OH Nucleoside	–97.4 (–98.8)	–118.2 (–96.9)	–72.7 (–143.7)
5′-OMe Nucleoside	–87.8 (–93.1)	–80.2 (–98.2)	–74.4 (–94.6)
5′-OH Pair	–99.4 (–99.9)	–122.6 (–97.3)	–79.0 (–144.2)
5′-OMe Pair	–131.5 (–136.0)	–128.5 (–126.0)	–75.8 (–91.5)
Trimer	–82.0 (–83.7)	–92.1 (–91.8)	–84.0 (–87.6)

[a] Representative reactant and transition structures for the 3 reaction steps are shown in Figures A2.1, 2.6 and 2.7.

Table A2.10. B3LYP/6-31G(d) pairing distances (Å) and hydrogen-bonding angles (deg.) at transition states and the reactant complexes (in parentheses) for modified G:C along the FapyG formation pathway.[a, b]

		Distance (Å)	Angle (deg.)	Distance (Å)	Angle (deg.)	Distance (Å)	Angle (deg.)	Distance (Å)
Model	**Step**	**O6···H–N**	**O6···H···N**	**N1–H···N**	**N1···H···N**	**N2–H···O**	**N2···H···O**	**C1′···C1′**
	1	1.885 (1.867)	178.4 (179.1)	1.879 (1.894)	178.9 (179.3)	1.808 (1.822)	179.6 (179.7)	10.696 (10.721)
5′-OH Pair	**2**	1.895 (1.897)	178.4 (178.3)	1.883 (1.874)	178.1 (178.6)	1.835 (1.811)	179.7 (179.5)	10.711 (10.679)
	3	1.916 (1.903)	177.2 (177.9)	1.856 (1.880)	178.0 (177.8)	1.830 (1.854)	178.6 (179.1)	10.441 (10.513)
	1	1.883 (1.855)	178.2 (179.0)	1.882 (1.900)	178.4 (178.4)	1.813 (1.837)	179.3 (179.9)	10.686 (10.725)
5′-OMe Pair	**2**	1.885 (1.884)	178.5 (178.7)	1.885 (1.883)	178.0 (178.4)	1.836 (1.831)	179.5 (179.7)	10.691 (10.685)
	3	1.916 (1.903)	177.2 (177.9)	1.856 (1.877)	178.0 (177.8)	1.830 (1.857)	178.6 (178.9)	10.447 (10.527)
	1	1.897 (1.879)	177.9 (177.5)	1.884 (1.906)	177.1 (176.3)	1.805 (1.824)	176.6 (176.7)	10.652 (10.658)
Trimer	**2**	1.913 (1.914)	176.2 (178.5)	1.884 (1.883)	177.3 (178.9)	1.821 (1.802)	177.8 (178.7)	10.584 (10.583)
	3	1.911 (1.901)	178.7 (177.3)	1.859 (1.874)	178.5 (176.7)	1.873 (1.906)	178.2 (177.5)	10.381 (10.426)

[a] Representative reactant and transition structures for the 3 reaction steps are shown in Figures A2.1, 2.6 and 2.7.

Table A2.11. B3LYP/6-31G(d) relative energy barriers (ΔE, kJ/mol) for the formation of OG through singlet intermediates calculated using the nucleobase and 5′-OMe nucleoside models.[a]

	Step 1	Step 2
Nucleobase	26.6	84.5
5′-OMe Nucleoside	33.1	95.9

[a] Representative transition state structures for the 2 reaction steps for the nucleobase model are shown in Figure A2.3.

Table A2.12. Key B3LYP/6-31G(d) geometrical parameters for the transition structures along the OG formation pathway calculated using different models.[a, b]

	Step 1			Step 2					
	Distance (Å)		Angle (deg.)	Distance (Å)				Angle (deg.)	
Model	**O8H···OW**	**O8···HOW**	**O8···H···OW**	**C8H···OW**	**C8···HOW**	**WO···HN7**	**WOH···N7**	**C8···H···OW**	**O···H···N7**
Nucleobase	1.241	1.122	150.2	1.612	1.223	1.386	1.169	130.1	141.8
5′-OH Nucleoside	1.232	1.131	147.4	1.606	1.219	1.414	1.154	129.7	140.3
5′-OMe Nucleoside	1.253	1.118	154.6	1.637	1.207	1.419	1.154	129.1	140.7
5′-OH Pair	1.232	1.131	147.0	1.849	1.150	1.468	1.137	124.9	142.5
5′-OMe Pair	1.235	1.128	147.4	1.655	1.204	1.415	1.156	128.8	141.4
Trimer	1.246	1.155	159.8	1.939	1.156	1.473	1.121	120.9	143.6

[a] Representative transition state structures of the 2 reaction steps are shown in Figures 2.9 and 2.10.

Table A2.13. Key B3LYP/6-31G(d) geometrical parameters for the reactants along the OG formation pathway calculated using different models.[a, b]

	Step 1			Step 2					
	Distance (Å)		Angle (deg.)	Distance (Å)				Angle (deg.)	
Model	**O8H···OW**	**O8···HOW**	**O8···H···OW**	**C8H···OW**	**C8···HOW**	**WO···HN7**	**WOH···N7**	**C8···H···OW**	**O···H···N7**
Nucleobase	1.975	0.982	154.6	2.612	1.125	0.982	1.945	111.4	149.8
5′-OH Nucleoside	3.712	0.972	113.9	4.102	1.120	0.983	1.905	70.4	180.0
5′-OMe Nucleoside	2.146	0.979	149.8	2.707	1.120	0.982	1.952	109.0	151.9
5′-OH Pair	2.015	0.981	150.5	2.681	1.126	0.989	1.926	111.6	155.6
5′-OMe Pair	2.023	0.981	150.3	4.591	1.121	0.977	2.059	64.4	149.0
Trimer	2.039	0.980	143.5	3.483	1.125	0.991	1.852	74.2	159.6

[a] Representative reactant structures of the 2 reaction steps are shown in Figures A2.4.

Table A2.14. B3LYP/6-31G(d) χ dihedral angle in the transition structures and the reactant complexes (in parentheses) along the OG formation pathway calculated using models that include 2′-deoxyribose.[a]

	Step 1	**Step 2**
5′-OH Nucleoside	–88.7 (–88.9)	–99.0 (–85.5)
5′-OMe Nucleoside	–88.7 (–78.0)	–91.8 (–85.4)
5′-OH Pair	–89.1 (–90.0)	–71.2 (–66.8)
5′-OMe Pair	–89.9 (–89.8)	–88.2 (–87.6)
Trimer	–91.8 (–98.0)	–81.7 (–77.5)

[a] Representative reactant and transition state structures of the 2 reaction steps are shown in Figures A2.4, 2.9 and 2.10.

Table A2.15. IEF-PCM-B3LYP/6-31G(d) pairing distances (Å) and hydrogen-bonding angles (degrees) between modified G:C in transition states along the OG formation pathway with nucleoside pair and trimer models.[a]

		Distance (Å)	Angle (deg.)	Distance (Å)	Angle (deg.)	Distance (Å)	Angle (deg.)	Distance (Å)
Model	**Step**	**O6···H–N**	**O6···H···N**	**N1–H···N**	**N1···H···N**	**N2–H···O**	**N2···H···O**	**C1′···C1′**
5′-OH Pair	**1**	1.861 (1.860)	179.2 (179.2)	1.893 (1.894)	179.0 (179.1)	1.830 (1.831)	179.7 (179.5)	10.727 (10.726)
	2	1.896 (1.869)	177.7 (178.7)	1.868 (1.891)	178.9 (179.2)	1.792 (1.821)	179.3 (179.5)	10.636 (10.661)
5′-OMe Pair	**1**	1.859 (1.862)	179.1 (179.0)	1.893 (1.894)	179.2 (179.1)	1.832 (1.828)	179.5 (179.4)	10.728 (10.719)
	2	1.885 (1.876)	178.1 (178.0)	1.877 (1.897)	178.9 (178.8)	1.809 (1.814)	179.2 (179.1)	10.665 (10.683)
Trimer	**1**	1.874 (1.878)	175.3 (178.5)	1.876 (1.873)	175.1 (177.8)	1.863 (1.875)	175.9 (176.7)	10.578 (10.651)
	2	1.910 (1.891)	178.6 (178.9)	1.875 (1.877)	179.1 (179.6)	1.783 (1.820)	179.6 (178.0)	10.622 (10.605)

[a] Representative reactant and transition state structures of the 2 reaction steps are shown in Figures A2.4, 2.9 and 2.10.

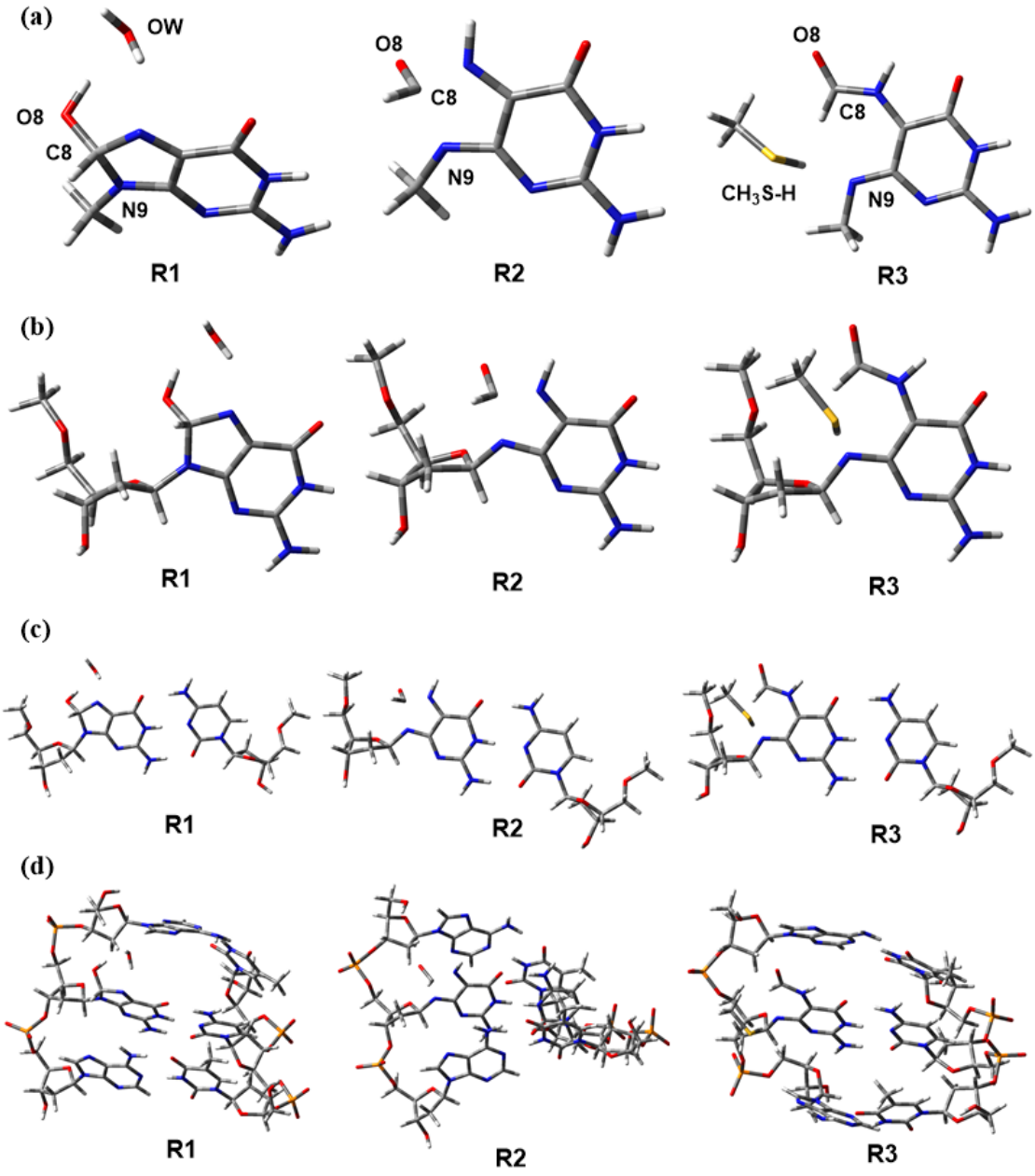

Figure A2.1. IEF-PCM-B3LYP/6-31G(d) representative reactant structures for the along the formation of FapyG for (a) nucleobase, (b) 5′-OMe nucleoside, (c) 5′-OMe nucleoside pair, and (d) trimer model. For clarity, atom numbers are only shown for the nucleobase model.

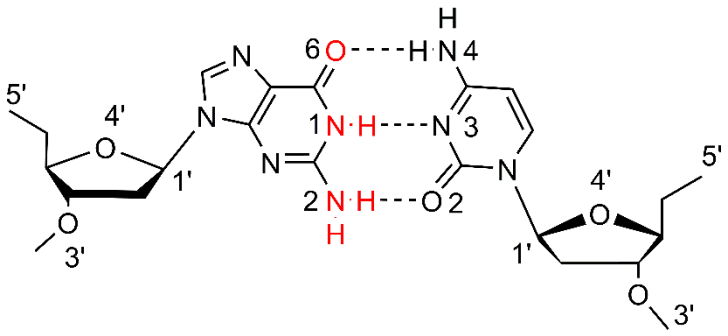

Figure A2.2. Schematic representation of a G:C pair and important pairing parameters.

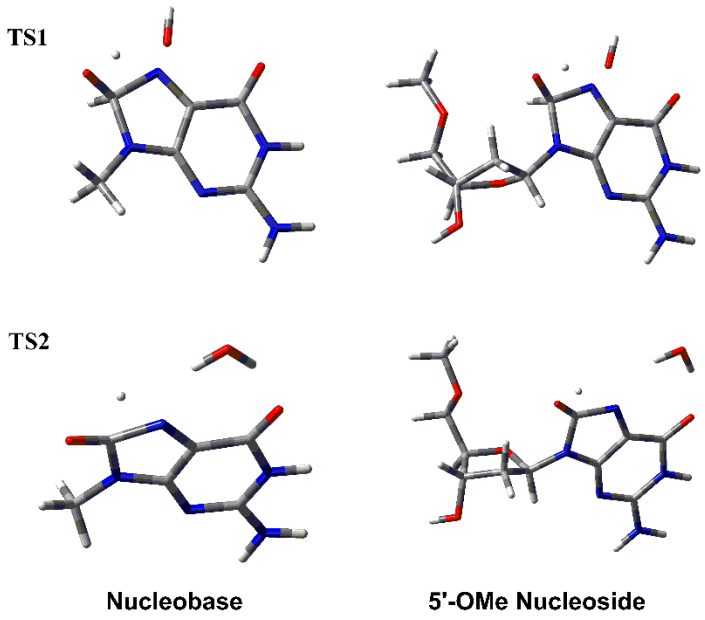

Figure A2.3. Transition structures along the OG formation pathway involving a singlet intermediate for the nucleobase and 5′-OMe nucleoside models.

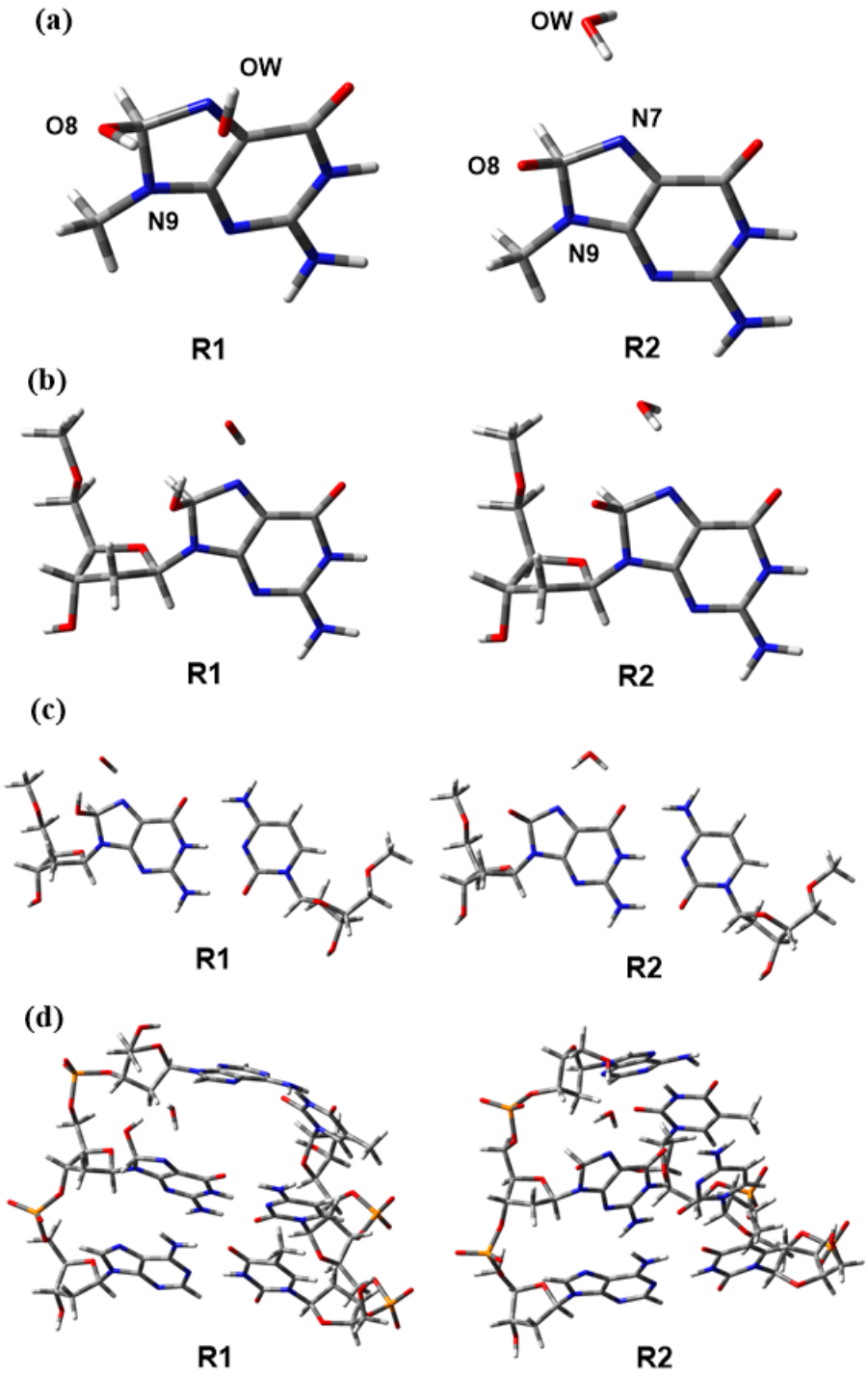

Figure A2.4. IEF-PCM-B3LYP/6-31G(d) representative reactant structures for the along the formation of OG for (a) nucleobase, (b) 5′-OMe nucleoside, (c) 5′-OMe nucleoside pair, and (d) trimer model. For clarity, atom numbers are only shown for the nucleobase model.

Appendix A3: Supplementary Information for Chapter 3

Table of Contents for Appendix A3

Table A3.1. Root-mean-square deviation (rmsd in Å; standard deviation in parentheses) in the DNA and enzyme backbone over the entire and clustered MD simulation trajectory.

System	rmsd (simulation)[a]			rmsd (cluster)[b]	
hOgg1	**DNA**	**Enzyme**	**Occ.**	**DNA**	**Enzyme**
anti-OG (Asp268^{-1})	2.721 (0.518)	1.410 (0.224)	100	2.721 (0.518)	1.410 (0.224)
anti-OG (Asp268)	2.635 (0.517)	1.043 (0.156)	100	2.635 (0.517)	1.043 (0.156)
syn-OG (Asp268^{-1})	2.759 (0.492)	1.208 (0.158)	100	2.759 (0.492)	1.208 (0.158)
syn-OG (Asp268)	2.344 (0.444)	1.410 (0.227)	95	2.342 (0.440)	1.409 (0.225)
FPG	**DNA**	**Enzyme**	**Occ.**	**DNA**	**Enzyme**
anti-OG (Glu3^{-1})	2.045 (0.589)	1.075 (0.125)	94	2.042 (0.589)	1.078 (0.124)
anti-OG (Glu3)	2.254 (0.387)	0.981 (0.106)	96	2.252 (0.385)	0.980 (0.101)
syn-OG (Glu3^{-1})	2.329 (0.459)	1.091 (0.113)	100	2.329 (0.459)	1.091 (0.113)
syn-OG (Glu3)	2.371 (0.530)	0.819 (0.110)	95	2.371 (0.528)	0.818 (0.099)

[a] rmsd calculated according to the position of the DNA (P, O3′, O5′, C3′, C4′, and C5′) and enzyme (C_α, C, and N) backbone atoms over the entire simulation trajectory with respect to the first frame. [b] Clustering was completed for the sampled trajectories based on the configurations of key active site residues, namely OG, Gly42, Lys249, Asp268, Gln315, and Phe319 for hOgg1, and OG, Pro2, and Glu3 for FPG. rmsd calculated according to the position of the DNA (P, O3′, O5′, C3′, C4′, and C5′) and enzyme (C_α, C, and N) backbone atoms for the cluster with the highest occupancy with respect to the first frame of the production phase.

Table A3.2. Per residue root-mean-square deviation (rmsd in Å; standard deviation in parentheses) for hOgg1 active site residues with respect to the reference crystal structure.[a]

	Gly42	Lys249	Asp268	Gln315	Phe319
anti-OG (Asp268^{-1})	0.083 (0.045)	0.566 (0.176)	0.161 (0.044)	0.367 (0.097)	0.179 (0.068)
anti-OG (Asp268)	0.103 (0.054)	0.748 (0.209)	0.293 (0.067)	0.397 (0.064)	0.217 (0.097)
syn-OG (Asp268^{-1})	0.070 (0.037)	0.623 (0.111)	0.997 (0.062)	0.993 (0.132)	0.172 (0.065)
syn-OG (Asp268)	0.078 (0.040)	0.785 (0.079)	0.694 (0.267)	0.900 (0.264)	0.263 (0.038)

[a] rmsd calculated according to the position of the active site residues (all heavy atoms) for the cluster with the highest occupancy with respect to the crystal structure (PDB ID: 3KTU).

Table A3.3. Average χ dihedral angle of OG (in degrees; standard deviation in parentheses) across the MD simulations and the reference crystal structures.[a]

hOgg1	
PDB ID: 3KTU	–124.3
anti-OG (Asp268^{-1})	–113.0 (11.3)
anti-OG (Asp268)	–103.7 (12.2)
syn-OG (Asp268^{-1})	61.8 (12.2)
syn-OG (Asp268)	76.2 (16.3)
FPG	
PDB ID: 1R2Y	101.1
anti-OG (Glu3^{-1})	–113.7 (11.0)
anti-OG (Glu3)	–101.6 (14.0)
syn-OG (Glu3^{-1})	57.3 (11.9)
syn-OG (Glu3)	67.4 (10.8)

[a] Average dihedral angle calculated for the cluster with the highest occupancy with respect to the first frame of the production phase.

Table A3.4. Hydrogen-bonding occupancies (%) for interactions between OG and hOgg1 active site residues or solvent, and corresponding average distance (Å) and angle (in degrees).[a, b]

Interaction	Asp268^{-1}		Asp268	
***anti*-OG···Residue**	**Occ.**	**distance (angle)**	**Occ.**	**distance (angle)**
N7–H···O(Gly42)	95	3.004 (159.9)	99	2.898 (159.8)
O4′···H–Oδ(Asp268)	–	–	27	2.820 (152.3)
N2–H···Oδ(Asp268)	97	2.965 (160.1)	92	3.098 (149.3)
N1–H···Oε(Gln315)	96	2.957 (151.1)	99	2.903 (157.5)
N2–H···Oε(Gln315)	91	3.004 (146.2)	96	2.945 (145.3)
O6···H(H_2O)	49	2.749 (159.3)	50	2.787 (159.1)
O4′···H(H_2O)	27	2.846 (157.1)	24	2.984 (159.2)
***syn*-OG···Residue**	**Occ.**	**distance (angle)**	**Occ.**	**distance (angle)**
N1–H···O(Gly42)	20	3.204 (139.4)	93	2.925 (143.1)
N2–H···O(Gly42)	77	3.006 (154.5)	88	3.010 (141.6)
O8···H–Oδ(Asp268)	–	–	89	2.681 (165.7)
O8···H–N(Asp268)	72	2.947 (137.9)	–	–
N7–H···Oε(Gln315)	41	3.319 (152.4)	17	3.298 (123.8)
O6···H–Nε(Gln315)	76	2.923 (160.7)	–	–
O6···H(H_2O)	36	2.845 (156.0)	28	2.815 (157.2)
O4′···H(H_2O)	13	3.066 (140.8)	14	2.937 (151.7)
N7–H···O(H_2O)	–	–	26	2.945 (150.0)

[a] Hydrogen-bonding occupancies calculated for the cluster with the highest occupancy. [b] Percent occupancy of each hydrogen bond is defined using a distance cut-off of < 3.4 Å and an angle cut-off of > 120°.

Table A3.5. Average interaction energies (ΔE in kcal/mol; standard deviations in parentheses) between various hOgg1 active site residues and *anti*- or *syn*-OG for different Asp268 protonation states.[a]

***anti*-OG**		
Residue	**Asp268^{-1}**	**Asp268**
Gly42	–6.5 (1.9)	–6.5 (1.6)
Asp268	–13.8 (1.2)	–14.9 (2.2)
Gln315	–13.0 (2.0)	–13.5 (2.0)
Phe319	–3.8 (0.6)	–4.5 (0.3)
H_2O	–7.1 (1.9)	–7.0 (2.1)
***syn*-OG**		
Residue	**Asp268^{-1}**	**Asp268**
Gly42	–7.1 (2.8)	–8.9 (1.7)
Asp268	–1.9 (0.5)	–10.2 (1.9)
Gln315	–4.9 (2.3)	–1.2 (1.5)
Phe319	–3.5 (0.5)	–2.8 (0.9)
H_2O	–3.9 (1.0)	–2.8 (3.5)

[a] Average linear interaction energies (LIE) and standard deviations (in parentheses) were calculated for the cluster with the highest occupancy.

Table A3.6. Hydrogen-bonding occupancies (%) for interactions between Asp268 and neighbouring residues for *anti*- or *syn*-OG bound in the hOgg1 active site.[a, b]

Interaction	*anti*-OG		*syn*-OG	
Asp268···Residue	**Asp268^{-1}**	**Asp268**	**Asp268^{-1}**	**Asp268**
Oδ··· H–N(Val269)	–	–	99	–
Oδ···H–N(His270)	99	72	–	98
Oδ···H–N(Met271)	100	89	–	18
O···H–N(Met271)	–	–	95	72
O···H–N(Trp272)	99	98	97	95
Oδ···H–N2(OG)	97	92	–	–
Oδ–H···O3′(OG)	–	75	–	–
Oδ–H···O4′(OG)	–	20	–	–
Oδ–H···O8(OG)	–	–	–	89
N–H···O8(OG)	–	–	72	–
Oδ···H(H_2O)	94	–	95	–

[a] Hydrogen-bonding occupancies calculated for the cluster with the highest occupancy. [b] Percent occupancy of each hydrogen bond is defined using a distance cut-off of < 3.4 Å and an angle cut-off of > 120°.

Table A3.7. Average distances (Å; standard deviation in parentheses) between the nucleophile or Asp/Glu and 2′-deoxyribose,[a] and between OG and Phe319 in the hOgg1 active site.[b]

hOgg1	Lys249(Nζ···C1′)	Asp268(Oδ1···O4′)	Asp268(Oδ2···O4′)	OG···Phe319
anti-OG (Asp268^{-1})	4.374 (0.389)	3.708 (0.317)	5.505 (0.257)	3.613 (0.319)
anti-OG (Asp268)	4.127 (0.437)	3.557 (0.341)	4.987 (0.273)	3.596 (0.278)
syn-OG (Asp268^{-1})	5.505 (0.784)	6.587 (0.329)	5.651 (0.307)	3.949 (0.376)
syn-OG (Asp268)	4.605 (0.504)	4.246 (0.612)	5.403 (0.418)	3.872 (0.563)
FPG	**Pro2(N···C1′)**	**Glu3(Oε1···O4′)**	**Glu3(Oε2···O4′)**	
anti-OG (Glu3^{-1})	4.219 (0.313)	6.452 (1.250)	7.454 (0.980)	–
anti-OG (Glu3)	5.316 (0.758)	4.340 (0.621)	4.606 (0.635)	–
syn-OG (Glu3^{-1})	3.998 (0.243)	4.802 (0.375)	6.758 (0.365)	–
syn-OG (Glu3)	4.542 (0.293)	4.516 (0.313)	5.097 (0.375)	–

[a] Average distances and standard deviations between the atoms measured for the cluster with the highest occupancy. [b] Average distances and standard deviations were evaluated between the centers of the mass of OG and Phe319 for hOgg1.

Table A3.8. Average number of water molecules in two spheres centered on a deoxyribose or Asp268 atom for OG bound in the hOgg1 active site.[a, b]

***anti*-OG**	Deoxyribose (O4′)	Asp268^{-1} (Oδ1)	Asp268^{-1} (Oδ2)
r = 3.4 Å	1	1	1
r = 6.0 Å	7	4	3
***anti*-OG**	Deoxyribose (O4′)	Asp268 (Oδ1)	Asp268 (Oδ2)
r = 3.4 Å	1	0	1
r = 6.0 Å	5	4	5
***syn*-OG**	Deoxyribose (O4′)	Asp268^{-1} (Oδ1)	Asp268^{-1} (Oδ2)
r = 3.4 Å	1	4	3
r = 6.0 Å	10	14	14
***syn*-OG**	Deoxyribose (O4′)	Asp268 (Oδ1)	Asp268 (Oδ2)
r = 3.4 Å	1	2	2
r = 6.0 Å	7	7	8

[a] Average number of water molecules calculated for the cluster with the highest occupancy. [b] A spherical solvent shell with a radius of 3.4 or 6.0 Å was centered on O4′ of deoxyribose, or Oδ1 or Oδ2 of Asp268.

Table A3.9. Average values of possible nucleophilic attack angles (in degrees; standard deviations in parentheses).[a]

System	**Angle (degree)**	
hOgg1	**∠(Nζ–C1′–O4′)**	**∠(Nζ–C1′–N9)**
anti-OG (Asp268^{-1})	153.9 (12.2)	62.7 (11.9)
anti-OG (Asp268)	149.9 (11.5)	59.2 (16.7)
syn-OG (Asp268^{-1})	109.6 (12.2)	78.3 (12.6)
syn-OG (Asp268)	136.0 (11.7)	74.4 (9.6)
FPG	**∠(N–C1′–O4′)**	**∠(N–C1′–N9)**
anti-OG (Glu3^{-1})	162.5 (7.0)	81.4 (6.5)
anti-OG (Glu3)	159.0 (10.3)	75.8 (5.3)
syn-OG (Glu3^{-1})	163.1 (8.3)	61.7 (5.6)
syn-OG (Glu3)	158.9 (7.0)	74.7 (4.8)

[a] Average angles and standard deviations calculated for the cluster with the highest occupancy.

Table A3.10. Per residue root-mean-square deviation (rmsd in Å; standard deviation in parentheses) for FPG active site residues with respect to the reference crystal structure.[a]

	anti-OG (Glu3^{-1})	*anti*-OG (Glu3)	*syn*-OG (Glu3^{-1})	*syn*-OG (Glu3)
Pro2	0.214 (0.112)	0.514 (0.060)	0.360 (0.087)	0.493 (0.050)
Glu3	0.209 (0.048)	0.580 (0.089)	0.975 (0.066)	0.273 (0.088)
Glu6	0.160 (0.089)	0.180 (0.057)	0.179 (0.055)	0.295 (0.128)
Glu78	0.407 (0.267)	0.614 (0.367)	0.666 (0.348)	0.553 (0.035)
Arg80	1.240 (0.113)	0.818 (0.285)	0.664 (0.058)	0.458 (0.335)
Ser220	0.108 (0.041)	0.308 (0.273)	0.388 (0.283)	0.599 (0.221)
Thr221	0.149 (0.049)	0.144 (0.050)	0.120 (0.038)	0.145 (0.040)
Val222	0.143 (0.049)	0.159 (0.052)	1.094 (0.170)	0.980 (0.061)
Arg223	0.924 (0.085)	0.995 (0.093)	1.027 (0.246)	1.067 (0.159)
Thr224	0.263 (0.066)	0.513 (0.059)	0.219 (0.060)	0.151 (0.047)
Tyr225	0.133 (0.042)	0.186 (0.048)	0.140 (0.045)	0.159 (0.053)

[a] rmsd calculated according to the position of the active site residues (all heavy atoms) for the cluster with the highest occupancy with respect to the crystal structure (PDB ID: 1R2Y).

Table A3.11. Hydrogen-bonding occupancies (%) for interactions between OG and FPG active site residues or solvent, and corresponding average interaction distance (Å) and angle (in degrees).[a, b]

Interaction	Glu3^{-1}		Glu3	
***anti*-OG⋯Residue**	**Occ.**	**distance (angle)**	**Occ.**	**distance (angle)**
N2–H⋯O(Ser220)	100	2.815 (152.8)	92	2.974 (142.5)
N1–H⋯O(Ser220)	92	3.019 (138.1)	98	2.849 (152.4)
O6⋯H–N(Val222)	57	3.083 (151.5)	48	3.088 (141.0)
O6⋯H–N(Arg223)	88	3.008 (154.0)	98	2.969 (159.9)
O8⋯H1(Arg223)	94	2.790 (155.0)	45	3.029 (142.5)
O6⋯H–N(Thr224)	100	2.963 (158.5)	97	3.020 (161.7)
N7–H⋯Oγ(Thr224)	98	2.905 (158.9)	59	3.005 (161.8)
O6⋯H–N(Tyr225)	95	2.948 (147.0)	89	2.986 (161.9)
N2–H⋯Oε(Glu6)[c]	33/15	3.166 (126.2)	43/26	3.020 (149.3)
O4′⋯H–Oε(Glu3)	–	–	22	3.115 (139.3)
O4′⋯H–O(H_2O)	39	2.858 (159.1)	44	2.957 (162.2)
***syn*-OG⋯Residue**	**Occ.**	**distance (angle)**	**Occ.**	**distance (angle)**
N7–H⋯O(Ser220)	93	2.970 (150.8)	100	2.827 (159.6)
O6⋯H–N(Val222)	18	3.144 (134.0)	20	3.075 (135.2)
O6⋯H–N(Arg223)	96	3.058 (159.7)	98	2.990 (159.5)
O6⋯H–N(Thr224)	99	2.952 (158.1)	94	2.910 (160.2)
N1–H⋯Oγ(Thr224)	99	2.916 (157.2)	98	2.964 (151.9)
O6⋯H–N(Tyr225)	100	2.923 (162.1)	85	2.945 (160.1)
N2–H⋯Oε(Glu78)[c]	48/46	2.934 (148.1)	55/36	2.955 (145.2)
O4′⋯H–Oε(Glu3)	–	–	17	2.953 (152.0)
O4′⋯H–O(H_2O)	27	2.978 (156.2)	51	2.933 (153.3)

[a] Hydrogen-bonding occupancies calculated for the cluster with the highest occupancy. [b] Percent occupancy of the hydrogen bond is defined using a distance cut-off of < 3.4 Å and an angle cut-off of > 120°. [c] Values correspond to hydrogen bonds with the two amino hydrogens of OG or two Oε atoms of Glu6 and Glu78.

Table A3.12. Average interaction energies (ΔE in kcal/mol; standard deviations in parentheses) between various FPG active site residues and *anti*- or *syn*-OG.[a]

***anti*-OG**		
Residue	**Glu3^{-1}**	**Glu3**
Glu3	–1.2 (2.0)	–1.9 (1.0)
Glu6	–16.5 (2.2)	–15.7 (2.3)
Met77	–1.0 (1.2)	–0.7 (1.0)
Glu78	11.5 (2.4)	9.9 (2.9)
Arg80	–2.0 (1.5)	–2.2 (2.1)
Ile175	0.9 (0.3)	0.5 (0.4)
Ser220	–9.1 (1.5)	–7.5 (1.2)
Thr221	–3.0 (1.3)	–2.6 (1.0)
Val222	–2.8 (2.1)	–4.3 (1.0)
Arg223	–31.1 (3.7)	–27.2 (4.1)
Thr224	–8.8 (1.9)	–8.3 (2.6)
Tyr225	–2.4 (1.0)	–2.3 (0.8)
***syn*-OG**		
Residue	**Glu3^{-1}**	**Glu3**
Glu3	5.9 (1.5)	1.2 (2.9)
Glu6	11.9 (2.3)	11.3 (3.3)
Met77	–0.5 (1.0)	–2.1 (2.4)
Glu78	–22.5 (3.0)	–20.2 (2.6)
Arg80	1.5 (1.4)	2.9 (2.5)
Ile175	–1.4 (0.7)	–1.2 (0.5)
Ser220	–4.8 (1.9)	–8.5 (3.5)
Thr221	–3.1 (0.7)	–2.2 (1.2)
Val222	–1.9 (0.9)	–3.4 (1.6)
Arg223	–2.8 (2.3)	–2.5 (2.4)
Thr224	–8.3 (2.1)	–5.7 (2.8)
Tyr225	–3.2 (0.8)	–3.0 (0.8)

[a] Average linear interaction energies (LIE) and standard deviations were calculated for the cluster with the highest occupancy.

Table A3.13. Hydrogen-bonding occupancies (%) for interactions between Glu3 and neighbouring residues or OG bound in the FPG active site.[a, b]

Interaction	*anti*-OG		*syn*-OG	
Glu3···Residue	**Glu3^{-1}**	**Glu3**	**Glu3^{-1}**	**Glu3**
O···H–N(Glu6)	65	66	38	45
N–H···Oε(Glu6)	93	79	85	100
Oε–H···Oε(Glu6)	–	–	–	55
O···H–N(Val7)	98	99	44	95
Oε···H–N(Gly173)	97	–	95	–
Oε···H–N(Asn174)	62	–	87	–
Oε···H–N(Ile175)	97	74	97	72
Oε···H–N(Tyr176)	53	62	65	42
Oε–H···N3(OG)	–	58	–	–
Oε–H···O3′(OG)	–	25	–	18
Oε–H···O4′(OG)	–	22	–	21
Oε···H(H_2O)	55	–	69	–

[a] Hydrogen-bonding occupancies calculated for the cluster with the highest occupancy. [b] Percent occupancy of the hydrogen bond is defined using a distance cutoff of < 3.4 Å and an angle cutoff of > 120°.

Table A3.14. Average number of water molecules in two spheres centered on a deoxyribose or Glu3 atom for OG bound in the FPG active site with anionic Glu3.[a, b]

***anti*-OG**	Deoxyribose (O4′)	Glu3^{-1} (Oε1)	Glu3^{-1} (Oε2)
r = 3.4 Å	2	1	1
r = 6.0 Å	6	3	3
***anti*-OG**	Deoxyribose (O4′)	Glu3 (Oε1)	Glu3 (Oε2)
r = 3.4 Å	1	0	0
r = 6.0 Å	4	2	1
***syn*-OG**	Deoxyribose (O4′)	Glu3^{-1} (Oε1)	Glu3^{-1} (Oε2)
r = 3.4 Å	2	0	1
r = 6.0 Å	5	2	2
***syn*-OG**	Deoxyribose (O4′)	Glu3 (Oε1)	Glu3 (Oε2)
r = 3.4 Å	1	0	0
r = 6.0 Å	4	1	1

[a] Average number of water molecules calculated for the cluster with the highest occupancy. [b]A spherical solvent shell with a radius of 3.4 or 6.0 Å was centered on O4′ of deoxyribose, or Oε1 or Oε2 of Glu3.

