MOLECULAR STANDARDS FOR ANALYTICAL ULTRACENTRIFUGATION: INVESTIGATING THE SUITABILITY OF DOUBLE-STRANDED DNA

MADUNI CHARUNI RANASINGHE
Master of Science in Physics Education, University of Colombo, Sri Lanka, 2021

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

MASTER OF SCIENCE

in

BIOCHEMISTRY

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

MOLECULAR STANDARDS FOR ANALYTICAL ULTRACENTRIFUGATION: INVESTIGATING THE SUITABILITY OF DOUBLE-STRANDED DNA

MADUNI CHARUNI RANASINGHE

Date of Defence: 13th September 2023

Dr. B. Demeler Thesis Supervisor	Professor	Ph.D.
Dr. T. Patel Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. S. Wetmore Thesis Examination Committee Member	Professor	Ph.D.
Dr. JD. Hamel Chair, Thesis Examination Committee	Assistant Professor	Ph.D.

DEDICATION

~~ In loving memory of my mother ~~

ABSTRACT

This thesis focuses on the development of molecular standards to validate analytical ultracentrifugation (AUC) instruments, alongside an exploration of multiple applications of AUC. The goal of the main study is to address the critical need to develop a robust molecular standard for validating AUC instruments, with a specific emphasis on evaluating double-stranded DNA molecules as a potential candidate. By extensive investigation into the hydrodynamic properties of different topologies of double stranded DNA at a wide range of temperatures, this study reveals the potential of linear and nicked double stranded DNA as a reliable standard for AUC, contributing to the accurate characterization of macromolecules in solution.

Supplementary research findings, as detailed in the attached publications presented in the attached appendix, further illustrate the versatility of AUC in various scientific domains. In one study, we employed AUC as a powerful technique to measure the sedimentation and diffusion coefficients of DNA minicircles to validate the elastic theory results obtained from our other collaborator. Also, it explores the impact of DNA supercoiling-induced shapes on minicircle hydrodynamics. In another collaborative effort, we study the oligomerization behavior of a de novo designed metalloprotein for photocatalytic hydrogen evolution, leveraging AUC for comprehensive characterization.

These collective studies highlight the indispensable role of AUC in the characterization of macromolecules in solution, with applications ranging from DNA dynamics to functional protein characterization, and they demonstrate the crucial importance of reliable molecular standards in enhancing the accuracy of AUC measurements. Together, they contribute to the advancement of analytical science and its applications across diverse research domains.

CONTRIBUTION OF AUTHORS

During the Masters program, I had the opportunity to make contributions to scientific research by collaborating on the publication of three manuscripts.

One of the main components of the thesis is the manuscript presented in Chapter 2, titled "Suitability of double-stranded DNA as a molecular standard for the validation of analytical ultracentrifugation instruments". It was published in the European Biophysics Journal, Volume 52, issue 4-5 on 27th July 2023. In this work, I have served as one of first authors with equal contribution with Dr. Jonathan M. Fogg, designing and performing AUC experiments, conducting AUC data analysis and interpretation, and writing the manuscript, all under the supervision of Dr. Borries Demeler. Collaborators Dr. Lynn Zechiedrich, Dr. Jonathan M. Fogg, and Dr. Daniel J. Catanese, Jr., contributed the DNA minicircle samples and gel electrophoresis data. Additionally, we have collectively submitted a Baylor College of Medicine invention disclosure titled "Using Minicircle DNA for Molecular Standards," with equal inventorship among all collaborators.

Our collaboration with Dr. Lynn Zechiedrich's group led to a second manuscript titled "DNA supercoiling-induced shapes alter minicircle hydrodynamic properties". This work was published in Nucleic Acids Research, Volume 51, Issue 8 on 8th May 2023 (included in Appendix A). In this study, we employed AUC as a powerful technique to measure the sedimentation and diffusion coefficients of DNA minicircles to validate the elastic theory results obtained from our other collaborator, Dr. Piotr Szymczak's group. In this study, I contributed designing and performing AUC experiments, conducting AUC data analysis and interpretation, and writing the AUC section of the manuscript. Our collaborators, Radost Waszkiewicz, Jr., Dr. Maciej Lisicki,

Dr. Piotr Szymczak and Dr. Maria L Ekiel-Jeżewska contributed methods for the theoretical prediction of hydrodynamic properties of DNA minicircle topoisomers based on elastic theory. Dr. Lynn Zechiedrich, Dr. Jonathan M. Fogg, Dr. Daniel J. Catanese, Jr. contributed the DNA minicircles samplesand gel electrophoresis data.

In addition, as a co-author, I have contributed to a third manuscript titled "Photocatalytic hydrogen evolution by a de novo designed metalloprotein that undergoes Ni-mediated oligomerization shift" which was published in Chemistry – A European Journal, Volume 29 on 28th November 2022 (included in Appendix B). This manuscript details the design and characterization of a functional H_2 evolving protein, and AUC was employed to study the oligomerization behavior of the 4SCC peptide upon Ni^{II} addition. For this collaborative research study, I contributed by designing and conducting AUC experiments, the analysis and interpretation of AUC results, and writing the AUC section of the manuscript.

ACKNOWLEDGMENTS

First and foremost, I want to express my gratitude to everyone who has helped in my educational journey and supported me throughout my life. I am grateful to all of my teachers and mentors for the guidance and encouragement. I would want to thank the University of Lethbridge for all the opportunities and help that made my experience as an international student very enriching and fulfilling.

I am grateful to my supervisor, Dr. Borries Demeler, for the support and encouragement throughout my M.Sc. Program. Thank you for the numerous opportunities you have provided me to collaborate on research projects, contribute to publications. and attend international conferences, which played a significant role in improving my writing, presentation and collaboration skills. Without his guidance, this accomplishment would not have been possible.

I would like to thank my committee members, Dr. Trushar Patel and Dr. Stacey Wetmore, for the essential advice and insightful feedback throughout the process. I really appreciate both of them being so supportive and helpful. Their constructive feedback has motivated me to continuously improve and excel in my work. Also, thank you Trushar for arranging the lab presentation meetings, I learnt a lot from them.

I am thankful for the contribution and support of all our collaborators throughout the research projects. Special thanks to Dr. Lynn Zechiedrich for the opportunity and learning experience you provided me at the GRC and BPS conferences.

Thank you to all the members of the Demeler lab, whose support has made my research journey truly remarkable. Thank you, Dr. Emre Brookes, for your help with supercomputers and implementing software advancements when I wanted to analyse large datasets. Thank you so much, Dr. Amy Henrickson, for mentoring me in lab experiments your excellent teaching skills

and for the support whenever I had questions or needed to learn. Thank you, Dr. Saeed Mortezazadeh, for always encouraging me and helping me with coding skills. I am blessed to have amazing colleagues in our lab who have become my friends and made the lab feel like a second home.

Thank you very much Deede. I appreciate your care and support from the since the first day I arrived, during the difficult times of the Covid-19 pandemic, which also happened to be my first experience with winter. Your kindness and generosity made me feel at home, even in a new country. I am grateful for everything you have done.

Neha, I really appreciate everything you did, especially when I first arrived. Your kindness and thoughtfulness were a huge comfort to me.

To my family, a special thank you to my dad, who has always been there for me. I am speechless to express my appreciation for how amazing you are, I wholeheartedly owe you everything. Thank you so much for everything you do! Thank you my sister and two brothers for all of our love, care, and support, as well as the endless hours you spent with me whenever I needed it; it meant everything to me.

Lastly, I am grateful to my amazing husband, Kaushika. Your support, love and care have been the pillar of strength in my life ever since we came together. I appreciate you and all that you do. Thank you!

TABLE OF CONTENTS

INDEX OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2DSA	Two-Dimensional Spectrum Analysis
2DSA-MC	Two-Dimensional Spectrum Analysis Monte Carlo
AUC	Analytical Ultracentrifugation
bp	Base Pair
BSA	Bovine Serum Albumin
cryoET	Electron Cryotomography
DNA	Deoxyribonucleic Acid
dsDNA	Double-stranded DNA
GA	Genetic Algorithms
GA-MC	Genetic Algorithms Monte Carlo
GMP	Good Manufacturing Practices
Lk	Linking number
LOD	Limits of Detectability
LOQ	Limits of Quantification
PSV	Partial Specific Volume
RNA	Ribonucleic Acid
SV	Sedimentation Velocity

Chapter 1: Introduction

The field of biochemistry has made significant progress in the development of analytical techniques, and among them, analytical ultracentrifugation (AUC) stands out as a powerful tool for characterizing macromolecules. In this introductory chapter, I provide an overview of AUC, which will serve as a foundation for comprehending subsequent sections of this thesis. By investigating the potential of double-stranded DNA as a molecular standard to validate AUC instruments, this thesis aims to enhance the accuracy of AUC measurements in the characterization of biological macromolecules.

1.1 AUC Background

1.1.1 What is AUC?

AUC has a long history dating back to the early 20th century. Theodor Svedberg, a notable Swedish physical chemist, made significant contributions to the understanding of colloids and molecular compounds during this period. In 1923, Svedberg invented the analytical ultracentrifuge, a high-speed centrifuge specially designed to measure the molecular weight of biopolymers [1]. In recognition of his groundbreaking research on disperse systems, Svedberg was awarded the Nobel Prize in Chemistry in 1926. Over the years, AUC has been improved through advances in centrifuge design, optics, and data analysis techniques [2, 3, 4], enhancing the versatility of AUC as a powerful tool for biophysical research.

AUC is a powerful biophysical separation technique that characterizes molecules and molecular interactions in the solution environment based on the size and shape of molecules, and the density difference between the molecule and the solvent [5]. A key advantage of AUC is that

it does not require labeling or immobilizing the molecules being studied [6]. Additionally, the analysis is based on first principles, without the need for external reference standards [7].

By adjusting solution conditions such as concentration, pH, ionic strength, buffer type, ligands, oxidation state, and temperature, AUC enables the replication of cellular conditions closely resembling those found *in vivo* [8], facilitating the exploration of their behavior under a physiologically relevant solution environment [9]. Moreover, AUC enables the analysis of macromolecules across a wide size range, ranging from 10^2 to 10^8 Daltons. This versatility is achieved by adjusting the rotor speed of the AUC experiment, as well as the viscosity and density of the solvent. The ultracentrifuge rotor, capable of reaching speeds up to 60,000 rpm, generates a force of approximately 300,000g [10]. High speeds are essential for efficient separation of small molecules, and slower rotor speeds are suitable for analyzing larger molecules.