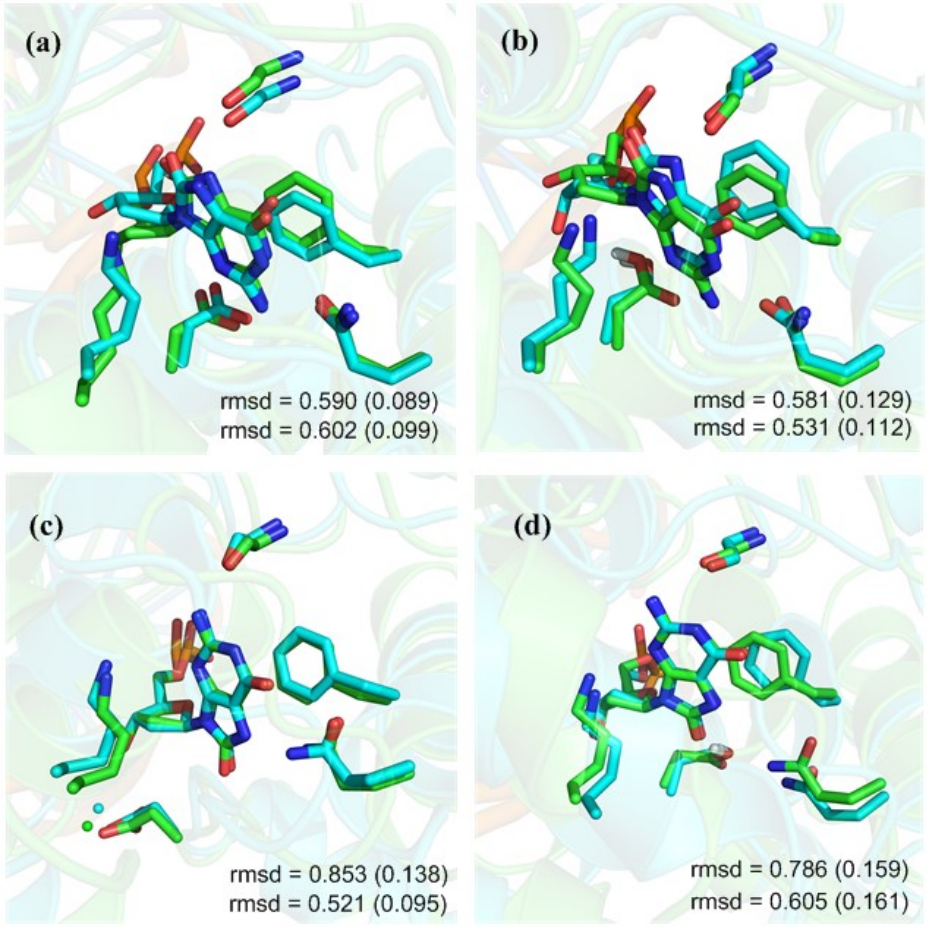

Figure A3.1. Overlays of MD representative structures from replicas for (a – b) *anti*- or (c – d) *syn*-OG bound in the hOgg1 active site with (a and c) anionic or (b and d) neutral Asp268. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

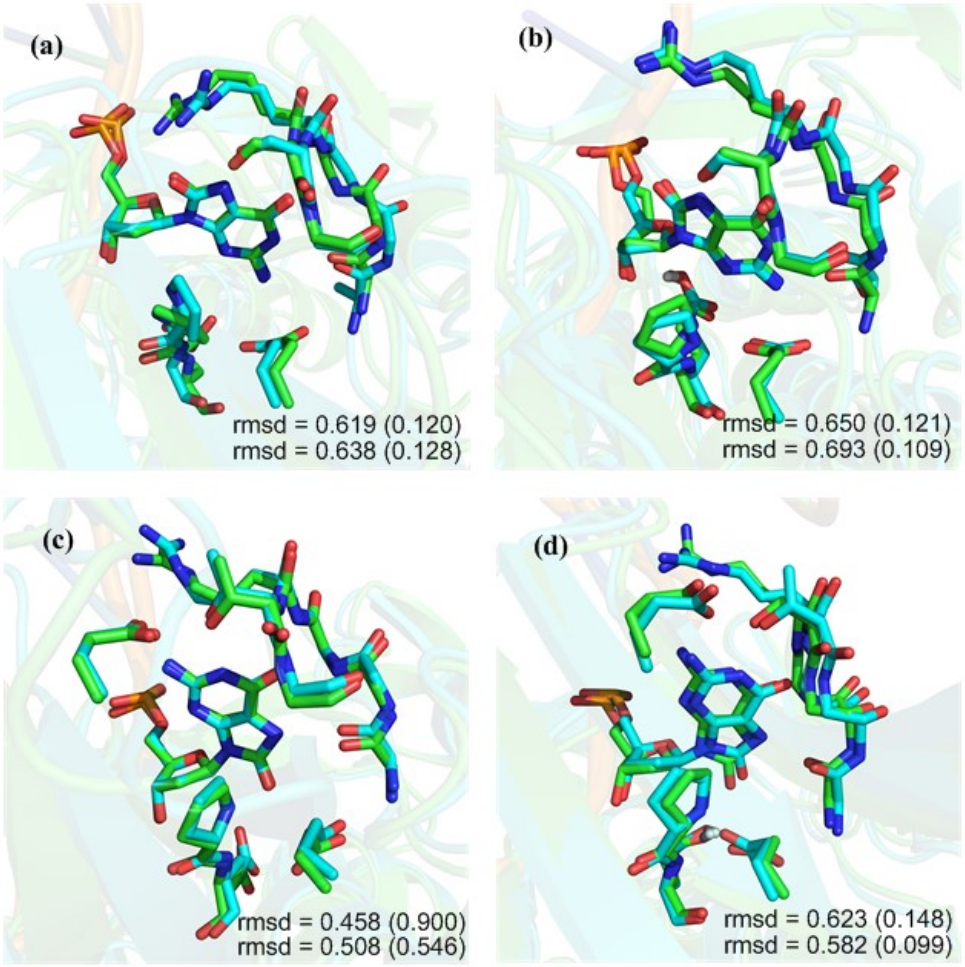

Figure A3.2. Overlays of MD representative structures from replicas for (a – b) *anti*- or (c – d) *syn*-OG bound in the FPG active site with (a and c) anionic or (b and d) neutral Glu3. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

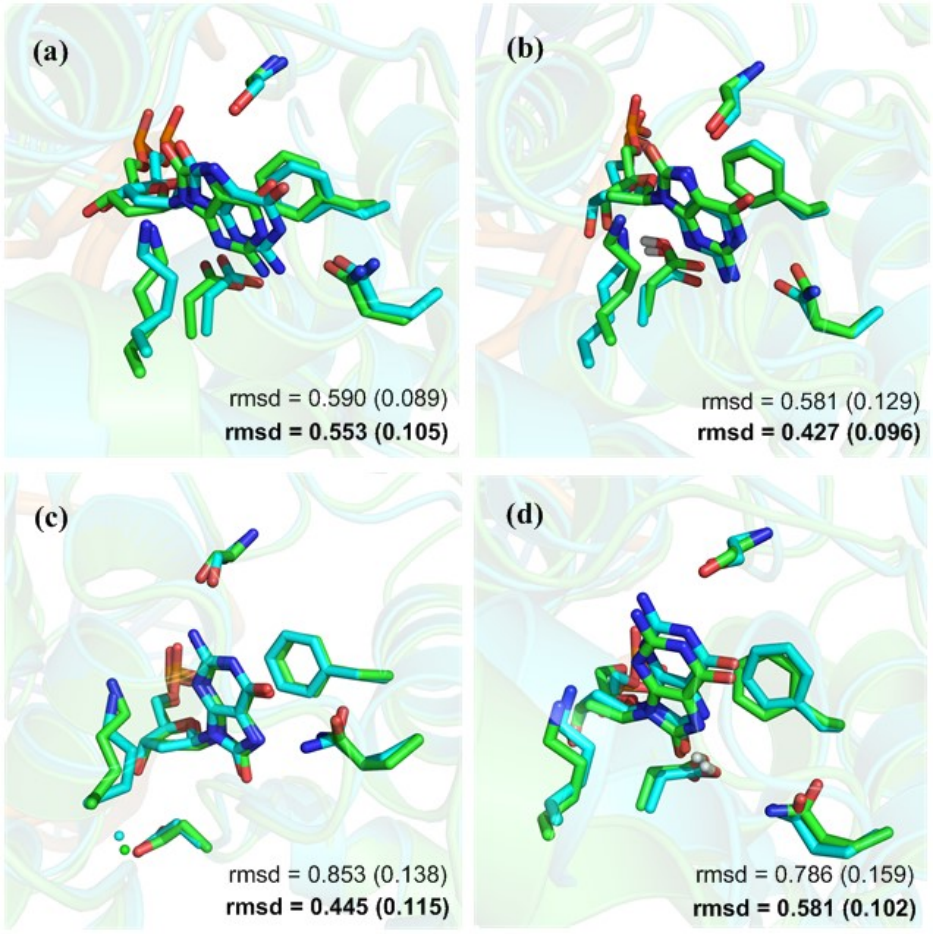

Figure A3.3. Overlays of MD representative structures of 20 ns (cyan) and 100 ns (green) simulations for (a – b) *anti*- or (c – d) *syn*-OG bound in the hOgg1 active site with (a and c) anionic or (b and d) neutral Asp268. rmsd (Å, bold for 100 ns simulation) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided. rmsd values between the representative structures of the 20 and 100 ns simulations for *anti* or *syn* with anionic or neutral Asp268 are a) 0.401, b) 0.586, c) 0.380, and d) 0.301 Å, respectively.

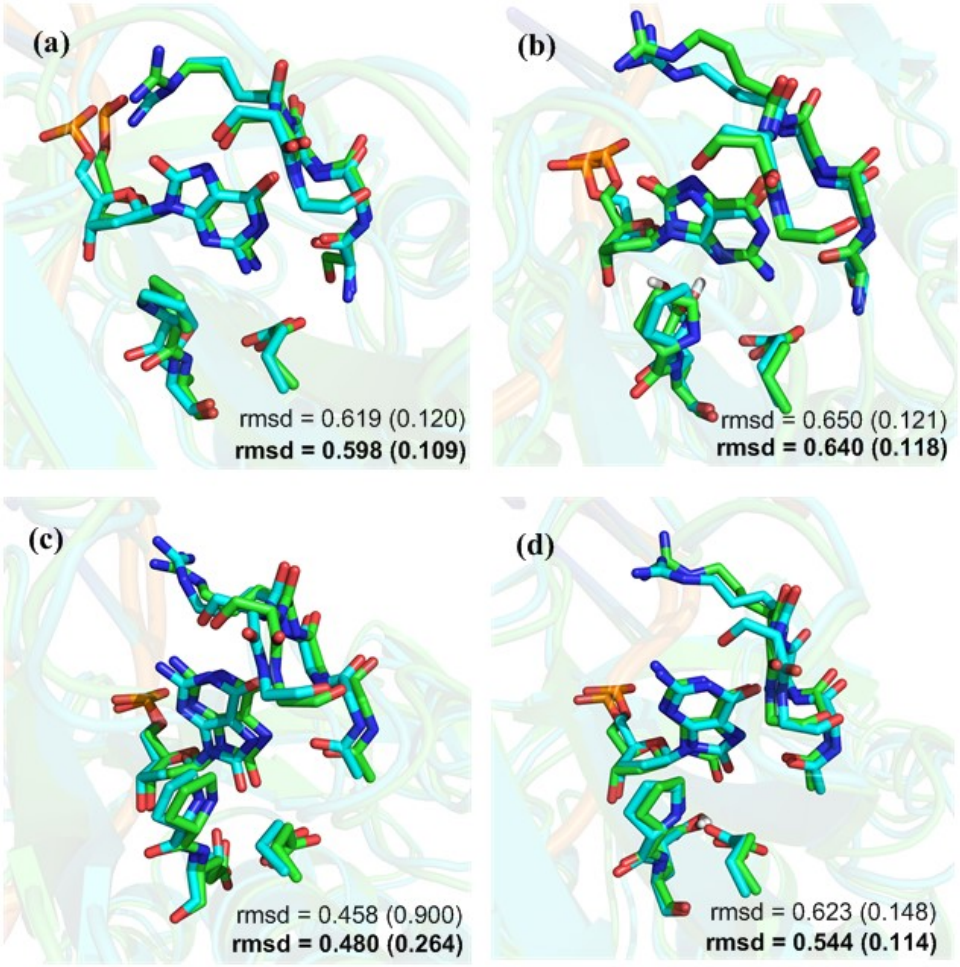

Figure A3.4. Overlays of MD representative structures of 20 ns (cyan) and 100 ns (green) simulations for (a – b) *anti*- or (c – d) *syn*-OG bound in the FPG active site with (a and c) anionic or (b and d) neutral Glu3. rmsd (Å, bold for 100 ns simulation) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided. rmsd values between the representative structures of the 20 and 100 ns simulations for *anti* or *syn* with anionic or neutral Glu3 are a) 0.325, b) 0.464, c) 0.471, and d) 0.377 Å, respectively.

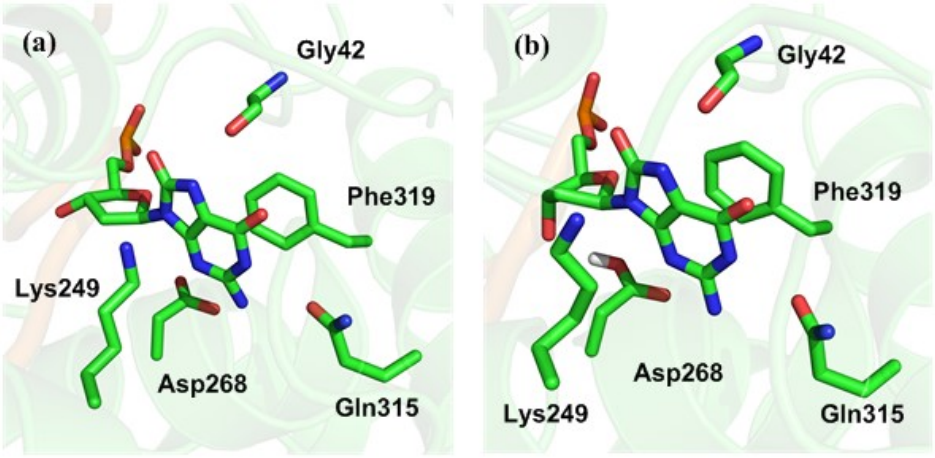

Figure A3.5. MD representative structure of *anti*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only the hydrogen atoms of neutral Asp268 (Oδ–H) is shown.

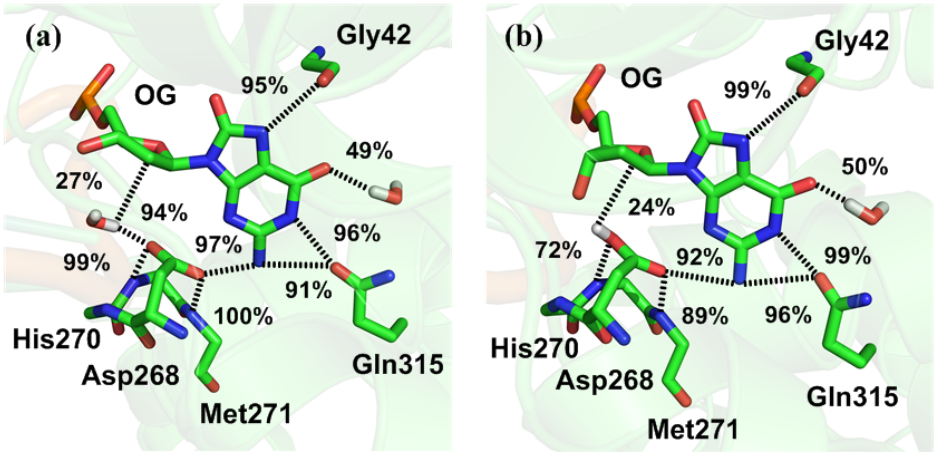

Figure A3.6. MD representative structure of *anti*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only the hydrogen atom of neutral Asp268 (Oδ–H) and water molecules are shown.

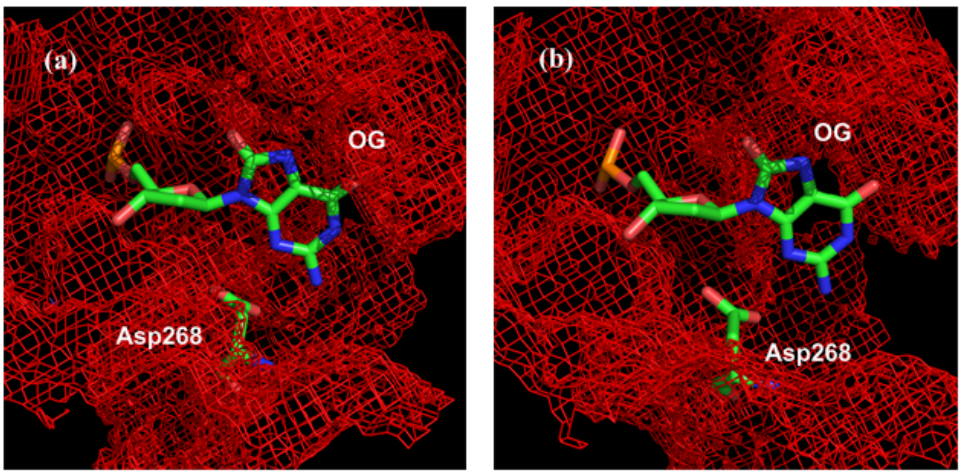

Figure A3.7. Water density distribution around *anti*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268.

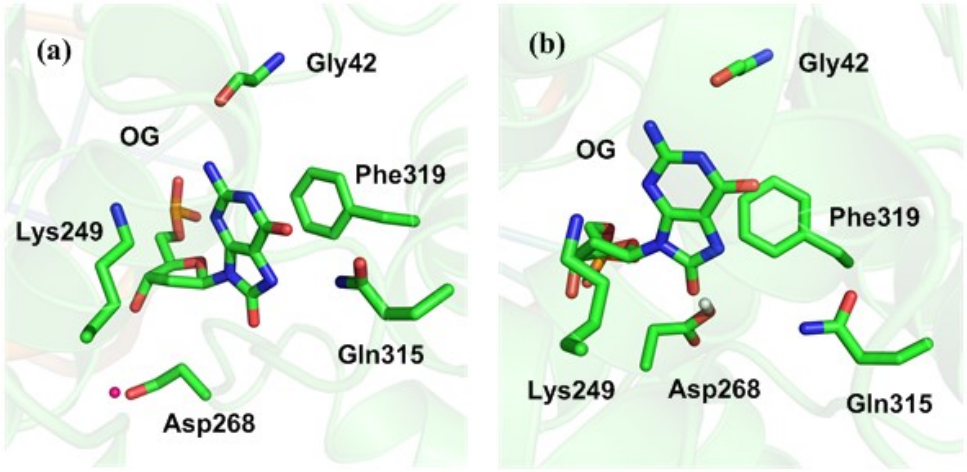

Figure A3.8. MD representative structure of *syn*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only hydrogen atom of neutral Asp268 (Oδ–H) is shown. Calcium ion is shown in pink.

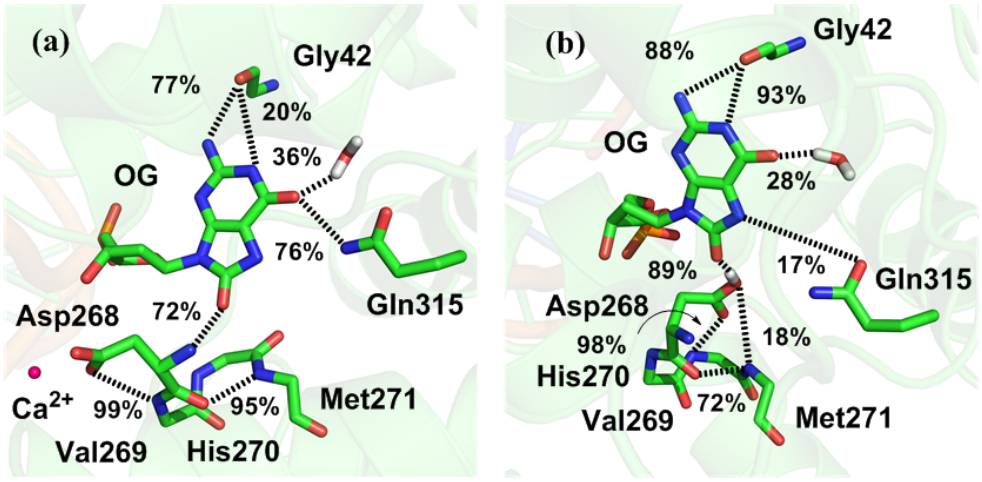

Figure A3.9. MD representative structure of *syn*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only hydrogen atoms of neutral Asp268 (Oδ–H) and water molecules are shown. Calcium ion is shown in pink.

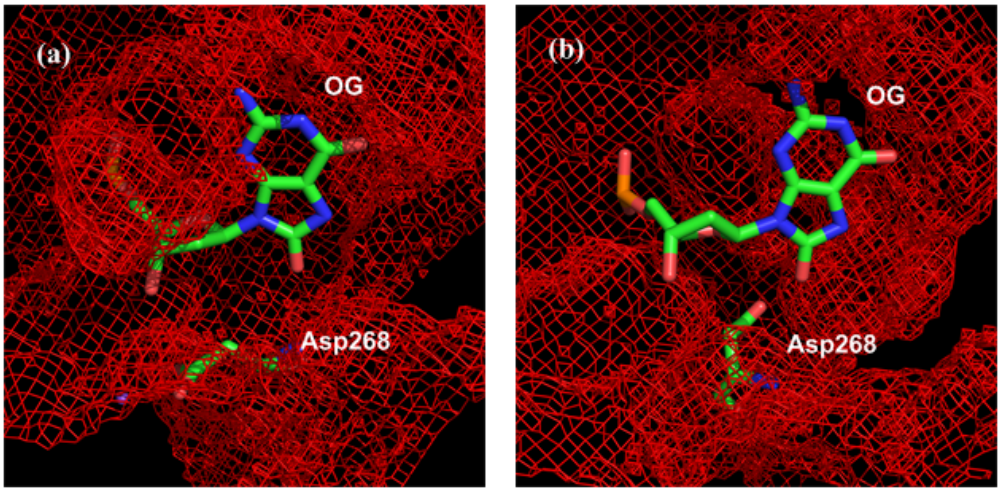

Figure A3.10. Water density distribution around *syn*-OG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268.

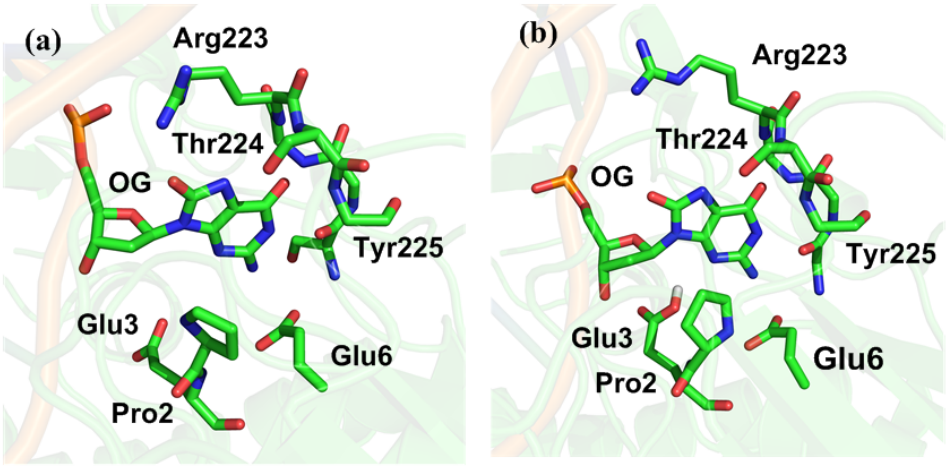

Figure A3.11. MD representative structure of *anti*-OG bound in the FPG active site with (a) anionic or (b) neutral Glu3. For clarity, only hydrogen atom of neutral Glu3 (Oε–H) is shown.

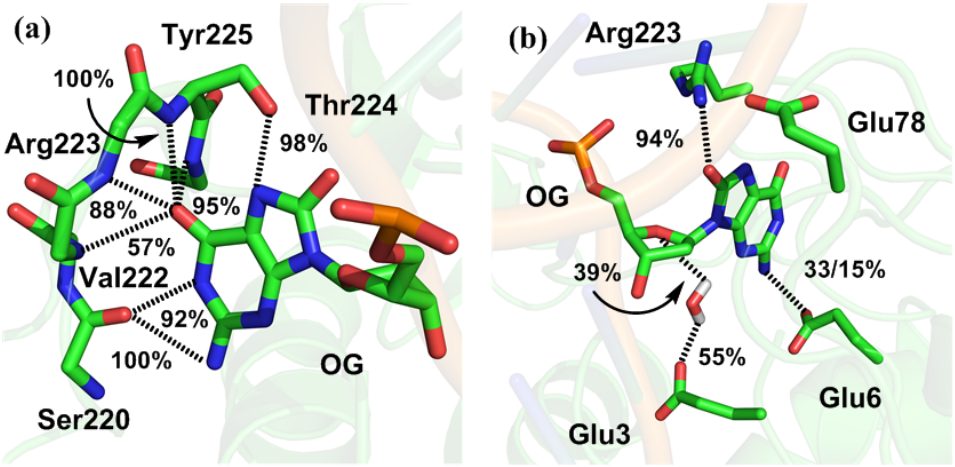

Figure A3.12. Major interactions between *anti*-OG and the FPG active site with anionic Glu3.

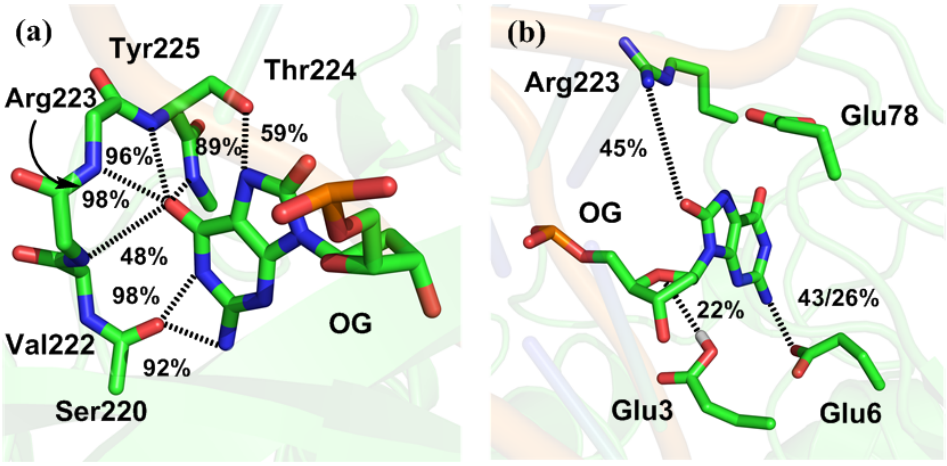

Figure A3.13. Major interactions between *anti*-OG and the FPG active site with neutral Glu3.

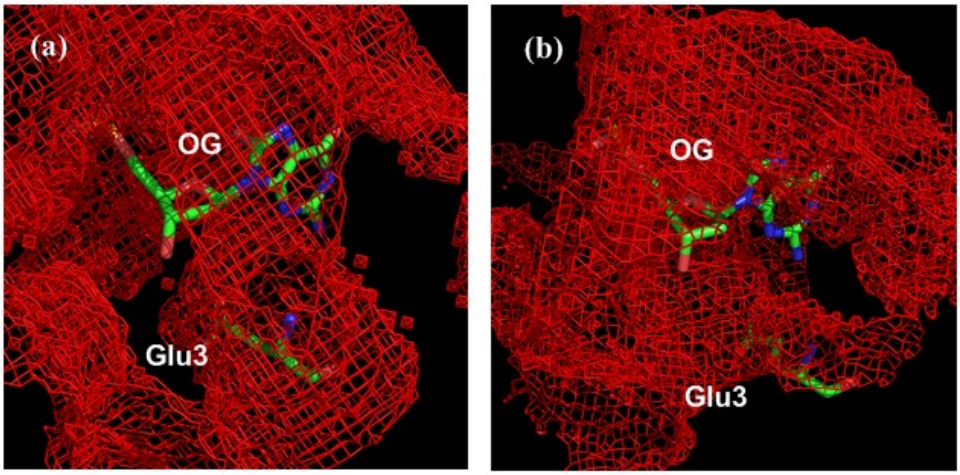

Figure A3.14. Water density distribution around *anti*-OG bound in the FPG active site with (a) anionic or (b) neutral Glu3.

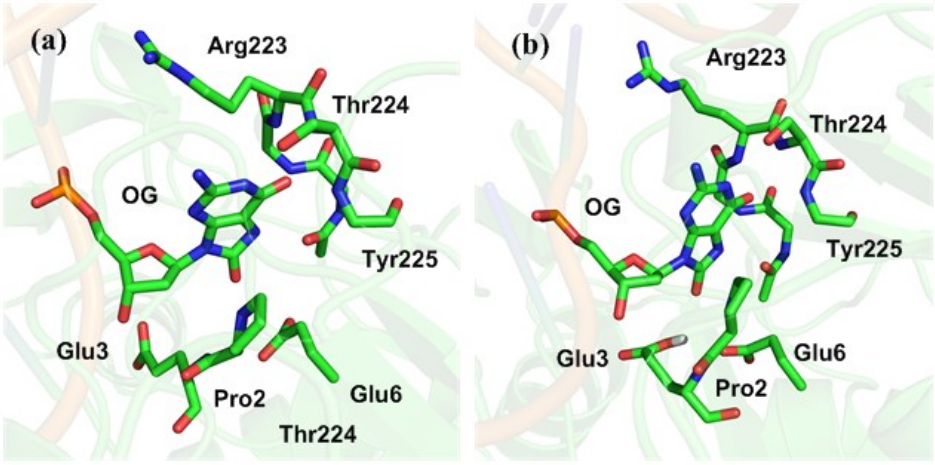

Figure A3.15. MD representative structure of *syn*-OG bound in the FPG active site with (a) anionic or (b) neutral Glu3. For clarity, only hydrogen atom of neutral Glu3 (Oε–H) is shown.

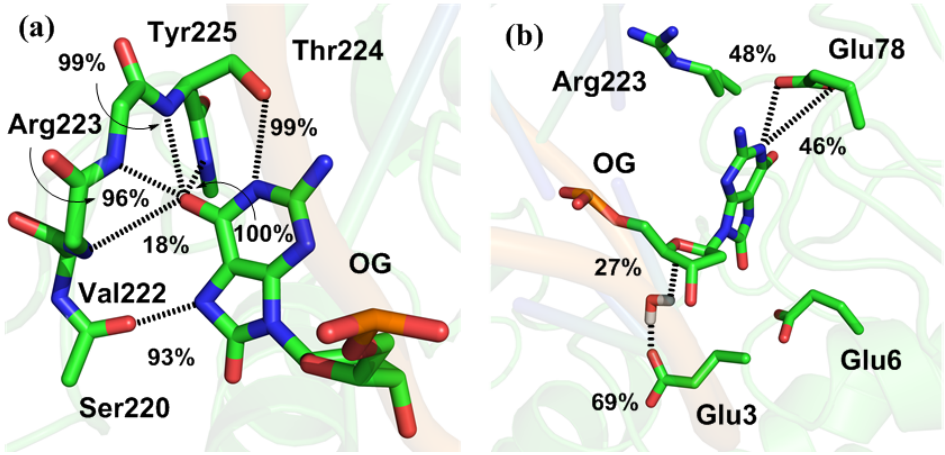

Figure A3.16. Major interactions between *syn*-OG and the FPG active site with anionic Glu3. For clarity, only hydrogen atoms of water molecules are shown.

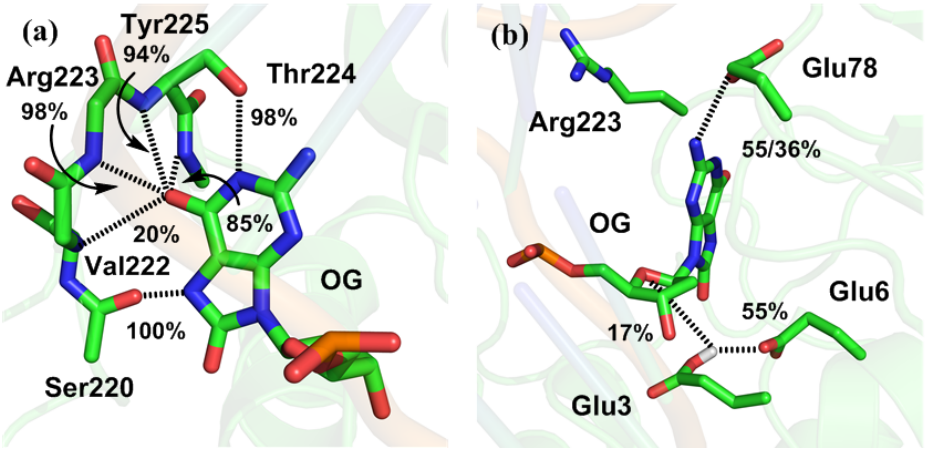

Figure A3.17. Major interactions between *syn*-OG and the FPG active site with neutral Glu3. For clarity, only hydrogen atom of neutral Glu3 (Oε–H) is shown.

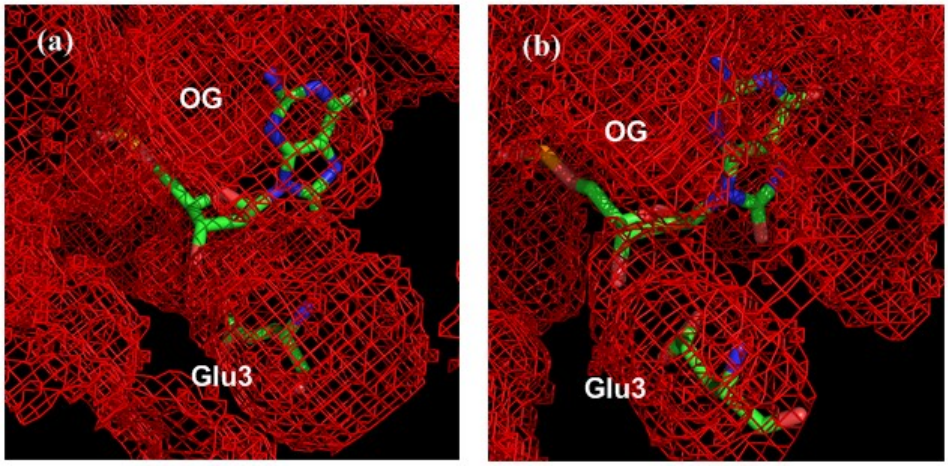

Figure A3.18. Water density distribution around *syn*-OG bound in the FPG active site with (a) anionic or (b) neutral Glu3.

Appendix A4: Supplementary Information for Chapter 4

Table of Contents of Appendix A4

Table A4.1. Root-mean-square deviation (rmsd in Å; standard deviation in parentheses) in the DNA and enzyme backbone over the entire and clustered MD simulation trajectory.

System	rmsd (simulation)[a]			rmsd (cluster)[b]	
hOGG1	**DNA**	**Enzyme**	**Occ.**	**DNA**	**Enzyme**
anti-FapyG (Asp268^{-1})	2.135 (0.508)	1.436 (0.322)	88	2.124 (0.491)	1.422 (0.300)
anti-FapyG (Asp268)	2.099 (0.489)	1.209 (0.103)	84	1.921 (0.422)	1.231 (0.191)
syn-FapyG (Asp268^{-1})	2.313 (0.526)	1.257 (0.141)	98	2.316 (0.540)	1.256 (0.141)
syn-FapyG (Asp268)	2.985 (0.539)	1.481 (0.209)	87	3.021 (0.564)	1.460 (0.187)
FPG	**DNA**	**Enzyme**	**Occ.**	**DNA**	**Enzyme**
anti-FapyG (Glu3^{-1})	1.795 (0.337)	1.131 (0.162)	97	1.794 (0.337)	1.131 (0.162)
anti-FapyG (Glu3)	1.830 (0.222)	0.901 (0.100)	95	1.832 (0.223)	0.900 (0.098)
syn-FapyG (Glu3^{-1})	1.897 (0.236)	1.215 (0.097)	96	1.900 (0.238)	1.216 (0.097)
syn-FapyG (Glu3)	2.022 (0.378)	0.925 (0.099)	98	2.020 (0.374)	0.924 (0.099)

[a] rmsd calculated according to the position of the DNA (P, O3′, O5′, C3′, C4′, and C5′) and enzyme (C_α, C, and N) backbone atoms over the entire simulation trajectory with respect to the first frame. [b] Clustering was completed for the sampled trajectories based on the configurations of key active site residues, namely FapyG, Gly42, Lys249, Asp268, Gln315, and Phe319 for hOgg1, and FapyG, Pro2, and Glu3 for FPG. rmsd calculated according to the position of the DNA (P, O3′, O5′, C3′, C4′, and C5′) and enzyme (C_α, C, and N) backbone atoms for the cluster with the highest occupancy with respect to the first frame of the production phase.

Table A4.2. Per residue root-mean-square deviation (rmsd in Å; standard deviation in parentheses) for hOgg1 active site residues with respect to the reference crystal structure.[a]

	Gly42	Lys249	Asp268	Gln315	Phe319
anti-FapyG (Asp268^{-1})	0.208 (0.089)	0.104 (0.053)	0.096 (0.045)	0.149 (0.046)	0.521 (0.236)
anti-FapyG (Asp268)	0.139 (0.080)	0.093 (0.050)	0.082 (0.032)	0.184 (0.064)	0.395 (0.177)
syn-FapyG (Asp268^{-1})	0.088 (0.039)	0.102 (0.043)	0.129 (0.055)	0.197 (0.069)	0.203 (0.082)
syn-FapyG (Asp268)	0.109 (0.091)	0.099 (0.049)	0.085 (0.041)	0.231 (0.071)	0.252 (0.230)

[a] rmsd calculated according to the position of the active site residues (all heavy atoms) for the cluster with the highest occupancy with respect to the crystal structure (PDB ID: 3KTU).