1.1.2 Sedimentation-diffusion process

Sedimentation

In AUC, the application of centrifugal force enables the separation of molecules based on their buoyant mass and frictional properties [7]. During the centrifugation process, macromolecules experience three forces: centrifugal force, buoyancy force, and frictional force (viscous drag). The centrifugal force arises from the rotational movement, and as the molecules start to move, the buoyancy force and viscous drag act in opposition to the centrifugal force. The buoyancy force has an impact on whether a molecule will sediment or float, based on the relative densities of the solute and the solvent. If the partial specific volume of the molecule is equal to the inverse of the density of the solvent, the sedimentation is zero. It occurs when the molecule has the same density as the buffer solution. If the molecule is denser than the solution, it will sediment, while a lower density causes the molecule to float.

The sedimentation coefficient, a parameter that quantifies the rate of sedimentation, is directly proportional to the mass of the molecule and inversely proportional to its frictional properties because frictional forces act as a resistance, slowing down sedimentation. The shape and size of the molecule affect the friction experienced during sedimentation [11]. Solvent properties, such as density and viscosity, also impact sedimentation. For example, the higher solvent viscosity observed in sucrose solutions increases friction, resulting in slower sedimentation rates [12].

The sedimentation coefficient (s) is given by Equ. 1.1.

$$s = \frac{M(1-\bar{v}\rho)}{Nf}$$

Equ. 1.1

where M is the molar mass, f is the frictional coefficient, $\bar{v}$ is its partial specific volume (PSV), ρ is the solvent density, and N is Avogadro's number.

Diffusion

In the sedimentation process of molecules, a moving boundary is formed, resulting in a concentration gradient that induces molecular diffusion. Therefore, within this concentration gradient, sedimentation and diffusion take place simultaneously [11]. The changes in the concentration profile are monitored to measure the sedimentation and diffusion transport over the course of experiment [11, 13, 14].

The diffusion rate of macromolecules is inversely related to their friction, which is influenced by their size and shape [15, 16]. Macromolecules with higher friction exhibit slower diffusion rates. Additionally, solvent viscosity plays a role in diffusion, as higher viscosity

restricts the movement of macromolecules, leading to slower diffusion rates [17]. However, the diffusion process remains unaffected by changes in solvent density. The diffusion coefficient (D) is given by Equ. 1.2 where T is the temperature in Kelvin, f is the frictional coefficient, R is the universal gas constant, and N is Avogadro's number.

$$D = \frac{RT}{Nf}$$

Equ. 1.2

Frictional coefficient

The frictional coefficient (f) can be calculated using the measured diffusion coefficient (D), the temperature in Kelvin (T), the universal gas constant (R) and Avogadro's number (N).

$$f = \frac{R\,T}{N\,D}$$

Equ. 1.3

Partial specific volume ($\bar{v}$)

The partial specific volume (PSV) represents the inverse of the density of the sedimenting molecule, indicating the volume of solvent displaced by 1 gram of solute. It includes the volume occupied by the macromolecule along with the transiently bound water molecules and ions.

The PSV is sensitive to the solution conditions [18]. Factors such as pH, salt concentration, and temperature have a significant impact on the hydration and ion binding of the macromolecule, thereby influencing the PSV value. Therefore, the PSV of a molecule can be highly sensitive to the specific solution conditions employed in the AUC experiment [14]. Understanding and controlling these solution conditions are important for the accurate

determination of the PSV and proper analysis of AUC data [19]. If the molar mass is known, the Svedberg equation (Equ. 1.4) can be used to determine the PSV (Equ. 1.5).

$$\frac{s}{D} = \frac{M(1-\bar{v}\rho)}{RT} \qquad \bar{v} = \frac{1}{\rho}\left(1-\frac{sRT}{DM}\right)$$

Equ. 1.4 Equ. 1.5

where s and D are the measured sedimentation and diffusion coefficients, T is the temperature in Kelvin, M is the molar mass, ρ is the solvent density, and R is the universal gas constant.

By combining the PSV with the measured sedimentation and diffusion coefficients, the molar mass of the molecule can be determined.

$$M = \frac{sRT}{D(1-\bar{v}\rho)}$$

Equ. 1.6

The volume of the molecule can be calculated using the sequence-derived molar mass, the partial specific volume and Avogadro's number.

$$V = \frac{M\bar{v}}{N}$$

Equ. 1.7

Assuming the molecule is a sphere, the radius of the hypothetical minimal sphere (r_0) *c*an be calculated using the actual volume (V) of the molecule.

$$r_0 = \left(\frac{3V}{4\pi}\right)^{1/3}$$

Equ. 1.8

Using the Stokes-Einstein equation, the frictional coefficient of the hypothetical minimal sphere (f_0) can be calculated.

$$f_0 = 6\pi\eta r_0$$

Equ. 1.9

Then, the ratio of the actual frictional coefficient and the frictional coefficient of the hypothetical minimal sphere can be calculated.

$$\phi = \frac{f}{f_0}$$

Equ. 1.10

The ratio of $\varphi = f/f_0$ is called the frictional ratio, and defines the anisotropy of the molecule.

Sedimentation velocity experiments

Otto Lamm, a Ph.D. student of The Svedberg, made a significant contribution in 1929 by developing a partial differential equation, which describes the time-dependent behavior of the concentration gradient of the solute during sedimentation velocity experiments [20, 21]. This equation serves as a fundamental tool for understanding and analyzing the behavior of solutes undergoing sedimentation in solution.

The analysis of AUC experimental data involves modeling the data using finite element solutions of the Lamm equation (Equ. 1.11). This modeling approach enables the direct determination of sedimentation (s) and diffusion (D) coefficients, and partial concentrations of the solutes (C) [21].

$$\left(\frac{\partial C}{\partial t}\right)_r = \frac{-1}{r}\frac{\partial}{\partial r}\left[s\omega^2 r^2 C - D r \frac{\partial C}{\partial r}\right]_t$$

Equ. 1.11

where r is the radius, t is the time, and ω is the angular velocity of the rotor.

1.1.3 Instrumentation

AUC experiments require specialized instrumentation for data acquisition and analysis. In this study, Optima AUC instruments with titanium rotors were used (see Figure 1.1).

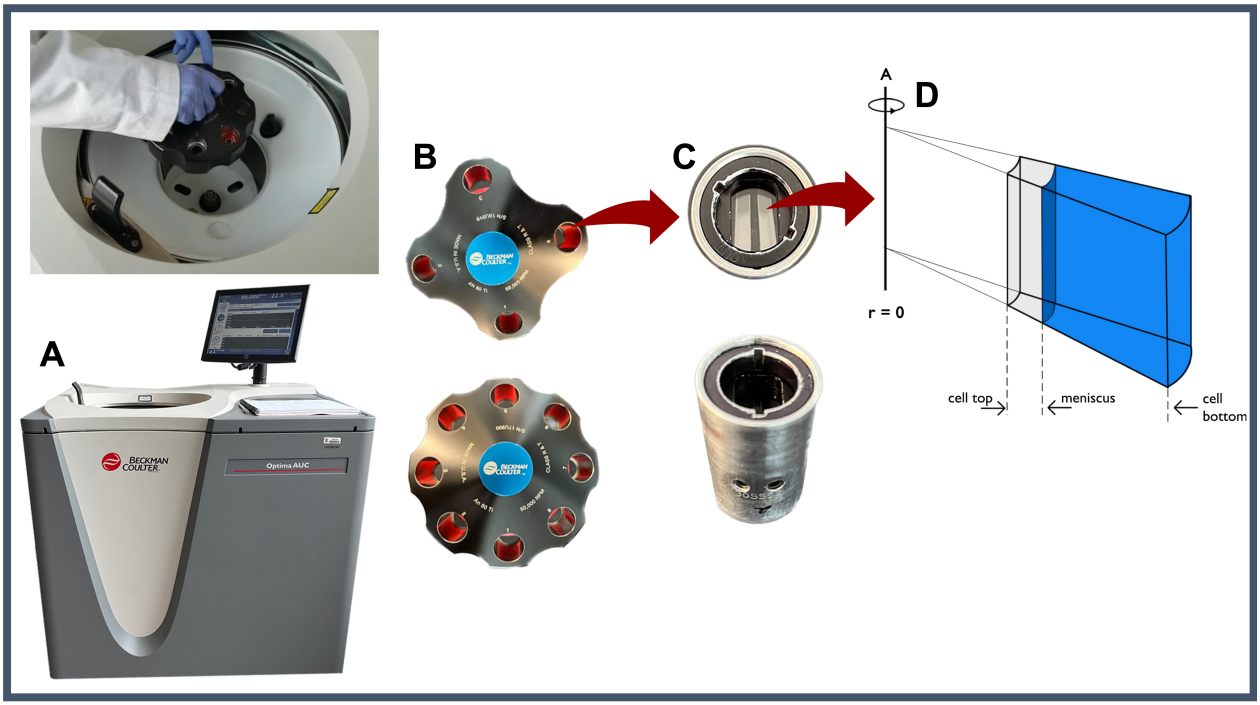


Figure 1.1. Instrumentation for AUC experiments. (A) Optima AUC machine and the rotor placement inside it. (B) Overhead view of a 4-hole and a 8-hole rotor (C) Overhead view and side view of an AUC cell (D) Representation of the sample column inside the AUC cell.

The rotors are available in two configurations: the four-hole rotor has a maximum rotor speed of 60 krpm and delivers higher force, while the eight-hole rotor has a maximum rotor speed of 50 krpm but has capacity to accommodate more samples. Each cell has a housing and a centrepiece with two sector-shaped compartments. The top and bottom of the centrepiece are capped with quartz or sapphire windows for sample placement.

Molecules can be observed using various optical systems that detect different properties of the molecules, such as refractive index, UV or visible absorbance, or fluorescent light emission.

The availability of multiple detectors significantly extends the range of applications in AUC, allowing for the analyses of diverse macromolecules.

The UV/vis absorbance detector is widely used to study biopolymers and materials that absorb in the UV or visible range, and it provides access to dilute concentration regimes due to their high sensitivity.

Fluorescence optics provide a highly sensitive method with exquisite selectivity, enabling precise characterization of fluorescent molecules in impure solutions, facilitating measurements of binding experiments. The intensity of fluorescence emission can be measured using a confocal microscope setup with an excitation laser, typically at 488 nm. It is particularly useful to study molecules with intrinsic fluorophores or eGFP fusions in impure cell extracts. Additionally, fluorescently labeled proteins can be used to study the order of assembly in multi-domain protein complexes.

Rayleigh interference optics offer fast data acquisition and are particularly suitable for studying non-absorbing molecules such as carbohydrates. This detection method relies on measuring refractive index differences at 675 nm, providing valuable insights into high-concentration studies and the behavior of non-absorbing species.

The multi-wavelength UV/vis detection system provides an advanced capability to acquire an additional spectral dimension in addition to the hydrodynamic separation, enabling independent characterization of spectral properties of molecules in mixtures in AUC experiments. By measuring multiple wavelengths, a comprehensive understanding of the molecular composition and hydrodynamic properties can be obtained. It enables the characterization of protein-nucleic acid interactions [22] andprotein binding properties (e.g., metallo-porphyrins)

[23], as well as assessing purity and composition in applications such as nanoparticle loading and viral vector analysis.

In this particular study, molecules were observed with absorbance optics at 260 nm to measure sample sedimentation. This optical setup measures the intensity of light as it passes through the sample, which indicates the concentration at each radial position. Multiple measurements are taken over time, resulting in a sedimentation velocity experiment comprising numerous scans, each recorded at different time points (see Figure 1.2).

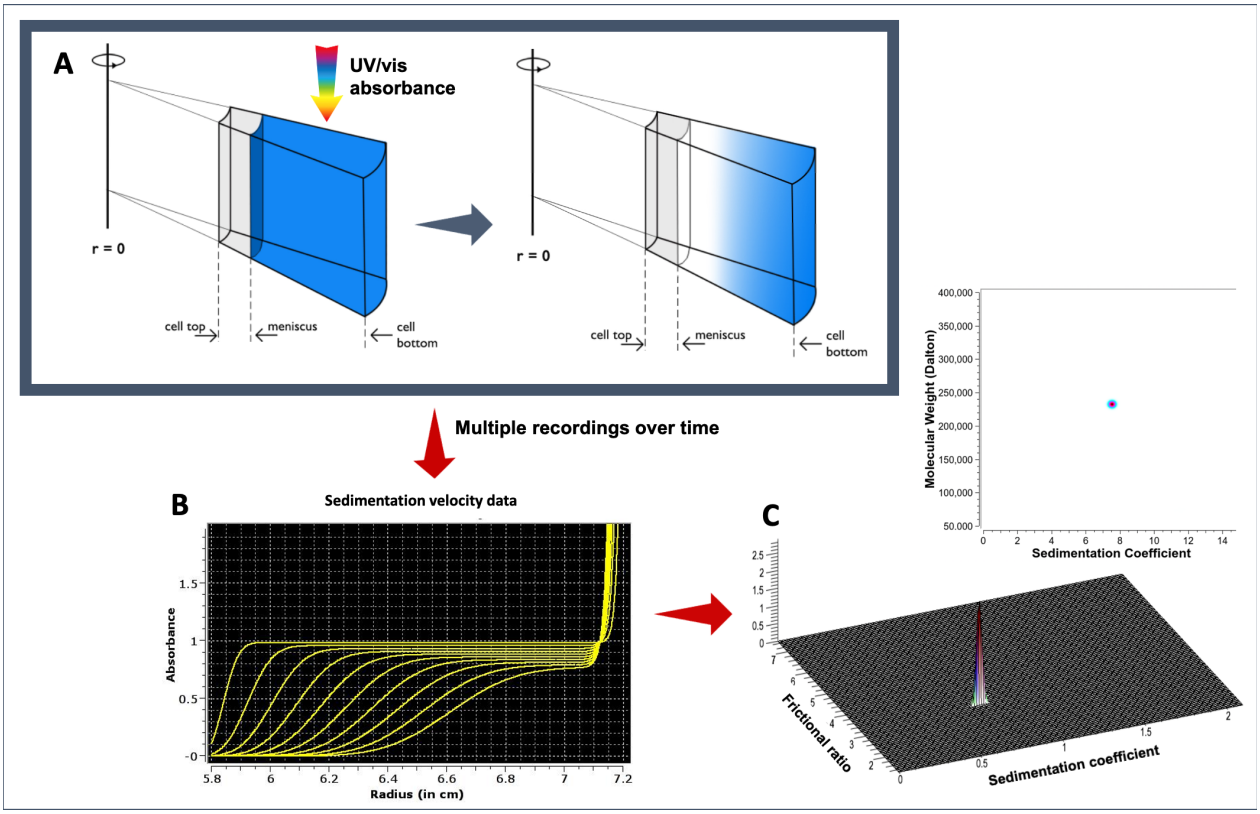


Figure 1.2. Representation of the process from AUC experiment to data interpretation. (A) The change in concentration gradient is because of the centrifugal force. (B) Multiple scans obtained over time. (C) Three-dimensional view of the observed single species from global Monte-Carlo analysis, along with sedimentation coefficient and molecular weight distribution.

All analysis results and visualizations are saved in the database, where the user can subsequently retrieve a report from the USLIMS website [24].

1.1.4 Analysis approach

The optimization methods used in the analysis of sedimentation velocity data obtained from AUC experiments are described below.

Two-dimensional Spectrum Analysis (2DSA) [25, 26]

2DSA provides a degenerate, linear fit to the experimental data, and is used to identify time- and radially invariant noise, as well as the boundary conditions (meniscus, bottom of the cell, total concentration) of the partial differential equation. It generates a grid of all possible solutions of the Lamm equation within a range of user-defined parameter values and parameter resolution, which describe the sedimentation and diffusion behavior of macromolecules in the solution. Each point on the grid corresponds to a particularcombination of sedimentation coefficient (s) and diffusion coefficient (D), with the amplitude at that point indicating the partial concentration of the corresponding species.

2DSA is particularly useful for identifying regions that contain signal, allowing the separation of true signal from systematic noise contributions. It achieves this by iteratively refining the grid, adding it back to the starting grids, and repeating the process until no further changes occur. This iterative refinement ensures that the final grid is equivalent to the original fine-grained grid. However, 2DSA can also introduce false positives and may require further refinement to remove these false signals.

Genetic Algorithms (GA) Analysis [27, 28]

GA analysis finds a parsimonious solution and provides a parsimonious regularization of the 2DSA spectrum, satisfying Occam's Razor law. Occam's Razor suggests that the simplest solution

describing the data well is the preferred solution. It has excellent convergence properties, and resolves solutes reliably. Instead of regularizing the solution and introducing an infinite number of solutes with different probabilities, the solution space is reduced to find the solution that minimizes the number of possible solutes. The iterative 2DSA refinement is used to initialize a genetic algorithm analysis. GA can also be used to fit discrete, non-linear models in reversible association, non-ideality and co-sedimenting solutes.

Monte Carlo Analysis (MC) [27]

The Monte Carlo analysis is a stochastic approach that can be used to measure the effect noise has on the fitted parameters. With the Monte Carlo approach, the statistical confidence limits of each measured parameter can be determined. Once a Monte Carlo analysis is completed, a frequency distribution of parameter values for each parameter is obtained, and statistics for the distribution can be calculated. The Monte Carlo analysis can be used for both 2DSA and GA optimization algorithms. The GA model is used to initialize a GA-MC analysis.

The AUC data analysis workflow is shown in Figure 1.3.

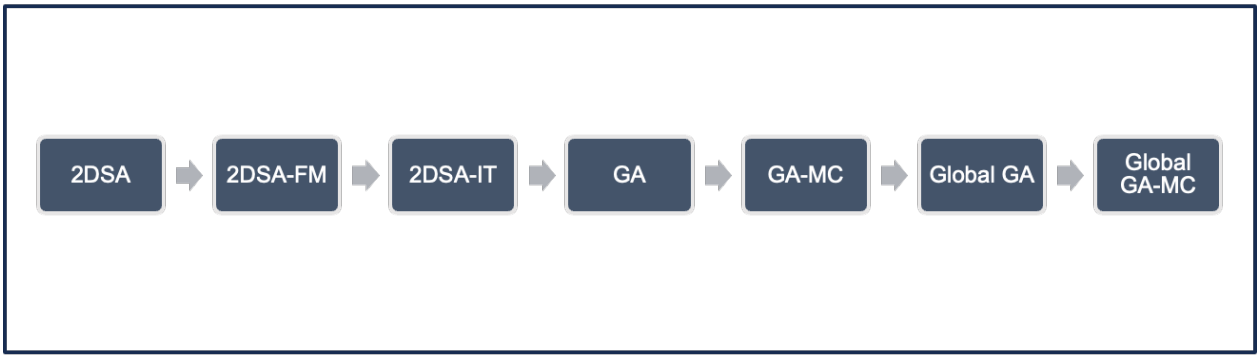


Figure 1.3. An illustration of the AUC data analysis refinement workflow in Ultrascan-III.

Data analysis requires complex modeling and optimization algorithms performed on supercomputers to accelerate the analysis, and to obtain higher resolution. Furthermore, the UltraScan Science Gateway [29] allows users to easily perform data analysis on a remote supercomputer without the need for expertise in programming supercomputers in Linux or high-end computing.

The utilization of parallel threads on supercomputers for a global GA-MC analysis significantly enhances the efficiency of an AUC analysis in this study. This parallelization technique allows simultaneous processing of multiple threads, enabling faster and more efficient data analysis.

1.1.5 What can be learned from AUC?

AUC can be used to study different types of molecules and experimental systems, such as proteins, nucleic acids, lipids, carbohydrates, synthetic polymers and nanoparticles [30, 31, 32, 33, 34, 35, 36].