Table A4.3. Average χ dihedral angle of FapyG (∠(O4′–C1′–N9–C4) in degrees; standard deviation in parentheses) across the MD simulations and the reference crystal structures.[a]

hOgg1	
PDB ID: 3KTU	–124.3
anti-FapyG (Asp268^{-1})	–112.2 (16.3)
anti-FapyG (Asp268)	–109.1 (16.0)
syn-FapyG (Asp268^{-1})	64.7 (18.6)
syn-FapyG (Asp268)	88.6 (24.9)
FPG	
PDB ID: 1R2Y[b]	101.1
PDB ID: 1XC8[c]	–64.3
anti-FapyG (Glu3^{-1})	–114.6 (15.8)
anti-FapyG (Glu3)	–99.4 (20.5)
syn-FapyG (Glu3^{-1})	69.9 (17.1)
syn-FapyG (Glu3)	94.2 (19.6)

[a] Average dihedral angle calculated for the cluster with the highest occupancy with respect to the first frame of the production phase. [b] Crystal structure of OG bound in the FPG active site used in this study. [c] Crystal structure of modified FapyG bound in the FPG active site.

Table A4.4. Average dihedral angle of FapyG opened-imidazole ring (∠(C4–C5–C8–O8) and ∠(C5–N7–C8–O8) in degrees; standard deviation in parentheses) across the MD simulations and the reference crystal structures.[a]

hOgg1	∠C4–C5–C8–O8	∠C5–N7–C8–O8
anti-FapyG (Asp268^{-1})	43.7 (20.5)	12.8 (19.1)
anti-FapyG (Asp268)	34.7 (18.9)	3.2 (11.0)
syn-FapyG (Asp268^{-1})	–47.5 (8.0)	3.0 (9.6)
syn-FapyG (Asp268)	–57.2 (50.5)	4.9 (10.5)
FPG		
PDB ID: 1XC8[b]	–54.2	35.3
anti-FapyG (Glu3^{-1})	118.1 (16.7)	–3.4 (10.6)
anti-FapyG (Glu3)	–101.5 (40.2)	4.6 (10.8)
syn-FapyG (Glu3^{-1})	–46.6 (9.4)	2.5 (8.4)
syn-FapyG (Glu3)	–49.7 (8.3)	2.9 (9.3)

[a] Average dihedral angle calculated for the cluster with the highest occupancy with respect to the first frame of the production phase. [b] Crystal structure of modified FapyG bound in the FPG active site.

Table A4.5. Hydrogen-bonding occupancies (%) for interactions between FapyG and hOgg1 active site residues or solvent, and corresponding average distance (Å) and angle (in degrees).[a, b]

Interaction				
***anti*-FapyG···Residue**	**Asp268^{-1}**	**distance (angle)**	**Asp268**	**distance (angle)**
N7–H···O(Gly42)	22	3.132 (128.1)	57	3.212 (132.0)
N2–H···Oδ(Asp268)	74	2.932 (165.8)	85	3.033 (159.0)
N2–H···Oε(Gln315)	85	3.102 (143.0)	86	3.025 (145.2)
N1–H···Oε(Gln315)	97	2.915 (152.2)	95	2.872 (152.0)
O3′···H–Oδ(Asp268)	–	–	68	2.889 (143.0)
O4′···H–Oδ(Asp268)	–	–	39	2.940 (131.9)
O6···H(H_2O)	59	2.663 (159.1)	61	2.802 (161.4)
O4′···H(H_2O)	27	3.109 (137.8)	31	2.819 (156.8)
***syn*-FapyG···Residue**	**Asp268^{-1}**	**distance (angle)**	**Asp268**	**distance (angle)**
N2–H···O(Gly42)	75	2.991 (157.7)	72	2.967 (155.8)
N1–H···O(Gly42)	33	3.195 (141.0)	30	3.064 (143.4)
N9–H···Oδ(Asp268)	39	2.969 (145.1)	60	2.886 (156.3)
N7–H···Oε(Gln315)	87	2.959 (151.1)	76	2.898 (153.2)
O6···H–Nε(Gln315)	74	2.961 (157.6)	69	2.971 (162.7)
O3′···H–Oδ(Asp268)	–	–	80	3.022 (130.4)
O4′···H–Oδ(Asp268)	–	–	34	3.051 (129.9)
O6···H(H_2O)	15	3.192 (139.3)	18	2.869 (152.4)
O4′···H(H_2O)	37	3.001 (148.1)	8	2.994 (134.6)

[a] Hydrogen-bonding occupancies calculated for the cluster with the highest occupancy. [b] Percent occupancy of each hydrogen bond is defined using a distance cut-off of < 3.4 Å and an angle cut-off of > 120°.

Table A4.6. Average interaction energies (ΔE in kcal/mol; standard deviations in parentheses) between various hOgg1 active site residues and *anti*- or *syn*-FapyG for different Asp268 protonation states.[a]

***anti*-FapyG**		
Residue	**Asp268^{-1}**	**Asp268**
Gly42	–2.0 (2.4)	–2.8 (2.4)
Asp268	–10.5 (3.1)	–13.7 (2.6)
Gln315	–13.4 (2.2)	–12.1 (2.8)
Phe319	–3.5 (1.1)	–4.1 (0.5)
H_2O	–3.8 (0.8)	–2.9 (0.7)
***syn*-FapyG**		
Residue	**Asp268^{-1}**	**Asp268**
Gly42	–5.1 (2.2)	–5.0 (3.0)
Asp268	–5.0 (3.2)	–5.6 (2.4)
Gln315	–11.3 (2.9)	–11.9 (5.1)
Phe319	–3.8 (0.7)	–3.9 (0.8)
H_2O	–1.8 (0.8)	–1.1 (1.0)

[a] Average linear interaction energy (LIE) and standard deviations (in parentheses) calculated for the cluster with the highest occupancy.

Table A4.7. Average distances (Å; standard deviation in parentheses) between the nucleophile or Asp/Glu and 2′-deoxyribose,[a] and between FapyG and Phe319 in the hOgg1 active site.[b]

System	Distance (Å)			
hOgg1	**Lys249(Nζ···C1′)**	**Asp268(Oδ1···O4′)**	**Asp268(Oδ2···O4′)**	**FapyG···Phe319**[b]
anti-FapyG (Asp268^{-1})	4.219 (0.448)	4.123 (0.708)	4.755 (0.756)	4.028 (0.363)
anti-FapyG (Asp268)	4.238 (0.437)	3.699 (0.318)	4.898 (0.340)	4.009 (0.344)
syn-FapyG (Asp268^{-1})	4.958 (0.399)	4.814 (0.791)	5.189 (0.772)	4.087 (0.331)
syn-FapyG (Asp268)	4.419 (0.511)	3.621 (0.326)	4.510 (0.375)	4.138 (0.412)
FPG	**Pro2(N···C1′)**	**Glu3(Oε1···O4′)**	**Glu3(Oε2···O4′)**	
anti-FapyG (Glu3^{-1})	4.221 (0.532)	4.744 (0.745)	6.746 (0.727)	–
anti-FapyG (Glu3)	4.326 (0.592)	3.809 (0.450)	5.392 (0.981)	–
syn-FapyG (Glu3^{-1})	4.871 (0.539)	4.300 (0.622)	6.112 (0.318)	–
syn-FapyG (Glu3)	4.797 (0.566)	3.894 (0.662)	4.201 (0.631)	–

[a] Average distances and standard deviations between the atoms measured for the cluster with the highest occupancy. [b] Average distances and standard deviations were evaluated between the centers of the mass of FapyG and Phe319 for hOgg1.

Table A4.8. Average number of water molecules in two spheres centered on a deoxyribose or Asp268 atom or FapyG bound in the hOgg1 active site.[a]

***anti*-FapyG**	Deoxyribose (O4′)	Asp268^{-1} (Oδ1)	Asp268^{-1} (Oδ2)
r = 3.4 Å	2	2	1
r = 6.0 Å	4	5	5
***anti*-FapyG**	Deoxyribose (O4′)	Asp268 (Oδ1)	Asp268 (Oδ2)
r = 3.4 Å	1	1	1
r = 6.0 Å	5	3	4
***syn*-FapyG**	Deoxyribose (O4′)	Asp268^{-1} (Oδ1)	Asp268^{-1} (Oδ2)
r = 3.4 Å	2	2	1
r = 6.0 Å	7	6	6
***syn*-FapyG**	Deoxyribose (O4′)	Asp268 (Oδ1)	Asp268 (Oδ2)
r = 3.4 Å	1	1	2
r = 6.0 Å	7	6	8

[a] Average number of water molecules calculated for the cluster with the highest occupancy. [b] A spherical solvent shell with a radius of 3.4 or 6.0 Å was centered on O4′ of deoxyribose, or Oδ1 or Oδ2 of Asp268.

Table A4.9. Average values of possible nucleophilic attack angles (in degrees; standard deviations in parentheses).[a]

System	Angle (degree)	
hOgg1	**∠(Nζ–C1′–O4′)**	**∠(Nζ–C1′–N9)**
anti-FapyG (Asp268^{-1})	159.0 (11.9)	63.8 (11.0)
anti-FapyG (Asp268)	155.4 (12.1)	64.5 (8.9)
syn-FapyG (Asp268^{-1})	109.1 (13.9)	70.2 (7.3)
syn-FapyG (Asp268)	146.6 (13.1)	73.9 (10.7)
FPG	**∠(N–C1′–O4′)**	**∠(N–C1′–N9)**
anti-FapyG (Glu3^{-1})	155.8 (12.2)	71.6 (11.1)
anti-FapyG (Glu3)	159.0 (9.9)	76.9 (9.3)
syn-FapyG (Glu3^{-1})	152.4 (10.2)	68.8 (13.5)
syn-FapyG (Glu3)	148.6 (9.6)	86.4 (9.6)

[a] Average angles and standard deviations calculated for the cluster with the highest occupancy.

Table A4.10. Per residue root-mean-square deviation (rmsd in Å; standard deviation in parentheses) for FPG active site residues with respect to the reference crystal structure.[a]

	***anti*-FapyG (Glu3^{-1})**	***anti*-FapyG (Glu3)**	***syn*-FapyG (Glu3^{-1})**	***syn*-FapyG (Glu3)**
Pro2	0.335 (0.126)	0.286 (0.112)	0.253 (0.088)	0.312 (0.109)
Glu3	0.256 (0.086)	0.489 (0.103)	0.957 (0.113)	0.200 (0.063)
Glu6	0.178 (0.062)	0.185 (0.063)	0.225 (0.089)	0.280 (0.081)
Glu78	0.695 (0.310)	0.301 (0.290)	0.884 (0.249)	0.929 (0.217)
Arg80	0.459 (0.042)	0.467 (0.053)	0.730 (0.061)	0.855 (0.219)
Ser220	0.634 (0.040)	0.129 (0.043)	0.150 (0.039)	0.177 (0.063)
Thr221	0.216 (0.041)	0.130 (0.037)	0.179 (0.049)	0.161 (0.050)
Val222	0.109 (0.046)	0.784 (0.341)	0.702 (0.433)	0.484 (0.358)
Arg223	0.960 (0.155)	0.894 (0.260)	0.910 (0.211)	0.669 (0.415)
Thr224	0.189 (0.133)	0.170 (0.060)	0.190 (0.044)	0.229 (0.071)
Tyr225	0.140 (0.044)	0.137 (0.049)	0.128 (0.052)	0.139 (0.048)

[a] rmsd calculated according to the position of the active site residues (all heavy atoms) for the cluster with the highest occupancy with respect to the crystal structure (PDB ID: 1R2Y).

Table A4.11. Hydrogen-bonding occupancies (%) for interactions between FapyG and FPG active site residues or solvent, and corresponding average interaction distance (Å) and angle (in degrees).[a, b]

Interaction		Model		
***anti*-FapyG···Residue**	**Glu3^{-1}**	**distance (angle)**	**Glu3**	**distance (angle)**
N2–H···O(Ser220)	87	3.019 (139.2)	89	2.979 (141.2)
N1–H···O(Ser220)	93	2.840 (152.7)	94	2.826 (152.3)
O6···H–N(Val222)	19	3.123 (143.9)	61	3.164 (149.6)
O6···H–N(Arg223)	99	3.020 (151.3)	98	3.146 (154.0)
O8···H–N(Arg223)	–	–	39	2.935 (131.4)
O6···H–N(Thr224)	78	2.999 (144.4)	88	3.022 (149.3)
O8···H–N(Thr224)	81	3.115 (134.5)	33	3.045 (154.0)
O8···H–Oγ(Thr224)	37	2.735 (156.3)	–	–
O6···H–N(Tyr225)	78	3.144 (159.5)	82	2.978 (155.4)
N2–H···Oε(Glu6)	53	3.052 (132.2)	60	3.084 (136.0)
N2–H···Oε(Glu6)	44	3.038 (132.2)	57	3.046 (124.7)
N7–H···O(H_2O)	41	2.995 (160.8)	30	2.937 (158.1)
N7–H···Oγ(Thr224)	–	–	38	3.040 (149.7)
N9–H···O(H_2O)	57	3.029 (151.5)	52	3.000 (154.6)
O4′···H–Oε(Glu3)	–	–	26	3.068 (135.3)
O3′···H–Oε(Glu3)	–	–	60	2.860 (160.0)
O4′···H–O(H_2O)	39	2.901 (149.5)	18	2.964 (152.2)
***syn*-FapyG···Residue**	**Glu3^{-1}**	**distance (angle)**	**Glu3**	**distance (angle)**
N7–H···O(Ser220)	98	2.979 (142.9)	99	3.042 (135.2)
N1–H···Oγ(Thr224)	80	3.004 (152.3)	88	2.902 (159.8)
N2–H···Oγ(Thr224)	76	3.212 (135.2)	46	2.957 (123.7)
O6···H–N(Val222)	42	3.201 (131.7)	39	3.201 (131.7)
O6···H–N(Arg223)	98	3.128 (147.0)	98	3.068 (154.9)
O6···H–N(Thr224)	97	2.887 (160.9)	98	2.965 (159.8)
O6···H–N(Tyr225)	99	2.965 (163.9)	97	2.965 (163.2)
N1–H···Oε(Glu78)	33	3.043 (142.6)	69	3.049 (142.6)
N2–H···Oε(Glu78)	29	2.946 (148.3)	43	3.051 (143.9)
O4′···H–Oε(Glu3)	–	–	31	3.096 (137.9)
O8···H–N(Glu6)	19	3.130 (132.7)	22	3.270 (134.0)
O4′···H–O(H_2O)	10	3.059 (148.5)	19	3.187 (149.7)

[a] Hydrogen-bonding occupancies calculated for the cluster with the highest occupancy. [b] Percent occupancy of the hydrogen bond is defined using a distance cut-off of < 3.4 Å and an angle cut-off of > 120°.

Table A4.12. Average interaction energies (ΔE in kcal/mol; standard deviations in parentheses) between various FPG active site residues and *anti*- or *syn*-FapyG.[a]

***anti*-FapyG**		
Residue	**Glu3^{-1}**	**Glu3**
Glu3	–2.1 (1.4)	–10.2 (3.3)
Glu6	–16.3 (2.9)	–16.0 (2.0)
Met77	–4.0 (1.5)	–4.8 (1.3)
Glu78	–0.9 (0.4)	–05 (0.3)
Ile175	0.6 (0.6)	0.4 (0.3)
Ser220	–9.8 (2.1)	–10.2 (2.3)
Thr221	–3.8 (0.6)	–2.5 (0.7)
Val222	–2.6 (0.7)	–2.9 (0.8)
Arg223	–4.9 (0.6)	–5.0 (0.7)
Thr224	–2.0 (0.3)	–3.3 (0.5)
Tyr225	–4.9 (2.2)	–4.4 (1.7)
***syn*-FapyG**		
Residue	**Glu3^{-1}**	**Glu3**
Glu3	–3.8 (0.7)	–4.9 (1.7)
Glu6	–1.7 (0.7)	–1.6 (0.5)
Met77	–6.7 (1.3)	–6.0 (1.2)
Glu78	–17.2 (8.4)	–19.0 (6.6)
Ile175	–1.1 (0.4)	–1.6 (0.7)
Ser220	–3.9 (0.9)	–4.4 (1.0)
Thr221	–2.7 (0.6)	–2.1 (0.6)
Val222	–2.9 (0.8)	–2.3 (1.6)
Arg223	–3.2 (0.7)	–2.9 (0.4)
Thr224	–6.0 (1.1)	–3.9 (0.8)
Tyr225	–5.0 (2.2)	–5.2 (1.9)

[a] Average linear interaction energies (LIE) and standard deviations were calculated for the cluster with the highest occupancy.

Table A4.13. Average number of water molecules in two spheres centered on a deoxyribose or Glu3 atom for FapyG bound in the FPG active site with anionic Glu3.[a, b]

***anti*-FapyG**	Deoxyribose (O4′)	$Glu3^{-1}$ (Oε1)	$Glu3^{-1}$ (Oε2)
r = 3.4 Å	1	1	0
r = 6.0 Å	5	5	4
***anti*-FapyG**	Deoxyribose (O4′)	Glu3 (Oε1)	Glu3 (Oε2)
r = 3.4 Å	1	0	0
r = 6.0 Å	4	1	1
***syn*-FapyG**	Deoxyribose (O4′)	$Glu3^{-1}$ (Oε1)	$Glu3^{-1}$ (Oε2)
r = 3.4 Å	1	0	0
r = 6.0 Å	3	1	1
***syn*-FapyG**	Deoxyribose (O4′)	Glu3 (Oε1)	Glu3 (Oε2)
r = 3.4 Å	1	0	0
r = 6.0 Å	4	1	1

[a] Average number of water molecules calculated for the cluster with the highest occupancy. [b] A spherical solvent shell with a radius of 3.4 or 6.0 Å was centered on O4′ of deoxyribose, or Oε1 or Oε2 of Glu3.

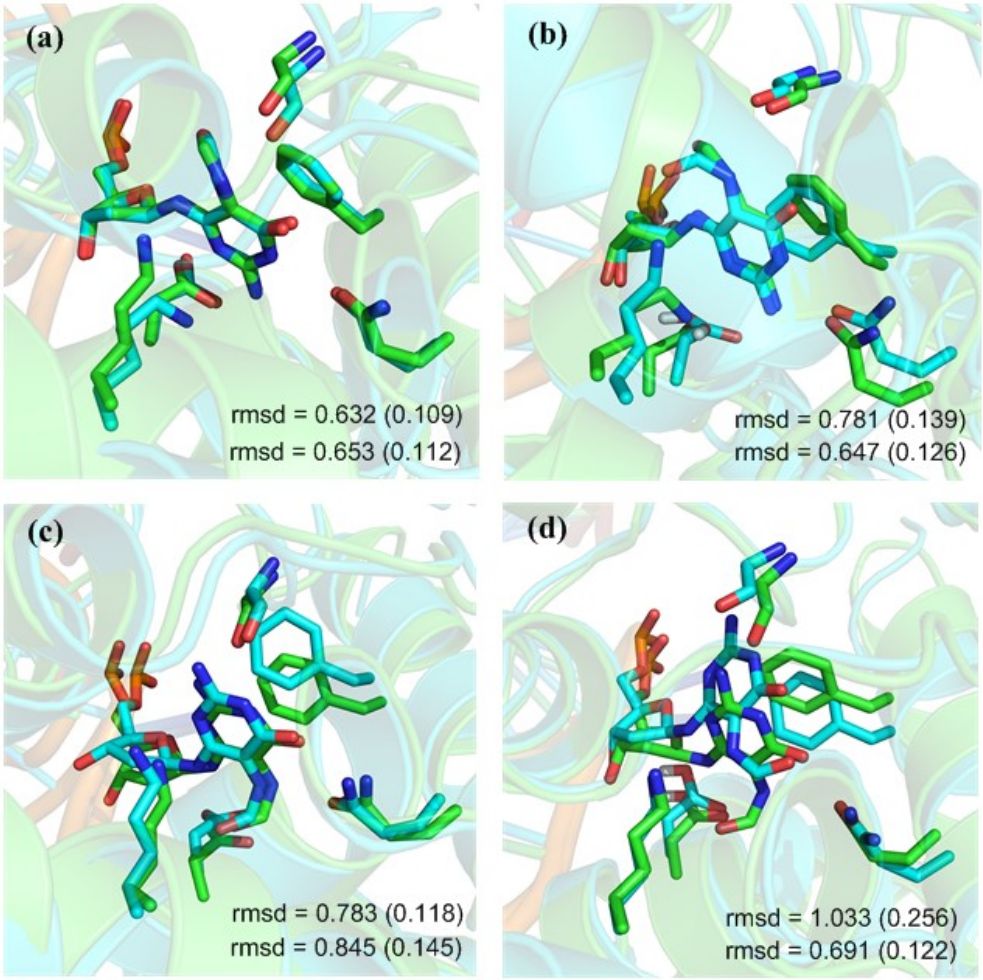

Figure A4.1. Overlays of MD representative structures from replicas for (a – b) *anti*- or (c – d) *syn*-FapyG bound in the hOgg1 active site with (a and c) anionic or (b and d) neutral Asp268. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

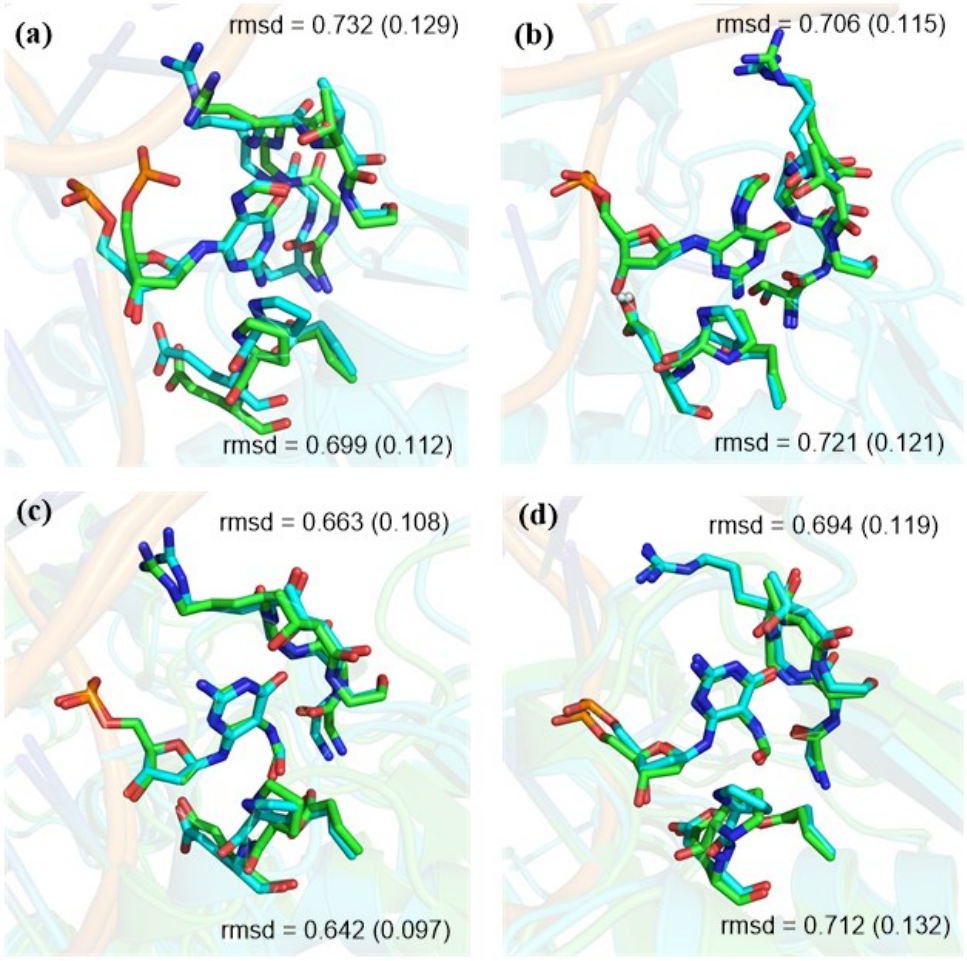

Figure A4.2. Overlays of MD representative structures from replicas for (a – b) *anti*- or (c – d) *syn*-FapyG bound in the FPG active site with (a and c) anionic or (b and d) neutral Glu3. rmsd (Å) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

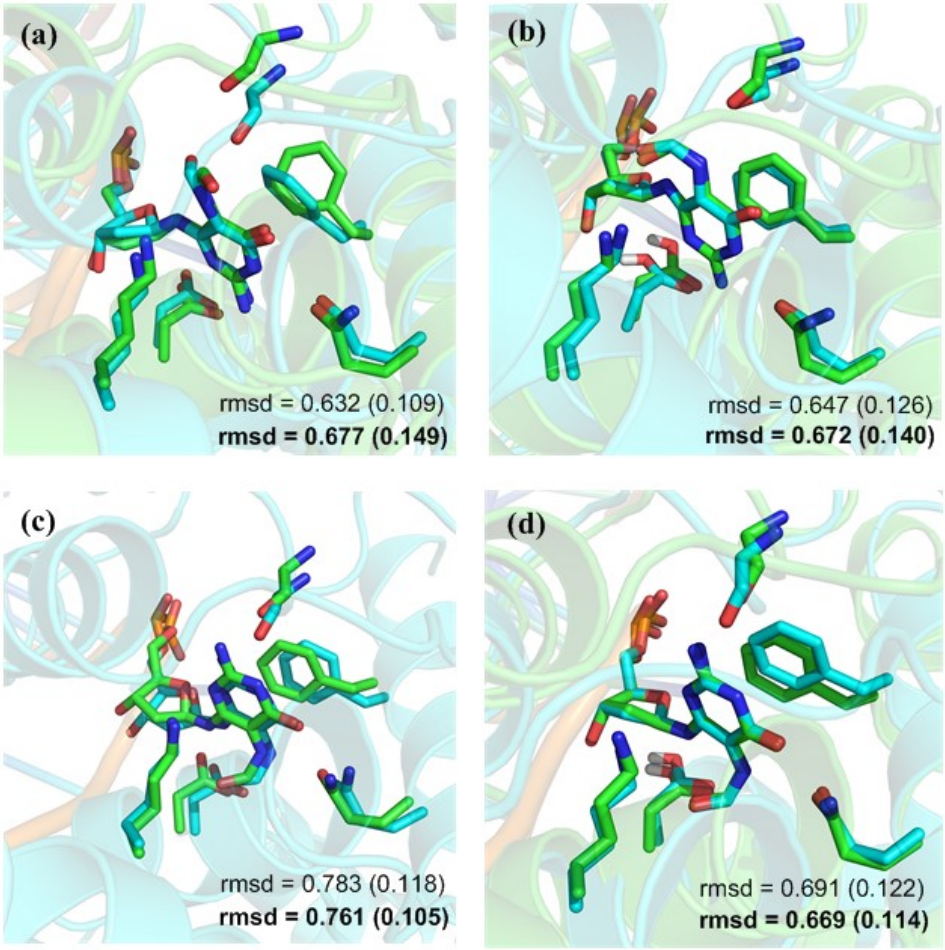

Figure A4.3. Overlays of MD representative structures of 20 ns (cyan) and 100 ns (green) simulations for (a – b) *anti*- or (c – d) *syn*-FapyG bound in the hOgg1 active site with (a and c) anionic or (b and d) neutral Asp268. rmsd (Å, bold for 100 ns simulation) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

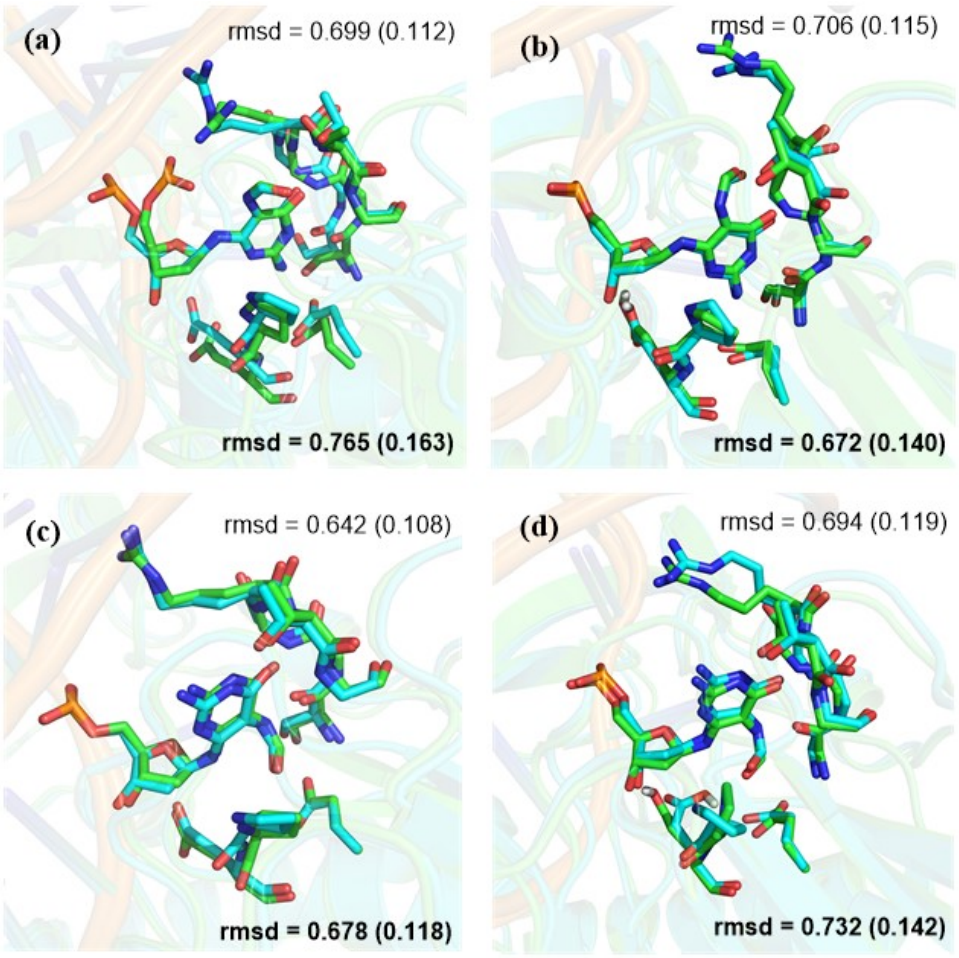

Figure A4.4. Overlays of MD representative structures of 20 ns (cyan) and 100 ns (green) simulations for (a – b) *anti*- or (c – d) *syn*-FapyG bound in the FPG active site with (a and c) anionic or (b and d) neutral Glu3. rmsd (Å, bold for 100 ns simulation) of active site backbone (standard deviation in parentheses) with respect to the crystal structure are provided.

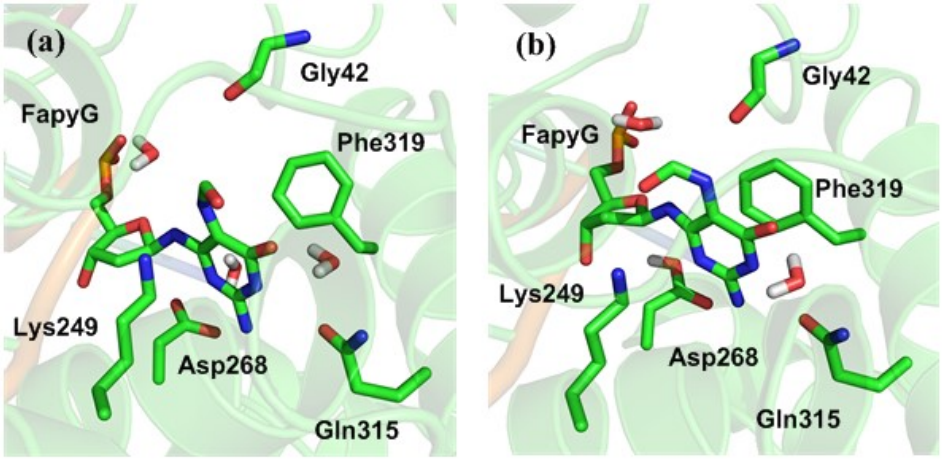

Figure A4.5. MD representative structure of *anti*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only the hydrogen atoms of neutral Asp268 (Oδ–H) is shown.

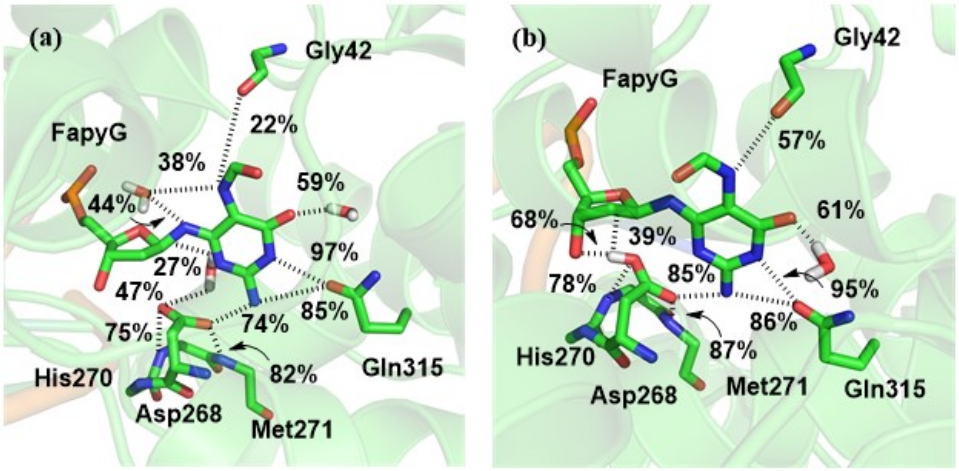

Figure A4.6. MD representative structure of *anti*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only the hydrogen atom of neutral Asp268 (Oδ–H) and water molecules are shown.

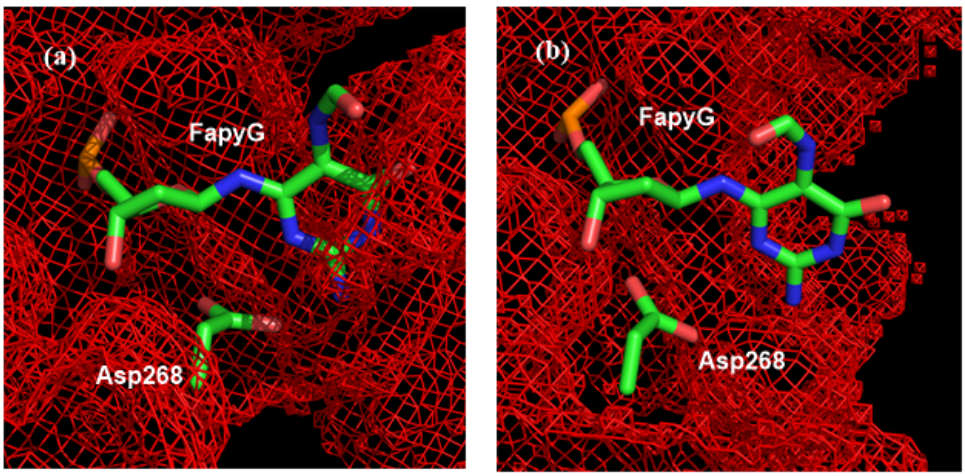

Figure A4.7. Water density distribution around *anti*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268.

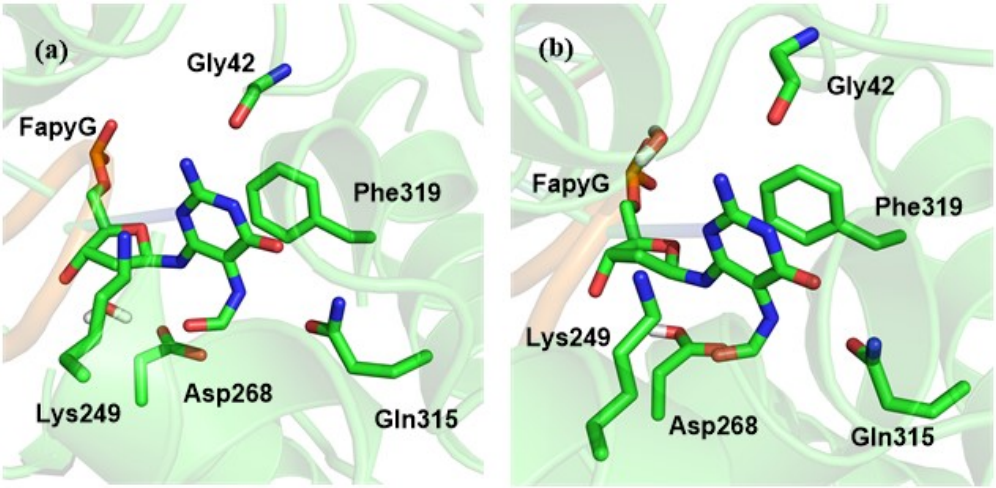

Figure A4.8. MD representative structure of *syn*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only hydrogen atom of neutral Asp268 (Oδ–H) is shown. Calcium ion is shown in pink.

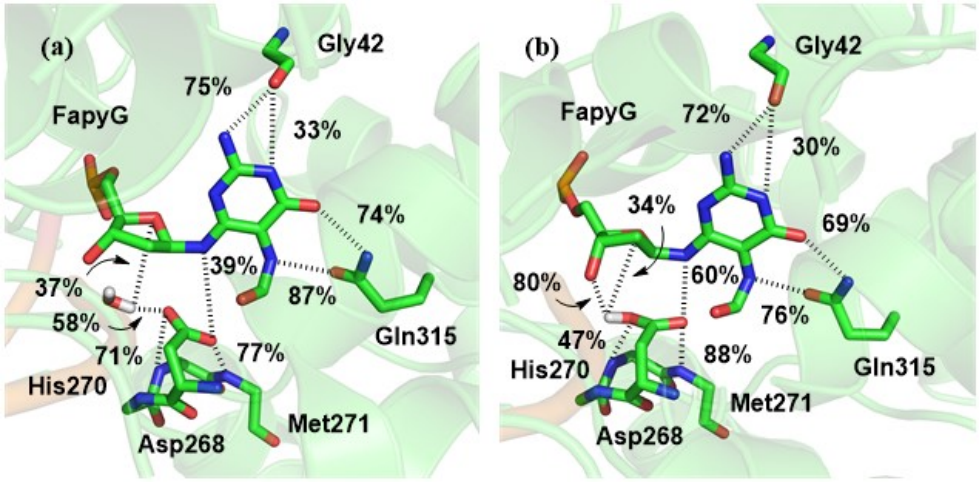

Figure A4.9. MD representative structure of *syn*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268. For clarity, only hydrogen atoms of neutral Asp268 (Oδ–H) and water molecules are shown. Calcium ion is shown in pink.

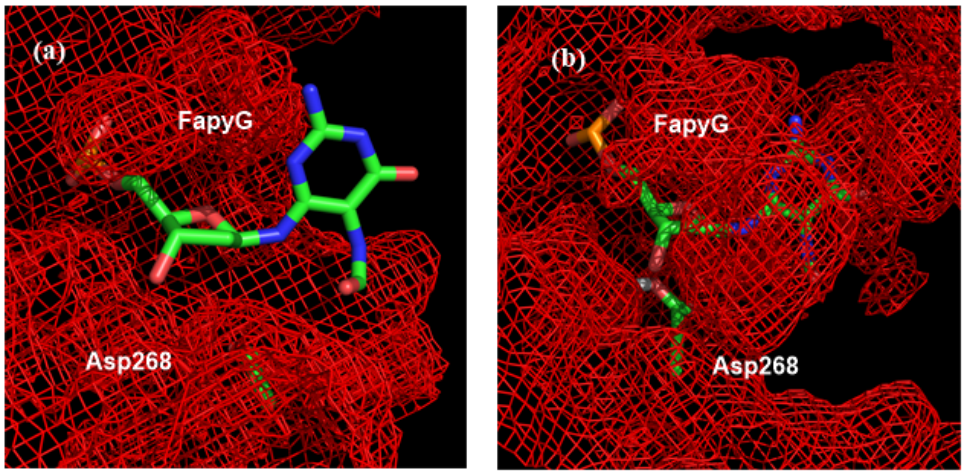

Figure A4.10. Water density distribution around *syn*-FapyG bound in the hOgg1 active site with (a) anionic or (b) neutral Asp268.