AUC has a wide range of applications. The composition analysis can be used to assess the protein purity, as well as to identify contaminants and their partial concentrations [37, 38], which is essential for quality control and determining the efficacy of purification processes. Additionally, AUC offers the advantage to examine the thermodynamics and (slow) kinetics of reactions, providing valuable insights into oligomerization, hetero- interactions, ligand binding, and measurements of equilibrium constants (Kd) [39, 40, 41, 42, 43]. Our previous study, to which I contributed as a co-author, examined the oligomerization behavior and Ni^{II} binding of the 4SCC peptide using AUC [44] (included in Appendix B). These insights gained from AUC

studies are crucial for elucidating the complex structure-function relationships of macromolecules [44, 45].

The diverse capabilities of AUC contribute to a deeper understanding of macromolecular structure and function, leading to advances in various scientific disciplines such as structural biology, drug discovery, biomaterials engineering, and therapeutic interventions [46].

1.2 The need for reliable molecular standards for AUC

AUC is a first principle method that does not require any standards. However, the accuracy of the measurements is only assured if the instruments are properly calibrated [47, 48]. Instrumental errors such as incorrect rotor speed and temperature recordings, errors in radial positioning and timing, detection system imperfections, and issues in chromatic aberration can introduce biases and uncertainties, and can distort the obtained results from the analysis, affecting the observed sedimentation coefficient, diffusion coefficient and partial concentration values [49].

In order to address these challenges effectively, it is necessary to implement rigorous quality control measures and calibration protocols [50,51]. Additionally, instrument validation should be frequently performed using a well-characterized and widely accepted reference standard [52] to assess whether the instrument is correctly calibrated and provides accurate measurements. Previous research has highlighted significant deviations in the measured sedimentation coefficients across different laboratories, emphasizing the importance of having a reference standard to establish comparability across studies, instruments, and laboratories [53].

However, the absence of a universally accepted standard poses a significant obstacle to the validation of AUC instruments [49]. Efforts are being made to address this issue, with ongoing research focusing on the development of a reliable and widely recognized molecular standard.

The introduction section of Chapter 2 further provides a comprehensive overview of the significance and rationale behind the need for reliable molecular standards in AUC. It highlights the limitations associated with existing protein-based standards and emphasizes the requirements for an ideal AUC molecular standard.

1.3 The requirements of an AUC molecular standard

Several key criteria must be met to establish a suitable AUC molecular standard [52]. Firstly, the molecular standard should exhibit long-term stability, maintaining its integrity and avoiding degradation or aggregation over time. This stability is crucial for consistent and reproducible measurements. Additionally, the standard should demonstrate predictable behavior across the entire temperature range accessible with AUC instruments. This temperature range is important for studying macromolecules under various physiological or experimental conditions.

A standard should have a homogeneous composition to ensure consistent and predictable behavior in AUC experiments. It must also have a precisely known molar mass as a fundamental requirement. Furthermore, the standard should exhibit an absence of concentration-dependent non-ideality, meaning that its behavior does not deviate from ideal solution behavior at the measured concentration.

Accessibility and affordability are practical considerations for an ideal AUC molecular standard. It should be readily available to the scientific community at a low cost, facilitating its widespread use and adoption. Another desirable attribute of a molecular standard for AUC is the

ability to modulate size and shape. Size and shape both affect the two primary parameters measured by AUC. The availability of an orthogonal method, such as mass spectrometry or electrophoresis, to confirm molecular properties further strengthens the reliability of the standard, as it provides an additional means of validation.

However, it is important to acknowledge that, despite the importance and requirements for a universal AUC molecular standard, currently there is no existing reference material with approved values that can be used to validate AUC instruments. Because AUC is a first-principles method, results from a molecular standard with known properties can be used to demonstrate that an AUC instrument is working properly.

Further details on these requirements and the challenges associated with AUC molecular standards are discussed in the introduction section of Chapter 2.

1.4 Double-stranded DNA as a potential molecular standard

Double-stranded DNA exhibits unique properties that make it a potential candidate to be used as a reliable and reproducible molecular standard in AUC experiments [52]. The exceptionally stable double helix structure of DNA renders it resistant to degradation [54]. Additionally, its molecular weight can be precisely determined from its nucleotide sequence [55]. The ability to customize DNA size and anisotropy by changing the sequence length further enhances its appeal as a molecular standard [56, 57].

The possibility of homogeneous production of DNA [58] offers a significant advantage because it eliminates the variability often associated with proteins, which are less stable than DNA and can also include post-translational modifications, variability in protein folding, and batch-to-batch variations [59, 60]. Furthermore, DNA can be produced recombinantly in

sufficient quantities, and with high purity, offering practical advantages for experimental use [61]. Its optical properties include a high extinction coefficient at 260 nm [62], enabling sensitive detection and quantification under dilute and ideal solution conditions.

The utilization of DNA as a molecular standard in AUC holds significant potential. However, to fully evaluate this potential, it is critical to systematically investigate the suitability of DNA as a molecular standard for AUC. By comprehensively investigating the suitability of DNA as a reference material for AUC, this research aims to bridge a significant gap in the field and contributes to the advancement and standardization of the AUC technique.

The significance of this project lies in its potential to validate the correct operation of AUC instruments, enable reliable comparisons across experiments and laboratories, and contribute to quality control in biopharma Good manufacturing practices (GMP) related environments. One such field that greatly benefits from these advancements is the biopharmaceutical industry. AUC plays a critical role in characterizing liquid drugs, specifically gene therapy injectables (drugs administered in liquid form), where accurate measurement and characterization are essential for ensuring patient safety. Having a reference standard available to validate AUC instruments becomes crucial in this context, as it ensures that drugs are measured correctly and meet the declared composition. This is of utmost importance in the biopharmaceutical industry, where any errors in measurement can have severe consequences for patient health.

Furthermore, the optimization and standardization of AUC facilitated by this research will contribute to the establishment of best practices and protocols, promoting the reproducibility of scientific results when AUC is used in diverse scientific fields.

1.5 Objectives and scope of the study

In our previous publication on DNA topology, conducted in collaboration with Lynn Zechiedrich's group, AUC was used to measure sedimentation and diffusion coefficients of DNA minicircles to validate the elastic theory results [63] (included in appendix A). These DNA minicircles, which are small circular DNA molecules with a known molar mass and well-defined topology, served as a model system for studying the influence of DNA topology on sedimentation behavior.

In continuation of this collaboration, the current study investigated the suitability of double-stranded DNA as a molecular standard for the validation of AUC instruments by examining the hydrodynamic properties of various topological states of DNA over the temperature range accessible by an AUC instrument.

Based upon the current literature, we hypothesized that distinct DNA topological states such as linear, nicked and supercoiled forms exhibit unique hydrodynamic properties, which can be influenced by temperature variations. Our study focused on three specific objectives:

1. Assess the stability, reproducibility, and homogeneity of DNA as a molecular standard.

This study aimed to evaluate the hydrodynamic properties of DNA, and its stability, reproducibility, and homogeneity, which are necessary for its suitability as a molecular standard.

2. Compare the hydrodynamic properties of different topological states of DNA.

DNA molecules can exist in linear and circular forms, each with distinct structural and dynamic characteristics. Therefore, understanding the impact of DNA topology on its hydrodynamic behavior is essential for evaluating its suitability as a molecular standard in AUC.

3. Determine the effect of temperature on the hydrodynamic properties of DNA.

Temperature exerts an effect on the solution behavior of macromolecules, making it essential to investigate its impact on the hydrodynamic properties of DNA [64, 65]. Therefore, a reference standard for all temperatures accessible with the AUC instrument must be available, and careful measurements of the reference standard over the accessible temperature range are required. Furthermore, the stability of the standard when exposed to a range of temperature fluctuations must be determined. Understanding the behavior of DNA under varying storage temperatures is also essential to guarantee its stability and suitability as a reliable standard material.

To achieve these objectives, we used DNA minicircles with 333 bp and 339 bp lengths, including different topological states such as linear, nicked circular, and negatively supercoiled DNA (see Figure 1.4). These diverse topological states represent distinct structural forms of DNA and enable a comprehensive examination of their impact on AUC measurements.

All DNA minicircle samples were prepared by our collaborators Dr. Jonathan M. Fogg, Dr. Daniel J. Catanese, Jr and Dr. Lynn Zechiedrich, Baylor College of Medicine. Information on the preparation of samples is provided in Chapter 2.

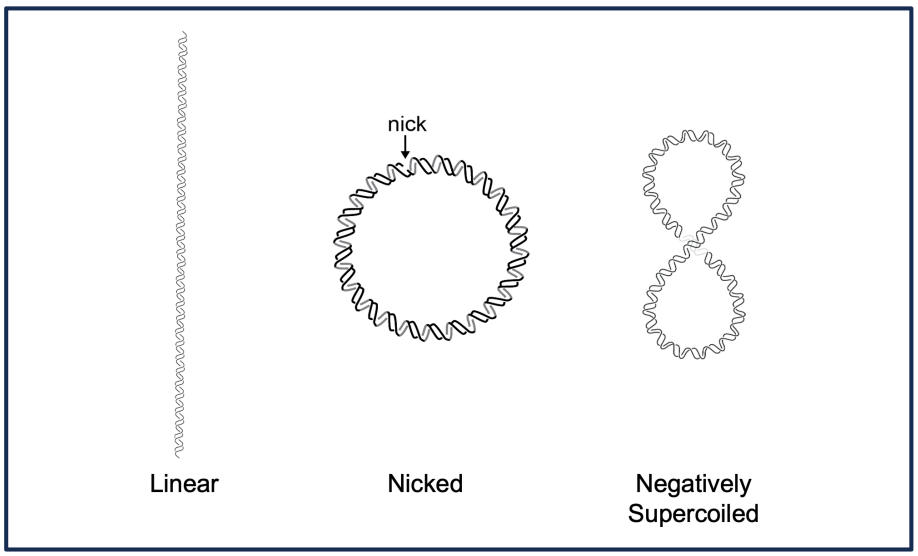


Figure 1.4. The DNA minicircle topologies used in the study.