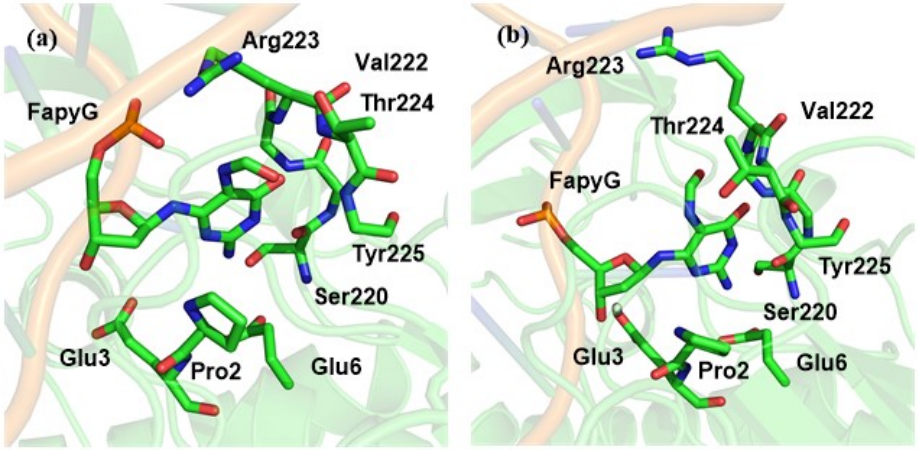

Figure A4.11. MD representative structure of *anti*-FapyG bound in the FPG active site with (a) anionic or (b) neutral Glu3. For clarity, only hydrogen atom of neutral Glu3 (Oε–H) is shown.

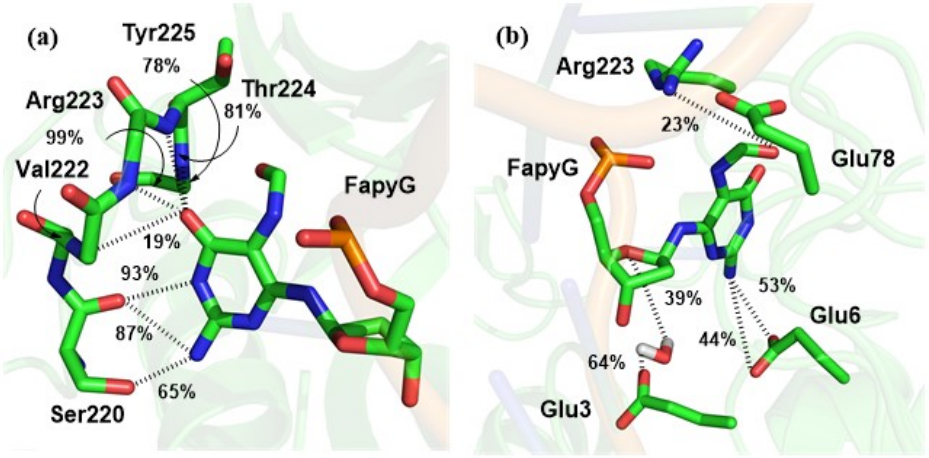

Figure A4.12. Major interactions between *anti*-FapyG and the FPG active site with anionic Glu3. For clarity, only hydrogen atoms of water are shown.

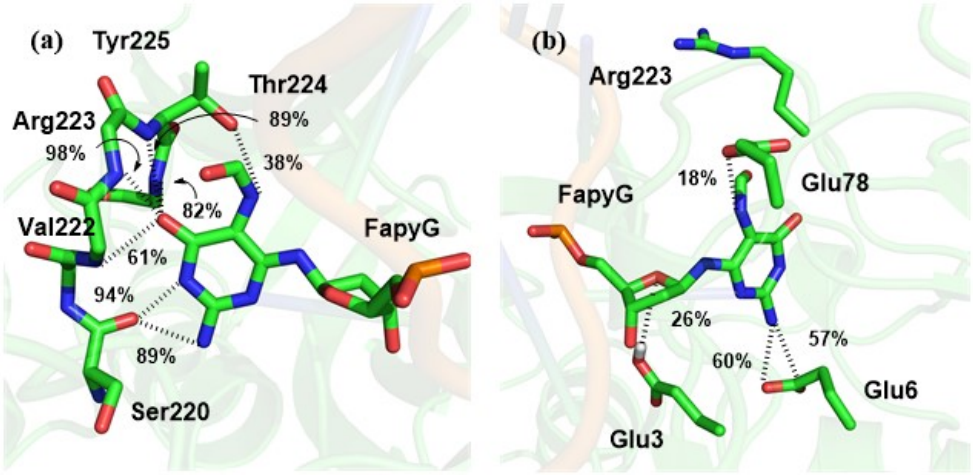

Figure A4.13. Major interactions between *anti*-FapyG and the FPG active site with neutral Glu3.

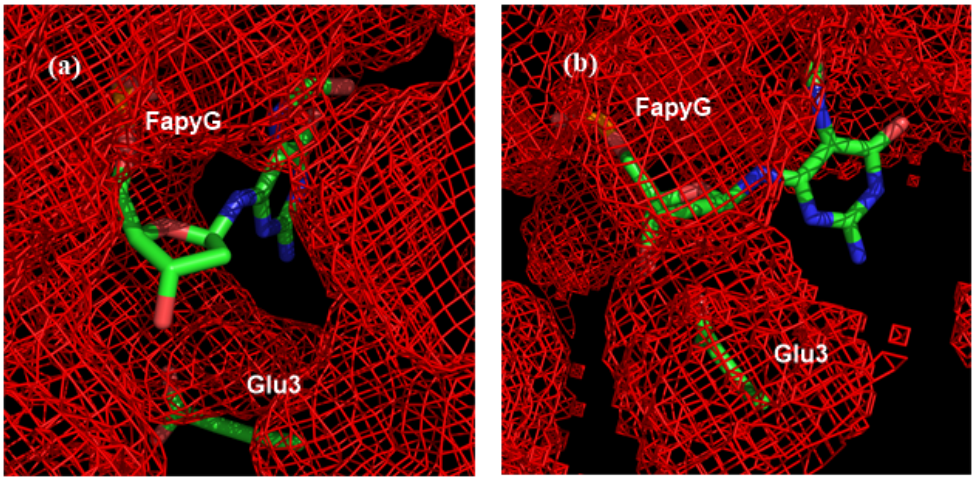

Figure A4.14. Water density distribution around *anti*-FapyG bound in the FPG active site with (a) anionic or (b) neutral Glu3.

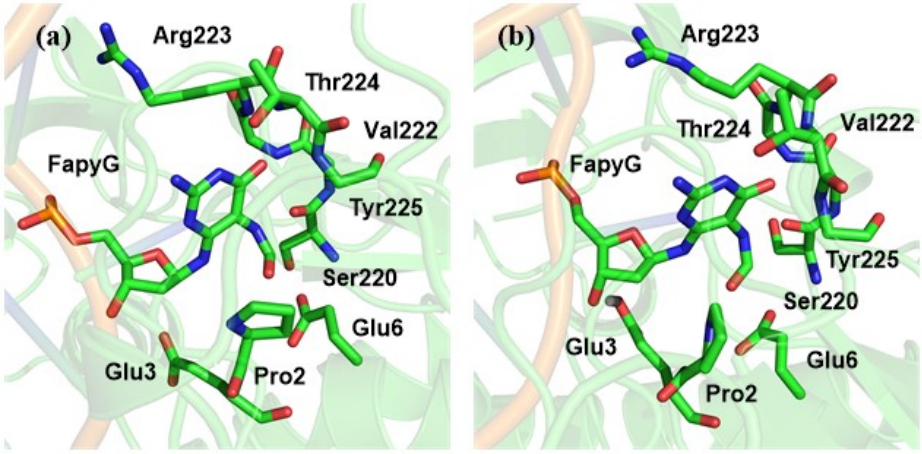

Figure A4.15. MD representative structure of *syn*-FapyG bound in the FPG active site with (a) anionic or (b) neutral Glu3. For clarity, only hydrogen atom of neutral Glu3 (Oε–H) is shown.

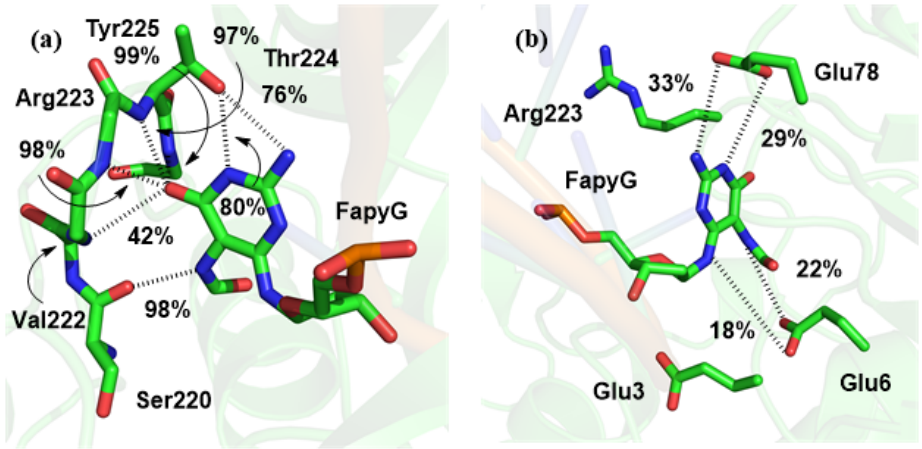

Figure A4.16. Major interactions between *syn*-FapyG and the FPG active site with anionic Glu3.

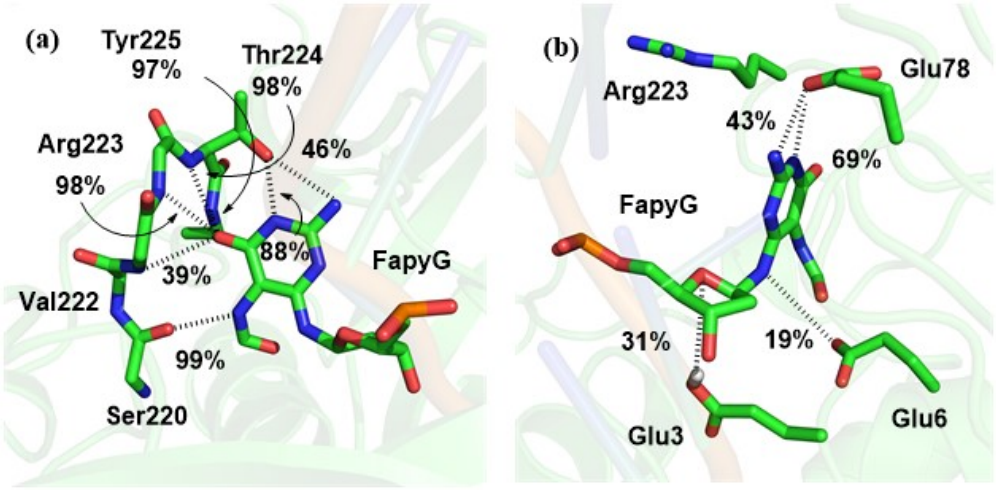

Figure A4.17. Major interactions between *syn*-FapyG and the FPG active site with neutral Glu3. For clarity, only hydrogen atom of neutral Glu3 (Oε–H) is shown.

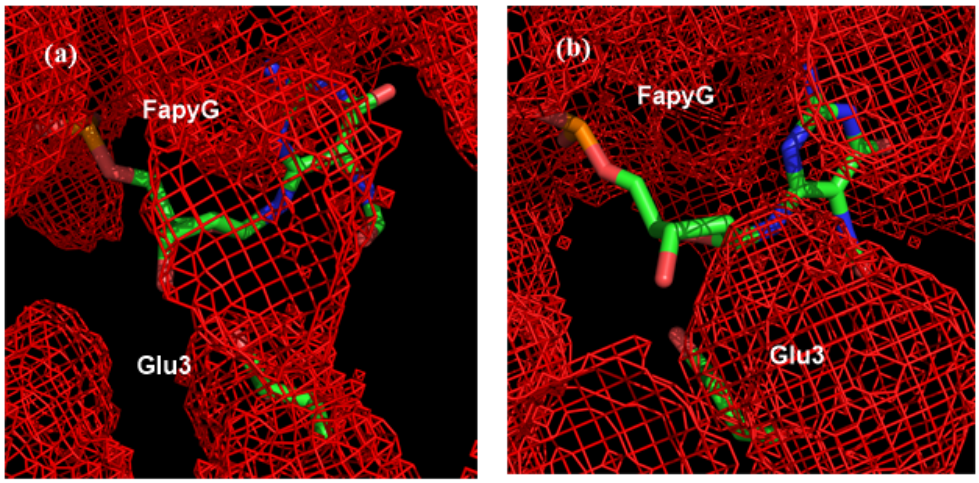

Figure A4.18. Water density distribution around *syn*-FapyG bound in the FPG active site with (a) anionic or (b) neutral Glu3.

Appendix A5: Supplementary Information for Chapter 5

Table of Contents for Appendix A5

Table A5.1. Relative Gibbs free energies (kJ/mol) for the stationary points characterized along the deglycosylation and β-elimination steps for the lowest-energy O-base pathway.[a, b]

Transition State	$\Delta G^{\ddagger}$ [c]	$\Delta G^{\ddagger}$ [d]
TS(a)	134.1	148.6
TS(b)	–4.0	–10.3
TS(b′)	27.6	23.9
TS(c)	21.0	21.5
TS(c′)	84.9	80.7

[a] Energies reported relative to a common (*anti*) reactant complex. [b] Values include unscaled correction to the Gibbs energy. [c] M06-2X/6-311+G(2df,2p)//PCM-B3LYP/6-31G(d,p). [d] M06-2X/6-311+G(2df,2p)//PCM-B3LYP/6-31G(d)

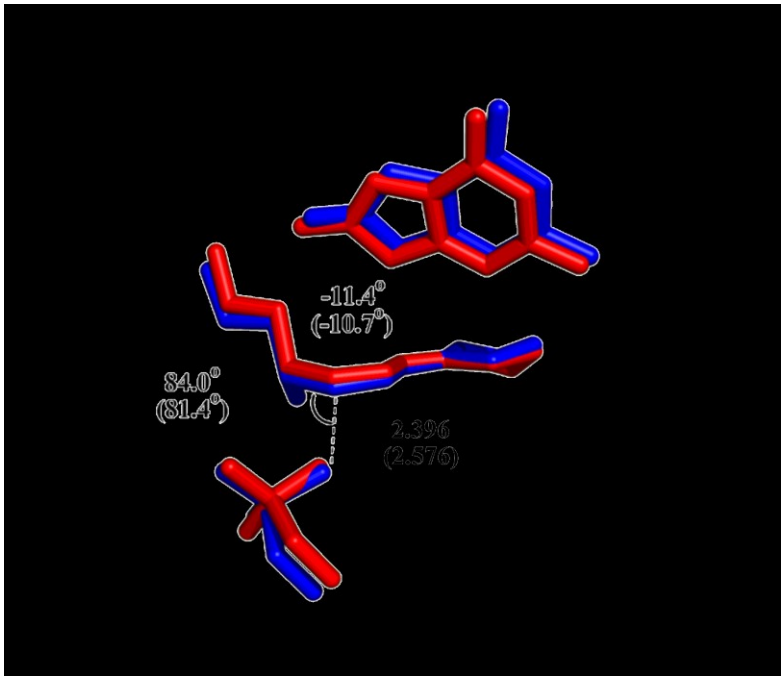

Figure A5.1. Overlay of the B3LYP/6-31G(d,p) transition state for 3′-PO_4 departure (N-base pathway) calculated using a model with the phosphate neutralized by a sodium (red) or potassium (blue) ion. Hydrogen atoms are not shown for clarity. C3′···O–PO_3 distance, angle (∠(O···C3′–C4′)) and dihedral angle (∠(C1′–C2′–C3′–C4′)) are shown for the potassium ion model with the corresponding value for the sodium ion model provided in parentheses.

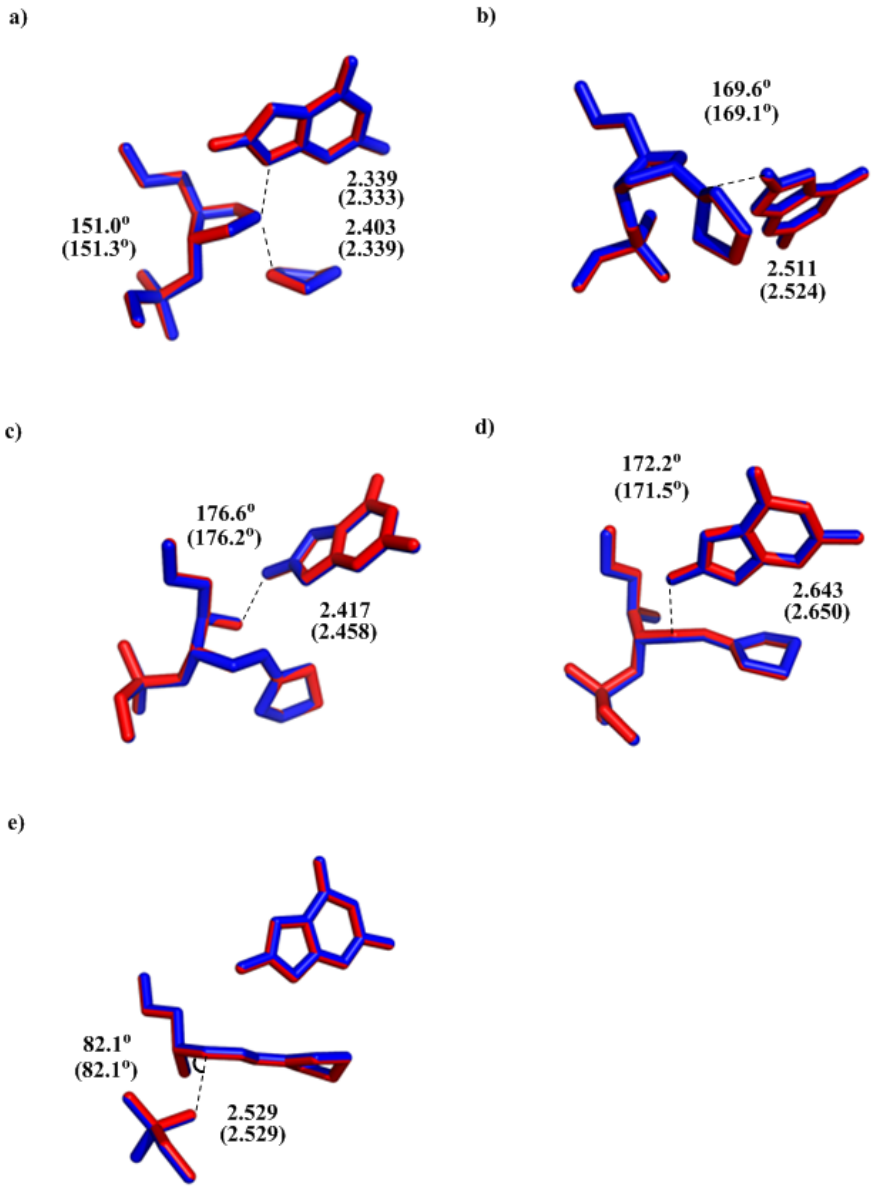

Figure A5.2. Overlays of B3LYP/6-31G(d) (red) and B3LYP/6-31G(d,p) (blue) optimized structures along the reaction pathway including the a) (*anti*) deglycosylation (TS(a)), b) Nα–H abstraction (TS(b)), c) base-assisted ring opening (TS(B′)), d) C2′–H abstraction (TS(c)) and e) 3′-PO_4 elimination reaction (TS(c′)) for the lowest-energy O-base pathway. Hydrogen atoms are not shown for clarity. Important B3LYP/6-31G(d,p) distances and angles are shown with the corresponding values B3LYP/6-31G(d) provided in parentheses.

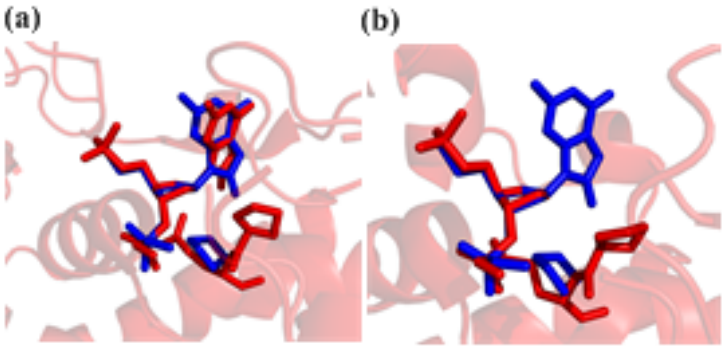

Figure A.5.3. Overlay of the *syn* reactant complex calculated in the present work (blue) with experimental crystal structures corresponding to (a) dOG bound in the active site (PDB ID: 1R2Y) or (b) an abasic-site bound in the active site (PDB ID: 3TWM). Hydrogen atoms and the DNA backbone are not shown for clarity.

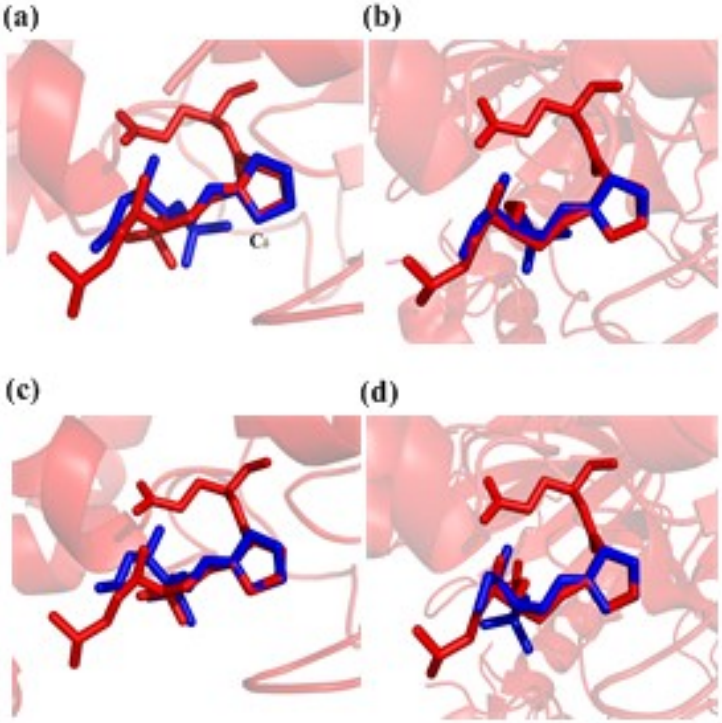

Figure A5.4. Overlay of the ring-opened Schiff base (IC(b)) calculated in the present work (blue) for the O-base (a and b) and N-base (c and d) pathway with crystal structures corresponding to a Schiff base intermediate (a and c; PDB ID: 1L1Z) or a borohydride-trapped abasic-site complex (b and d; PDB ID: 1K82). Hydrogen atoms and DNA backbone are not shown for clarity.

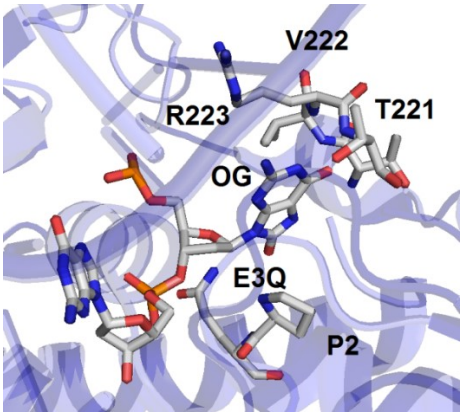

Figure A5.5. Important FPG active site residues identified in the crystal structure of the E3Q mutant (PDB ID: 1R2Y)

Appendix A6: Supplementary Information for Chapter 6

Table of Contents for Appendix A6

Table A6.1. Relative energies (ΔE) and Gibbs free energies (ΔG) for stationary points characterized along the proton abstraction and phosphate elimination reactions in the β-lyase step for the O-base pathway (kJ/mol).[a]

		syn			*anti*		
Reaction Step	**Stationary Point**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**
C2′–H Abstraction	**RC1**	0.0	0.0	0.0	0.0	0.0	0.0
	TS1	32.5	31.9	22.0	60.1	58.9	62.4
	IC1	25.3	24.9	–0.6	41.0	40.2	31.1
3′-PO$_4$ Elimination	**TS2**	48.4	47.4	19.3	–	–	–
	IC2	–7.6	–7.4	–18.7	–	–	–

[a] Energies reported relative to the corresponding (O-base or N-base) reactant complex. [b] Unscaled relative energies obtained with IEF-PCM-B3LYP/6-31G(d). [c] IEF-PCM-B3LYP-D3/6-311+G(2df,2p) relative energies including scaled (0.9806) zero-point vibrational energy correction. [d] SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) relative free energies including unscaled thermal corrections.

Table A6.2. Relative energies (ΔE) and Gibbs free energies (ΔG) for stationary points characterized along the proton abstraction reaction in the β-lyase step for the N-base pathway (kJ/mol).[a]

		syn			*anti*		
Reaction Step	**Stationary Point**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**
C2′–H Abstraction	**RC1**	0.0	0.0	0.0	0.0	0.0	0.0
	TS1	54.5	53.5	51.5	89.7	88.1	83.4
	IC1	–8.0	–7.9	–2.6	17.4	17.1	7.4

[a] Energies reported relative to the corresponding (O-base or N-base) reactant complex. [b] Unscaled relative energies obtained with IEF-PCM-B3LYP/6-31G(d). [c] IEF-PCM-B3LYP-D3/6-311+G(2df,2p) relative energies including scaled (0.9806) zero-point vibrational energy correction. [d] SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) relative free energies including unscaled thermal correction.

Table A6.3. Relative energies (ΔE) and Gibbs free energies (ΔG) for stationary points characterized along the proton abstraction reaction in the δ-lyase step for the O-base and N-base pathways (kJ/mol).[a]

		O-base			*N-base*		
Reaction Step	**Stationary Point**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**
C4′–H Abstraction	**RC2**	0.0	0.0	0.0	0.0	0.0	0.0
	TS3	38.4	37.7	34.7	44.4	43.6	48.0
	IC3	–34.9	–34.2	–53.8	–104.6	–102.6	–102.8

[a] Energies reported relative to the corresponding (O-base or N-base) reactant complex. [b] Unscaled relative energies obtained with IEF-PCM-B3LYP/6-31G(d). [c] IEF-PCM-B3LYP-D3/6-311+G(2df,2p) relative energies including scaled (0.9806) zero-point vibrational energy correction. [d] SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) relative free energies including unscaled thermal corrections.

Table A6.4. Relative energies (ΔE) and Gibbs free energies (ΔG) for stationary points characterized along the phosphate elimination reaction and the enol-keto rearrangement in the δ-lyase step (kJ/mol).[a]

		Direct			*Assisted (N-base)*		
Reaction Step	**Stationary Point**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**	**ΔE[b]**	**ΔE[c]**	**ΔG[d]**
5′-PO$_4$ Elimination	RC3/RC3′	0.0	0.0	0.0	0.0	0.0	0.0
	TS4′				26.7	26.2	21.9
	TS4/IC4′	160.3	157.2	188.6	7.6	7.4	18.8
	TS4″				164.3	161.1	185.3
Enol-Keto Rearrangement	IC4/IC4″	81.5	79.9	78.9	74.0	72.5	92.0
	TS5	173.2	169.8	200.6		–	
	P	151.9	148.9	176.8		–	

[a] Energies reported relative to the corresponding (O-base and N-base) reactant complex. [b] Unscaled relative energies obtained with IEF-PCM-B3LYP/6-31G(d). [c] IEF-PCM-B3LYP-D3/6-311+G(2df,2p) relative energies including scaled (0.9806) zero-point vibrational energy correction. [d] SMD-M06-2X/6-311+G(2df,2p)//IEF-PCM-B3LYP/6-31G(d) relative free energies including unscaled thermal corrections.

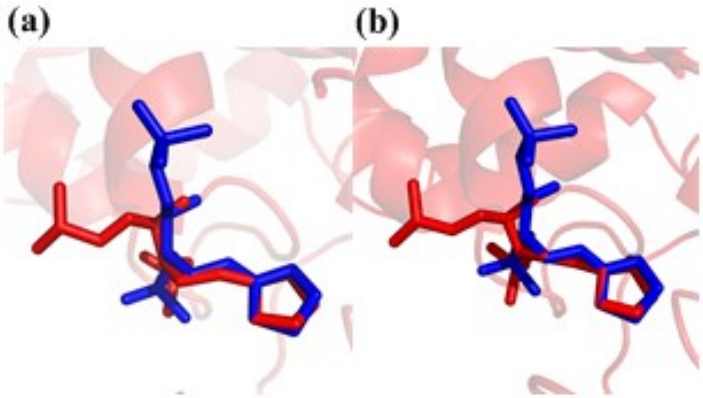

Figure A6.1. Overlay of the initial 1-(4-hydroxy-3,5-diyl dimethyl bis(phosphate)pentylidene) pyrrolidinium model (blue) and the crystal structure (red) of (a) the borohydride-trapped abasic site (PDB ID: 1K82) or (b) the Schiff base intermediate (PDB ID: 1L1Z).

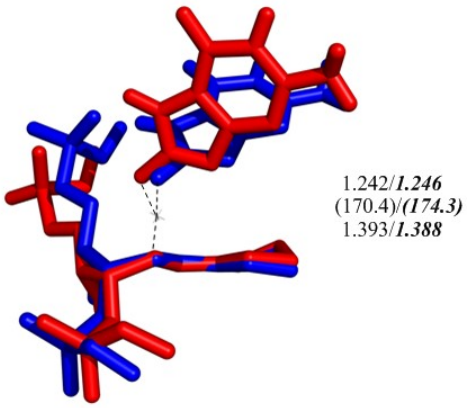

Figure A6.2. Overlay of the optimized transition structures for the *syn* (red) and *anti* (blue) orientation of OG^- in the C2′–H abstraction step along the O-base pathway. Important distances (Å) and angles (degree, in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) level of theory are provided. Italicized bold values correspond to the *anti* OG complex.

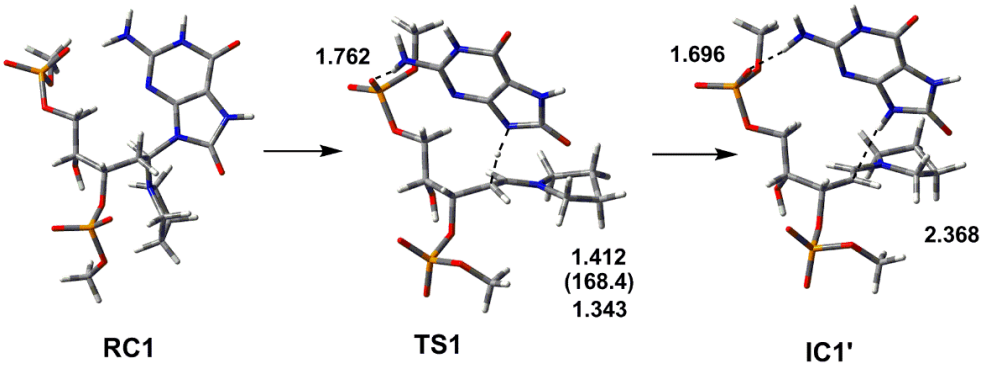

Figure A6.3. Optimized structures along the C2′–H abstraction step for the *syn* N-base pathway. Important distances (Å) and angles (degree, in parentheses) obtained with IEF-PCM-B3LYP/6-31G(d) level of theory are provided.

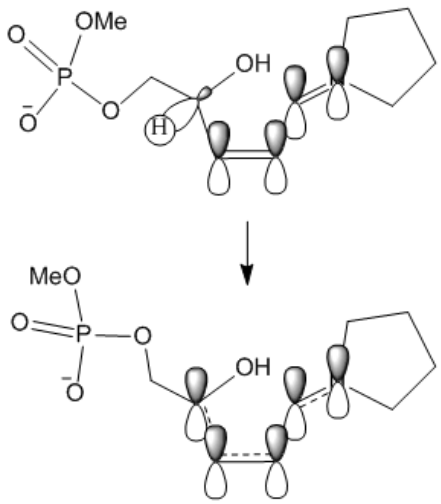

Figure A6.4. Molecular orbital interpretation of the δ-elimination reaction, which depicts bending of the phosphate group to allow the p-orbital of C4′ to contribute to the conjugated π-system and result in a more stable intermediate.

Appendix A7: Supplementary Information for Chapter 7

Table of Contents

Table A7.1. Root-mean-square deviation (rmsd in Å; standard deviation in parentheses) in the enzyme and DNA backbone over the entire or clustered MD simulation trajectory.[a]

System	rmsd	
	DNA	Enzyme
Lys249 Asp268	2.870 (1.450)	1.440 (0.147)
Lys249 $Asp268^{-1}$	2.400 (0.509)	1.265 (0.189)

[a] rmsd calculated according to the position of the DNA (P, O3′, O5′, C3′, C4′, and C5′) and enzyme (C_α, C, and N) backbone atoms over the entire simulation trajectory with respect to the first frame.

Table A7.2. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for cationic lysine-crosslink (group I, Figure 7.3).[a]

Reaction Step			ΔG	
			Asp	H_2O
β-elimination[b]	C2′–H abstraction	R	0.0	0.0
		TS1	–1.8	
		IC1	–81.4	105.2
	3′-PO_4 departure	R2	0.0	
		TS2	76.8	2.9
		IC2	31.3	
δ-elimination	simultaneous abstraction and elimination	R3	–	0.0
		TS3	–	107.7
		IC3	–	–39.3
Hydrolysis	step 1	R4	0.0	0.0
		TS4	61.5	75.6
		IC4	5.9	21.9
	step 2	R5	0.0	0.0
		TS5	55.9	61.6
		P	50.5	26.1

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections. [b] β-elimination with water molecule includes simultaneous abstraction and elimination.

Table A7.3. Gibbs energies (kJ/mol) for stationary points characterized along the β-elimination reaction facilitated by OG^{-1}.

		ΔG			
	Reaction Step	**neutral lys-crosslink[a]**	**cationic lys-crosslink[b]**	**cationic pro-crosslink[c]**	**cationic pro-crosslink[d]**
C2′–H abstraction	R1	0.0	0.0	0.0	0.0
	TS1	207.4	17.7	21.5	22.0
	IC1	5.9	–25.8	–3.8	–0.6
3′-PO_4 departure	R2	–	0.0	0.0	0.0
	TS2	–	107.3	80.7	19.9
	P	–	23.9	44.1	–18.1

[a] Current chapter, C2′–H abstraction by neutral OG
[b] Data for D-N_rC_s pathway from ref. 12.
[c] Data for O-base elimination pathway from ref. 26.
[d] Data for O-base elimination pathway from ref. 27.

Table A7.4. Gibbs energies (kJ/mol) for stationary points characterized along the reactions in Figure A7.2.[a]

	ΔG
Reaction Step	**H_2O**
R1	0.0
TS1	48.4
IC1	–22.7
R2	0.0
TS2	230.6
P	43.1

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections.

Table A7.5. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for cationic proline-crosslink (group II, Figure 7.5).[a]

Reaction Step			ΔG		
			H_2O (R1)	H_2O (R2-P1)	H_2O (R2-P2)[b]
β-elimination	simultaneous abstraction and elimination	R		0.00	
		TS1		105.7	
		IC1		–6.0	
δ-elimination	simultaneous abstraction and elimination	R2	–		0.0
		TS2	–		121.9
		IC2	–		–18.2
Hydrolysis	step 1	R3	0.0	0.0	0.0
		TS3	56.5	48.1	20.1
		IC3	19.8	10.7	–29.5
	step 2	R4	0.0	0.0	0.0
		TS4	44.8	61.0	55.8
		IC4	30.9	26.4	41.7
δ-elimination	simultaneous abstraction and elimination	R5	0.0		–
		TS5	175.7		–
		IC5	–6.1		–
Enol-keto	tautomerization	R6		0.0	0.0
		TS6		136.2	148.2
		P		–24.2	–7.8

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections. [b] In R2-P2 pathway the hydrolysis takes place after enol-keto rearrangement.

Table A7.6. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for neutral lysine-crosslink (group III, R1-P1 and R1-P2, Figure 7.5).[a]

Reaction Step			ΔG		
			Asp	OG^{-1}	H_2O
β-elimination	simultaneous abstraction and elimination	R	0.0	0.0	0.0
		TS1	233.8	115.8	258.9 (241.4)[b]
		IC1	29.8	28.8	47.0 (19.0)[b]
δ-elimination	simultaneous abstraction and elimination	R3		–	0.0
		TS3		–	235.2
		IC3		–	–24.5
Hydrolysis	step 1	R4		0.0	0.0
		TS4		122.2	121.3
		IC4		9.2	9.4
	step 2	R5		0.0	0.0
		TS5		55.9	61.6
		P		50.5	26.1

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections. [b] Values in parentheses correspond to the reactions catalyzed by 2 water molecules.

Table A7.7. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for neutral lysine-crosslink (group III, R2, Figure 7.5).[a]

Reaction Step			ΔG	
			Asp	H_2O
β-elimination	C2′–H abstraction/ crosslink protonation	R	0.0	0.0
		TS1	75.8	143.6
		IC1	7.3	–1.7
	3′-PO_4 departure	R3	0.0	0.00
		TS3	76.8	122.4
		IC3	31.7	107.5
Hydrolysis	step 1	R4	0.0	
		TS4	61.5	
		IC4	5.9	
	step 2	R5	0.0	
		TS5	55.9	
		P	50.5	

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections.

Table A7.8. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for neutral lysine-crosslink (group III, R3-P1, Figure 7.6).[a]

Reaction Step			ΔG	
			H_2O	Asp^{-1}
Hydrolysis	step 1	R	0.0	
		TS1	117.5	
		IC1	–6.8	
β-elimination	simultaneous abstraction and elimination	R2	0.0	0.0
		TS2	261.0	187.1
		IC2	54.5	13.1
Hydrolysis	step 2	R3	0.0	
		TS3	55.9	
		P	50.5	

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections.

Table A7.9. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for neutral lysine-crosslink (group III, R3-P2, Figure 7.6).[a]

Reaction Step			ΔG
			H_2O
Hydrolysis[b]	step 1	R	0.0
		TS1	117.5
		IC1	–6.8
	C2′–H abstraction/ crosslink protonation	R2	0.0
		TS2	206.2
		IC2	6.1
β-elimination	3′-PO_4 departure	R3	0.0
		TS3	84.1
		P	–42.8

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections. [b] C2′–H abstraction and crosslink protonation completes the crosslink hydrolysis.

Table A7.10. Gibbs energies (kJ/mol) for stationary points characterized along the pathways for neutral lysine-crosslink (group III, R3-P3, Figure 7.6).[a]

Reaction Step			ΔG		
			Asp^{-1}	**Lys**	**H_2O**
Hydrolysis	step 1	R		0.0	
		TS1		117.5	
		IC1		–6.8	
	step 2	R2		0.0	
		TS2		67.6	
		IC2		15.2	
β-elimination[b]	C2′–H abstraction	R3	0.0	0.0	0.0
		TS3	13.1	93.3 (61.2)[c]	
		IC3	–19.6	83.4 (45.2)[c]	233.5 (191.4)[d]
	3′-PO_4 departure	R4	0.0	0.0	
		TS4	32.1	91.0 (49.0)[c]	54.3 (26.5)[d]
		IC4	–56.1	–93.7 (–79.8)[c]	
δ-elimination	simultaneous abstraction and elimination	R5		–	0.00
		TS5		–	175.7
		P		–	–6.1

[a] Energies were obtained with SMD-M06-2X/6-311++(2df,2p) and include unscaled thermal corrections. [b] β-elimination with water molecule includes simultaneous abstraction and elimination. [c] Values in parentheses correspond to the *endo* structure (*pro*-R proton abstraction). [d] Values in parentheses correspond to the reactions catalyzed by 2 water molecules.

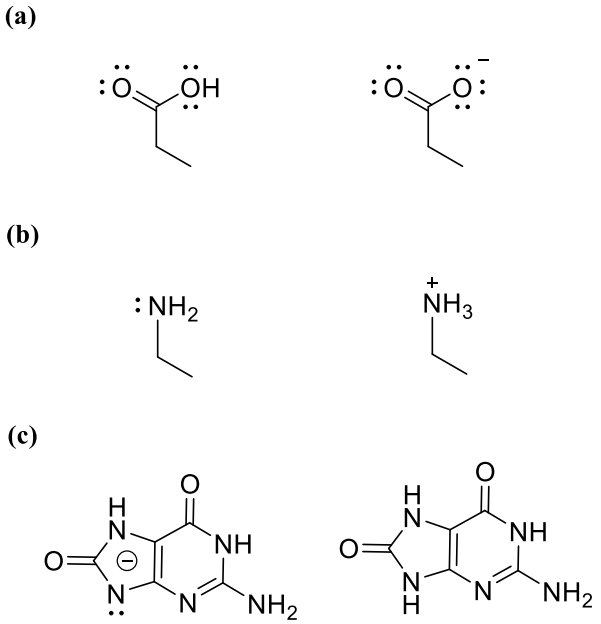

Figure A7.1. 2D representation of the modelled acids and bases in Chapter 7: (a) neutral and anionic Asp or Glu, (b) neutral and cationic Lys, and (c) neutral and anionic OG.

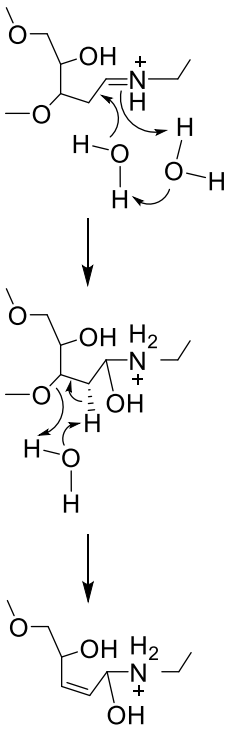

Figure A7.2. Elimination reaction mechanism from the cationic lysine-crosslink (group I). The Gibbs energies are provided in Table A7.4.

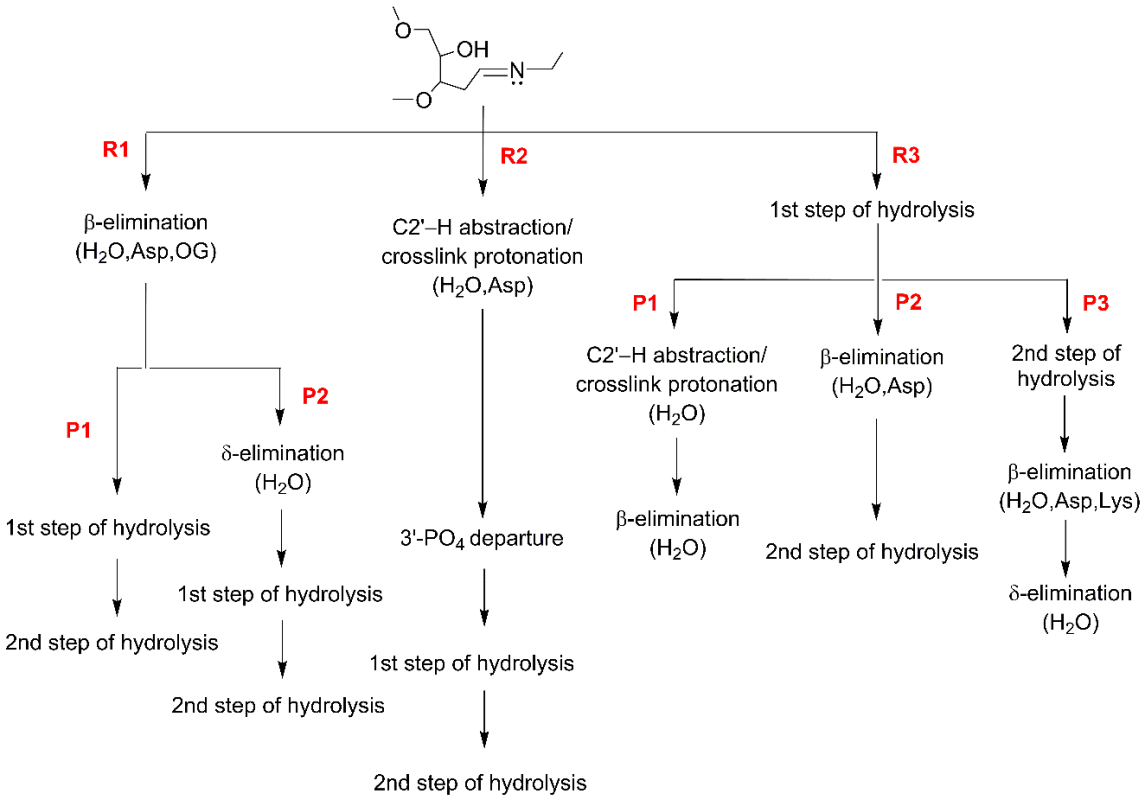

Figure A7.3. General reaction pathways considered for neutral lysine-crosslink (group III). Individual routes are labeled as R1, R2, and R3, while pathways within each route are labeled with P. Detailed reaction mechanisms are shown in Figures 7.5 – 7.6 and Gibbs energies are provided in Tables A7.6 – A7.10.

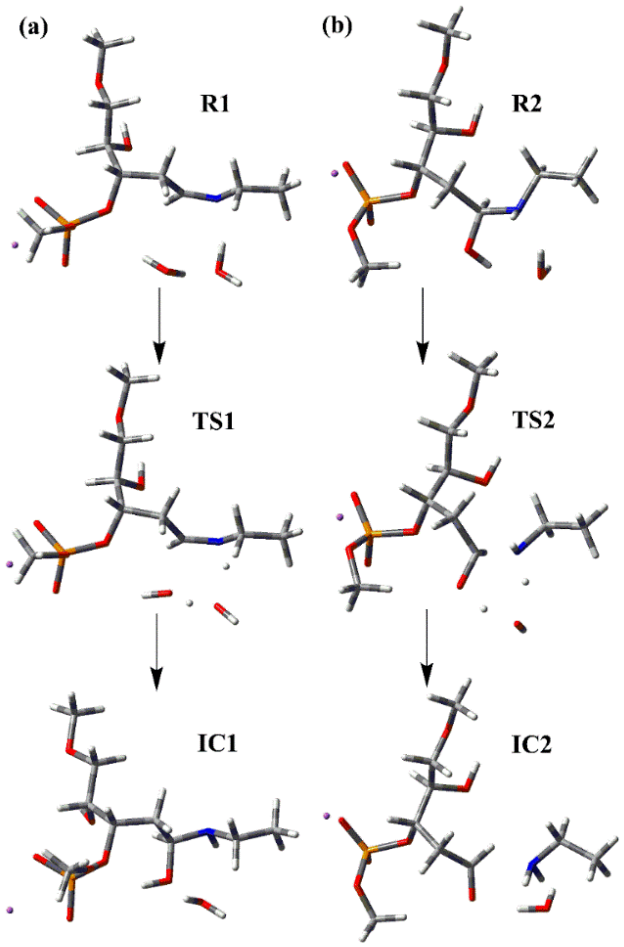

Figure A7.4. Structure of stationary points along the (a) first and (b) second step of the neutral lysine-crosslink hydrolysis.

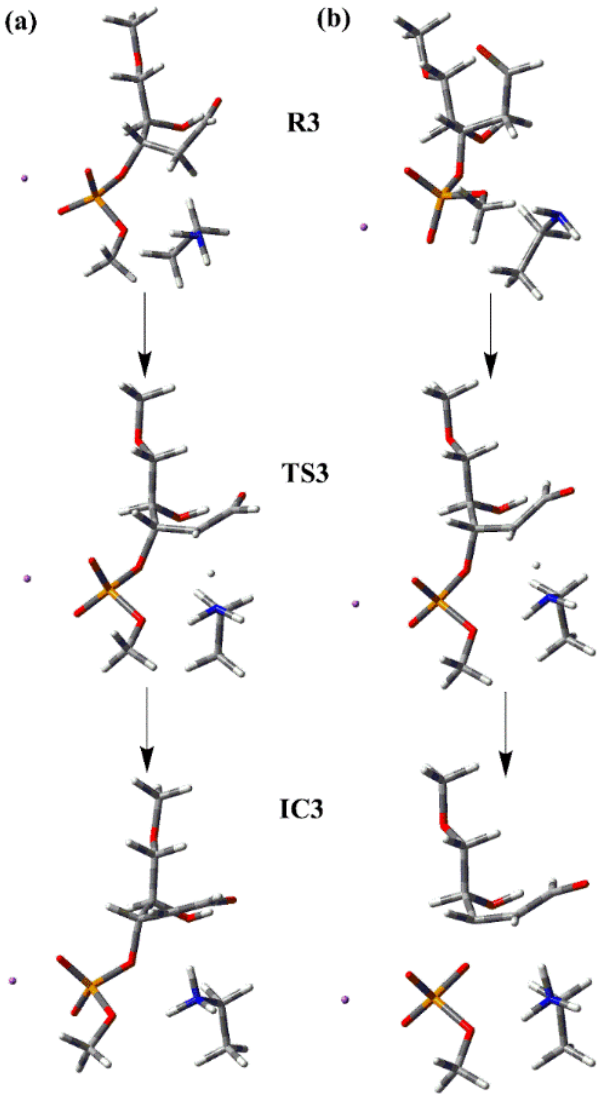

Figure A7.5. Structure of stationary points along C2′–H abstraction by lysine on the (a) endo- and (b) exo- ring-opened keto-abasic site.

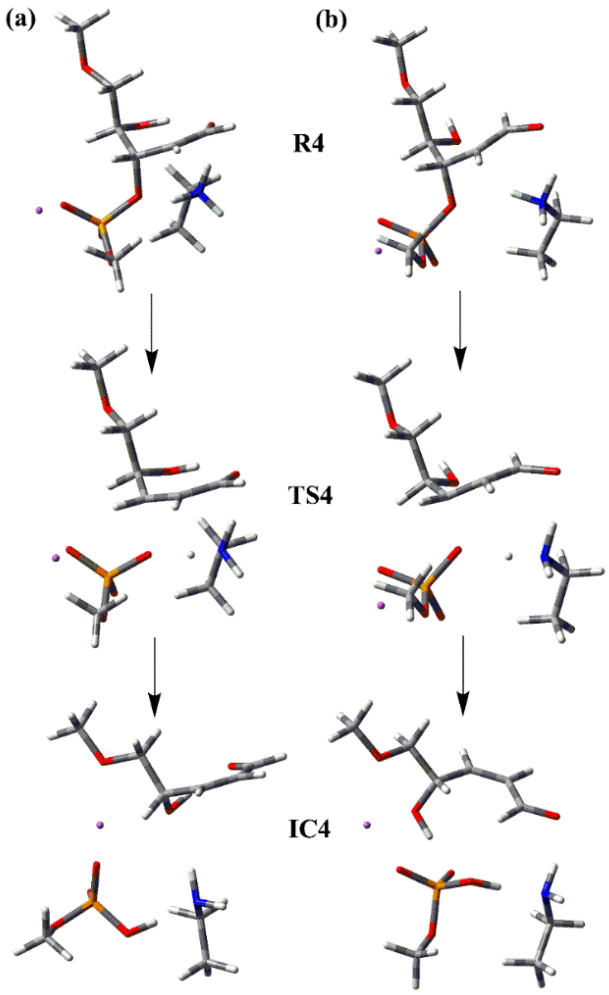

Figure A7.6. Structure of stationary points along the 3′-PO_4 protonation and departure on (a) *endo*- and (b) *exo*-ring-opened keto-abasic site.

Transition State Structures (electronic energy in a.u.)

Cationic Lysine Crosslink (Figure 7.3)

Figure 7.3 left C2′–H abstraction and elimination by water, cationic lysine crosslink (–1441.13091)

1	-1.222	942.14135	1.59960
7	-3.573	81-1.54569	-0.74983
6	-4.971	87-1.30380	-0.38955
6	-5.413	26-2.05141	0.87060
1	-5.591	33-1.59725	-1.24340
1	-5.094	62-0.22368	-0.26059
1	-6.471	52-1.85373	1.07267
8	-1.112	613.47988	-0.05574
6	-0.956	002.30164	-0.83063
6	-0.272	421.30621	0.10320
8	-1.035	901.22235	1.33115
6	-0.054	89-0.10080	-0.47710
6	-1.214	63-1.07752	-0.47478
6	-2.550	48-0.74973	-0.39847
1	-0.319	312.48323	-1.70898
1	-1.933	221.94083	-1.18302
1	0.723	33 1.69963	0.32944
1	-0.957	77-2.04084	-0.91629
6	-1.791	504.52312	-0.74261
1	-1.865	325.36431	-0.05040
1	-1.233	664.83606	-1.63575
1	-2.801	164.21061	-1.04523
1	-5.288	68-3.13329	0.75058
1	-4.831	29-1.73255	1.74205
1	-1.113	79-1.36344	1.30500
1	0.324	66 0.02372	-1.49764
8	1.023	92 -0.74700	0.29304
15	2.622	80 -0.51888	-0.13946
8	2.872	08 0.93414	-0.47489
8	3.432	65 -1.11274	0.98826
8	2.726	52 -1.36604	-1.52117
11	4.611	39 0.95275	1.11194
6	2.705	99 -2.80418	-1.47444
1	1.741	70 -3.16595	-1.10144
1	2.852	73 -3.14951	-2.49909
1	3.509	88 -3.17998	-0.83577
1	-2.855	420.21573	-0.00399
8	-0.618	93-1.13894	2.21346
1	0.324	63 -1.18994	1.87048
1	-3.351	82-2.46596	-1.11423
1	-0.783	28-0.13087	2.22515

Figure 7.3 middle C2′–H abstraction by Asp⁻, cationic lysine crosslink (–1632.67588)

8	-3.020	003.21421	-0.94574
6	-2.700	071.88313	-0.55609
6	-1.256	831.92794	-0.05611
8	-1.137	972.96815	0.91953
6	-0.828	500.61933	0.62213
8	-1.142	84-0.44143	-0.30156

6 0.67055 0.54034 0.94481
6 1.24699 1.51632 1.80111
6 2.81815 0.14450 -1.65447
8 1.85371 0.96844 -1.41784
15 -1.69575-1.893120.23219
8 -2.34082-2.57401-0.95905
8 -2.57739-1.723961.45541
8 -0.35374-2.683450.72197
6 0.60808 -3.13939-0.24978
1 -2.801351.19265 -1.40241
1 -3.372161.55209 0.25311
1 -0.592532.12755 -0.90948
1 -1.438120.48845 1.52350
1 0.96545 -0.484851.17521
1 1.28138 0.74129 -0.15633
1 0.77063 2.48836 1.90051
1 1.29508 -2.33191-0.52148
1 1.17044 -3.947270.22513
6 -4.370673.37103 -1.34050
1 -4.509034.42083 -1.61089
1 -4.608582.74154 -2.21076
1 -5.061753.11480 -0.52283
1 0.10241 -3.52242-1.14082
11 -4.28736-2.818510.30479
7 2.40065 1.38303 2.42218
1 -1.636153.71783 0.54749
6 3.24850 0.17645 2.40958
1 2.77695 -0.584423.04475
1 3.28013 -0.204001.38358
6 4.64871 0.50745 2.91365
1 5.25906 -0.400132.91745
1 5.13873 1.24144 2.26512
1 4.62784 0.90160 3.93680
6 3.44554 0.30047 -3.04930
1 3.73360 1.35348 -3.16712
1 2.64867 0.13340 -3.78649
6 4.63232 -0.62396-3.31460
1 5.02953 -0.47206-4.32530
1 5.44078 -0.44245-2.59860
1 4.33994 -1.67390-3.21444
8 3.24807 -0.72189-0.86896
1 2.74298 2.17735 2.95097

Figure 7.3 middle Protonation of 3′-PO_4 by Asp–H, cationic lysine crosslink (–1632.68266)

8 -4.364712.30767 -0.43157
6 -3.116152.22277 0.24024
6 -2.276911.21447 -0.56220
8 -2.204691.64103 -1.91939
6 -0.901001.12291 0.02845
8 -0.45198-0.89957-0.40303
6 0.14789 1.87667 -0.48649
6 1.35485 1.95993 0.19725
15 -1.25770-2.007520.38067
8 -2.77793-1.854570.23675
8 -0.76555-3.419700.04640

8 -0.97345-1.697351.99146
6 0.32027 -2.011502.50916
1 -3.251081.88363 1.27828
1 -2.623243.20815 0.25226
1 -2.758940.23190 -0.48128
1 -0.863190.80331 1.06354
1 1.01755 -1.20342-0.91814
1 0.03297 2.34708 -1.45802
1 1.48539 1.41134 1.12518
1 1.11278 -1.492231.95708
1 0.33777 -1.687563.55473
6 -5.215523.30898 0.09930
1 -6.133633.29772 -0.49321
1 -5.464003.10655 1.15143
1 -4.754324.30602 0.03317
1 0.50653 -3.090922.46456
11 -2.90230-3.99955-0.52538
7 2.41675 2.64888 -0.19990
1 -3.112271.91538 -2.13804
6 3.69089 2.71264 0.52016
1 3.73261 1.84132 1.17766
1 4.49215 2.59656 -0.21751
6 3.86407 4.01771 1.29846
1 4.84037 4.02770 1.79480
1 3.81269 4.88632 0.63270
1 3.08704 4.12335 2.06272
6 5.52071 -1.21483-0.63975
1 5.61159 -0.13280-0.49808
6 4.24572 -1.56828-1.40159
1 4.23462 -1.10741-2.39758
1 4.18194 -2.64933-1.58137
6 2.96744 -1.15588-0.68708
8 1.88965 -1.43714-1.40315
8 2.95591 -0.628630.42110
1 6.40327 -1.56177-1.18759
1 5.52698 -1.678340.35144
1 2.34662 3.17583 -1.06443

Figure 7.3 right C2′–H abstraction by OG⁻, cationic lysine crosslink (–1982.07640)

8 0.46602 -2.85247-1.73884
6 0.44861 -1.77153-0.81911
6 1.77781 -1.04386-0.92369
8 1.82306 -0.34171-2.17066
6 2.08090 -0.090490.25168
8 3.34311 0.56833 -0.10028
6 1.06013 0.98583 0.58093
6 0.50848 1.76110 -0.46625
7 -2.434890.44896 0.30675
6 -2.01920-0.271111.36859
7 -3.04449-1.002991.91980
6 -4.18445-0.736121.16832
6 -5.51637-1.197961.26580
8 -6.03663-1.987832.06177
7 -6.29780-0.602310.23188
6 -5.833020.29001 -0.70410

7 -6.743590.70730 -1.65661
7 -4.590620.70442 -0.76526
6 -3.764560.16360 0.18294
8 -0.82487-0.317691.85972
15 4.78005 -0.054390.41078
8 5.84810 0.50905 -0.50606
8 4.72248 -1.565200.51792
8 4.91438 0.52517 1.93207
6 5.09277 1.93514 2.12351
1 0.30981 -2.142000.20882
1 -0.38723-1.09121-1.04042
1 2.57260 -1.80136-0.90075
1 2.24024 -0.698631.14712
1 0.00233 0.33353 1.14165
1 1.36409 1.57413 1.45438
1 0.41997 1.32698 -1.45883
1 -2.96485-1.618632.71549
1 -7.25737-0.926840.18929
1 -7.699690.83355 -1.34457
1 -6.401641.49804 -2.18965
1 5.99387 2.28578 1.61076
1 4.22399 2.49056 1.75270
6 -0.75117-3.57272-1.73927
1 -0.64807-4.38198-2.46733
1 -0.96778-4.00644-0.74993
1 -1.60108-2.93558-2.02916
1 5.19428 2.09670 3.19885
11 6.45499 -1.70560-1.02662
7 0.03178 2.98143 -0.33863
1 0.09474 3.42336 0.57471
1 2.63260 0.19828 -2.12334
6 -0.701083.72176 -1.37271
1 -0.209144.69266 -1.49974
1 -0.583423.16663 -2.30714
6 -2.178713.90010 -1.02560
1 -2.670094.48010 -1.81445
1 -2.674512.93046 -0.93068
1 -2.298004.44788 -0.08361

Figure 7.3 R1 First step of cationic lysine crosslink hydrolysis after β-elimination (–863.84861)

1 -1.252220.24501 -1.25833
7 2.78431 -1.20725-0.59806
6 3.53430 -1.868080.47824
6 4.91776 -2.27539-0.02009
1 2.98658 -2.749130.84157
1 3.63613 -1.171911.31559
1 5.48033 -2.755450.78622
8 0.11593 3.90221 0.58118
6 0.02599 3.11428 -0.58906
6 -0.520201.72777 -0.21327
8 -0.658011.01550 -1.42865
6 0.34448 1.07603 0.84361
6 1.18725 0.03433 0.76831
6 1.51812 -0.73107-0.45496
1 1.01237 2.99859 -1.06715

1 -0.655743.56982 -1.32403
1 -1.507211.89107 0.25077
1 0.25580 1.56206 1.81337
1 1.69377 -0.272231.67921
1 1.11269 -0.31343-1.36849
6 0.59963 5.20760 0.32305
1 0.62718 5.73616 1.27920
1 1.61348 5.18594 -0.10543
1 -0.059395.75249 -0.37027
1 4.84610 -2.99006-0.84876
8 0.42806 -2.14287-0.41696
1 0.00488 -2.258130.48959
1 -0.46737-1.83216-0.85179
8 -1.59103-2.139381.24591
8 -1.92169-1.33417-0.83183
6 -2.31595-1.611590.36888
6 -3.75618-1.241410.72068
1 -4.01762-1.748861.65425
1 -3.76848-0.162280.93300
6 -4.76701-1.55390-0.38773
1 -5.77411-1.23573-0.09603
1 -4.49585-1.04101-1.31459
1 -4.80254-2.62906-0.59813
1 5.48297 -1.40302-0.36446
1 2.96276 -1.60964-1.51199

Figure 7.3 R2 δ-elimination by water, cationic lysine crosslink (–1325.38534)

8 1.71253 -1.087550.15699
6 0.73510 -1.30937-0.92141
6 -0.65431-0.92109-0.47728
8 -0.886090.43208 -0.31687
6 -1.73111-1.77713-0.66085
6 -3.12656-1.52440-0.60413
6 -3.70568-0.27827-0.47331
1 0.78810 -2.37170-1.16533
1 1.05450 -0.72910-1.79240
1 -1.45040-2.81072-0.85491
1 -3.77890-2.39013-0.69572
1 -3.077660.60056 -0.37448
7 -5.02503-0.03898-0.46694
1 -0.037420.87998 -0.10608
6 -5.635941.27393 -0.27049
6 -6.298431.42218 1.10083
1 -6.371621.43571 -1.06700
1 -4.848762.02235 -0.40187
1 -6.761012.41123 1.18834
1 -5.563441.30994 1.90459
15 2.79440 0.16644 0.08204
8 1.68902 1.38762 0.18042
8 3.50691 0.21313 -1.24888
8 3.65498 0.06523 1.31430
6 2.11570 2.72653 -0.15203
1 2.51205 2.75599 -1.17014
1 1.23347 3.36282 -0.07343
1 2.87396 3.06673 0.55991

1 -0.55497-1.542391.02218
8 0.01252 -1.664021.99941
1 -0.04888-2.588862.31559
1 0.94085 -1.505001.63739
11 5.57963 0.14386 -0.13814
1 -5.65584-0.82965-0.54323
1 -7.082230.66989 1.24319

Figure 7.3 R2 First step of cationic lysine crosslink hydrolysis after δ-elimination (–748.11175)

1 -1.225670.80410 -1.10921
7 2.92102 -0.08441-0.59771
6 3.77023 -0.701290.42968
6 5.20531 -0.81823-0.07496
1 3.38074 -1.692450.70222
1 3.74583 -0.076501.32706
1 5.83825 -1.264340.69796
6 -1.277483.50634 0.24291
6 -0.694102.32095 -0.01414
8 -0.865591.72392 -1.23767
6 0.18544 1.69830 0.98575
6 1.14231 0.75730 0.84142
6 1.59442 0.16295 -0.43327
1 -1.103464.02745 1.17791
1 -1.910223.98164 -0.49881
1 0.04973 2.09251 1.99065
1 1.67130 0.44193 1.73539
1 1.13224 0.58202 -1.31969
1 5.26264 -1.45811-0.96359
8 0.75133 -1.41374-0.53290
1 0.38966 -1.678650.36176
1 -0.20241-1.18797-0.91721
8 -1.27281-1.799481.12354
8 -1.65159-0.83067-0.87268
6 -2.03194-1.245160.29933
6 -3.49335-1.000190.66538
1 -3.72464-1.599231.55136
1 -3.574630.05482 0.96499
6 -4.47769-1.28267-0.47464
1 -5.50447-1.06155-0.16283
1 -4.24197-0.67124-1.34995
1 -4.43851-2.33480-0.77939
1 5.61258 0.16559 -0.33029
1 3.16631 -0.36434-1.54143

Cationic Proline Crosslink (Figure 7.4)

Figure 7.4 β-elimination by water, proline crosslink (–1518.54891)

1 -0.127492.71369 1.58028
7 -3.48034-0.51585-0.09508
6 -4.743770.18824 0.20472
1 -5.025070.82978 -0.64073
1 -4.632540.81769 1.09250
8 0.10705 3.84600 -0.22079
6 -0.095422.60108 -0.87102

6	0.44484	1.55754	0.10335
8	-0.18190	1.75790	1.39361
6	0.28316	0.09330	-0.33712
6	-1.05659	-0.58895	-0.14110
6	-2.27198	0.04998	0.01098
1	0.46360	2.54957	-1.81678
1	-1.16198	2.45253	-1.09347
1	1.52144	1.72823	0.20126
1	-1.05614	-1.61894	-0.49107
6	-0.40689	4.95464	-0.94702
1	-0.20664	5.84646	-0.34963
1	0.08904	5.05276	-1.92247
1	-1.49042	4.85921	-1.10692
1	-0.84093	-0.73951	1.66352
1	0.58053	0.03856	-1.39037
8	1.26702	-0.71033	0.41500
15	2.81301	-0.93091	-0.18196
8	3.34275	0.36506	-0.75149
8	3.59405	-1.56163	0.94622
8	2.55391	-1.94594	-1.42415
11	5.24129	0.14398	0.65115
6	2.20942	-3.31668	-1.15495
1	1.24997	-3.37945	-0.63013
1	2.12754	-3.81182	-2.12394
1	2.98621	-3.79819	-0.55483
1	-2.29835	1.09977	0.28712
8	-0.20953	-0.54749	2.48732
1	0.65508	-0.84094	2.06288
1	-0.14977	0.46921	2.39394
6	-3.69973	-1.92877	-0.43506
1	-3.17001	-2.18686	-1.35900
1	-3.30563	-2.56801	0.36789
6	-5.22621	-2.04324	-0.56075
1	-5.58493	-3.04404	-0.30779
1	-5.53444	-1.82503	-1.58933
6	-5.74862	-0.95490	0.39355
1	-6.77251	-0.64412	0.17083
1	-5.71964	-1.31185	1.42932

Figure 7.4 R1 First step of hydrolysis after β-elimination, proline crosslink (–941.26554)

1	1.55392	-0.16030	1.23085
6	-4.86157	-0.29969	0.85052
6	-3.41008	-0.45358	1.30836
7	-2.64069	-0.08148	0.10859
6	-3.44364	-0.27305	-1.12009
6	-4.80407	-0.78448	-0.60840
1	-3.14054	0.19961	2.14472
1	-2.97067	-0.99163	-1.79860
1	-3.54336	0.68055	-1.65208
1	-5.63516	-0.41940	-1.21797
1	-5.16014	0.75423	0.89458
1	-5.55668	-0.87679	1.46633
8	1.72210	3.78557	-0.51190
6	1.41072	2.98780	0.61296
6	1.50032	1.50726	0.21259

8	1.26548	0.76759	1.39703
6	0.58275	1.20419	-0.95197
6	-0.56944	0.51905	-1.00583
6	-1.28475	-0.10477	0.13551
1	0.39519	3.20296	0.98362
1	2.11803	3.16955	1.43717
1	2.52902	1.33856	-0.14782
1	0.93523	1.63993	-1.88512
1	-1.05479	0.42502	-1.97432
1	-0.87070	0.13313	1.10881
6	1.67472	5.17120	-0.22570
1	1.93284	5.69996	-1.14669
1	0.67004	5.48324	0.09884
1	2.39474	5.44616	0.56045
1	-3.19335	-1.49172	1.59710
1	-4.82415	-1.87964	-0.63694
8	-0.76953	-1.78270	0.14929
1	-0.48937	-2.05765	-0.75778
1	0.25876	-1.82691	0.57476
8	1.33172	-2.60191	-1.36886
8	1.59949	-1.87825	0.74315
6	2.05740	-2.31141	-0.40190
6	3.57447	-2.43127	-0.50807
1	3.80332	-2.98448	-1.42381
1	3.97184	-1.41462	-0.64135
6	4.23366	-3.07898	0.71520
1	5.32175	-3.11757	0.59328
1	4.00677	-2.51319	1.62298
1	3.87536	-4.10459	0.86016

Figure 7.4 R1 Second step of hydrolysis after β-elimination, proline crosslink (–749.27386)

1	-1.02467	1.99972	0.00447
6	3.50926	-2.00238	0.43337
6	3.00362	-0.76962	1.20058
7	2.30466	0.08633	0.17547
6	2.41442	-0.58923	-1.16395
6	3.63299	-1.50075	-1.01443
1	2.31785	-1.01666	2.01414
1	2.51221	0.17581	-1.93493
1	1.50095	-1.16208	-1.34928
1	3.63207	-2.30866	-1.75134
1	2.77218	-2.81155	0.48950
1	4.45026	-2.37898	0.84326
8	-4.19583	-0.88644	-0.44363
6	-3.73948	0.30625	0.16814
6	-2.29061	0.56653	-0.25917
8	-1.96558	1.84147	0.25774
6	-1.39297	-0.56697	0.20444
6	-0.09583	-0.56773	0.53570
6	0.88861	0.57589	0.53762
1	-3.79288	0.22882	1.26680
1	-4.35484	1.16580	-0.13867
1	-2.28011	0.57027	-1.36455
1	-1.91204	-1.52305	0.26238
1	0.31784	-1.51080	0.89350

1 1.02557 0.94270 1.57242
6 -5.51884-1.21931-0.06912
1 -5.77879-2.14646-0.58675
1 -5.60510-1.377731.01723
1 -6.23220-0.43338-0.36181
1 3.82667 -0.182411.61847
1 4.55674 -0.92492-1.14492
8 0.59334 1.59538 -0.34513
1 1.53531 2.23000 -0.33935
8 2.83667 2.49216 -0.25679
1 2.98107 2.99654 0.56013
1 2.80345 1.12613 0.06438

Figure 7.4 R2 δ-elimination by water, proline crosslink (–1402.90305)

8 2.30591 -1.091580.06434
6 1.28823 -1.19445-0.99647
6 -0.07005-0.78912-0.48019
8 -0.243580.55602 -0.21525
6 -1.18213-1.59703-0.67143
6 -2.56662-1.31213-0.52936
6 -3.09295-0.06441-0.24547
1 1.29374 -2.23628-1.32073
1 1.60479 -0.56165-1.83139
1 -0.94242-2.62108-0.95202
1 -3.24212-2.15070-0.66570
1 -2.420820.77380 -0.09791
7 -4.392840.22522 -0.12043
1 0.63053 0.95729 -0.01436
6 -5.48238-0.75133-0.27170
1 -5.40244-1.26933-1.23408
1 -5.41185-1.506030.52337
6 -6.759930.09673 -0.15237
1 -7.58699-0.466920.28658
1 -7.075330.43878 -1.14440
6 -4.897341.56964 0.21325
6 -6.322301.29652 0.70625
1 -4.902202.20495 -0.68297
1 -4.256062.03994 0.96423
1 -6.973922.16653 0.59133
1 -6.305141.02087 1.76690
15 3.43539 0.12061 0.05356
8 2.37991 1.37804 0.21997
8 4.15413 0.21560 -1.27150
8 4.28916 -0.081681.27892
6 2.84961 2.71354 -0.06241
1 3.25574 2.76580 -1.07557
1 1.98638 3.37343 0.03152
1 3.61162 3.00495 0.66694
1 0.05154 -1.514480.97949
8 0.64097 -1.712001.92541
1 0.55050 -2.649092.19488
1 1.56452 -1.569341.54411
11 6.21747 0.04453 -0.14240

Figure 7.4 R2-P1 First step of hydrolysis after δ-elimination before enol-keto rearrangement, proline crosslink (–825.52888)

1	-1.58121	0.72063	-1.09059
6	4.66795	-0.36582	-0.91804
6	3.19835	-0.33252	-1.34068
7	2.50757	0.09245	-0.11066
6	3.31291	-0.21947	1.09120
6	4.58712	-0.87890	0.53037
1	2.98861	0.37264	-2.15137
1	2.77435	-0.89020	1.76976
1	3.53966	0.70492	1.63530
1	5.47108	-0.63397	1.12535
1	5.09008	0.64542	-0.94538
1	5.27322	-1.00582	-1.56557
6	-1.90569	3.50907	0.04928
6	-1.25089	2.33899	-0.06583
8	-1.28161	1.65389	-1.25455
6	-0.44447	1.82964	1.05359
6	0.56175	0.93106	1.06818
6	1.16452	0.25696	-0.10357
1	-1.84084	4.10097	0.95554
1	-2.48638	3.90055	-0.77896
1	-0.69157	2.28752	2.00913
1	1.02466	0.71678	2.02763
1	0.77134	0.57250	-1.06417
1	2.85091	-1.32850	-1.64882
1	4.47321	-1.96866	0.53250
8	0.41558	-1.35065	-0.15252
1	0.12145	-1.59821	0.75367
1	-0.62227	-1.19887	-0.56549
8	-1.88686	-1.77496	1.32961
8	-1.90490	-0.93866	-0.75602
6	-2.49310	-1.28621	0.36570
6	-3.99453	-1.03103	0.42516
1	-4.38037	-1.51438	1.32754
1	-4.13362	0.05182	0.55485
6	-4.75278	-1.49483	-0.82459
1	-5.82119	-1.27119	-0.73201
1	-4.37110	-0.99402	-1.71854
1	-4.64745	-2.57567	-0.97255

Figure 7.4 R2-P2 First step of hydrolysis after enol-keto rearrangement, proline crosslink (–786.22260)

6	2.58415	-2.55472	-0.96379
6	1.33071	-1.68959	-0.82286
7	1.84792	-0.43085	-0.23845
6	3.18326	-0.62673	0.37311
6	3.42570	-2.14075	0.25587
1	0.82462	-1.47817	-1.76814
1	3.18141	-0.28196	1.41075
1	3.93108	-0.05365	-0.18698
1	4.48757	-2.37628	0.14768
1	3.11026	-2.31053	-1.89391
1	2.34790	-3.62172	-0.97726

6	-0.36957	3.28439	-0.40587
8	-0.81914	4.40598	-0.58935
6	0.98300	3.08701	0.16091
6	1.59689	1.90454	0.36439
6	1.03315	0.58932	0.00629
1	-0.96819	2.38677	-0.64328
1	1.49738	4.00696	0.42960
1	2.59217	1.90743	0.79998
1	0.09245	0.55712	-0.54213
1	0.60251	-2.13460	-0.13344
1	3.05661	-2.64625	1.15494
8	0.14920	0.16161	1.67456
1	-0.17367	0.99534	2.05802
1	-0.78130	-0.39379	1.31613
8	-1.84435	0.18731	-0.94670
8	-1.82179	-1.07474	0.92590
6	-2.33062	-0.68249	-0.19384
6	-3.61273	-1.40063	-0.62454
1	-4.12732	-0.75689	-1.34512
1	-3.30216	-2.29817	-1.17973
6	-4.53912	-1.80500	0.52514
1	-5.41787	-2.34043	0.14765
1	-4.01589	-2.45237	1.23406
1	-4.89329	-0.92602	1.07616

Figure 7.4 R2-P2 Enol-keto rearrangement after δ-elimination, proline crosslink (–557.53855)

6	-3.83827	0.83201	-0.50482
6	-2.51779	0.39529	-0.26312
8	-2.17352	-0.83922	-0.31352
6	-1.53560	1.42564	0.15884
6	-0.18183	1.31314	0.20095
6	0.51539	0.10282	-0.13227
1	-4.08896	1.87689	-0.35385
1	-4.39224	0.33666	-1.30406
1	-1.96081	2.39526	0.40421
1	0.39945	2.18341	0.48838
1	-0.07449	-0.75943	-0.42965
1	-3.42843	-1.52040	0.18004
8	-4.40511	-1.44823	0.58798
1	-4.42420	-0.31093	0.38073
1	-4.33816	-1.59934	1.55069
7	1.80710	-0.03229	-0.09291
6	2.50429	-1.30115	-0.44830
1	1.92865	-2.15053	-0.07709
1	2.56019	-1.35691	-1.54081
6	2.78691	1.01278	0.32343
1	2.60606	1.93246	-0.23697
1	2.63603	1.21040	1.38960
6	3.88954	-1.13548	0.17860
1	3.86593	-1.43249	1.23256
1	4.63747	-1.74615	-0.33141
6	4.15387	0.37375	0.04312
1	4.48016	0.61501	-0.97365
1	4.91353	0.73892	0.73731

Figure 7.4 R2-P2 Second step of hydrolysis, proline crosslink (–594.23477)