1.6 Thesis organization

This introductory chapter has emphasized the importance of reliable molecular standards in AUC measurements and highlighted the significance of this study in advancing macromolecular characterization. The subsequent chapters are built upon this foundation, explaining the experimental design, methodologies, and data analysis techniques used to assess the suitability of double-stranded DNA as a molecular standard.

I present my first-author manuscript in Chapter 2, which is an essential component of the thesis. It comprehensively details the results and discussion sections of the study.

Finally, Chapter 3 serves as a summary of the main findings of the thesis, highlighting their significance in AUC research. It reiterates the objectives of the study and offers suggestions for future studies to further explore the utilization of DNA as an AUC molecular standard.

In addition, the appendix contains two co-authored manuscripts that provide supplementary information for readers interested in accessing experimental results from my other

research projects conducted during my MSc study. These publications demonstrate further applications of the AUC technique.

1.7 References

1 Harding SE. The Svedberg Lecture 2017. From nano to micro: the huge dynamic range of the analytical ultracentrifuge for characterising the sizes, shapes and interactions of molecules and assemblies in Biochemistry and Polymer Science. Eur Biophys J. 2018 Oct;47(7):697-707. doi: 10.1007/s00249-018-1321-3.

2 Liu, J., Yadav, S., Andya, J., Demeule, B., & Shire, S. J. (2015). Analytical Ultracentrifugation and Its Role in Development and Research of Therapeutical Proteins. Methods in enzymology, 562, 441–476. https://doi.org/10.1016/bs.mie.2015.04.008

3 Savelyev, A., Gorbet, G. E., Henrickson, A., & Demeler, B. (2020). Moving analytical ultracentrifugation software to a good manufacturing practices (GMP) environment. PLoS computational biology, 16(6), e1007942. https://doi.org/10.1371/journal.pcbi.1007942

4 Mortezazadeh, S., & Demeler, B. (2023). Systematic noise removal from analytical ultracentrifugation data with UltraScan. European biophysics journal : EBJ, 10.1007/s00249-023-01631-6. Advance online publication. https://doi.org/10.1007/s00249-023-01631-6

5 Demeler, B., Brookes, E., & Nagel-Steger, L. (2009). Analysis of heterogeneity in molecular weight and shape by analytical ultracentrifugation using parallel distributed computing. Methods in enzymology, 454, 87–113. https://doi.org/10.1016/S0076-6879(08)03804-4

6 Harding, S. E., Gillis, R. B., & Adams, G. G. (2016). Assessing sedimentation equilibrium profiles in analytical ultracentrifugation experiments on macromolecules: from simple average molecular weight analysis to molecular weight distribution and interaction analysis. Biophysical reviews, 8(4), 299–308. https://doi.org/10.1007/s12551-016-0232-8

7 Demeler B. (2019). Measuring molecular interactions in solution using multi-wavelength analytical ultracentrifugation: combining spectral analysis with hydrodynamics. The biochemist, 41(2), 14–18. https://doi.org/10.1042/bio04102014

8 Cao, W., & Demeler, B. (2008). Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution for multicomponent reacting systems. Biophysical journal, 95(1), 54–65. https://doi.org/10.1529/biophysj.107.123950

9 Howlett, G. J., Minton, A. P., & Rivas, G. (2006). Analytical ultracentrifugation for the study of protein association and assembly. Current opinion in chemical biology, 10(5), 430–436. https://doi.org/10.1016/j.cbpa.2006.08.017

10 Stoutjesdyk, M., Brookes, E., Henrickson, A., & Demeler, B. (2020). Measuring compressibility in the optima AUC™ analytical ultracentrifuge. European biophysics journal : EBJ, 49(8), 711–718. https://doi.org/10.1007/s00249-020-01482-5

11 Demeler, B., Nguyen, T. L., Gorbet, G. E., Schirf, V., Brookes, E. H., Mulvaney, P., El-Ballouli, A. O., Pan, J., Bakr, O. M., Demeler, A. K., Hernandez Uribe, B. I., Bhattarai, N., & Whetten, R. L. (2014). Characterization of size, anisotropy, and density heterogeneity of nanoparticles by sedimentation velocity. Analytical chemistry, 86(15), 7688–7695. https://doi.org/10.1021/ac501722r

12 Momen-Heravi, F., Balaj, L., Alian, S., Trachtenberg, A. J., Hochberg, F. H., Skog, J., & Kuo, W. P. (2012). Impact of biofluid viscosity on size and sedimentation efficiency of the isolated microvesicles. Frontiers in physiology, 3, 162. https://doi.org/10.3389/fphys.2012.00162
13 Cole, J. L., Lary, J. W., P Moody, T., & Laue, T. M. (2008). Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. Methods in cell biology, 84, 143–179. https://doi.org/10.1016/S0091-679X(07)84006-4
14 Edwards, G. B., Muthurajan, U. M., Bowerman, S., & Luger, K. (2020). Analytical Ultracentrifugation (AUC): An Overview of the Application of Fluorescence and Absorbance AUC to the Study of Biological Macromolecules. Current protocols in molecular biology, 133(1), e131. https://doi.org/10.1002/cpmb.131
15 Chan, T. C., Li, H. T., & Li, K. Y. (2015). Effects of Shapes of Solute Molecules on Diffusion: A Study of Dependences on Solute Size, Solvent, and Temperature. The journal of physical chemistry. B, 119(51), 15718–15728. https://doi.org/10.1021/acs.jpcb.5b10550
16 de Sancho, D., Sirur, A., & Best, R. B. (2014). Molecular origins of internal friction effects on protein-folding rates. Nature communications, 5, 4307. https://doi.org/10.1038/ncomms5307
17 von Bülow, S., Siggel, M., Linke, M., & Hummer, G. (2019). Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. Proceedings of the National Academy of Sciences of the United States of America, 116(20), 9843–9852. https://doi.org/10.1073/pnas.1817564116
18 Murphy, L. R., Matubayasi, N., Payne, V. A., & Levy, R. M. (1998). Protein hydration and unfolding--insights from experimental partial specific volumes and unfolded protein models. Folding & design, 3(2), 105–118. https://doi.org/10.1016/S1359-0278(98)00016-9
19 Savelyev A. (2021). Assessment of the DNA partial specific volume and hydration layer properties from CHARMM Drude polarizable and additive MD simulations. Physical chemistry chemical physics : PCCP, 23(17), 10524–10535. https://doi.org/10.1039/d1cp00688f
20 Cao W and Demeler B. Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution of the Lamm equation. (2005) Biophys J. 89(3):1589-602
21 Cao, W., & Demeler, B. (2008). Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution for multicomponent reacting systems. Biophysical journal, 95(1), 54–65. https://doi.org/10.1529/biophysj.107.123950
22 Zhang, J., Pearson, J. Z., Gorbet, G. E., Cölfen, H., Germann, M. W., Brinton, M. A., & Demeler, B. (2017). Spectral and Hydrodynamic Analysis of West Nile Virus RNA-Protein Interactions by Multiwavelength Sedimentation Velocity in the Analytical Ultracentrifuge. Analytical chemistry, 89(1), 862–870. https://doi.org/10.1021/acs.analchem.6b03926
23 Johnson, C. N., Gorbet, G. E., Ramsower, H., Urquidi, J., Brancaleon, L., & Demeler, B. (2018). Multi-wavelength analytical ultracentrifugation of human serum albumin complexed with porphyrin. European biophysics journal : EBJ, 47(7), 789–797. https://doi.org/10.1007/s00249-018-1301-7
24 Demeler, B., Gorbet, G.E. (2016). Analytical Ultracentrifugation Data Analysis with UltraScan-III. In: Uchiyama, S., Arisaka, F., Stafford, W., Laue, T. (eds) Analytical Ultracentrifugation. Springer, Tokyo. https://doi.org/10.1007/978-4-431-55985-6_8
25 Brookes E, Cao W, Demeler B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. Eur Biophys J. 2010 Feb;39(3):405-14. doi: 10.1007/s00249-009-0413-5. Epub 2009 Feb 27.

PMID: 19247646.
26 Kim H, Brookes E, Cao W, Demeler B. Two-dimensional grid optimization for sedimentation velocity analysis in the analytical ultracentrifuge. Eur Biophys J. 2018 Oct;47(7):837-844. doi: 10.1007/s00249-018-1309-z. Epub 2018 May 17. PMID: 29777290; PMCID: PMC6158092.
27 Brookes, E., Demeler, B. Genetic Algorithm Optimization for Obtaining Accurate Molecular Weight Distributions from Sedimentation Velocity Experiments. In: Wandrey, C., Cölfen, H. (eds) Analytical Ultracentrifugation VIII. Progress in Colloid and Polymer Science, vol 131. Springer, Berlin, Heidelberg. https://doi.org/10.1007/2882_004
28 Brookes, E. H., & Demeler, B. (2007, July). Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments. In Proceedings of the 9th annual conference on Genetic and evolutionary computation (pp. 361-368).
29 Memon, S., Riedel, M., Janetzko, F., Demeler, B., Gorbet, G., Marru, S., Grimshaw, A., Gunathilake, L., Singh, R., Attig, N., & Lippert, T. (2014). Advancements of the UltraScan scientific gateway for open standards-based cyberinfrastructures. Concurrency and Computation: Practice and Experience, 26(13), 2280-2291.
30 Mrozowich, T., Park, S. M., Waldl, M., Henrickson, A., Tersteeg, S., Nelson, C. R., De Klerk, A., Demeler, B., Hofacker, I. L., Wolfinger, M. T., & Patel, T. R. (2023). Investigating RNA-RNA interactions through computational and biophysical analysis. Nucleic acids research, 51(9), 4588–4601. https://doi.org/10.1093/nar/gkad223
31 Valderrama, O. J., & Nischang, I. (2021). Reincarnation of the Analytical Ultracentrifuge: Emerging Opportunities for Nanomedicine. Analytical chemistry, 93(48), 15805–15815. https://doi.org/10.1021/acs.analchem.1c03116
32 Urban, M. J., Holder, I. T., Schmid, M., Fernandez Espin, V., Garcia de la Torre, J., Hartig, J. S., & Cölfen, H. (2016). Shape Analysis of DNA-Au Hybrid Particles by Analytical Ultracentrifugation. ACS nano, 10(8), 7418–7427. https://doi.org/10.1021/acsnano.6b01377
33 Cole, J. L., & Hansen, J. C. (1999). Analytical ultracentrifugation as a contemporary biomolecular research tool. Journal of biomolecular techniques : JBT, 10(4), 163–176.
34 Hansen JC, Lebowitz J, Demeler B. Analytical ultracentrifugation of complex macromolecular systems. Biochemistry. 1994 Nov 15;33(45):13155-63. doi: 10.1021/bi00249a001. PMID: 7947722.
35 Demeler, B., Nguyen, T. L., Gorbet, G. E., Schirf, V., Brookes, E. H., Mulvaney, P., El-Ballouli, A. O., Pan, J., Bakr, O. M., Demeler, A. K., Hernandez Uribe, B. I., Bhattarai, N., & Whetten, R. L. (2014). Characterization of size, anisotropy, and density heterogeneity of nanoparticles by sedimentation velocity. Analytical chemistry, 86(15), 7688–7695. https://doi.org/10.1021/ac501722r
36 Patel, T. R., Winzor, D. J., & Scott, D. J. (2016). Analytical ultracentrifugation: A versatile tool for the characterisation of macromolecular complexes in solution. Methods (San Diego, Calif.), 95, 55–61. https://doi.org/10.1016/j.ymeth.2015.11.006
37 Henrickson, A., Gorbet, G. E., Savelyev, A., Kim, M., Hargreaves, J., Schultz, S. K., Kothe, U., & Demeler, B. (2022). Multi-wavelength analytical ultracentrifugation of biopolymer mixtures and interactions. Analytical biochemistry, 652, 114728. https://doi.org/10.1016/j.ab.2022.114728