6 -2.65094-1.61624-0.21003
6 -2.11527-0.46135-1.07077
7 -1.147280.28418 -0.18758
6 -1.15914-0.351581.17700
6 -2.49660-1.092051.22665
1 -1.61513-0.79260-1.98375
1 -1.044980.43098 1.92767
1 -0.31463-1.042591.25517
1 -2.49443-1.887191.97729
1 -2.03579-2.51214-0.35193
1 -3.68134-1.87289-0.46983
6 3.22285 -0.151960.27111
8 4.31860 -0.564760.62885
6 2.28577 -1.02313-0.46934
6 1.04456 -0.72534-0.89346
6 0.27947 0.56512 -0.76258
1 2.89001 0.87542 0.47622
1 2.68776 -2.01179-0.68708
1 0.51615 -1.49759-1.45282
1 0.03649 0.90342 -1.79197
1 -2.908530.23600 -1.35531
1 -3.30788-0.397021.47240
8 0.85049 1.52449 0.00606
1 0.00717 2.29622 0.13888
8 -1.209342.76105 0.25611
1 -1.391103.28656 -0.53974
1 -1.455081.36890 -0.04170

Figure 7.4 Enol-keto rearrangement after δ-elimination (–420.91922)

6 1.99507 1.14852 0.35251
6 0.76477 0.46134 0.21905
8 0.68984 -0.812290.40915
6 -0.417211.20552 -0.26842
6 -1.704120.79679 -0.30564
6 -2.24492-0.510860.11641
1 2.03300 2.19968 0.08268
1 2.60819 0.88627 1.21708
1 -0.202652.22292 -0.58826
1 -2.462101.49025 -0.66572
1 -1.53074-1.240770.52573
1 1.91287 -1.231750.07284
8 2.95120 -1.11883-0.38184
1 2.76457 0.05071 -0.34922
1 2.88801 -1.41288-1.30928
8 -3.43762-0.765410.01402

Neutral Lysine crosslink (Figures 7.5 and 7.6)

Figure 7.5 R1 C2′–H abstraction and elimination by Asp–H, neutral lysine crosslink (–1632.62130)

1 0.93415 2.50117 -1.98117
7 3.75301 0.51633 0.96171

6 4.98698 0.47491 0.18173
6 5.80440 -0.765290.55587
1 5.56561 1.37834 0.41624
1 4.79041 0.47925 -0.90386
1 6.75594 -0.778160.01227
8 0.45081 4.26400 -0.92175
6 0.55299 3.39244 0.19146
6 0.07504 2.01925 -0.34352
8 0.82955 1.66843 -1.48549
6 0.21985 1.02574 0.75724
6 1.37371 0.24226 1.09879
6 2.66171 0.34622 0.33300
1 -0.075983.73571 1.02513
1 1.59637 3.32652 0.53437
1 -0.992882.10967 -0.58306
1 1.54542 0.27010 2.18233
6 1.03869 5.53701 -0.68959
1 0.92183 6.11196 -1.61082
1 0.53383 6.06351 0.13240
1 2.10767 5.44587 -0.44979
1 6.01784 -0.778301.63010
1 5.24793 -1.675460.30789
1 1.02544 -0.901830.97050
1 -0.502371.16596 1.55717
8 -1.14147-0.38871-0.11032
15 -2.74152-0.226810.20540
8 -3.148601.17704 -0.20977
8 -3.54473-1.34033-0.43692
8 -2.68109-0.384671.83106
11 -4.662460.24902 -1.73108
6 -3.91743-0.510722.54694
1 -4.50000-1.354732.16428
1 -3.65962-0.687523.59295
1 -4.507980.40984 2.47005
1 2.60772 0.20624 -0.75065
8 0.32331 -2.247800.66160
1 -0.71643-1.297300.13949
6 1.05184 -3.02284-0.11007
8 2.25981 -2.87803-0.29471
6 0.26474 -4.15968-0.76941
1 -0.52172-3.70093-1.38471
1 -0.26739-4.699320.02454
6 1.12338 -5.11001-1.60164
1 1.89190 -5.58629-0.98434
1 0.50778 -5.89755-2.05123
1 1.63713 -4.57334-2.40551

Figure 7.5 R1 C2′–H abstraction and elimination by water, neutral lysine crosslink (–1440.63892)

1 1.29558 2.17399 -1.63595
7 3.84842 -1.270730.62303
6 5.13465 -1.06112-0.03057
6 5.73042 -2.39357-0.49582
1 5.81456 -0.605320.70273
1 5.06042 -0.36351-0.88349
1 6.72856 -2.24255-0.92242

8 1.09589 3.60399 -0.06479
6 1.10356 2.42463 0.72660
6 0.42795 1.34647 -0.14270
8 1.10765 1.25075 -1.38804
6 0.42564 0.04451 0.60361
6 1.48120 -0.894520.70269
6 2.82948 -0.692940.11500
1 0.55303 2.57757 1.66648
1 2.13754 2.13448 0.96845
1 -0.617361.63962 -0.29154
1 1.50812 -1.430791.65555
6 1.82720 4.67080 0.51876
1 1.76908 5.51436 -0.17315
1 1.39669 4.96714 1.48607
1 2.88196 4.39838 0.67120
1 5.81375 -3.092350.34337
1 5.09688 -2.85691-1.26067
1 0.99519 -1.93693-0.03072
1 -0.305470.02517 1.40442
8 -1.08376-0.82279-0.63100
15 -2.57289-0.60388-0.13970
8 -2.818830.84572 0.30044
8 -3.60609-1.08803-1.15816
8 -2.58911-1.584251.19361
11 -4.516661.01729 -1.21793
6 -3.79875-1.665241.94391
1 -4.62969-2.006671.31436
1 -3.63627-2.390452.74605
1 -4.05659-0.693732.38436
1 2.89194 -0.05551-0.77528
8 0.32073 -2.79591-0.69355
1 0.73303 -2.85459-1.57468
1 -0.47316-2.08824-0.78060

Figure 7.5 R1 C2′–H abstraction by OG, neutral lysine crosslink (–1981.56437)

8 -0.643573.69947 0.63400
6 -0.534282.31189 0.92898
6 -1.835481.63078 0.51242
8 -1.879891.60578 -0.92006
6 -1.974210.24597 1.14218
8 -3.72699-0.065891.03002
6 -1.17266-0.834540.66819
6 -1.22908-2.099731.36275
7 2.40377 -0.10218-0.17337
6 2.37411 -0.210001.15468
7 3.61559 -0.209561.72210
6 4.52325 -0.102270.67424
6 5.93763 -0.045640.64530
8 6.75861 -0.076621.56723
7 6.37240 0.06848 -0.70534
6 5.54944 0.11704 -1.80716
7 6.17310 0.29368 -3.02734
7 4.24178 0.06329 -1.74697
6 3.73835 -0.03824-0.48085
8 1.31436 -0.297971.92934

15 -4.608170.10009 -0.27248
8 -3.76912-0.09933-1.55917
8 -5.405851.40326 -0.30359
8 -5.71583-1.11413-0.17149
6 -5.22873-2.45701-0.23103
1 -0.353752.17038 2.00754
1 0.30505 1.86640 0.37705
1 -2.660902.24806 0.89294
1 -2.054320.31592 2.22658
1 -0.92209-0.84741-0.38987
1 0.44169 -0.429871.38136
1 -1.64222-2.042882.38064
1 3.81952 -0.307602.70606
1 7.37501 0.16841 -0.81873
1 7.06522 -0.17127-3.15087
1 5.54050 0.13100 -3.80210
1 -4.46812-2.634620.53764
1 -6.08314-3.11725-0.05596
6 0.54444 4.40393 0.93048
1 0.37331 5.45360 0.67484
1 0.80264 4.33533 1.99998
1 1.39874 4.02764 0.34621
1 -4.79649-2.67235-1.21474
11 -4.063222.25561 -1.94948
7 -0.84004-3.281731.00232
1 -2.398670.79426 -1.19272
6 -0.29130-3.41845-0.34499
1 -1.00879-3.06264-1.10491
1 0.60869 -2.79297-0.47341
6 0.06449 -4.87835-0.63054
1 0.47214 -4.99862-1.64204
1 0.81027 -5.242170.08637
1 -0.82227-5.51688-0.53826

Figure 7.5 R1-P1 First step of hydrolysis after β-elimination, neutral lysine crosslink (–671.83454)

1 2.19578 1.48252 0.38150
7 -2.26832-0.15361-1.05892
6 -3.32676-1.14411-0.85193
6 -4.13951-0.885340.42201
1 -2.89469-2.14907-0.84053
1 -3.97753-1.09358-1.73121
1 -4.95545-1.612070.50312
8 3.96460 0.34200 0.04799
6 3.01279 -0.69114-0.14687
6 1.84021 -0.360680.78981
8 1.38546 0.95724 0.51484
6 0.75748 -1.393730.67667
6 -0.49043-1.317260.17986
6 -1.10253-0.14609-0.43892
1 3.44215 -1.676810.08774
1 2.65939 -0.70199-1.18953
1 2.22927 -0.409901.82186
1 1.06127 -2.354841.09004
1 -0.453520.67503 -0.70009
6 5.04223 0.29551 -0.87511

1 5.69765 1.13797 -0.64316
1 5.60964 -0.64109-0.77732
1 4.68618 0.38664 -1.91152
1 -3.50697-0.957421.31099
8 -1.539931.14087 1.26622
1 -0.635181.48771 1.35193
1 -2.596860.87054 -1.15996
1 -4.571290.12025 0.40250
1 -1.10055-2.213220.24791
8 -2.672202.34367 -0.53001
1 -3.589212.48667 -0.24683
1 -2.086251.82484 0.48955

Figure 7.5 R1-P2 δ-elimination by water, neutral lysine crosslink (–1324.89110)

8 1.98030 -1.211780.59992
6 0.49833 -1.71755-0.80814
6 -0.70669-1.11442-0.38917
8 -0.898580.22942 -0.74265
6 -1.91811-1.95507-0.28825
6 -3.22347-1.60162-0.28172
6 -3.80324-0.26902-0.39893
1 0.65138 -2.77672-0.63516
1 1.15189 -1.24799-1.53392
1 -1.69275-3.01421-0.17642
1 -3.95500-2.40174-0.18535
1 -3.113240.57534 -0.49980
7 -5.07788-0.11703-0.37703
1 -0.076340.70354 -0.48443
6 -5.587251.24302 -0.48041
6 -6.422431.60571 0.75230
1 -6.230761.29312 -1.37051
1 -4.780381.98363 -0.61974
1 -6.865812.60150 0.63857
1 -5.801491.60704 1.65526
15 2.81383 0.05340 0.13479
8 1.59232 1.20553 0.15709
8 3.32811 -0.07758-1.29871
8 3.89302 0.46774 1.12935
6 1.96578 2.56949 -0.06546
1 2.40386 2.69492 -1.06303
1 1.05708 3.17170 0.00822
1 2.68355 2.89531 0.69428
1 -0.42920-1.105840.95141
8 0.07566 -0.909812.12854
1 -0.09707-1.706622.66267
1 1.03650 -1.006231.72622
11 5.45832 0.42540 -0.55253
1 -7.230220.88051 0.89723

Figure 7.5 R1-P2 First step of hydrolysis after δ-elimination, neutral lysine crosslink (–556.09829)

1 2.96537 1.78230 -0.39221
7 -1.59300-0.24281-0.93596
6 -2.49131-1.31250-0.49706
6 -3.03588-1.075940.91673

1 -1.97714-2.27559-0.56194
1 -3.31400-1.34672-1.21885
1 -3.71914-1.884931.19762
6 4.10602 -0.291770.20518
6 2.79050 -0.09508-0.01186
8 2.25337 1.12399 -0.33142
6 1.83873 -1.195360.09490
6 0.50211 -1.22356-0.11774
6 -0.32415-0.12647-0.59181
1 4.48280 -1.275480.45678
1 4.82711 0.51772 0.13625
1 2.30088 -2.124510.41714
1 0.16963 0.75209 -0.97607
1 -2.22156-1.022841.64449
8 -0.541521.18368 1.19461
1 0.30672 1.65249 1.12619
1 -2.026110.74252 -0.97778
1 -3.58268-0.128870.96230
1 0.00358 -2.168200.07567
8 -2.101632.22957 -0.35593
1 -2.956352.30889 0.09639
1 -1.305921.79691 0.51766

Figure 7.5 R1-P2 Second step of hydrolysis after δ-elimination, neutral lysine crosslink (–556.11721)

1 1.74618 0.38650 -2.15973
7 -1.140590.23330 -0.00332
6 -1.667791.40897 0.73079
6 -2.234112.45400 -0.22552
1 -0.857771.82738 1.33734
1 -2.443001.04347 1.41116
1 -2.623203.30531 0.34209
6 3.62746 0.54938 -0.62653
6 2.35490 0.22514 -0.33512
8 1.36581 0.18224 -1.28961
6 1.93417 -0.093351.02964
6 0.76119 -0.522711.53339
6 -0.54097-0.919630.86046
1 4.38189 0.57242 0.15022
1 3.94541 0.78174 -1.63903
1 2.73959 0.05062 1.74641
1 -1.30533-1.022081.65424
1 -1.462812.82883 -0.90792
8 -0.42952-2.016760.06707
1 -1.45565-2.04733-0.47297
1 -0.407400.50088 -0.67085
1 -3.053032.04033 -0.82337
1 0.74985 -0.678832.61059
8 -2.56855-1.64008-1.00728
1 -1.91849-0.33913-0.55674
1 -2.40309-1.69720-1.96229

Figure 7.5 R2 C2′–H abstraction and protonation of neutral crosslink by water (–1517.09492)

8 1.42750 3.74379 -0.05280
6 0.49320 2.75463 -0.45387

6 1.17236 1.39341 -0.32367
8 1.52204 1.17744 1.05608
6 0.31361 0.26221 -0.92888
8 1.08081 -0.92785-1.23187
6 -0.93338-0.06800-0.11075
6 -2.09146-0.51484-0.87791
8 -2.929121.88098 0.33804
15 2.11747 -1.61190-0.14235
8 1.49833 -1.501191.24662
8 3.50992 -1.04930-0.32612
8 2.13070 -3.13713-0.68570
6 0.95028 -3.94057-0.54200
1 0.20206 2.91504 -1.50632
1 -0.423212.80248 0.15284
1 2.09646 1.43321 -0.91829
1 0.00553 0.62120 -1.91755
1 -0.70191-0.685350.75906
1 -1.471430.87896 0.27817
1 -2.17297-0.22483-1.92498
1 0.12758 -3.52404-1.13298
1 1.20128 -4.93461-0.91710
6 0.88039 5.04851 -0.09740
1 1.66420 5.73882 0.22551
1 0.56143 5.31974 -1.11609
1 0.01411 5.14742 0.57416
1 0.65213 -4.005440.50870
11 3.80078 0.50037 1.26543
7 -3.17322-1.00367-0.36244
1 1.32553 0.23322 1.29507
6 -3.35815-1.263461.07449
1 -2.50707-1.845511.43788
1 -3.34171-0.277841.55462
6 -4.67230-1.994501.32056
1 -4.81311-2.147812.39460
1 -5.51680-1.405410.94729
1 -4.68540-2.972810.82902
1 -4.091701.39240 -0.36736
1 -3.223812.26461 1.17829
8 -4.872060.88859 -0.93072
1 -4.04962-0.69497-0.84228
1 -5.681961.00470 -0.41146

Figure 7.6 R3 First step of hydrolysis, neutral lysine crosslink (–1517.10307)

1 -0.692432.61098 1.76743
7 -3.36757-0.652670.25039
6 -4.17098-0.66568-0.97439
6 -5.65799-0.72604-0.63469
1 -3.88493-1.53074-1.59016
1 -3.954950.23439 -1.56078
1 -6.25576-0.75044-1.55156
8 0.08941 4.12801 0.47710
6 0.01859 3.07258 -0.47144
6 0.26295 1.78446 0.31582
8 -0.656621.70959 1.40112
6 0.13939 0.52418 -0.55173

6 -1.299820.11426 -0.89787
6 -2.06159-0.507070.24069
1 0.77717 3.20151 -1.25889
1 -0.975403.06445 -0.94667
1 1.29592 1.80885 0.68904
1 0.67398 0.71208 -1.49006
1 -1.845520.99285 -1.26334
1 -1.59066-0.540311.21271
6 -0.274605.38929 -0.05844
1 -0.201096.11629 0.75387
1 0.40210 5.68866 -0.87209
1 -1.305145.37990 -0.44371
1 -5.88914-1.62595-0.05410
8 -1.47398-2.542800.04388
1 -0.58854-2.513080.44211
1 -3.61107-1.421790.99681
1 -5.961520.14628 -0.04628
1 -1.27434-0.60844-1.71870
8 0.75505 -0.592140.12371
8 -3.24984-2.740901.72473
1 -3.86978-3.406881.38625
1 -2.22170-2.773040.85997
15 2.32949 -0.96022-0.09839
8 3.24296 0.23833 0.11641
8 2.62627 -2.151620.78906
8 2.26687 -1.34901-1.68623
11 4.23953 -0.884921.89578
6 3.50152 -1.49569-2.39810
1 4.08554 -2.33257-1.99720
1 3.24301 -1.70740-3.43781
1 4.09438 -0.57690-2.34477

Figure 7.6 R3-P1 C2′–H abstraction by Asp⁻ after first step of hydrolysis, neutral lysine crosslink (–1708.56102)

1 -1.313362.02169 1.98854
7 -3.467120.10296 -0.75606
6 -4.75324-0.04926-0.06619
6 -5.62741-1.24961-0.46122
1 -5.331230.87302 -0.22920
1 -4.53618-0.087601.00927
1 -6.58088-1.226890.08376
8 -1.626563.75731 0.76941
6 -1.466862.76362 -0.24950
6 -0.449531.75819 0.29885
8 -0.923511.24674 1.54735
6 -0.189240.65917 -0.71599
6 -1.05831-0.42908-1.01983
6 -2.35992-0.73076-0.30223
1 -1.096113.22930 -1.17643
1 -2.420972.25979 -0.45117
1 0.49609 2.28968 0.45225
1 -1.14002-0.62025-2.09720
6 -2.680354.65861 0.49653
1 -2.735855.35976 1.33439
1 -2.497485.22527 -0.42991

1 -3.645774.13863 0.39795
1 -5.85909-1.22214-1.53431
1 -5.11541-2.19059-0.25138
1 0.30047 1.08795 -1.58807
8 1.37975 0.01751 0.01863
15 2.75465 0.69258 -0.35436
8 2.71240 2.22832 -0.40320
8 3.88820 0.16783 0.53986
8 3.06403 0.25908 -1.93064
11 4.37338 2.30501 1.14040
6 3.20208 -1.13191-2.22142
1 2.28970 -1.68394-1.97164
1 3.39692 -1.22126-3.29526
1 4.04323 -1.56808-1.66777
1 -2.26047-0.542700.77093
8 0.12465 -2.80768-0.86669
1 -0.22789-1.73814-0.78307
6 0.51444 -3.405490.25829
8 1.06247 -4.500910.23384
6 0.20663 -2.669111.55503
1 -0.88508-2.555301.61087
1 0.60896 -1.654551.45898
6 0.74378 -3.375232.79793
1 1.83365 -3.470432.75310
1 0.48773 -2.811143.70225
1 0.32993 -4.384602.89308
1 -3.58573-0.02831-1.76102
8 -2.71907-2.13408-0.49042
1 -1.90367-2.61136-0.72205

Figure 7.6 R3-P1 Second step of hydrolysis after β-elimination, neutral lysine crosslink (–671.86007)

1 -1.260450.37705 1.72802
7 1.71306 -0.196290.12887
6 2.92355 0.61719 0.39765
6 3.35202 0.50022 1.85710
1 2.70502 1.65818 0.13702
1 3.71304 0.26048 -0.27136
1 4.24967 1.10204 2.03134
8 -3.09758-0.063190.85776
6 -2.384020.28390 -0.32701
6 -1.397251.38050 0.08548
8 -0.627230.90674 1.20602
6 -0.514421.81757 -1.06090
6 0.53387 1.20297 -1.62874
6 1.16080 -0.14515-1.31289
1 -3.070180.65040 -1.10529
1 -1.81975-0.57356-0.71499
1 -1.976452.25173 0.42279
1 -0.808812.77152 -1.49574
1 2.07135 -0.25113-1.92951
6 -3.90850-1.216850.70124
1 -4.38559-1.407771.66575
1 -4.68680-1.05669-0.05899
1 -3.31048-2.092890.41134
1 2.56623 0.86250 2.52931

8	0.29353	-1.19510	-1.46069
1	0.85217	-2.06388	-0.98143
1	0.92872	0.09704	0.73625
1	3.58144	-0.53798	2.11975
1	1.00318	1.72423	-2.46176
8	1.69059	-2.72493	-0.18606
1	1.87004	-1.30179	0.24180
1	1.11545	-3.12559	0.48583

Figure 7.6 R3-P1 C2′–H abstraction and elimination by water after first step of hydrolysis, neutral lysine crosslink (–1517.07436)

1	1.11116	2.42345	-1.47460
7	3.70039	-1.25765	0.88883
6	5.10536	-1.14809	0.48173
6	5.49983	-1.85679	-0.82109
1	5.71471	-1.54376	1.30505
1	5.33764	-0.07808	0.41198
1	6.57763	-1.75384	-0.99904
8	0.44870	3.84353	-0.10083
6	0.50632	2.74333	0.79461
6	0.16368	1.50842	-0.06467
8	1.02418	1.49187	-1.20115
6	0.29393	0.28062	0.78127
6	1.39599	-0.56240	0.97056
6	2.74779	-0.40776	0.24660
1	-0.21693	2.85995	1.61390
1	1.51592	2.65423	1.22483
1	-0.88331	1.58851	-0.37181
1	1.47614	-0.95261	1.98897
6	0.90620	5.05877	0.47460
1	0.83905	5.82371	-0.30225
1	0.28043	5.35418	1.32841
1	1.94924	4.97372	0.81238
1	5.27408	-2.92908	-0.76169
1	4.95731	-1.44327	-1.67430
1	0.96575	-1.66981	0.39139
1	-0.50219	0.20951	1.51641
8	-1.10339	-0.85639	-0.64323
15	-2.59738	-0.79542	-0.17534
8	-2.98404	0.60737	0.33364
8	-3.59744	-1.31446	-1.21593
8	-2.56425	-1.83725	1.11850
11	-4.67863	0.68497	-1.18464
6	-3.79250	-2.10771	1.78479
1	-4.53922	-2.50756	1.08760
1	-3.58900	-2.85436	2.55828
1	-4.19609	-1.20381	2.26043
1	3.12362	0.62057	0.30958
8	0.66012	-2.58116	-0.51680
1	1.31318	-2.21913	-1.16277
1	-0.21877	-2.05723	-0.71403
1	3.38899	-2.22779	0.87576
8	2.59229	-0.72177	-1.16737
1	2.07461	0.02855	-1.53023

Figure 7.6 R3-P2 C2′–H abstraction and protonation of neutral crosslink by water after fist step of hydrolysis (–1517.07089)

1	1.68869	-2.45664	-1.59415
7	3.42806	0.59950	0.48503
6	2.62079	0.58157	1.74776
6	3.00674	-0.54796	2.69638
1	1.57331	0.52422	1.44294
1	2.77488	1.55139	2.23187
1	2.42226	-0.45579	3.61755
8	-2.21561	-2.15796	0.25302
6	-1.42278	-2.27259	-0.93786
6	0.05633	-1.85766	-0.78786
8	0.74890	-2.67508	-1.72864
6	0.33058	-0.34640	-1.06869
6	1.72343	0.08176	-1.42270
6	2.94104	-0.37861	-0.79171
1	-1.92095	-1.64545	-1.68011
1	-1.44487	-3.31676	-1.27863
1	0.39265	-2.07992	0.23510
1	1.86255	0.18706	-2.49873
1	3.80710	-0.22633	-1.44440
6	-1.99776	-3.20384	1.19228
1	-2.66075	-3.01953	2.04150
1	-2.23652	-4.18203	0.75069
1	-0.96214	-3.22039	1.55192
1	2.80254	-1.52696	2.25797
8	2.94567	-1.68889	-0.23591
1	3.85381	-1.95295	-0.00671
1	4.06785	-0.49819	2.97327
1	4.41574	0.43300	0.70194
1	2.01704	1.92905	-1.20636
1	-0.31010	-0.05294	-1.90497
8	-0.18670	0.35056	0.14002
15	-1.55838	1.21470	0.12975
8	-2.46197	0.81986	-1.03280
8	-2.20873	1.10256	1.50077
8	-0.93991	2.71655	-0.10275
11	-3.72940	-0.39241	0.50745
6	-1.86223	3.80397	-0.19656
1	-2.45080	3.90151	0.72380
1	-1.27117	4.71143	-0.34266
1	-2.54095	3.67274	-1.04774
8	2.53750	2.77556	-0.95916
1	3.34461	1.54709	0.05251
1	3.04799	2.96903	-1.76244

Figure 7.6 R3-P2 3′-PO_4 elimination (-1305.49111)

1	-3.206859	-0.383647	-1.673907
8	-3.612033	-1.349510	0.206108
6	-2.802804	-0.292266	0.695459
6	-1.803480	0.015420	-0.433772
8	-2.505554	0.290293	-1.636536
6	-0.925249	1.160994	-0.064556
6	-1.103127	2.440088	-0.607779

6	-0.126329	3.403567	-0.437648
1	-2.275029	-0.588643	1.613561
1	-3.421342	0.593124	0.912222
1	-1.145051	-0.862502	-0.540780
1	-1.958185	2.642115	-1.242804
1	-0.297612	4.453801	-0.675315
6	-4.702685	-1.658263	1.059214
1	-5.261967	-2.466729	0.582641
1	-4.354374	-1.993894	2.046567
1	-5.364510	-0.790017	1.194901
8	1.087690	3.122203	-0.052615
1	1.240641	2.085182	-0.223087
1	-0.414639	1.055712	0.885821
8	1.079018	0.711334	-0.727313
15	1.742497	-0.557380	-0.041155
8	0.860375	-1.800204	-0.177958
8	3.179794	-0.773532	-0.512053
8	1.738978	-0.207298	1.583373
6	2.767268	0.633612	2.108956
1	2.666281	1.666918	1.751802
1	2.662193	0.628990	3.197576
1	3.758065	0.258150	1.833884
11	2.617019	-2.918461	-1.129914

Figure 7.6 R3-P3 Second step of hydrolysis, neutral lysine crosslink (–1517.12872)

1	-1.33502	2.40787	1.42654
7	-3.13749	-1.33289	-0.14203
6	-3.83759	-0.22306	0.55733
6	-5.35091	-0.34110	0.40763
1	-3.46530	0.72916	0.17235
1	-3.53937	-0.28228	1.60772
1	-5.84174	0.47037	0.95465
8	-0.73087	3.99614	0.13081
6	-0.42099	2.91135	-0.73406
6	-0.03828	1.74138	0.17270
8	-1.07558	1.51916	1.12494
6	0.24579	0.45128	-0.60911
6	-1.01058	-0.29722	-1.09585
6	-1.56451	-1.34676	-0.11386
1	0.41121	3.17208	-1.40556
1	-1.29870	2.66327	-1.35282
1	0.89669	2.01005	0.68457
1	0.84047	0.73951	-1.48357
1	-1.77392	0.44781	-1.34863
1	-1.35403	-1.06034	0.92732
6	-1.23344	5.13165	-0.55245
1	-1.45045	5.89163	0.20214
1	-0.49486	5.53048	-1.26309
1	-2.15642	4.89430	-1.10275
1	-5.65647	-0.26975	-0.64276
8	-1.18805	-2.60787	-0.44695
1	-1.89382	-3.29910	0.19111
1	-3.43702	-1.40163	-1.11887
1	-5.71770	-1.29216	0.80802
1	-0.77413	-0.84406	-2.01396

8 1.02588 -0.429770.22331
8 -2.94131-3.714320.78458
1 -3.33656-2.318420.33273
1 -2.73716-3.678881.73323
15 2.64489 -0.54326-0.00989
8 3.21017 -1.280661.18798
8 2.76514 -1.41748-1.38245
8 3.25404 0.81817 -0.29521
6 2.29433 -2.78045-1.36360
1 1.23913 -2.82284-1.07484
1 2.41803 -3.16402-2.37915
1 2.89375 -3.37975-0.67059
11 4.61766 0.52605 1.57550

Figure 7.6 R3-P3 C2′–H abstraction by Lys after hydrolysis, AP-site *pro*-S (–1440.68395)

6 -3.12895-1.129990.18018
8 -4.02770-1.44251-0.87967
6 -1.83194-0.67702-0.49158
8 -2.090820.40636 -1.38070
6 -0.72957-0.317950.50750
8 0.51029 -0.22201-0.34883
6 -0.879680.88884 1.39049
6 -2.005421.76291 1.43993
8 -2.277062.55185 2.36844
1 -2.95113-2.015330.81116
1 -3.55003-0.336480.81108
1 -1.45839-1.53631-1.07216
1 -2.913870.16176 -1.84001
1 -0.56142-1.210841.11784
1 -0.369420.78485 2.35297
1 0.12606 1.96383 0.71766
1 -2.637211.77826 0.52650
15 1.88404 -0.975300.02798
6 4.14020 0.18424 0.79986
8 2.74749 0.35850 0.51091
8 2.53625 -1.57099-1.20928
8 1.72492 -1.956271.18052
1 4.67435 -0.18124-0.08344
1 4.27603 -0.517791.63011
6 1.57150 3.67567 -1.99604
6 0.47432 3.01925 -1.16687
7 0.94215 2.61199 0.18771
1 1.94758 4.58746 -1.51767
1 1.17843 3.95013 -2.98064
1 -0.372263.69587 -1.02009
1 0.09545 2.11019 -1.63891
1 1.17758 3.41869 0.76937
1 1.75837 1.98430 0.14571
1 2.41385 2.99184 -2.14868
1 4.53362 1.16155 1.08831
6 -5.34058-1.72669-0.42966
1 -5.94837-1.92957-1.31518
1 -5.35730-2.609440.22712
1 -5.77220-0.875950.11841
11 2.56806 -3.64593-0.16063

Figure 7.6 R3-P3 C2′–H abstraction by Lys after hydrolysis, AP-site *pro*-R (–1440.69676)

6	-3.07147	-1.07799	0.179052
8	-3.94804	-1.52402	-0.844738
6	-1.84736	-0.43501	-0.486554
8	-2.17084	0.76472	-1.164495
6	-0.67743	-0.28670	0.506247
8	0.53407	-0.16471	-0.336094
6	-0.72219	0.84939	1.521714
6	-1.84246	1.67564	1.831092
8	-2.83151	1.98869	1.128555
1	-2.75922	-1.93279	0.804919
1	-3.57915	-0.35096	0.828191
1	-1.50167	-1.13523	-1.256314
1	-2.53580	1.35794	-0.459663
1	-0.56695	-1.23797	1.036774
1	-0.13242	0.62087	2.413587
1	0.12448	1.89071	0.936053
1	-1.74893	2.17710	2.818822
15	1.90448	-0.95606	-0.006874
6	4.10897	0.07728	1.026062
8	2.74081	0.30467	0.665384
8	2.58041	-1.38610	-1.297111
8	1.70947	-2.07574	1.003947
1	4.69710	-0.21300	0.148817
1	4.17990	-0.70317	1.791726
6	1.42750	3.75555	-1.888228
6	0.38492	3.13396	-0.964122
7	0.95622	2.64564	0.318589
1	1.91406	4.62123	-1.423153
1	0.95232	4.09367	-2.815582
1	-0.39568	3.85745	-0.709572
1	-0.10695	2.27581	-1.428599
1	1.26091	3.41694	0.914921
1	1.75603	2.01888	0.175142
1	2.20156	3.02669	-2.154244
1	4.49487	1.01576	1.430374
6	-5.14845	-2.05955	-0.329353
1	-5.75841	-2.37063	-1.18222
1	-4.96384	-2.93548	0.314589
1	-5.70894	-1.31489	0.257881
11	2.57534	-3.58301	-0.562592

Figure 7.6 R3-P3 Protonation of 3′-PO_4 by Lys (–1440.70495)

6	-2.96227	-1.02961	0.13911
8	-3.37709	-1.59555	-1.09138
6	-1.57208	-0.40029	-0.06390
8	-1.63616	0.72549	-0.91948
6	-0.96034	-0.16964	1.30816
8	0.91092	0.26067	0.75052
6	-1.34463	0.81198	2.23631
6	-1.88707	2.08767	1.98741
8	-2.10110	2.66289	0.87730
1	-2.92791	-1.80647	0.92195
1	-3.67202	-0.25465	0.47238

1 -0.94947-1.14498-0.56377
1 -1.993261.47363 -0.38188
1 -0.62930-1.105851.74600
1 -1.121650.61064 3.28179
1 -0.383552.87319 0.09212
1 -2.136372.66473 2.89751
15 1.88871 -0.852180.18808
6 2.77756 -1.284822.64147
8 3.11267 -0.928201.30237
8 2.57623 -0.43151-1.11417
8 1.20481 -2.222140.08953
1 3.70364 -1.271943.22343
1 2.33866 -2.289002.68695
6 2.09356 2.71401 -2.23620
6 0.66308 2.88031 -1.74290
7 0.58072 2.71220 -0.25450
1 2.75194 3.48654 -1.82045
1 2.11225 2.81408 -3.32666
1 0.25807 3.86736 -1.98281
1 -0.001772.11845 -2.15225
1 1.20661 3.36237 0.22732
1 0.82313 1.71341 0.05963
1 2.47924 1.72538 -1.96460
1 2.07122 -0.565733.07525
6 -4.67432-2.15309-1.01685
1 -4.90622-2.56763-2.00166
1 -4.72973-2.95887-0.26736
1 -5.42934-1.39269-0.76207
11 1.85610 -2.36972-2.10669

Figure 7.6 R3-P3 C2′–H abstraction by Asp⁻ after hydrolysis, AP-site (–1573.45489)

8 -3.79599-2.26393-0.97741
6 -2.39516-2.08326-0.86780
6 -2.11993-0.71487-0.25075
8 -2.57362-0.662071.11219
6 -0.62958-0.31663-0.39692
8 -0.392640.96334 0.22566
6 0.39333 -1.280500.19732
6 0.40162 -1.611401.58478
6 3.72105 -0.95347-0.13270
8 2.73388 -0.10914-0.21756
15 -1.073412.40700 -0.27724
8 -1.774872.92301 0.96516
8 -1.769422.22284 -1.59386
8 0.28169 3.32070 -0.42679
6 1.21854 2.97339 -1.45501
1 -1.92289-2.14759-1.86370
1 -1.95279-2.87701-0.24259
1 -2.717990.02810 -0.78502
1 -0.43460-0.22261-1.47239
1 0.63654 -2.13499-0.43967
1 1.61362 -0.681600.06337
1 1.14959 -2.378991.85781
1 1.92875 3.80158 -1.53135
1 1.75617 2.05515 -1.19377

6 -4.13708-3.53280-1.49535
1 -5.22837-3.58013-1.54533
1 -3.72613-3.68601-2.50644
1 -3.77316-4.34803-0.84927
1 0.70916 2.84469 -2.41648
11 -1.386541.09590 2.32242
1 -1.91068-1.137231.65706
6 5.07659 -0.34511-0.52009
1 5.22325 0.55630 0.08908
1 4.99241 0.01364 -1.55489
6 6.25630 -1.30398-0.37197
1 7.19477 -0.82116-0.66958
1 6.35870 -1.642550.66428
1 6.11630 -2.19534-0.99165
8 3.62748 -2.136950.21604
8 -0.25822-1.082112.52078

Figure 7.6 R3-P3 Protonation of 3′-PO_4 by Asp–H, AP-site (–1573.43835)

6 -3.19932-0.98528-0.34726
8 -2.09389-0.68539-1.02049
1 -1.43877-0.15781-0.45200
1 1.20759 -1.94281-1.07346
8 4.40160 -1.01559-0.38262
6 3.30638 -1.310290.47669
6 2.06878 -0.55675-0.03683
8 1.62749 -1.06324-1.27813
6 1.01861 -0.538871.07167
8 -0.253090.61028 0.34130
6 0.34410 -1.695211.53108
6 0.15401 -2.914790.85480
15 -0.024282.17510 0.16590
8 -0.728272.68151 -1.09794
8 1.44959 2.58132 0.23997
8 -0.723482.86356 1.49932
6 -2.127222.66859 1.68821
1 3.55108 -1.005801.50923
1 3.10241 -2.391700.48219
1 2.36247 0.48157 -0.20435
1 1.28733 0.16313 1.85730
1 -0.15614-1.619442.49451
1 -0.42135-3.661531.44393
1 -2.701233.16583 0.89626
1 -2.385221.60386 1.69706
1 -2.387213.11570 2.65284
8 -3.45165-0.613220.78980
6 -4.12950-1.87053-1.16505
1 -3.59547-2.81075-1.35756
1 -4.26116-1.40227-2.14840
6 -5.47323-2.13274-0.48922
1 -6.09354-2.79210-1.10632
1 -5.33479-2.605440.48793
1 -6.02229-1.19925-0.32725
8 0.56070 -3.28639-0.28105
6 5.57200 -1.71112-0.01550
1 6.35690 -1.42964-0.72413

1 5.90329 -1.449491.00350
1 5.43008 -2.80342-0.05748
11 1.31277 3.25583 -1.95046

Figure 7.6 R3-P3 C2′–H abstraction and elimination by water after hydrolysis, AP-site (–1381.89555)

1 2.26004 -1.189641.56024
8 3.01321 -2.20755-0.14263
6 2.51372 -1.05928-0.81901
6 1.27850 -0.63201-0.02573
8 1.63885 -0.465681.36079
6 0.62082 0.63604 -0.55755
6 1.22678 1.94915 -0.39667
6 2.54168 2.23326 0.08853
1 2.23505 -1.30247-1.85521
1 3.27429 -0.26816-0.83929
1 0.53659 -1.43391-0.10271
1 0.79658 2.73403 -1.01843
6 4.25247 -2.66588-0.66146
1 4.55069 -3.52932-0.06217
1 4.15400 -2.97265-1.71252
1 5.02822 -1.88967-0.59182
1 0.68903 2.11882 1.41721
1 0.21163 0.44941 -1.55100
8 -0.852350.63278 0.27872
15 -2.17624-0.22496-0.18650
8 -1.79222-1.60236-0.69075
8 -3.16032-0.193810.96776
8 -2.705810.61750 -1.47897
11 -3.25772-2.533870.86303
6 -3.181621.95862 -1.28904
1 -2.382452.60284 -0.90668
1 -3.501932.31949 -2.26852
1 -4.027971.97520 -0.59511
1 3.03176 1.41943 0.66445
8 0.19041 1.62225 2.18315
1 -0.529261.15085 1.49317
1 0.80645 0.85225 2.32765
8 3.14721 3.30827 -0.06425

Figure 7.6 R3-P3 C4′–H abstraction and elimination by water after β-elimination, AP-site (–1103.88255)

8 -1.18833-1.218460.09586
6 0.02752 -1.111031.00095
6 1.18405 -0.446540.32077
8 1.10315 0.95873 0.27429
6 2.45346 -1.047500.39540
6 3.75085 -0.561340.22591
6 4.15209 0.78480 -0.01962
1 0.24340 -2.145941.27091
1 -0.34278-0.565081.87178
1 2.39158 -2.119740.59235
1 4.56173 -1.287700.27929
1 3.33120 1.52486 -0.08557

1 0.19825 1.18519 -0.02105
15 -2.49921-0.113340.24427
8 -1.651321.16031 -0.42259
8 -2.732250.12460 1.70145
8 -3.52975-0.60385-0.71858
6 -2.368002.39323 -0.57593
1 -1.688493.10318 -1.05233
1 -3.246792.24448 -1.21221
1 -2.681962.78315 0.39858
1 0.82706 -1.05798-1.22870
8 0.06205 -1.38020-1.93056
1 0.26922 -2.29829-2.18824
1 -0.74174-1.38003-1.05783
8 5.33241 1.16439 -0.16438

Figure A7.2 First step (–1517.58042)

1	-0.135810	1.013431	2.736688
7	-2.967974	-1.071495	-0.553519
6	-4.230859	-1.310767	0.191205
6	-5.282069	-1.951750	-0.707373
1	-4.608648	-0.367848	0.610164
1	-3.993147	-1.968289	1.033699
1	-6.204252	-2.116438	-0.141960
8	-0.418539	3.752066	-0.574327
6	-0.966093	3.276074	0.639159
6	-0.092750	2.129979	1.144435
8	-0.771683	1.580254	2.270882
6	0.205301	1.068750	0.072904
6	-0.999203	0.424476	-0.633155
6	-1.904403	-0.450769	0.234738
1	-2.001415	2.925004	0.495803
1	-0.993652	4.066836	1.403742
1	0.879494	2.539512	1.452608
1	0.826718	1.524841	-0.699210
1	-1.605622	1.226240	-1.069687
1	-2.298885	0.099250	1.090949
6	-1.113655	4.879260	-1.084229
1	-0.617273	5.169278	-2.013023
1	-2.166727	4.640790	-1.296618
1	-1.081285	5.721793	-0.378210
1	-5.524887	-1.307154	-1.560317
8	-1.127879	-1.566802	0.830375
1	-0.214959	-1.221520	1.020473
1	-1.107573	-2.528396	0.003372
1	-4.944380	-2.919070	-1.095454
1	-3.157129	-0.495694	-1.373155
1	-0.615148	-0.178497	-1.464520
8	-1.456119	-3.177326	-0.908523
1	-1.797129	-4.042210	-0.612080
1	-2.255641	-2.547623	-1.078611
8	0.972357	0.036136	0.755867
15	2.316581	-0.636848	0.039929
8	2.250846	-0.420383	-1.458368
8	3.593229	-0.171163	0.704205

8	1.924355	-2.156097	0.488784
11	4.432402	0.496161	-1.386769
6	2.967748	-3.140962	0.595030
1	3.774028	-2.776974	1.237129
1	2.513058	-4.029608	1.035918
1	3.364297	-3.391456	-0.394992

Figure A7.2 Second step (–1517.50864)

1	-0.710315	0.504848	2.465771
7	-3.885773	-0.189335	-0.652926
6	-5.237092	-0.095894	0.008835
6	-6.358473	-0.345907	-0.989217
1	-5.298261	0.903267	0.446517
1	-5.233812	-0.835743	0.810287
1	-7.315637	-0.296841	-0.462519
8	1.410789	2.855955	-0.431319
6	0.250937	2.749609	0.391970
6	0.166219	1.317696	0.937990
8	-0.813509	1.333639	1.970091
6	-0.138314	0.346962	-0.194435
6	-1.436123	-0.122206	-0.585707
6	-2.662951	0.108376	0.263353
1	-0.645695	3.012078	-0.189257
1	0.320157	3.436807	1.246116
1	1.147945	1.074026	1.354985
1	0.541425	0.477047	-1.031498
1	-1.603675	-0.005653	-1.658351
1	-2.828135	1.132091	0.607733
6	1.512015	4.138410	-1.051273
1	2.417027	4.131841	-1.662766
1	0.643857	4.330950	-1.695158
1	1.585467	4.934511	-0.298771
1	-6.371739	0.410649	-1.780646
8	-2.714775	-0.798864	1.326328
1	-3.163315	-0.391022	2.085767
1	0.166428	-2.060321	0.109862
1	-6.277734	-1.337259	-1.447445
1	-3.852725	0.449892	-1.454633
1	-1.245164	-1.579883	-0.538124
8	-0.784243	-2.645430	-0.337732
1	-1.274673	-2.997937	0.427703
1	-3.760040	-1.133929	-1.036303
8	0.848389	-1.131782	0.476004
15	2.462812	-1.036081	0.130371
8	2.637332	-0.564736	-1.305643
8	3.120144	-0.161908	1.180016
8	2.952280	-2.570067	0.344689
11	3.501805	1.475619	-0.465599
6	2.755100	-3.541262	-0.697170
1	3.187823	-3.191335	-1.638159
1	3.261232	-4.451687	-0.370797
1	1.688761	-3.750818	-0.839367

Reactant Complexes (electronic energy in a.u.)