38 Valderrama, O. J., & Nischang, I. (2021). Reincarnation of the Analytical Ultracentrifuge: Emerging Opportunities for Nanomedicine. Analytical chemistry, 93(48), 15805–15815. https://doi.org/10.1021/acs.analchem.1c03116
39 Cole, J. L., Lary, J. W., P Moody, T., & Laue, T. M. (2008). Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. Methods in cell biology, 84, 143–179. https://doi.org/10.1016/S0091-679X(07)84006-4
40 Attri, A. K., Lewis, M. S., & Korn, E. D. (1991). The formation of actin oligomers studied by analytical ultracentrifugation. The Journal of biological chemistry, 266(11), 6815–6824.
41 Hu, J., Hernandez Soraiz, E., Johnson, C. N., Demeler, B., & Brancaleon, L. (2019). Novel combinations of experimental and computational analysis tested on the binding of metalloprotoporphyrins to albumin. International journal of biological macromolecules, 134, 445–457. https://doi.org/10.1016/j.ijbiomac.2019.05.060
42 Spinnrock, A., & Cölfen, H. (2018). Control of Molar Mass Distribution by Polymerization in the Analytical Ultracentrifuge. Angewandte Chemie (International ed. in English), 57(27), 8284–8287. https://doi.org/10.1002/anie.201713149
43 Yang, T. C., Catalano, C. E., & Maluf, N. K. (2015). Analytical Ultracentrifugation as a Tool to Study Nonspecific Protein-DNA Interactions. Methods in enzymology, 562, 305–330. https://doi.org/10.1016/bs.mie.2015.04.009
44 Prasad, P., Hunt, L. A., Pall, A. E., Ranasinghe, M., Williams, A. E., Stemmler, T. L., Demeler, B., Hammer, N. I., & Chakraborty, S. (2023). Photocatalytic Hydrogen Evolution by a De Novo Designed Metalloprotein that Undergoes Ni-Mediated Oligomerization Shift. Chemistry (Weinheim an der Bergstrasse, Germany), 29(14), e202202902. https://doi.org/10.1002/chem.202202902
45 Chou, C. Y., Hsieh, Y. H., & Chang, G. G. (2011). Applications of analytical ultracentrifugation to protein size-and-shape distribution and structure-and-function analyses. Methods (San Diego, Calif.), 54(1), 76–82. https://doi.org/10.1016/j.ymeth.2010.11.002
46 Demeler, B. (2005). UltraScan: a comprehensive data analysis software package for analytical ultracentrifugation experiments. Modern analytical ultracentrifugation: techniques and methods, 10, 210-229.
47 Edwards, G. B., Muthurajan, U. M., Bowerman, S., & Luger, K. (2020). Analytical Ultracentrifugation (AUC): An Overview of the Application of Fluorescence and Absorbance AUC to the Study of Biological Macromolecules. Current protocols in molecular biology, 133(1), e131. https://doi.org/10.1002/cpmb.131
48 Zhao, H., Nguyen, A., To, S. C., & Schuck, P. (2021). Calibrating analytical ultracentrifuges. European biophysics journal : EBJ, 50(3-4), 353–362. https://doi.org/10.1007/s00249-020-01485-2
49 Ghirlando, R., Balbo, A., Piszczek, G., Brown, P. H., Lewis, M. S., Brautigam, C. A., Schuck, P., & Zhao, H. (2013). Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. Analytical biochemistry, 440(1), 81–95. https://doi.org/10.1016/j.ab.2013.05.011
50 Stoutjesdyk, M., Henrickson, A., Minors, G., & Demeler, B. (2020). A calibration disk for the correction of radial errors from chromatic aberration and rotor stretch in the Optima AUC™ analytical ultracentrifuge. European biophysics journal : EBJ, 49(8), 701–709. https://doi.org/10.1007/s00249-020-01434-z

51 Zhao, H., Balbo, A., Metger, H., Clary, R., Ghirlando, R., & Schuck, P. (2014). Improved measurement of the rotor temperature in analytical ultracentrifugation. Analytical biochemistry, 451, 69–75. https://doi.org/10.1016/j.ab.2014.02.006
52 Savelyev, A., Gorbet, G. E., Henrickson, A., & Demeler, B. (2020). Moving analytical ultracentrifugation software to a good manufacturing practices (GMP) environment. PLoS computational biology, 16(6), e1007942. https://doi.org/10.1371/journal.pcbi.1007942
53 Zhao, H., Ghirlando, R., Alfonso, C., Arisaka, F., Attali, I., Bain, D. L., Bakhtina, M. M., Becker, D. F., Bedwell, G. J., Bekdemir, A., Besong, T. M., Birck, C., Brautigam, C. A., Brennerman, W., Byron, O., Bzowska, A., Chaires, J. B., Chaton, C. T., Cölfen, H., Connaghan, K. D., … Schuck, P. (2015). A multilaboratory comparison of calibration accuracy and the performance of external references in analytical ultracentrifugation. PloS one, 10(5), e0126420. https://doi.org/10.1371/journal.pone.0126420
54 Bochman, M. L., Paeschke, K., & Zakian, V. A. (2012). DNA secondary structures: stability and function of G-quadruplex structures. Nature reviews. Genetics, 13(11), 770–780. https://doi.org/10.1038/nrg3296
55 Fogg, J. M., Randall, G. L., Pettitt, B. M., Sumners, W. L., Harris, S. A., & Zechiedrich, L. (2012). Bullied no more: when and how DNA shoves proteins around. Quarterly reviews of biophysics, 45(3), 257–299. https://doi.org/10.1017/S0033583512000054
56 Fogg, J. M., Kolmakova, N., Rees, I., Magonov, S., Hansma, H., Perona, J. J., & Zechiedrich, E. L. (2006). Exploring writhe in supercoiled minicircle DNA. Journal of physics. Condensed matter : an Institute of Physics journal, 18(14), S145–S159. https://doi.org/10.1088/0953-8984/18/14/S01
57 Zadegan, R. M., & Norton, M. L. (2012). Structural DNA nanotechnology: from design to applications. International journal of molecular sciences, 13(6), 7149–7162. https://doi.org/10.3390/ijms13067149
58 Carole A Foy , Helen C Parkes, Emerging Homogeneous DNA-based Technologies in the Clinical Laboratory, Clinical Chemistry, Volume 47, Issue 6, 1 June 2001, Pages 990–1000, https://doi.org/10.1093/clinchem/47.6.990
59 Chen, L., & Kashina, A. (2021). Post-translational Modifications of the Protein Termini. Frontiers in cell and developmental biology, 9, 719590. https://doi.org/10.3389/fcell.2021.719590
60 Hunter, D. J. B., Bhumkar, A., Giles, N., Sierecki, E., & Gambin, Y. (2018). Unexpected instabilities explain batch-to-batch variability in cell-free protein expression systems. Biotechnology and bioengineering, 115(8), 1904–1914. https://doi.org/10.1002/bit.26604
61 Tan, S. C., & Yiap, B. C. (2009). DNA, RNA, and protein extraction: the past and the present. Journal of biomedicine & biotechnology, 2009, 574398. https://doi.org/10.1155/2009/574398
62 Cavaluzzi, M. J., & Borer, P. N. (2004). Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. Nucleic acids research, 32(1), e13. https://doi.org/10.1093/nar/gnh015
63 Waszkiewicz, R., Ranasinghe, M., Fogg, J. M., Catanese, D. J., Ekiel-Jeżewska, M. L., Lisicki, M., Demeler, B., Zechiedrich, L., & Szymczak, P. (2023). DNA supercoiling-induced shapes alter minicircle hydrodynamic properties. Nucleic acids research, 51(8), 4027–4042. https://doi.org/10.1093/nar/gkad183
64 Driessen, R. P., Sitters, G., Laurens, N., Moolenaar, G. F., Wuite, G. J., Goosen, N., & Dame, R. T. (2014). Effect of temperature on the intrinsic flexibility of DNA and its interaction with

architectural proteins. Biochemistry, 53(41), 6430–6438. https://doi.org/10.1021/bi500344j
65 Kriegel, F., Matek, C., Dršata, T., Kulenkampff, K., Tschirpke, S., Zacharias, M., Lankaš, F., & Lipfert, J. (2018). The temperature dependence of the helical twist of DNA. Nucleic acids research, 46(15), 7998–8009. https://doi.org/10.1093/nar/gky599

Chapter 2: Suitability of double-stranded DNA as a molecular standard for the validation of analytical ultracentrifugation instruments

2.1 Overview

This chapter presents the manuscript I contributed as one of the primary authors, sharing first authorship with Dr. Jonathan M. Fogg, and co-authorship with Dr. Daniel J. Catanese, Jr., Dr. Lynn Zechiedrich, and Dr. Borries Demeler. It was published in the European Biophysics Journal, Volume 52, issue 4-5 on 27th July 2023. This manuscript details the suitability of DNA as a molecular standard to validate AUC instruments. With the permission of my other authors, I have included this publication in my thesis and have reformatted it to according to the formatting requirements of my thesis.

Invention Disclosure: All the authors collectively submitted a Bayor College of Medicine invention disclosure titled "Using Minicircle DNA for Molecular Standards" with equal inventorship among all the collaborators, reiterating the strong partnership and shared contributions.

2.2 Author list and affiliations

Maduni Ranasinghe[1]*, Jonathan M. Fogg[2]*, Daniel J. Catanese, Jr.[3], Lynn Zechiedrich[2], Borries Demeler[1,4,#]

*Contributed equally

#Corresponding author email address: demeler@gmail.com

[1]Department of Chemistry and Biochemistry, University of Lethbridge, Alberta, T1K3M4, Canada

[2]Department of Molecular Virology and Microbiology, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Department of Pharmacology and Chemical Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
[3]Department of Biosciences, Rice University, 6100 Main St., Houston, TX 77005, USA
[4]Department of Chemistry and Biochemistry, University of Montana, Missoula, MT 59812, USA

2.3 Abstract

To address the current lack of validated molecular standards for analytical ultracentrifugation (AUC), we investigated the suitability of double-stranded DNA molecules. We compared the hydrodynamic properties of linear and circular DNA as a function of temperature. Negatively supercoiled, nicked, and linearized 333 and 339 bp minicircles were studied. We quantified the hydrodynamic properties of these DNAs at five different temperatures, ranging from 4 to 37 °C. To enhance the precision of our measurements, each sample was globally fitted over triplicates and four rotor speeds. The exceptional stability of DNA allowed each sample to be sedimented repeatedly over the course of several months without aggregation or degradation, and with excellent reproducibility. The sedimentation and diffusion coefficient distributions of linearized and nicked minicircle DNA demonstrated a highly homogeneous sample, and increased with temperature, indicating a decrease in friction. The sedimentation of linearized DNA was the slowest; supercoiled DNA sedimented the fastest. With increasing temperature, the supercoiled samples shifted to slower sedimentation, but sedimented faster than nicked minicircles. These results suggest that negatively supercoiled DNA becomes less compact at higher temperatures. The supercoiled minicircles, as purified from bacteria, displayed heterogeneity. Therefore, supercoiled DNA isolated from bacteria is unsuitable as a molecular standard. Linear and nicked samples are well suited as a molecular standard for AUC and have exceptional colloidal stability in an AUC cell. Even after sixty experiments at different speeds

and temperatures, measured over the course of 4 months, all topological states of DNA remained colloidal, and their concentrations remained essentially unchanged.

Keywords

Analytical ultracentrifugation; molecular standards; double-stranded DNA; DNA minicircle; hydrodynamics; global analysis

2.4 Introduction

Analytical ultracentrifugation (AUC) is a widely used technique to study biological macromolecules in the solution phase [1]. Hydrodynamic analyses have provided important information about the composition, interactions, and structure of biological macromolecules [2]. AUC has been considered essential for characterizing macromolecules in solution since its invention by Theodor Svedberg [3] nearly 100 years ago. A particular strength of AUC is the possibility to study molecules in a physiologically relevant solution environment, where size and shape distributions can be characterized as a function of temperature, sample concentration, and buffer conditions [4]. AUC data analysis is based on first principles to derive hydrodynamic properties such as sedimentation and diffusion coefficients, and partial concentration of multiple components in a mixture. The diffusion coefficient directly yields the frictional coefficient and the hydrodynamic radius of the molecule. If the partial specific volume (PSV) is known, the frictional ratio and molar mass can be determined as well. Experimental data can be modelled by finite element solutions of the Lamm equation [5, 6] to yield the sedimentation and diffusion coefficients, and partial concentrations directly.

Unlike other solution-based composition characterization methods such as size exclusion chromatography, gel electrophoresis, or field flow fractionation, AUC does not require reference

standards. This assertion is only true, however, if the instrument can be trusted to report accurate measurements, a question that has been investigated previously [7, 8]. Failures by the instrument to accurately report temperature, radial positioning, rotor speed, time, or wavelength affect the reliability of all parameters determined in an AUC measurement, and could limit the value of the measurements produced by the instrument. To address this limitation, a well-characterized molecular standard should be used to confirm the accuracy of the instrument, giving the investigator the necessary confidence that the instrument can be trusted for measurements of other systems. The requirements for an AUC molecular standard are multifold:

- stability for an extended period;
- resistance to degradation or aggregation;
- predictable behavior over the temperature range accessible with the instrument;
- homogeneous in composition;
- precisely known molar mass;
- no concentration-dependent non-ideality;
- ideal solution behavior at concentrations suitable for measurement in an AUC instrument;
- available at low cost to the scientific community;
- high degree of purity;
- properties can be readily modulated to explore a range of size and shape factors; and
- available orthogonal method (e.g., mass spectrometry, electrophoresis, sequencing) to confirm molecular properties of the standard such as molar mass and composition.

Although commercial standards with characterized hydrodynamic properties do not exist, bovine serum albumin (BSA) has been used previously in a multi-laboratory study to establish

reproducibility of AUC experiments [8]. BSA fits some of the criteria above. Importantly, in a suitable buffer it is stable at 4 °C for one year and, if sufficiently dilute, BSA can be observed under ideal solution conditions. Whereas BSA has a precisely known monomer mass, higher order oligomers occur that must be removed by extensive purification before a homogeneous preparation is obtained. Additionally, the stability of BSA at temperatures higher than 20 ºC during AUC has not been explored. Finally, with only one size and one shape, BSA does not offer variability in size and shape factors. These factors limit the utility of BSA for use in the validation of AUC instruments.

Some or all of these limitations exist for other proteins as well. Many proteins are less stable than BSA, and more prone to aggregation and degradation over extended periods of time. In addition, many proteins form oligomeric states that vary depending on experimental conditions. Furthermore, the standard method used to assess proteins, SDS polyacrylamide gel electrophoresis, is unsuitable as an orthogonal quantification method because SDS modifies both the conformation and the oligomerization state of proteins in solution [9, 10]. We propose that double-stranded DNA (hereafter referred to as DNA because we are not testing single-stranded DNA) satisfies the requirements for a molecular standard listed above, and here we investigate this proposition. Some of its unique additional benefits include:

- high stability under many conditions;
- precisely known molar mass and homogenous when not synthesized chemically (contaminating incomplete products would be problematic);
- high extinction coefficient at 260 nm for measurement under dilute and ideal solution conditions, and the ultraviolet (UV) extinction coefficient can be precisely calculated from the DNA sequence to support accurate quantification; and

- molar mass can be easily modulated by changing the length of the DNA molecule.

To investigate its suitability as a molecular standard, we chose two different lengths of DNAs (333 bp and 339 bp). Comparing DNA molecules of molar mass differing by less than 2 % tests the precision that can be obtained from AUC. We compared different topological states (negatively supercoiled, nicked [open-circle], and linearized) of these DNA molecules to elucidate the hydrodynamic differences imposed by conformation and temperature.

Negatively supercoiled DNA can be readily extracted from bacterial cells in high quantities using standard molecular biology techniques or using commercially available kits. If it satisfies the other requirements listed above, it should be suitable as a molecular standard. Supercoiled DNA is, however, structurally diverse [11] and likely fluctuates between different conformations. Although we are beginning to understand how to model the sedimentation of supercoiled DNA [12] we do not yet fully understand how to account for the conformational heterogeneity. Nicked and linearized samples require additional processing steps, but removing supercoiling may yield useful molecular standards.

We investigated DNA over a range of temperatures. Temperature is important for DNA nanoparticle formation, stability, flexibility, and interactions with architectural proteins [13]. Understanding how the solution behavior of DNA changes with temperature should yield insight into DNA at the nanoscale. Yet the effect of temperature on sedimentation and diffusion transport of DNA conformation and dynamics is a largely unexplored topic. To corroborate the AUC results with an orthogonal approach, we determined the composition of all DNA samples by polyacrylamide gel electrophoresis, which provides a very sensitive composition assessment [12].

2.5 Materials and methods

2.5.1 Chemicals and reagents

100 bp DNA ladder was purchased from New England Biolabs (Ipswich, MA, USA). Calcium chloride and sodium chloride were purchased from Sigma-Aldrich (St. Louis, MO, USA). Acrylamide was purchased from Fisher Scientific (Pittsburgh, PA, USA). SYBR Gold was purchased from ThermoFisher Scientific (Waltham, MA, USA). All other chemicals were purchased from VWR International (West Chester, PA, USA).

2.5.2 Generation and purification of minicircle DNA

Minicircle DNA was generated via λ-integrase mediated site-specific recombination as described previously [14, 15]. Plasmids pMC333 [15], and pMC339-BbvC1 [14] were used to generate 333 bp and 339 bp minicircles, respectively. Supercoiled, linearized, and nicked samples of each minicircle were prepared as described in [12]. Individual minicircle topoisomers (used as markers for electrophoresis) were prepared as described previously [11].

2.5.3 Gel electrophoresis

DNA samples were analyzed by gel electrophoresis through 5 % polyacrylamide gels (acrylamide: bis-acrylamide = 29:1) in Tris-acetate buffer (pH 8.2) containing 150 mM NaCl and 10 mM $CaCl_2$ at 40 V (~2 V/cm) for 25 hours. Buffer was continuously recirculated during electrophoresis. Electrophoresis was performed using a PROTEAN II xi cell (Bio-Rad, Hercules, CA, USA). Gels were run at 4 °C (in a walk-in cold room), room temperature (~23 °C) (in the laboratory), or 37 °C. The electrophoretic apparatus (including the gel and running buffer) was pre-chilled (for 24 hours) or preheated (for 5 hours) and the temperature of the recirculating water bath water was verified before electrophoresis. For electrophoresis at 37 °C, the temperature was

maintained by circulating 37 °C water from a recirculating water bath through the central core of the electrophoretic apparatus. When electrophoresis was complete, gels were subsequently stained with SYBR Gold then visualized using a FOTO/ANALYST Investigator imaging system and quantified using ImageQuant TL, version 8.1 (GE Healthcare Life Sciences, Marlborough, MA, USA). The identity of the topoisomers was determined by comparison to known purified topoisomer markers. The topoisomer distributions were determined by quantitation of the digital images of fluorescently stained gels using image analysis software as described previously [12].