Cationic Lysine Crosslink (Figure 7.3)

Figure 7.3 left C2′–H abstraction and elimination by water, cationic lysine crosslink (–1441.16292)

1	-1.26556	2.28612	1.66644
7	-3.41947	-1.43779	-0.50923
6	-4.85197	-1.12206	-0.69272
6	-5.73427	-2.20512	-0.07669
1	-5.01820	-1.03922	-1.77064
1	-5.02708	-0.14519	-0.23753
1	-6.78181	-1.96063	-0.27397
8	-1.23356	3.38119	-0.14717
6	-0.99442	2.11991	-0.74255
6	-0.28681	1.28797	0.32799
8	-1.03395	1.34569	1.54998
6	-0.07428	-0.18234	-0.07958
6	-1.17070	-1.17569	0.40199
6	-2.57097	-0.76097	0.18383
1	-0.35204	2.20878	-1.63053
1	-1.94790	1.66056	-1.05877
1	0.70957	1.71800	0.47537
1	-0.97201	-2.15389	-0.04576
6	-1.93520	4.28592	-0.98862
1	-2.06740	5.20984	-0.42196
1	-1.36564	4.49938	-1.90348
1	-2.92125	3.88721	-1.26858
1	-5.52738	-3.18619	-0.51609
1	-5.58797	-2.26450	1.00540
1	-1.03927	-1.25657	1.49434
1	0.00380	-0.23623	-1.17126
8	1.12823	-0.69493	0.51003
15	2.56289	-0.53517	-0.32060
8	2.63806	0.83089	-0.96601
8	3.65111	-0.89982	0.65973
8	2.39345	-1.63171	-1.51375
11	4.69255	1.17167	0.19493
6	2.52875	-3.02988	-1.20387
1	1.72969	-3.35866	-0.53007
1	2.45127	-3.56714	-2.15098
1	3.49872	-3.22942	-0.74099
1	-2.92923	0.14655	0.66361
8	0.12179	-0.63373	3.30467
1	0.83482	-0.81686	2.66861
1	-3.09581	-2.29727	-0.95615
1	-0.23627	0.21207	2.97387

Figure 7.3 middle C2′–H abstraction by Asp^{-}, cationic lysine crosslink (–1632.67588)

8	-3.02000	3.21421	-0.94574
6	-2.70007	1.88313	-0.55609
6	-1.25683	1.92794	-0.05611
8	-1.13797	2.96815	0.91953
6	-0.82850	0.61933	0.62213

8	-1.14284-0.44143-0.30156
6	0.67055 0.54034 0.94481
6	1.24699 1.51632 1.80111
6	2.81815 0.14450 -1.65447
8	1.85371 0.96844 -1.41784
15	-1.69575-1.893120.23219
8	-2.34082-2.57401-0.95905
8	-2.57739-1.723961.45541
8	-0.35374-2.683450.72197
6	0.60808 -3.13939-0.24978
1	-2.801351.19265 -1.40241
1	-3.372161.55209 0.25311
1	-0.592532.12755 -0.90948
1	-1.438120.48845 1.52350
1	0.96545 -0.484851.17521
1	1.28138 0.74129 -0.15633
1	0.77063 2.48836 1.90051
1	1.29508 -2.33191-0.52148
1	1.17044 -3.947270.22513
6	-4.370673.37103 -1.34050
1	-4.509034.42083 -1.61089
1	-4.608582.74154 -2.21076
1	-5.061753.11480 -0.52283
1	0.10241 -3.52242-1.14082
11	-4.28736-2.818510.30479
7	2.40065 1.38303 2.42218
1	-1.636153.71783 0.54749
6	3.24850 0.17645 2.40958
1	2.77695 -0.584423.04475
1	3.28013 -0.204001.38358
6	4.64871 0.50745 2.91365
1	5.25906 -0.400132.91745
1	5.13873 1.24144 2.26512
1	4.62784 0.90160 3.93680
6	3.44554 0.30047 -3.04930
1	3.73360 1.35348 -3.167122
1	2.64867 0.13340 -3.786492
6	4.63232 -0.62396-3.314599
1	5.02953 -0.47206-4.325296
1	5.44078 -0.44245-2.598599
1	4.33994 -1.67390-3.214442
8	3.24807 -0.72189-0.86896
1	2.74298 2.17735 2.950965

Figure 7.3 middle Protonation of 3′-PO_4 by Asp–H, cationic lysine crosslink (–1632.70145)

8	-3.76235-1.82096-1.87961
6	-3.22253-1.03720-0.82133
6	-1.70747-1.21136-0.89171
8	-1.24690-0.85643-2.19300
6	-0.97964-0.355250.15171
8	0.44580 -0.746920.10976
6	-1.100721.11359 -0.06493
6	-1.396721.96266 0.94098
15	1.02849 -1.861361.17606
8	0.05062 -3.005271.36241

8	2.41093	-2.23564	0.68104
8	1.07417	-1.03731	2.58199
6	2.02574	0.03585	2.71516
1	-3.61507	-1.37298	0.15147
1	-3.48896	0.02184	-0.95676
1	-1.46840	-2.26260	-0.67747
1	-1.36530	-0.62268	1.14131
1	1.57636	0.01637	-1.03716
1	-0.91534	1.48268	-1.07080
1	-1.58740	1.57140	1.94089
1	1.95196	0.73643	1.87815
1	1.78771	0.54684	3.65090
6	-5.15572	-1.63590	-2.05518
1	-5.46085	-2.26133	-2.89796
1	-5.71691	-1.94152	-1.15957
1	-5.39800	-0.58578	-2.27794
1	3.04323	-0.36464	2.76361
11	1.71656	-4.45177	0.55116
7	-1.43807	3.33755	0.87885
1	-1.90644	-1.23310	-2.80157
6	-2.25806	4.10569	1.81149
1	-2.09098	3.69112	2.81273
1	-1.87315	5.13167	1.82865
6	-3.75719	4.11264	1.48905
1	-4.30885	4.70714	2.22698
1	-3.94331	4.54545	0.49894
1	-4.16312	3.09471	1.49591
6	4.62618	3.12741	-1.99761
1	3.98844	3.96040	-1.68576
6	3.78888	1.92798	-2.43506
1	3.15870	2.17562	-3.29905
1	4.42596	1.09791	-2.76666
6	2.87940	1.39597	-1.34219
8	2.15842	0.35062	-1.76871
8	2.80977	1.86086	-0.21754
1	5.26168	3.47053	-2.82068
1	5.26990	2.87005	-1.15083
1	-1.38138	3.72746	-0.05590

Figure 7.3 right C2′–H abstraction by OG⁻, cationic lysine crosslink (–1982.12816)

8	1.83424	-4.25012	0.184582
6	1.30297	-3.04938	0.717357
6	1.90823	-1.87268	-0.044972
8	1.42340	-1.87215	-1.381144
6	1.67638	-0.53927	0.688355
8	2.38864	0.49972	-0.023758
6	0.22314	-0.07791	0.876149
6	-0.69931	-0.02179	-0.362495
7	-1.93572	0.70899	0.013471
6	-1.97482	2.08931	0.293122
7	-3.29708	2.36514	0.573035
6	-4.06430	1.20461	0.469195
6	-5.45009	0.96350	0.652612
8	-6.35338	1.74058	0.960625
7	-5.73044	-0.41833	0.423374

6	-4.81346-1.376030.077415
7	-5.25966-2.66263-0.038295
7	-3.52986-1.11678-0.091528
6	-3.20556 0.177870.119074
8	-1.02180 2.864920.277771
15	3.87266 0.98399 0.487013
8	4.51485 1.71018 -0.67827
8	4.65934 -0.175011.067336
8	3.52974 1.99598 1.723574
6	2.84924 3.22971 1.44204
1	1.54603 -2.973861.791827
1	0.20513 -3.041500.617149
1	2.99834 -2.01928-0.048843
1	2.10811 -0.642851.689985
1	-0.26189 -0.73181 1.610493
1	0.26515 0.92000 1.321556
1	-1.04674-1.02574-0.617716
1	-3.632663.28780 0.806801
1	-6.70061-0.676310.569481
1	-6.22820-2.81958-0.284503
1	-4.60488-3.29278-0.483378
1	3.41112 3.82314 0.713922
1	1.83953 3.04321 1.060089
6	1.29858 -5.400290.806368
1	1.76734 -6.270150.338292
1	1.51270 -5.416981.887297
1	0.20719 -5.466210.670602
1	2.78416 3.77372 2.386994
11	6.22727 0.12554 -0.646634
7	-0.039430.52103 -1.532725
1	0.39230 1.41440 -1.305046
1	1.15944 -0.94838-1.609044
6	-0.86898 0.63109-2.739082
1	-1.284496 -0.363132 -2.945926
1	-1.723977 1.310127 -2.591545
6	-0.031463 1.106375 -3.923604
1	-0.650941 1.182329 -4.823681
1	0.399027 2.095927 -3.728737
1	0.790853 0.411749 -4.124556

Figure 7.3 R1 First step of cationic lysine crosslink hydrolysis after β-elimination (–863.87127)

1	-0.65611-0.94341-0.96086
7	2.30117 1.52571 0.16185
6	2.90769 2.75837 0.67342
6	3.66745 3.49370 -0.43323
1	3.59237 2.46262 1.47527
1	2.13748 3.39436 1.11105
1	4.15691 4.37797 -0.01338
8	-4.147311.05375 0.52572
6	-3.408520.56205 -0.57540
6	-1.953870.31927 -0.13882
8	-1.28352-0.22764-1.24881
6	-1.359961.60252 0.40583
6	-0.074992.02683 0.45145
6	1.04351 1.22295 0.03788

1	-3.424	761.27701	-1.41433
1	-3.819	22-0.39128	-0.93832
1	-1.990	69-0.38945	0.70722
1	-2.105	952.26016	0.84923
1	0.11879	3.00170	0.88853
1	0.87558	0.22811	-0.37238
6	-5.493	521.34895	0.19525
1	-5.974	251.72072	1.10346
1	-5.556	422.11941	-0.58821
1	-6.027	870.45264	-0.15357
1	4.43849	2.85311	-0.87297
8	3.76721	-0.63977	-0.53272
1	3.26542	-1.37479	-0.04302
1	3.54208	-0.81280	-1.46089
8	2.40292	-2.56589	0.67635
8	0.75845	-1.75440	-0.61297
6	1.22062	-2.58024	0.23570
6	0.28057	-3.65965	0.79248
1	0.89287	-4.51235	1.10444
1	-0.16350	-3.25341	1.71383
6	-0.82945	-4.09718	-0.16599
1	-1.47703	-4.84781	0.30183
1	-1.45093	-3.24747	-0.46502
1	-0.41259	-4.53595	-1.07983
1	2.98747	3.81625	-1.22778
1	2.96451	0.75097	-0.12921

Figure 7.3 R2 δ-elimination by water, cationic lysine crosslink (–1325.42594)

8	1.83114	-1.10961	-0.29699
6	0.60850	-1.28196	-1.02912
6	-0.58877	-1.39565	-0.05124
8	-0.86393	-0.17231	0.59406
6	-1.77601	-1.92664	-0.78000
6	-3.01557	-1.40443	-0.95264
6	-3.44051	-0.13893	-0.43601
1	0.72220	-2.20619	-1.60190
1	0.46039	-0.44143	-1.71257
1	-1.59904	-2.90019	-1.23386
1	-3.74367	-1.97987	-1.52049
1	-2.749	320.45654	0.15111
7	-4.633	290.33794	-0.64530
1	-0.043	900.14210	1.03357
6	-5.151	051.61634	-0.12886
6	-6.300	221.39816	0.85427
1	-5.478	742.20323	-0.99260
1	-4.313	492.13726	0.33977
1	-6.672	182.37121	1.18841
1	-5.966	630.83337	1.72965
15	2.50700	0.37199	-0.05640
8	1.66455	0.75443	1.31053
8	2.17669	1.31730	-1.19130
8	3.97300	0.17687	0.24581
6	2.00380	1.97156	2.00198
1	1.81380	2.84162	1.36508
1	1.36881	2.01680	2.88784

1	3.05489	1.94864	2.30305
1	-0.30279	-2.17512	0.67979
8	1.56713	-3.21134	1.69090
1	1.96466	-4.02214	1.33797
1	1.93153	-2.50707	1.12428
11	4.54282	1.60170	-1.50915
1	-5.28731	-0.21859	-1.19501
1	-7.13208	0.86227	0.38527

Figure 7.3 R2 First step of cationic lysine crosslink hydrolysis after δ-elimination (–748.13965)

1	1.59036	1.22165	0.67508
7	-2.42851	-0.58553	-0.05968
6	-3.85472	-0.54790	-0.41090
6	-4.56283	-1.79854	0.10318
1	-3.94252	-0.47518	-1.50191
1	-4.29999	0.35302	0.02001
1	-5.61602	-1.77104	-0.19097
6	1.22100	3.72715	-0.22793
6	0.37913	2.74814	0.18659
8	0.72661	1.71750	0.97592
6	-1.01388	2.79824	-0.24999
6	-1.93244	1.78600	-0.26658
6	-1.60488	0.42380	0.02448
1	0.87034	4.53416	-0.86150
1	2.27457	3.70666	0.03094
1	-1.34357	3.76405	-0.62651
1	-2.93340	2.01797	-0.61747
1	-0.58071	0.15901	0.28538
1	-4.11520	-2.70547	-0.31628
8	-0.95004	-2.92482	0.19663
1	-0.08055	-2.44469	0.09668
1	-0.94135	-3.22828	1.11796
8	1.03555	-1.18034	0.12418
8	2.71349	0.30371	0.23963
6	2.25655	-0.86304	0.04039
6	3.24884	-1.97306	-0.33001
1	3.10983	-2.78426	0.39748
1	2.92215	-2.38911	-1.29277
6	4.71518	-1.54888	-0.39270
1	5.35628	-2.40020	-0.65065
1	4.86795	-0.76644	-1.14274
1	5.04902	-1.14782	0.56947
1	-4.51293	-1.85878	1.19515
1	-1.99128	-1.52647	0.08024

Cationic Proline Crosslink (Figure 7.4)

Figure 7.4 β-elimination by water, proline crosslink (–1518.58314)

1	-0.28833	2.93684	1.43710
7	-3.31985	-0.61029	0.24280
6	-4.66691	0.03541	0.12485
1	-4.62237	0.73133	-0.71887
1	-4.88280	0.58898	1.03954
8	-0.29312	3.64669	-0.55742

6	-0.375322.27392	-0.89220
6	0.31043	1.52931	0.25269
8	-0.240071.96475	1.50294
6	0.19354	-0.003610.13768
6	-0.96313	-0.642650.95958
6	-2.29871	-0.029700.76826
1	0.14203	2.05860	-1.83836
1	-1.431801.97184	-1.00497
1	1.37786	1.76888	0.20780
1	-0.97176	-1.713570.75166
6	-0.927014.50210	-1.49777
1	-0.809865.52314	-1.12909
1	-0.459324.41802	-2.48836
1	-1.998154.27017	-1.59001
1	-0.69650	-0.496892.01691
1	0.07666	-0.27080	-0.91921
8	1.36816	-0.643540.65743
15	2.68021	-0.86036	-0.34545
8	2.93778	0.39869	-1.14355
8	3.79796	-1.361210.53664
8	2.13766	-1.99465	-1.38210
11	5.14835	0.40796	-0.24368
6	2.04179	-3.35966	-0.93790
1	1.29795	-3.45765	-0.13934
1	1.72783	-3.94674	-1.80302
1	3.00994	-3.71647	-0.57712
1	-2.444050.98622	1.12783
8	0.97584	0.23361	3.46224
1	1.45845	-0.231892.75721
1	0.65487	1.03079	2.99937
6	-3.35723	-2.00304	-0.30538
1	-2.51153	-2.15479	-0.97700
1	-3.27376	-2.692020.54037
6	-4.72764	-2.09307	-0.98543
1	-5.10015	-3.11937	-0.991829
1	-4.65712	-1.74933	-2.022425
6	-5.60500	-1.13825	-0.159476
1	-6.50098	-0.81323	-0.692092
1	-5.91572	-1.611400.777687

Figure 7.4 R1 First step of hydrolysis after β-elimination, proline crosslink (–941.28578)

1	-0.978871.40500	0.66273
6	3.71573	-2.499060.56638
6	2.66626	-1.428770.86358
7	1.64184	-1.64544	-0.19900
6	2.18485	-2.47612	-1.30580
6	3.67652	-2.61137	-0.96797
1	2.18297	-1.489311.83970
1	1.99456	-1.99042	-2.26535
1	1.67248	-3.44417	-1.29320
1	4.08912	-3.55092	-1.34326
1	3.43327	-3.451801.02867
1	4.70056	-2.216440.94580
8	-4.56341	-0.67270	-0.58931
6	-3.69392	-0.325940.47002

6	-2.38768	0.25100	-0.11174
8	-1.58416	0.63762	0.96997
6	-1.78430	-0.72044	-1.09229
6	-0.56367	-1.30869	-1.13340
6	0.45016	-1.13139	-0.13269
1	-3.45718	-1.20258	1.09437
1	-4.13987	0.44184	1.11934
1	-2.67436	1.12888	-0.71748
1	-2.46775	-0.96366	-1.90375
1	-0.34250	-1.95909	-1.97505
1	0.23108	-0.56240	0.76224
6	-5.80362	-1.18408	-0.13291
1	-6.40227	-1.41369	-1.01778
1	-5.66836	-2.10193	0.45934
1	-6.33995	-0.44714	0.48355
1	3.04545	-0.40946	0.74060
1	4.23830	-1.78862	-1.42299
8	0.46677	0.10074	2.91228
1	0.93116	0.73435	2.30909
1	-0.45486	0.17619	2.60723
8	1.74168	1.61189	0.98343
8	-0.09312	2.52892	0.03573
6	1.15781	2.43989	0.22671
6	2.05996	3.44165	-0.51426
1	2.52222	4.07963	0.25218
1	2.88949	2.87405	-0.95530
6	1.35999	4.29813	-1.56847
1	2.06262	4.99962	-2.03447
1	0.92632	3.67629	-2.35936
1	0.54264	4.87330	-1.12357

Figure 7.4 R1 Second step of hydrolysis after β-elimination, proline crosslink (–749.29491)

1	0.99492	1.90436	0.18839
6	-3.91571	-1.59316	-0.78344
6	-2.48315	-1.14542	-1.15401
7	-2.15756	-0.12070	-0.14510
6	-2.73876	-0.60678	1.12042
6	-4.12173	-1.11343	0.68196
1	-1.78939	-2.00029	-1.09969
1	-2.77543	0.20299	1.85274
1	-2.14548	-1.42817	1.55415
1	-4.48785	-1.90805	1.33827
1	-4.01738	-2.67801	-0.87861
1	-4.65436	-1.13834	-1.45000
8	4.43984	-0.69951	0.30360
6	3.68859	0.26433	-0.40974
6	2.43776	0.62219	0.40490
8	1.84875	1.72310	-0.26311
6	1.54485	-0.58487	0.57967
6	0.23016	-0.72100	0.37011
6	-0.75748	0.32454	-0.11897
1	3.38957	-0.12050	-1.39843
1	4.27582	1.18254	-0.56810
1	2.79364	0.91509	1.41018
1	2.08980	-1.45634	0.94119

1 -0.19610-1.704550.55592
1 -0.517840.59492 -1.15826
6 5.60245 -1.10781-0.39153
1 6.11340 -1.842810.23580
1 5.35456 -1.57143-1.35913
1 6.28108 -0.26080-0.57712
1 -2.40353-0.72066-2.16057
1 -4.84869-0.295880.71735
8 -0.658691.48310 0.71262
1 -1.334032.12250 0.37192
8 -2.911322.56644 -0.43556
1 -2.755572.85745 -1.34823
1 -2.934731.57425 -0.48873

Figure 7.4 R2 δ-elimination by water, proline crosslink (–1402.84536)

8 2.40342 -1.07539-0.37312
6 1.13566 -1.21111-1.03563
6 0.00553 -1.382720.00942
8 -0.22947-0.194900.73298
6 -1.22582-1.88952-0.66268
6 -2.49144-1.40219-0.69381
6 -2.91323-0.20293-0.02971
1 1.21306 -2.10157-1.66524
1 0.94511 -0.33465-1.66084
1 -1.06040-2.81981-1.20374
1 -3.23585-1.96599-1.24691
1 -2.182460.35552 0.54526
7 -4.127990.25977 -0.06358
1 0.61835 0.10456 1.12712
6 -5.28048-0.35820-0.77887
1 -5.01370-0.53804-1.82243
1 -5.50312-1.31666-0.29909
6 -6.416950.65784 -0.59957
1 -7.391980.16647 -0.57832
1 -6.415411.37207 -1.42935
6 -4.548491.49467 0.65654
6 -6.071601.36806 0.72042
1 -4.239712.35788 0.05703
1 -4.050721.53734 1.62678
1 -6.553332.34332 0.81815
1 -6.363830.75375 1.57864
15 3.10328 0.39075 -0.11250
8 2.34051 0.72495 1.31246
8 2.71976 1.38304 -1.18952
8 4.58220 0.17596 0.09819
6 2.72101 1.91745 2.02437
1 2.50750 2.80791 1.42412
1 2.12918 1.93884 2.94052
1 3.78515 1.87942 2.27420
1 0.34286 -2.195130.68086
8 2.27164 -3.214891.60097
1 2.64543 -4.014431.19958
1 2.58823 -2.495721.02450
11 5.03572 1.64973 -1.67062

Figure 7.4 R2-P1 First step of hydrolysis after δ-elimination before enol-keto rearrangement, proline crosslink (–825.54604)

1	-1.66266	0.11570	-0.99251
6	4.15719	0.29348	-0.54503
6	2.72508	0.20123	-1.07695
7	1.91801	0.89311	-0.03323
6	2.70136	1.07687	1.21586
6	3.96646	0.24152	0.98024
1	2.57091	0.69675	-2.03854
1	2.11278	0.74842	2.07480
1	2.92873	2.14310	1.32798
1	4.81899	0.63895	1.53618
1	4.62125	1.24082	-0.84252
1	4.77708	-0.52329	-0.92187
6	-3.54291	2.58216	-0.50217
6	-2.37984	1.88891	-0.44387
8	-2.07502	0.97230	-1.38949
6	-1.44172	2.22247	0.62370
6	-0.12346	1.90023	0.77340
6	0.66754	1.21613	-0.20229
1	-3.77271	3.35780	0.21902
1	-4.25409	2.39717	-1.29992
1	-1.87536	2.82747	1.41620
1	0.37075	2.23142	1.68100
1	0.23984	0.98592	-1.16859
1	2.37016	-0.83840	-1.14152
1	3.79669	-0.78972	1.30635
8	1.38157	-2.73013	-1.38872
1	1.04728	-2.87954	-0.47161
1	0.65365	-2.15125	-1.68307
8	-0.21501	-2.69441	0.87503
8	-0.86759	-1.13383	-0.60032
6	-1.03844	-1.85297	0.45528
6	-2.35195	-1.65684	1.22447
1	-2.34265	-2.31780	2.09642
1	-2.38233	-0.62257	1.59536
6	-3.59145	-1.92155	0.35699
1	-4.51134	-1.77618	0.93475
1	-3.61923	-1.24721	-0.50428
1	-3.59254	-2.95061	-0.02081

Figure 7.4 R2-P2 Enol-keto rearrangement after δ-elimination, proline crosslink (–557.59199)

6	-3.79073	1.26491	0.04418
6	-2.53060	0.76519	0.02308
8	-2.22697	-0.55371	-0.01794
6	-1.41364	1.69073	0.04523
6	-0.06767	1.44867	0.02200
6	0.52227	0.15221	-0.03585
1	-3.95593	2.33450	0.07843
1	-4.65827	0.61412	0.02772
1	-1.71134	2.73446	0.08594
1	0.59211	2.30950	0.04954
1	-0.12582	-0.71656	-0.06406
1	-3.05669	-1.10378	-0.02669

8	-4.52244	-1.97039	-0.03651
1	-4.65906	-2.51268	0.75652
1	-4.63303	-2.58376	-0.78094
7	1.80811	-0.07211	-0.05744
6	2.39688	-1.43802	-0.11114
1	1.81508	-2.11452	0.51765
1	2.35335	-1.78802	-1.14846
6	2.87322	0.96617	-0.00943
1	2.69495	1.71680	-0.78270
1	2.83193	1.45321	0.97037
6	3.84043	-1.22111	0.34747
1	3.89404	-1.22525	1.44142
1	4.50567	-2.00070	-0.03008
6	4.17340	0.17423	-0.20707
1	4.41730	0.11300	-1.27277
1	5.01285	0.64939	0.30509

Figure 7.4 R2-P2 First step of hydrolysis after enol-keto rearrangement, proline crosslink (–786.22799)

6	2.09072	-2.89116	-0.73418
6	1.00774	-1.81199	-0.66702
7	1.75999	-0.56999	-0.32185
6	3.15836	-0.88146	0.07224
6	3.14551	-2.40207	0.27255
1	0.46640	-1.65151	-1.60095
1	3.42075	-0.32614	0.97400
1	3.82608	-0.58207	-0.74295
1	4.13365	-2.83817	0.10792
1	2.51774	-2.94197	-1.74208
1	1.69254	-3.87753	-0.48532
6	-0.14961	3.30855	-0.31321
8	-0.56254	4.45785	-0.34433
6	1.28175	3.03582	-0.03016
6	1.87202	1.82131	0.02018
6	1.16895	0.58304	-0.26158
1	-0.82400	2.45101	-0.48366
1	1.87727	3.92548	0.16160
1	2.93203	1.76926	0.25098
1	0.09988	0.57989	-0.52858
1	0.27260	-1.98099	0.12646
1	2.83319	-2.63692	1.29543
8	0.19437	0.29243	2.10709
1	-0.17891	1.11661	2.45553
1	-0.59243	-0.16683	1.68868
8	-1.70481	0.36487	-0.90712
8	-1.88003	-0.94283	0.93017
6	-2.31583	-0.44763	-0.15250
6	-3.72196	-0.87289	-0.61658
1	-4.31673	0.04452	-0.72381
1	-3.62194	-1.27052	-1.63571
6	-4.43533	-1.87535	0.28938
1	-5.43045	-2.12020	-0.10182
1	-3.86371	-2.80534	0.37516
1	-4.55383	-1.47424	1.30094

Figure 7.4 R2-P2 Second step of hydrolysis after enol-keto rearrangement, proline crosslink (–594.25799)

6	-3.10722	-1.13853	-0.15970
6	-1.92925	-0.76278	-1.08527
7	-1.04740	0.03438	-0.21496
6	-1.06019	-0.64675	1.09368
6	-2.54768	-0.98725	1.28538
1	-1.43118	-1.67583	-1.45103
1	-0.65315	0.01295	1.86342
1	-0.44955	-1.56441	1.07453
1	-2.67697	-1.89426	1.88253
1	-3.45433	-2.15498	-0.36552
1	-3.95554	-0.46627	-0.31885
6	3.33997	0.00832	0.30237
8	4.46946	-0.30691	0.65576
6	2.48182	-0.93420	-0.44187
6	1.22097	-0.75852	-0.87809
6	0.27560	0.41304	-0.75723
1	2.92653	1.00367	0.51920
1	2.97248	-1.88353	-0.65419
1	0.78242	-1.58752	-1.43250
1	0.06928	0.75950	-1.78710
1	-2.22928	-0.17544	-1.95971
1	-3.05814	-0.17346	1.80967
8	0.83402	1.43900	0.01831
1	0.15009	2.14777	0.08330
8	-1.57186	2.77675	0.16617
1	-1.84166	3.17707	-0.67567
1	-1.71638	1.80334	0.03920

Figure 7.4 Enol-keto rearrangement after δ-elimination (–420.96751)

6	-1.67878	1.67375	-0.008721
6	-0.63418	0.81532	-0.007934
8	-0.75871	-0.53659	-0.022358
6	0.73128	1.33447	0.00872
6	1.91940	0.68280	0.011129
6	2.16689	-0.76390	-0.00276
1	-1.51041	2.74377	0.006106
1	-2.70459	1.32095	-0.022609
1	0.77168	2.42126	0.020162
1	2.82931	1.28020	0.024811
1	1.28734	-1.42542	-0.017724
1	-1.71621	-0.79001	-0.021726
8	-3.39929	-1.27909	0.008598
1	-3.58158	-1.78681	0.815879
1	-3.57596	-1.90114	-0.715673
8	3.30460	-1.22400	0.002281

Neutral Lysine crosslink (Figures 7.5 and 7.6)

Figure 7.5 R1 C2′–H abstraction and elimination by Asp–H, neutral lysine crosslink (–1632.70369)

1	1.98461	-1.20866	-1.73337
7	3.37840	0.67730	1.44862
6	4.08176	-0.40397	2.13345
6	4.86540	0.13576	3.33297
1	4.78554	-0.84727	1.41559
1	3.39739	-1.20576	2.45790
1	5.44737	-0.66454	3.80369
8	2.67262	0.21598	-3.21307
6	2.21831	1.11220	-2.20640
6	0.99961	0.43870	-1.57760
8	1.36668	-0.84231	-1.07534
6	0.37106	1.26083	-0.45011
6	1.30560	1.67872	0.69811
6	2.11295	0.58526	1.34684
1	1.95337	2.08949	-2.63977
1	3.00594	1.26213	-1.45443
1	0.23360	0.32306	-2.36090
1	1.99578	2.44883	0.33709
6	3.91619	0.59750	-3.77930
1	4.17860	-0.16376	-4.51787
1	3.84432	1.57457	-4.27920
1	4.70651	0.65033	-3.01633
1	5.55438	0.92852	3.02166
1	4.18594	0.55263	4.08491
1	0.66467	2.14517	1.45750
1	-0.04329	2.16844	-0.90025
8	-0.73842	0.47685	0.06729
15	-2.23156	1.14395	0.22962
8	-2.55114	2.07109	-0.93007
8	-3.21003	0.00419	0.43003
8	-1.96800	1.98500	1.60506
11	-4.45451	0.77490	-1.38992
6	-3.01682	2.83343	2.09346
1	-3.89811	2.24201	2.36647
1	-2.62862	3.33259	2.98321
1	-3.29150	3.58229	1.34282
1	1.55613	-0.27225	1.74912
8	-0.44430	-1.96764	1.45157
1	-0.52775	-1.18971	0.85511
6	-0.87114	-3.09639	0.85130
8	-0.82407	-4.15503	1.44569
6	-1.39228	-2.92933	-0.56644
1	-0.57713	-2.49622	-1.16168
1	-2.17793	-2.16269	-0.53530
6	-1.90068	-4.23262	-1.17867
1	-2.72330	-4.65024	-0.58964
1	-2.26034	-4.06199	-2.19950
1	-1.10848	-4.98673	-1.21733

Figure 7.5 R1 C2′–H abstraction and elimination by water, neutral lysine crosslink (–1440.73927)

1	1.40530	1.79534	-1.17968
7	3.40620	-1.35159	0.16636
6	4.17038	-1.75200	-1.01107
6	5.19090	-2.83493	-0.65066
1	4.70345	-0.86380	-1.37678
1	3.51867	-2.10228	-1.82979
1	5.81004	-3.08517	-1.51936
8	1.81658	3.10834	0.43488
6	1.59606	1.90190	1.15810
6	0.49341	1.16767	0.40179
8	0.91672	0.98033	-0.95884
6	0.11365	-0.18741	1.01852
6	1.27669	-1.15061	1.30847
6	2.14302	-1.49625	0.12685
1	1.29054	2.11438	2.19425
1	2.51859	1.30443	1.17724
1	-0.41025	1.78861	0.41257
1	1.90939	-0.74547	2.10479
6	2.94023	3.84094	0.89845
1	3.03122	4.72422	0.26189
1	2.80441	4.16096	1.94160
1	3.86303	3.24678	0.82949
1	5.84786	-2.49418	0.15684
1	4.68648	-3.74795	-0.31489
1	0.81944	-2.07296	1.69152
1	-0.39188	0.02292	1.96884
8	-0.78947	-0.87745	0.13437
15	-2.39504	-0.57763	0.12144
8	-2.69687	0.86723	0.48564
8	-2.89924	-1.03327	-1.23719
8	-2.85283	-1.58835	1.31913
11	-3.19403	1.44500	-1.70193
6	-4.22922	-1.57415	1.72189
1	-4.88021	-1.87394	0.89283
1	-4.32671	-2.29609	2.53535
1	-4.52171	-0.58031	2.07703
1	1.62556	-1.90069	-0.75530
8	-1.26277	0.64241	-2.76840
1	-1.64129	-0.20832	-2.45238
1	-0.47298	0.75913	-2.19178

Figure 7.5 R1 C2′–H abstraction by OG, neutral lysine crosslink (–1981.60994)

8	-1.79388	-4.07204	-0.322649
6	-1.55016	-2.82244	-0.958651
6	-2.47238	-1.78945	-0.320816
8	-2.07095	-1.60005	1.050534
6	-2.53033	-0.47911	-1.134253
8	-3.69453	0.32984	-0.804823
6	-1.25598	0.36988	-1.10452
6	-1.34018	1.52958	-2.070299
7	2.65501	0.34291	-0.646316
6	2.45779	-0.97662	-0.993316
7	3.58579	-1.72110	-0.6186

6	4.49788	-0.86078	-0.019108
6	5.76965	-1.06717	0.539575
8	6.44900	-2.10232	0.664526
7	6.28826	0.17619	1.02430
6	5.64015	1.38230	0.949367
7	6.29235	2.46225	1.541557
7	4.45733	1.54977	0.418019
6	3.87251	0.40024	-0.063973
8	1.46400	-1.48513	-1.557003
15	-4.18350	0.67897	0.722266
8	-2.96574	0.85921	1.627286
8	-5.20862	-0.33295	1.192602
8	-4.96019	2.07896	0.464546
6	-4.20297	3.25014	0.11949
1	-1.77307	-2.90331	-2.036935
1	-0.49852	-2.52512	-0.850733
1	-3.48914	-2.20849	-0.325048
1	-2.72974	-0.77571	-2.170677
1	-1.06814	0.73698	-0.092791
1	-0.38543	-0.24997	-1.380096
1	-1.94354	1.34682	-2.969372
1	3.68015	-2.71742	-0.740268
1	7.17474	0.10545	1.509746
1	7.29401	2.49746	1.380638
1	5.84809	3.33498	1.277724
1	-3.64029	3.09356	-0.806258
1	-4.92609	4.05612	-0.023289
6	-0.91829	-5.08145	-0.785604
1	-1.15621	-5.99587	-0.234588
1	-1.04941	-5.27009	-1.863594
1	0.13535	-4.81726	-0.609475
1	-3.51194	3.51330	0.926109
11	-3.84470	-1.41839	2.656589
7	-0.78419	2.67311	-1.993292
1	-2.06612	-0.63002	1.243366
6	0.05503	2.98157	-0.825726
1	-0.51968	2.84252	0.103366
1	0.90390	6	2.28488	-0.777264
6	0.55848	8	4.42151	1	-0.904833
1	1.18264	4.66180	6	-0.036412
1	1.15907	7	4.57189	2	-1.808875
1	-0.27826	5	5.12989	-0.935565

Figure 7.5 R1-P1 First step of hydrolysis after β-elimination, neutral lysine crosslink (–671.87712)

1	2.15558	1.35081	0.09603
7	-2.45388	-0.37837	-0.67342
6	-3.39600	-1.47583	-0.49095
6	-4.56908	-1.02280	0.38475
1	-2.94711	-2.38266	-0.06687
1	-3.77374	-1.74023	-1.48758
1	-5.31563	-1.82056	0.46233
8	3.96084	0.26528	0.01434
6	3.06599	-0.82968	-0.09301
6	1.82519	-0.43540	0.72284
8	1.35139	0.81877	0.24459

6	0.79792	-1.52896	0.68854
6	-0.47406	-1.56340	0.24570
6	-1.21288	-0.44571	-0.34697
1	3.51878	-1.75311	0.29793
1	2.77981	-0.99568	-1.14293
1	2.15030	-0.32855	1.77268
1	1.18606	-2.44896	1.12607
1	-0.63172	0.45285	-0.52796
6	5.08516	0.16240	-0.84663
1	5.68639	1.06148	-0.69438
1	5.69114	-0.72258	-0.60553
1	4.77782	0.10453	-1.90082
1	-4.22682	-0.77150	1.39443
8	-0.84002	2.66953	1.17243
1	-0.10530	2.09848	0.88982
1	-2.81709	1.41687	-0.94353
1	-5.05260	-0.13785	-0.04267
1	-0.99848	-2.50662	0.36725
8	-2.80003	2.40993	-0.92103
1	-3.55548	2.63840	-0.35681
1	-1.45569	2.66174	0.40825

Figure 7.5 R1-P2 δ-elimination by water, neutral lysine crosslink (–1324.97355)

8	1.77825	-1.03072	-0.439011
6	0.46935	-1.19266	-1.03737
6	-0.58429	-1.55634	0.025555
8	-0.85749	-0.46541	0.887176
6	-1.82614	-2.04577	-0.670554
6	-3.02508	-1.45874	-0.842072
6	-3.47156	-0.14410	-0.377429
1	0.57526	-2.00192	-1.765646
1	0.18542	-0.27245	-1.554539
1	-1.69354	-3.03090	-1.11715
1	-3.78394	-2.00813	-1.397536
1	-2.75906	0.46146	0.190682
7	-4.65803	0.25522	-0.652414
1	-0.00981	-0.17025	1.275572
6	-5.04896	1.57084	-0.164703
6	-6.26966	1.46607	0.755447
1	-5.31427	2.18363	-1.03793
1	-4.22458	2.08607	0.358013
1	-6.60533	2.46259	1.063922
1	-6.02841	0.89061	1.65616
15	2.39094	0.44144	-0.062804
8	1.72969	0.57435	1.441573
8	1.86767	1.52303	-0.986995
8	3.89714	0.32405	0.013495
6	1.97841	1.78116	2.182408
1	1.56415	2.64735	1.65649
1	1.48523	1.66487	3.148747
1	3.05425	1.91498	2.334898
1	-0.16597	-2.40107	0.598433
8	2.16443	-3.44120	1.152966
1	2.45098	-4.08004	0.48261
1	2.20082	-2.58809	0.681542

11 4.08530 2.04308 -1.553641
1 -7.097280.96380 0.243422

Figure 7.5 R1-P2 First step of hydrolysis after δ-elimination, neutral lysine crosslink (–556.14155)

1 3.33290 2.25549 0.242671
7 -0.94843-0.27123-0.427175
6 -1.56925-1.58278-0.579583
6 -2.58105-1.825090.547215
1 -0.85099-2.41015-0.625217
1 -2.10149-1.56866-1.539909
1 -3.08374-2.786980.395585
6 4.79679 0.30867 0.321113
6 3.45897 0.34119 0.157447
8 2.73352 1.50050 0.122175
6 2.69478 -0.893460.004515
6 1.36751 -1.10439-0.171143
6 0.31430 -0.09443-0.248543
1 5.32757 -0.635080.34335
1 5.38371 1.21550 0.435494
1 3.32550 -1.778190.040774
1 0.63258 0.93979 -0.140782
1 -2.07636-1.853111.519709
8 -4.938980.87673 0.506481
1 -4.847790.91802 1.470359
1 -1.989771.20099 -0.408931
1 -3.34066-1.036250.56812
1 1.06561 -2.14332-0.258322
8 -2.567612.01517 -0.390831
1 -2.724232.21295 -1.327365
1 -4.121631.33014 0.185974

Figure 7.5 R1-P2 Second step of hydrolysis after δ-elimination, neutral lysine crosslink (–556.14148)

1 1.81695 0.03001 -2.216534
7 -1.139270.25789 0.007993
6 -1.555151.52518 0.627469
6 -2.244442.42805 -0.392885
1 -0.703032.05758 1.080567
1 -2.249461.28716 1.443323
1 -2.563813.36521 0.075818
6 3.68589 0.39857 -0.715684
6 2.40757 0.12934 -0.390093
8 1.42639 -0.05437-1.331134
6 1.97989 0.02317 1.006178
6 0.79914 -0.305501.56349
6 -0.53090-0.712990.941305
1 4.43463 0.53700 0.054615
1 4.01282 0.47472 -1.749137
1 2.78886 0.25317 1.696392
1 -1.25278-0.800391.769781
1 -1.566762.68145 -1.217395
8 -0.36479-1.955800.295229
1 -1.24853-2.18055-0.076642
1 -0.452660.43859 -0.725811
1 -3.126271.93558 -0.817254

1	0.78654 -0.299382.652193
8	-2.90526-1.71473-0.882035
1	-2.47063-0.83735-0.695516
1	-2.73375-1.86376-1.825514

Figure 7.5 R2 C2′–H abstraction and protonation of neutral crosslink by water (–1517.14795)

1	-0.949872.64016 1.012044
7	-3.02854-0.974950.982157
6	-3.64288-1.048162.305535
6	-4.613850.11643 2.514617
1	-2.88081-1.056583.100457
1	-4.19262-1.996222.365975
1	-5.111580.03000 3.486685
8	-0.858893.58093 -0.920793
6	-0.739882.22231 -1.323343
6	0.03106 1.52120 -0.205896
8	-0.645281.71563 1.033967
6	0.22252 0.02121 -0.456557
6	-1.05494-0.82736-0.427033
6	-1.76182-0.868500.900806
1	-0.205412.14408 -2.282894
1	-1.741621.78189 -1.445879
1	1.03650 1.96173 -0.161265
1	-1.75448-0.49254-1.199541
6	-1.682044.34965 -1.78152
1	-1.714555.36323 -1.374621
1	-1.271884.38444 -2.801491
1	-2.704473.94560 -1.827792
1	-4.081941.07372 2.483246
1	-5.382230.12570 1.733887
1	-0.78170-1.85589-0.697484
1	0.68128 -0.09226-1.446088
8	1.11207 -0.519460.544313
15	2.72153 -0.565660.289901
8	3.25110 0.74036 -0.287535
8	3.37416 -0.990251.590326
8	2.75457 -1.74455-0.8449
11	4.53248 1.02065 1.63616
6	3.98848 -1.99192-1.528434
1	4.76283 -2.32586-0.827582
1	3.79517 -2.78703-2.251813
1	4.33255 -1.09381-2.052106
1	-1.13094-0.852571.796223
8	-2.30685-3.48803-2.17673
1	-1.93901-3.05735-2.962415
1	-3.05653-2.90013-1.923663
8	-4.34522-1.73021-1.374565
1	-4.39756-0.93831-1.931171
1	-3.95651-1.40431-0.515788

Figure 7.6 R3 First step of hydrolysis, neutral lysine crosslink (–1517.14177)

1	0.31947 2.77959 -1.81986
7	3.42203 -0.38292-0.32711
6	4.20270 -0.246680.90553

6	5.69746	-0.34	1020.60765
1	3.90514	-1.05	1731.59408
1	3.98321	0.70190	1.41665
1	6.28256	-0.26	0681.53044
8	-0.58	5134.19277	-0.49041
6	-0.32	5903.16124	0.45215
6	-0.44	3451.84728	-0.32203
8	0.42478	1.88479	-1.45076
6	-0.12	1370.62512	0.54783
6	1.36405	0.39733	0.84840
6	2.18025	-0.08	770-0.32929
1	-1.05	0833.19736	1.28001
1	0.68549	3.28420	0.87163
1	-1.48	7241.73952	-0.64584
1	-0.64	9980.74920	1.50028
1	1.80393	1.31144	1.26789
1	1.64558	-0.24	281-1.26426
6	-0.36	6915.49390	0.02815
1	-0.57	8496.20066	-0.77785
1	-1.03	5435.70391	0.87590
1	0.67367	5.62514	0.36069
1	5.93873	-1.29	7410.13050
8	1.34701	-3.05	9980.31262
1	0.61850	-2.48	4590.02792
1	3.58207	-1.94	344-1.37396
1	6.01064	0.46129	-0.06938
1	1.43315	-0.36	6921.63317
8	-0.59	985-0.58	605-0.08797
8	3.39185	-2.85	020-1.72265
1	4.06924	-3.40	657-1.30687
1	1.96236	-3.05	881-0.45166
15	-2.14	125-1.10	0670.10556
8	-3.16	389-0.01	896-0.20789
8	-2.27	795-2.36	275-0.72051
8	-2.09	506-1.39	3371.71340
11	-3.97	406-1.36	262-1.94418
6	-3.33	005-1.64	4942.39475
1	-3.79	747-2.56	5432.025795
1	-3.08	730-1.76	7353.452368
1	-4.02	238-0.80	6292.269201

Figure 7.6 R3-P1 C2′–H abstraction by Asp⁻ after first step of hydrolysis, neutral lysine crosslink (–1708.63474)

1	3.19830	-0.24	3731.43632
7	1.86294	2.62931	-0.11007
6	2.10401	3.28539	1.17845
6	1.22441	4.49514	1.52970
1	3.15885	3.59552	1.19234
1	2.00110	2.51181	1.94881
1	1.51004	4.90616	2.50740
8	4.62102	-0.69	234-0.14310
6	3.47046	-0.30	597-0.89861
6	2.26165	-0.83	071-0.13141
8	2.25814	-0.28	6261.18982
6	0.93309	-0.54	715-0.85326

6	0.63100	0.87666	-1.34123
6	0.56050	1.99247	-0.27282
1	3.51680	-0.73508	-1.91267
1	3.41620	0.78797	-0.96585
1	2.34395	-1.92438	-0.06359
1	1.36123	1.16068	-2.11059
6	5.81584	-0.10769	-0.62492
1	6.62784	-0.44452	0.02557
1	6.03131	-0.42362	-1.65731
1	5.76777	0.99167	-0.60186
1	1.33900	5.29077	0.78256
1	0.16673	4.22148	1.55586
1	0.93319	-1.19583	-1.74031
8	-0.16402	-0.93528	0.00733
15	-0.72875	-2.46738	0.04319
8	0.39436	-3.49234	0.01338
8	-1.65096	-2.55146	1.24703
8	-1.53964	-2.62403	-1.36433
11	-0.27438	-4.17483	2.13325
6	-2.69301	-1.78881	-1.59696
1	-2.41315	-0.73322	-1.64501
1	-3.11330	-2.10207	-2.55558
1	-3.43763	-1.92990	-0.80790
1	0.29771	1.54310	0.69423
8	-2.63738	1.77593	-1.48227
1	-0.34501	0.81934	-1.83192
6	-3.58706	1.50741	-0.66745
8	-4.73062	1.11412	-0.98351
6	-3.24204	1.63415	0.83689
1	-2.55766	2.48076	0.97108
1	-2.65808	0.73862	1.09851
6	-4.45623	1.75049	1.75727
1	-5.14391	0.91719	1.58516
1	-4.16346	1.75165	2.81544
1	-5.01361	2.67508	1.56205
1	1.94976	3.31244	-0.86409
8	-0.39478	2.97569	2 -0.651683
1	-1.24985	2.51635	4 -0.91704

Figure 7.6 R3-P1 Second step of hydrolysis after β-elimination, neutral lysine crosslink (–671.88127)

1	1.35505	-0.08654	1.719271
7	-1.82834	0.12658	0.099302
6	-2.96055	-0.78603	0.318745
6	-3.61119	-0.52642	1.675368
1	-2.65165	-1.84192	0.252252
1	-3.69011	-0.61713	-0.483885
1	-4.46181	-1.19940	1.829198
8	3.24987	0.15872	0.827986
6	2.52272	-0.30131	-0.308185
6	1.48414	-1.29358	0.225211
8	0.71755	-0.65942	1.254373
6	0.61364	-1.84234	-0.880277
6	-0.46738	-1.32036	-1.477358
6	-1.18553	-0.00034	-1.224513
1	3.19344	-0.78968	-1.031378

1	2.00304	0.52963	-0.801339
1	2.03086	-2.13805	0.673259
1	0.95461	-2.80745	-1.253798
1	-2.00445	0.07148	-1.958409
6	4.12564	1.23336	0.530621
1	4.60832	1.52433	1.467028
1	4.89755	0.93142	-0.192446
1	3.58015	2.09589	0.120143
1	-2.89901	-0.69535	2.492043
8	-0.26933	1.06127	-1.403713
1	-0.76344	1.88411	-1.178379
1	-1.09386	-0.06208	0.787826
1	-3.97035	0.50630	1.745164
1	-0.91681	-1.92570	-2.263649
8	-2.04721	2.87905	-0.223302
1	-2.17359	1.94080	0.094812
1	-1.53555	3.29845	0.486754

Figure 7.6 R3-P1 C2′–H abstraction and elimination by water after first step of hydrolysis, neutral lysine crosslink (–1517.17383)

1	-0.52902	1.99383	1.683022
7	-3.33294	-1.04622	-1.350068
6	-4.64421	-1.22504	-0.717326
6	-4.74702	-2.31041	0.362809
1	-5.36403	-1.44669	-1.516763
1	-4.93738	-0.25279	-0.301091
1	-5.78020	-2.39254	0.723413
8	-0.38049	3.72629	0.60567
6	-0.45499	2.82837	-0.49433
6	0.26178	1.56070	-0.039447
8	-0.29958	1.15440	1.23713
6	0.20439	0.41333	-1.063782
6	-1.17072	-0.00902	-1.616212
6	-2.33493	-0.30541	-0.642941
1	0.03531	3.24976	-1.384158
1	-1.50802	2.62845	-0.7417
1	1.31623	1.79634	0.134887
1	-1.52248	0.75385	-2.320855
6	-1.12348	4.92178	0.409511
1	-1.01315	5.51835	1.317506
1	-0.73676	5.49117	-0.446893
1	-2.18857	4.70688	0.241067
1	-4.45821	-3.28829	-0.042532
1	-4.09190	-2.08899	1.208466
1	-0.97573	-0.90841	-2.211716
1	0.79713	0.76102	-1.920611
8	0.81188	-0.77835	-0.551893
15	2.39603	-0.97912	-0.157822
8	2.77515	-0.01216	0.953686
8	2.52725	-2.45973	0.096673
8	3.08149	-0.50670	-1.566142
11	1.13842	-0.35570	2.484779
6	4.46949	-0.14612	-1.560896
1	5.09690	-1.00889	-1.310113
1	4.71150	0.18978	-2.571663

1	4.65892	0.66141	-0.846958
1	-2.80059	0.64160	-0.33351
8	0.28274	-2.44830	1.83669
1	-0.41821	-2.18213	1.205021
1	1.04474	-2.67652	1.250012
1	-2.95509	-1.93781	-1.666518
8	-1.90389	-0.97280	0.568659
1	-1.43213	-0.27667	1.068827

Figure 7.6 R3-P2 C2′–H abstraction and protonation of neutral crosslink by water after fist step of hydrolysis (–1517.14848)

1	1.75328	-2.42412	-1.673884
7	3.18601	0.56505	0.550722
6	2.46352	0.21075	1.788916
6	2.99031	-1.02378	2.531111
1	1.40845	0.09862	1.538985
1	2.53575	1.08292	2.450476
1	2.45910	-1.14593	3.482767
8	-2.11357	-2.20301	0.251212
6	-1.34344	-2.34457	-0.948513
6	0.11201	-1.83967	-0.860724
8	0.80641	-2.53731	-1.890531
6	0.24683	-0.31034	-1.057064
6	1.64471	0.21012	-1.445908
6	2.94320	-0.14497	-0.691863
1	-1.88921	-1.79232	-1.717187
1	-1.30672	-3.40487	-1.232041
1	0.52440	-2.09657	0.124426
1	1.80635	-0.12160	-2.47781
1	3.75529	0.18742	-1.352577
6	-1.83288	-3.19120	1.237912
1	-2.50469	-3.00495	2.079279
1	-2.01642	-4.20003	0.842663
1	-0.79798	-3.12959	1.59434
1	2.85422	-1.93483	1.941696
8	3.01488	-1.58785	-0.594579
1	3.91877	-1.83805	-0.344482
1	4.05903	-0.91998	2.762533
1	4.18773	0.57281	0.737892
1	1.58701	1.30295	-1.466926
1	-0.40588	-0.02759	-1.889819
8	-0.23311	0.34038	0.134429
15	-1.61305	1.22140	0.123317
8	-2.50964	0.79263	-1.028942
8	-2.23613	1.11312	1.501337
8	-0.96039	2.69468	-0.145945
11	-3.70636	-0.48991	0.535521
6	-1.84738	3.81362	-0.251727
1	-2.41892	3.94817	0.673942
1	-1.22451	4.69420	-0.422439
1	-2.53895	3.68710	-1.092647
8	2.76682	3.32474	-0.356441
1	2.86125	2.42843	0.041442
1	3.51300	3.37158	-0.973344

Figure 7.6 R3-P2 3′-PO_4 elimination (-1305.52670)

1	-2.78396	-0.37227	-1.737628
8	-3.85940	-1.13694	-0.00727
6	-2.95365	-0.25941	0.6507
6	-1.67169	-0.27960	-0.179027
8	-1.96060	0.09783	-1.518407
6	-0.57857	0.64210	0.389307
6	-0.91458	2.10994	0.380494
6	-0.37094	3.06648	-0.385705
1	-2.76603	-0.59494	1.682976
1	-3.37519	0.75681	0.683877
1	-1.26187	-1.30026	-0.147677
1	-1.68205	2.42410	1.082231
1	-0.68889	4.10209	-0.294158
6	-5.16142	-1.11592	0.551699
1	-5.77434	-1.80577	-0.033723
1	-5.15210	-1.44535	1.601274
1	-5.60426	-0.10968	0.50442
8	0.60715	2.93469	-1.320086
1	0.85307	1.98606	-1.335067
1	-0.39829	0.31958	1.423618
8	0.63390	0.42126	-0.383352
15	1.80850	-0.59153	0.184141
8	1.21069	-1.87654	0.716449
8	2.81708	-0.72704	-0.936315
8	2.40713	0.23020	1.460376
6	3.15434	1.43323	1.21316
1	2.55702	2.15800	0.649827
1	3.40549	1.85134	2.190057
1	4.07080	1.20884	0.659198
11	2.50951	-3.03947	-0.864205

Figure 7.6 R3-P3 Second step of hydrolysis, neutral lysine crosslink (–1517.15476)

1	-1.23510	2.37641	1.50116
7	-3.14858	-1.29973	-0.22422
6	-3.87334	-0.11542	0.25837
6	-5.38252	-0.33881	0.19055
1	-3.60748	0.79599	-0.30103
1	-3.56018	0.06138	1.29335
1	-5.91838	0.53826	0.56976
8	-0.65564	3.98826	0.22261
6	-0.40858	2.91423	-0.67635
6	0.00922	1.72525	0.18944
8	-0.98184	1.49079	1.18605
6	0.25726	0.44904	-0.62739
6	-1.01090	-0.26746	-1.12531
6	-1.67290	-1.24287	-0.13027
1	0.38792	3.17738	-1.38905
1	-1.32318	2.68811	-1.24739
1	0.96919	1.98227	0.66022
1	0.85127	0.75013	-1.49854
1	-1.73442	0.49666	-1.42950
1	-1.44753	-0.94046	0.89837

6 -1.191175.13719 -0.41094
1 -1.352165.88806 0.36654
1 -0.496515.54047 -1.16235
1 -2.149944.91541 -0.90347
1 -5.71230-0.50983-0.84174
8 -1.16897-2.53217-0.41766
1 -1.60514-3.148740.21259
1 -3.40259-1.49326-1.19496
1 -5.67758-1.209160.78688
1 -0.76527-0.84963-2.02035
8 1.02935 -0.465620.17781
8 -3.12811-3.713981.20022
1 -3.41638-2.866660.76420
1 -2.98983-3.467672.12848
15 2.64899 -0.57365-0.03774
8 3.20356 -1.325901.15647
8 2.78806 -1.43380-1.41965
8 3.26403 0.78930 -0.30281
6 2.32876 -2.79871-1.41914
1 1.27921 -2.85448-1.11364
1 2.43625 -3.16309-2.443609
1 2.94265 -3.40803-0.747819
11 4.60856 0.47759 1.575412

Figure 7.6 R3-P3 C2′–H abstraction by Lys after hydrolysis, AP-site *pro*-S (–1440.71938)

6 -3.08614-0.82922-0.446742
8 -4.06431-0.27946-1.32357
6 -1.901890.13618 -0.490088
8 -2.344281.44588 -0.144312
6 -0.72770-0.284700.419235
8 0.48393 0.00962 -0.322712
6 -0.662230.44420 1.764624
6 -1.700020.06906 2.783735
8 -2.48604-0.856272.699297
1 -2.77895-1.83173-0.786168
1 -3.48299-0.914310.571943
1 -1.513950.13446 -1.518916
1 -3.182371.56193 -0.62645
1 -0.76720-1.365680.584143
1 0.31068 0.23139 2.232672
1 -0.679701.52983 1.621638
1 -1.685640.70275 3.696724
15 1.82928 -0.89247-0.155792
6 3.86037 -0.423821.461443
8 2.54231 -0.036771.05251
8 2.63889 -0.80243-1.435523
8 1.50213 -2.308530.292737
1 4.56682 -0.328500.629594
1 3.86597 -1.455141.832528
6 1.59893 3.24501 -1.930606
6 0.76223 3.54997 -0.681744
7 1.40164 3.22735 0.603396
1 2.54309 3.80583 -1.917466
1 1.06110 3.51806 -2.84827
1 0.50029 4.61671 -0.665236

1 -0.182812.99663 -0.724736
1 2.29368 3.72099 0.660074
1 1.64050 2.23501 0.611383
1 1.83997 2.17664 -1.982952
1 4.15344 0.25136 2.268014
6 -5.30413-0.96312-1.27346
1 -5.98360-0.44980-1.958575
1 -5.19723-2.01085-1.592257
1 -5.73203-0.94853-0.260056
11 2.54024 -3.12900-1.627303

Figure 7.6 R3-P3 C2′–H abstraction by Lys after hydrolysis, AP-site *pro*-R (–1440.71664)

6 -3.01240-1.091915-0.283214
8 -3.99556-0.869034-1.274681
6 -1.89945-0.051901-0.459939
8 -2.376471.274645-0.318316
6 -0.67098-0.3787710.436123
8 0.49858 -0.193115-0.393544
6 -0.503500.4980021.701032
6 -1.644020.4944192.685891
8 -2.799890.7879252.416868
1 -2.58618-2.105633-0.387107
1 -3.45682-1.0227590.722869
1 -1.54854-0.12689-1.493459
1 -2.756371.3354730.581861
1 -0.69996-1.4313590.736469
1 0.40307 0.1715312.217053
1 -0.337971.5385151.382424
1 -1.388770.2318423.731167
15 1.89456 -0.985036-0.102478
6 4.09857 0.0549510.896905
8 2.70796 0.24855 0.61016
8 2.54499 -1.369075-1.418457
8 1.68390 -2.1447510.859927
1 4.65814 -0.152074-0.021528
1 4.23927 -0.7687431.606612
6 1.22105 3.811106-2.13273
6 0.38298 3.706341-0.853183
7 1.10689 3.2443270.341675
1 2.05970 4.507085-1.998455
1 0.61800 4.171067-2.976284
1 -0.053064.686512-0.618343
1 -0.458653.022512-1.009292
1 1.88806 3.8757640.524909
1 1.52921 2.3341740.152081
1 1.63558 2.833543-2.408414
1 4.46209 0.9818541.345299
6 -5.07925-1.770862-1.173817
1 -5.78417-1.519889-1.970808
1 -4.75387-2.816266-1.300546
1 -5.59146-1.684557-0.20254
11 2.57182 -3.59144-0.747871

Figure 7.6 R3-P3 Protonation of 3′-PO_4 by Lys, AP-site *pro*-S (–1440.68316)

6 -3.09366-0.570580.452035
8 -3.92462-1.22341-0.505219
6 -1.68562-0.57129-0.139141
8 -1.701310.05557 -1.423608
6 -0.668360.08618 0.801084
8 0.69788 0.00281 0.16642
6 -0.855321.50714 1.198028
6 -1.480502.51699 0.480079
8 -1.318123.77646 0.676971
1 -3.11832-1.105191.41467
1 -3.440630.45812 0.619262
1 -1.37196-1.61784-0.253491
1 -2.51627-0.27425-1.842351
1 -0.60892-0.576101.6716
1 -0.319781.80711 2.102052
1 0.74691 2.31131 0.29943
1 -2.123972.21143 -0.36282
15 1.61553 -1.345190.149547
6 2.31929 -1.054502.692465
8 2.73137 -1.068801.32281
8 2.38802 -1.39214-1.154384
8 0.79583 -2.587780.473583
1 3.22750 -1.013673.298977
1 1.75309 -1.957612.943573
6 2.75644 2.29240 -1.69034
6 1.65715 3.29413 -1.361393
7 1.24372 3.21621 0.072994
1 3.66446 2.49384 -1.10969
1 3.01638 2.37698 -2.750582
1 1.97047 4.32404 -1.55263
1 0.75464 3.09836 -1.945276
1 0.43647 3.85860 0.31849
1 2.02178 3.34628 0.721355
1 2.43166 1.26474 -1.493662
1 1.70692 -0.169882.908298
6 -5.30252-1.14867-0.185378
1 -5.84850-1.65701-0.984403
1 -5.51999-1.647080.771297
1 -5.64543-0.10505-0.120533
11 1.51007 -3.48866-1.576814

Figure 7.6 R3-P3 Protonation of 3′-PO_4 by Lys, AP-site *pro*-R (–1440.71718)

6 -2.89657-1.14469-0.21512
8 -3.81775-1.07678-1.28593
6 -1.84469-0.04646-0.41151
8 -2.414431.25081 -0.39780
6 -0.64978-0.220270.57123
8 0.56075 0.01544 -0.18320
6 -0.633270.75683 1.77327
6 -1.825780.69368 2.69373
8 -2.973440.94025 2.35301
1 -2.40040-2.13122-0.19819

1	-3.41671	-1.02069	0.74846
1	-1.43104	-0.17287	-1.41658
1	-2.84444	1.35678	0.47548
1	-0.62027	-1.25292	0.93339
1	0.27631	0.56530	2.34784
1	-0.56523	1.77741	1.37339
1	-1.62556	0.42335	3.74864
15	1.90479	-0.88337	0.04690
6	3.90885	-0.35133	1.67023
8	2.59431	0.01490	1.23167
8	2.72960	-0.84382	-1.22479
8	1.55993	-2.28028	0.54059
1	4.62177	-0.29858	0.84026
1	3.91144	-1.36252	2.09351
6	1.25318	3.19928	-2.51372
6	0.41323	3.52918	-1.27369
7	1.07955	3.30383	0.01884
1	2.16218	3.81428	-2.54916
1	0.68956	3.38118	-3.43809
1	0.09955	4.58108	-1.31038
1	-0.50366	2.92940	-1.27236
1	1.94985	3.83708	0.04498
1	1.35573	2.32323	0.07281
1	1.56024	2.14628	-2.50489
1	4.19608	0.36455	2.44312
6	-4.85699	-2.02775	-1.16609
1	-5.51523	-1.89769	-2.02916
1	-4.46912	-3.05939	-1.16441
1	-5.44108	-1.87767	-0.24452
11	2.61459	-3.16341	-1.35163

Figure 7.6 R3-P3 C2′–H abstraction by Asp⁻ after hydrolysis, AP-site (–1573.46779)

8	-4.23130	-1.98547	-0.660942
6	-2.82897	-1.87540	-0.817398
6	-2.35106	-0.61723	-0.097889
8	-2.50035	-0.73740	1.32355
6	-0.91959	-0.20792	-0.514259
8	-0.51039	0.91871	0.279851
6	0.19304	-1.26106	-0.421996
6	0.55430	-1.85524	0.911149
6	3.95500	-1.01032	-0.36157
8	3.15924	-0.38148	-1.119297
15	-0.87888	2.49574	-0.185791
8	-1.60836	3.06876	1.016171
8	-1.48413	2.47642	-1.557935
8	0.60602	3.17257	-0.155768
6	1.57983	2.77797	-1.137908
1	-2.56659	-1.81545	-1.888162
1	-2.32598	-2.76485	-0.401786
1	-3.01117	0.20539	-0.392392
1	-0.98020	0.10981	-1.560368
1	-0.03277	-2.11523	-1.082612
1	1.13714	-0.84414	-0.815523
1	1.59397	-2.22824	0.952086
1	2.37917	3.52299	-1.100504

1	1.99375 1.78937 -0.911051
6	-4.75581-3.14004-1.286087
1	-5.83586-3.13675-1.117037
1	-4.56307-3.13870-2.370735
1	-4.32945-4.06230-0.860424
1	1.13790 2.77360 -2.139892
11	-1.820751.25783 2.310777
1	-1.75173-1.310061.632614
6	5.44398 -0.57651-0.465192
1	5.48827 0.49804 -0.237174
1	5.74027 -0.65858-1.520686
6	6.41546 -1.350360.425635
1	7.44688 -0.994070.300373
1	6.14469 -1.247711.481883
1	6.39357 -2.420410.192563
8	3.67438 -1.933760.449723
8	-0.17568-1.940631.899677

Figure 7.6 R3-P3 Protonation of 3′-PO_4 by Asp–H, AP-site (–1573.44252)

6	3.13616 -1.122550.360707
8	2.01449 -0.875381.038596
1	1.35508 -0.384870.469647
1	-1.27161-1.908001.029542
8	-4.43245-0.906330.465393
6	-3.32710-1.14171-0.405032
6	-2.10204-0.408680.150599
8	-1.61474-1.009281.332861
6	-1.02870-0.26343-0.949107
8	0.12079 0.53274 -0.328186
6	-0.48890-1.49765-1.563093
6	-0.40189-2.75535-0.99906
15	0.09592 2.15781 -0.191918
8	1.04018 2.52583 0.943034
8	-1.315452.70621 -0.062432
8	0.66706 2.67824 -1.636872
6	2.02212 2.35567 -1.984791
1	-3.57363-0.78215-1.419456
1	-3.11101-2.21696-0.463873
1	-2.415100.60538 0.416967
1	-1.425010.40512 -1.72198
1	-0.02524-1.37648-2.540842
1	0.08052 -3.50573-1.664887
1	2.72199 2.97650 -1.41419
1	2.24180 1.29997 -1.79713
1	2.13391 2.57043 -3.051031
8	3.34545 -0.76923-0.789116
6	4.13231 -1.905481.200674
1	3.65072 -2.850101.484459
1	4.27936 -1.357772.140313
6	5.46078 -2.155190.490976
1	6.13802 -2.728691.13315
1	5.31003 -2.71491-0.437238
1	5.95157 -1.211800.230528
8	-0.80951-3.197300.136099
6	-5.58039-1.622470.072528

1 -6.37398-1.394550.791269
1 -5.92249-1.33267-0.935822
1 -5.40493-2.710860.07046
11 -0.803593.26072 2.137952

Figure 7.6 R3-P3 C2′–H abstraction and elimination by water after hydrolysis, AP-site (–1381.98924)

1 1.99018 -1.745301.179577
8 3.25873 -1.55872-0.493978
6 2.73536 -0.23442-0.463883
6 1.29071 -0.385400.000533
8 1.28422 -1.077171.26389
6 0.51544 0.93241 0.150599
6 1.11116 1.90848 1.170438
6 2.13709 2.86769 0.62528
1 2.78534 0.23695 -1.454894
1 3.32095 0.37761 0.237382
1 0.75463 -0.98487-0.742514
1 0.30199 2.51968 1.594093
6 4.64895 -1.60286-0.774332
1 4.94656 -2.65394-0.757198
1 4.86851 -1.18135-1.76562
1 5.22852 -1.04941-0.020697
1 1.54400 1.37259 2.026254
1 0.47268 1.41349 -0.832927
8 -0.817200.64505 0.619108
15 -2.009960.07739 -0.362691
8 -1.44416-0.76929-1.483028
8 -2.99616-0.588300.581092
8 -2.630821.42067 -1.043725
11 -2.20984-2.73587-0.50238
6 -3.344402.35564 -0.220908
1 -2.682012.79490 0.53328
1 -3.709563.14177 -0.884981
1 -4.188461.86870 0.276269
1 2.53436 3.58754 1.372173
8 -1.19063-2.484041.597635
1 -1.83257-1.740471.635728
1 -0.32208-2.023681.538036
8 2.52086 2.90558 -0.527714

Figure 7.6 R3-P3 C4′–H abstraction and elimination by water after β-elimination, AP-site (–1103.95322)

8 1.18921 -0.95507-0.516523
6 -0.16137-0.92695-0.990324
6 -1.15107-0.931870.195953
8 -1.105990.27349 0.930248
6 -2.52892-1.23686-0.311768
6 -3.65148-0.49146-0.345199
6 -3.838920.89025 0.125162
1 -0.30607-1.82368-1.603418
1 -0.31840-0.03853-1.608788
1 -2.60209-2.24296-0.726409
1 -4.55315-0.93191-0.769211
1 -2.961781.39205 0.559054

1	-0.16087	0.51633	1.057824
15	2.13659	0.42355	-0.536707
8	1.58404	1.00397	0.945131
8	1.66971	1.34905	-1.622269
8	3.55318	-0.03699	-0.391316
6	2.01289	2.31986	1.303736
1	1.57258	2.55373	2.276548
1	3.10565	2.35496	1.384385
1	1.68224	3.05527	0.561841
1	-0.85274	-1.78541	0.832269
8	1.31717	-3.20457	1.235576
1	1.54934	-3.87524	0.575415
1	1.43309	-2.36712	0.742919
8	-4.92275	1.45782	0.032006

Figure A7.2 First step (–1517.59787)

1	-0.26344	0.66082	2.515382
7	-3.20521	-1.34610	-0.404281
6	-4.30681	-1.95604	0.360049
6	-5.66524	-1.43294	-0.100193
1	-4.12576	-1.73575	1.414174
1	-4.23400	-3.04006	0.226873
1	-6.45591	-1.93644	0.463946
8	-1.10579	3.31743	-0.767235
6	-1.51937	2.79179	0.477595
6	-0.45914	1.80106	0.953674
8	-0.98345	1.18669	2.128994
6	-0.06110	0.76342	-0.112164
6	-1.20816	-0.00609	-0.78277
6	-2.22392	-0.66762	0.104845
1	-2.49744	2.28911	0.38977
1	-1.63169	3.58533	1.231358
1	0.45461	2.36068	1.199828
1	0.47435	1.28124	-0.909186
1	-1.75247	0.71898	-1.404258
1	-2.29127	-0.39312	1.15011
6	-1.97897	4.32409	-1.257272
1	-1.57425	4.65999	-2.21447
1	-2.99589	3.93365	-1.411268
1	-2.03150	5.17770	-0.566159
1	-5.75087	-0.35537	0.069894
8	-0.86951	-2.31639	1.029384
1	-0.09050	-1.76432	1.227532
1	-0.59619	-2.76898	0.193658
1	-5.83154	-1.63304	-1.164171
1	-3.17855	-1.55937	-1.400079
1	-0.79118	-0.74899	-1.471968
8	0.11979	-3.02427	-1.42617
1	1.00650	-2.66736	-1.202962
1	0.27479	-3.93974	-1.707254
8	0.85624	-0.15705	0.529432
15	2.43873	-0.24851	0.064635
8	2.66817	0.60566	-1.164875
8	3.36096	0.01910	1.231585
8	2.42531	-1.85855	-0.28706

11	4.41047	1.71963	-0.054672
6	3.67588	-2.57800	-0.249225
1	4.18395	-2.39763	0.700386
1	3.42855	-3.63642	-0.345678
1	4.31746	-2.27537	-1.083309

Figure A7.2 Second step (–1517.62591)

1	-2.29440	-1.18772	2.351079
7	0.51984	-2.41053	-0.034453
6	0.81754	-3.85025	0.276177
6	1.87102	-3.97018	1.368824
1	1.15269	-4.30334	-0.660478
1	-0.12526	-4.31531	0.564897
1	2.06652	-5.02896	1.562579
8	-3.61976	1.44808	-0.296788
6	-3.45849	0.18660	0.312664
6	-2.03464	0.11629	0.874633
8	-1.81474	-1.15853	1.507183
6	-0.95337	0.32274	-0.187697
6	-0.84333	-0.71790	-1.320104
6	-0.65028	-2.21251	-1.00511
1	-3.62864	-0.63155	-0.40573
1	-4.17521	0.05030	1.139986
1	-1.90618	0.91051	1.618147
1	-1.18053	1.28089	-0.662107
1	-1.74517	-0.66037	-1.939379
1	-0.34058	-2.71947	-1.922327
6	-4.91936	1.64170	-0.836234
1	-4.93246	2.63773	-1.283751
1	-5.14749	0.89381	-1.609481
1	-5.68941	1.58624	-0.05308
1	2.81686	-3.50209	1.076865
8	-1.76836	-2.88612	-0.536732
1	-2.07019	-2.41934	0.275574
1	3.51596	-1.37491	-0.794779
1	1.53005	-3.51565	2.305389
1	1.38057	-1.96486	-0.453463
1	-0.01525	-0.39138	-1.957211
8	2.71326	-1.35032	-1.339637
1	2.47784	-0.38317	-1.415313
1	0.29944	-1.89524	0.827475
8	0.30833	0.43283	0.524835
15	1.42430	1.55709	0.077522
8	2.46087	1.63850	1.173165
8	1.97528	1.23698	-1.311531
8	0.56261	2.92270	-0.080144
11	4.19124	1.76934	-0.382387
6	0.18805	3.66945	1.094082
1	1.06880	3.89516	1.700545
1	-0.27019	4.59286	0.737667
1	-0.53743	3.11198	1.695282