2.5.4 AUC experimental design

DNA samples were analyzed by sedimentation velocity analytical ultracentrifugation (SV AUC) in a Beckman Coulter Optima AUC™ at the Canadian Center for Hydrodynamics at the University of Lethbridge in Alberta, Canada. All studies were conducted using an AN50Ti rotor. Approximately 450 µl of each DNA sample (A_{260} = 0.42–0.79, see Figure 2.1 (a) and (b)) in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM $CaCl_2$ buffer was loaded into each sector of a 12 mm double-sector epon charcoal centerpiece (Beckman Coulter, Indianapolis, USA). Data were collected in intensity mode at 260 nm using the UV/visible optics. Each sample was measured at five different temperatures of 4, 12, 20, 28, and 37 °C, in this order. At each of these temperatures, the samples were measured at four different rotor speeds of 15, 20, 26.7, and 35.6 krpm, in this order, and each of these experiments was performed in triplicate. After data collection at each temperature, speed, and replicate was completed, the AUC cells were shaken to redistribute the DNA samples thoroughly before the next run. The density and viscosity of the buffer at each temperature were calculated with UltraScan-III, version 4.0, release 6612 [16] (https://ultrascan.aucsolutions.com).

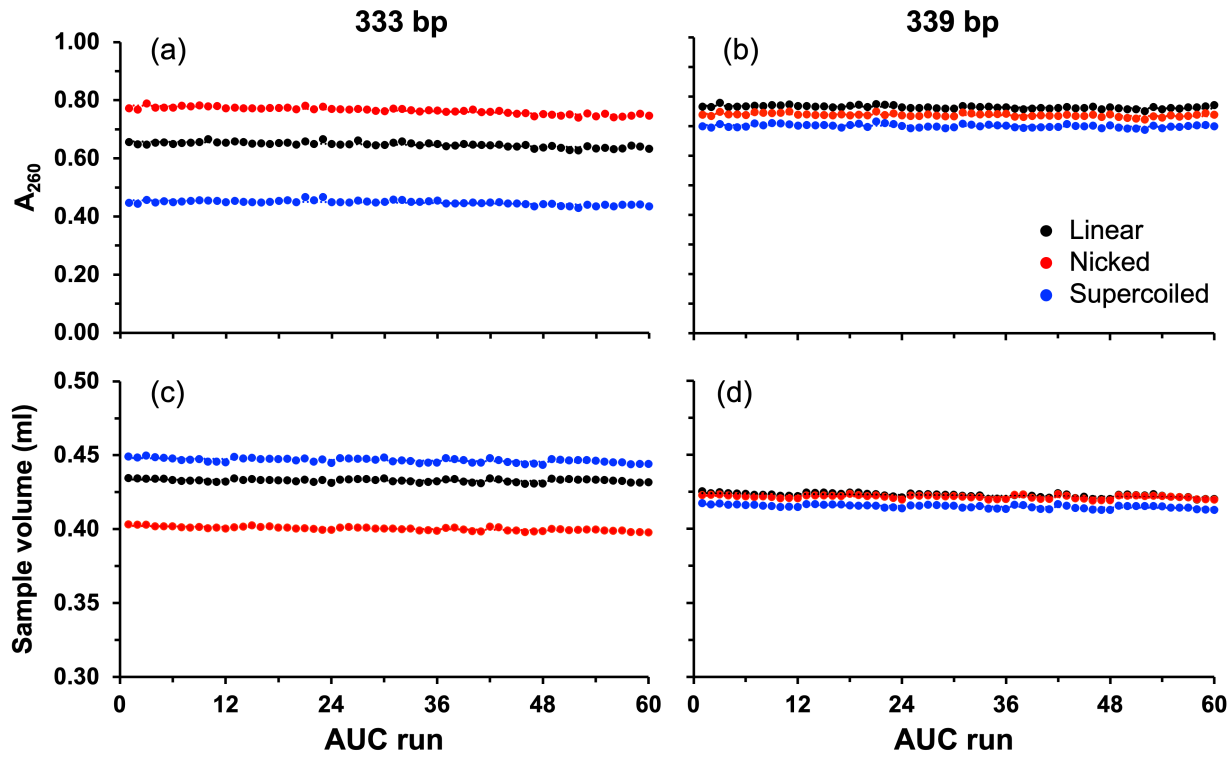


Figure 2.1 Effect of time on total concentration and volume of DNA. The total concentrations (top, for (a) 333 bp and (b) 339 bp minicircle DNA) determined directly from the fitted iterative 2DSA models, and volumes (bottom, (c) 333 bp and (d) 339 bp minicircle DNA) for each topoisomer (black=linearized, red=nicked, blue=negatively supercoiled) over all experiments performed over a period of four months and 60 runs. Optical density measurements were employed for quantification.

2.5.5 AUC Data analysis of individual datasets

All data were analyzed with UltraScan-III, version 4.0 (6612) [16] and were collected using the UltraScan data acquisition module [17]. The data were first imported and converted into UltraScan-III OpenAUC format [18]. Parallel distributed data analysis was performed on the UltraScan Science Gateway using XSEDE resources (Expanse, Bridges-2, Anvil) and on the Chinook cluster at the University of Lethbridge [19]. Two-dimensional spectrum analysis (2DSA) [20] was used to remove time- and radially- invariant noise, and to fit the meniscus

position. A final 2DSA refinement with up to ten iterations was obtained for each dataset. The models resulting from the final 2DSA fit were used to initialize a genetic algorithm analysis (GA) [21], which was followed by a Monte Carlo analysis (GA-MC) [22] to obtain 95 % confidence intervals for the sedimentation and diffusion coefficients, the partial concentration of each solute in a dataset, and to quantify the effect of experimental stochastic noise on the obtained parameter values.

The DNA loading concentrations and volumes for each sample were tracked across all experiments. Loading concentrations were obtained directly from the fitted iterative 2DSA models. The sample volumes were calculated from the fitted meniscus positions and the positions at the cell bottom. All centerpieces used have a cell bottom position of 7.2 cm when at rest. Due to rotor stretch, which varies as a function of rotor speed, this position moves outward during the experiment. Using the rotor stretch calibrations performed in UltraScan [7], the stretch value was calculated and subtracted from the meniscus position to obtain the channel positions adjusted for a rotor at rest. The volume V was then calculated according to Equ. 2.1, where α is the sector angle (2.5 °), h is the pathlength (1.2 cm), b the cell bottom at rest (7.2 cm), and r_m is the speed-dependent rotor stretch adjusted meniscus position.

$$V = \frac{\alpha}{360^o} h\pi\left(b^2 - r_m^2\right)$$

Equ. 2.1

2.5.6 Global analysis

The signal generated by each sample for each of the fitted parameters must be considered when selecting a rotor speed for a sedimentation velocity (SV) experiment. The observed signals for the sedimentation and the diffusion transport are dependent on the rotor speed. At high rotor

speeds, sedimentation resolution is favored. A fast-sedimenting particle, however, will spend less time diffusing before being pelleted, which limits the diffusion signal that can be observed [23, 24]. Hence, the diffusion signal is enhanced at lower rotor speeds. The optimal information from both transport parameters can be combined by performing a global analysis over both high-speed and low-speed experiments at the same temperature, and by including replicate measurements in the global fit to maximize precision. UltraScan supports the global fitting of datasets from multiple experimental conditions and can be used with either the 2DSA or GA analysis.

In a global analysis, multiple SV AUC datasets are globally fitted, which enhances signal-to-noise ratios and improves the confidence limits for the determined hydrodynamic parameters [16]. This global fit generates $2n + 1$ models for n datasets. The super-global model represents the best fit global model for all datasets. For each dataset, scaled and variable ratio models are also created. In a scaled model, the super-global model is scaled to match the total concentration of each individual dataset. The scaled model has the same set of solutes as the global model, as well as the same ratios between each solute as the best-fit global model, but solute concentrations are scaled to match to total concentration observed in an individual sample. A variable ratio model allows the ratio between solutes to vary to account for micro-variation between solutes, as would be observed in a reversible association. In the variable ratio model, the solutes remain identical to those present in the super-global model.

For the homogeneous linearized and nicked minicircle DNA used in this study, global models were created for each sample at each temperature by combining corresponding GA-MC models for that sample performed at different speeds. This combined model was used to initialize a global GA fit. The resulting GA model was then used to initialize a global GA-MC analysis to obtain the sedimentation and diffusion coefficients for linearized and nicked minicircle DNA, and

to determine the confidence intervals of each parameter. Heterogeneity was not observed in the SV experiments of linearized and nicked samples. The SV data from negatively supercoiled samples was not appropriate for a global GA-MC analysis due to heterogeneity in composition. Instead, for negatively supercoiled samples, a global 2DSA, followed by a 2DSA-MC [20] was used to globally fit all replicates and speeds for a single temperature. In all cases, the super-global model results are reported. RMSD values were checked by inspecting the variable ratio models, to ensure the obtained fit reports random residuals for each individual dataset. By applying global analysis, the obtained model must fit many datasets under multiple conditions equally well, reducing the chance for over-fitting, and increasing the signal to noise ratio.

2.5.7 Hydrodynamic properties

SV experiments were fitted to the Lamm equation [25] to obtain sedimentation (*s*, Equ. 2.2) and diffusion (*D*, Equ. 2.3) coefficients. *s* depends on the molar mass *M*, the frictional coefficient *f,* and the buoyancy of the particle, which is a function of its partial specific volume (PSV), $\bar{v}$, and solvent density ρ. Both *s* and *D* are inversely proportional to the frictional coefficient *f* of the solute.

$$s = \frac{M(1-\bar{v}\rho)}{Nf} \qquad \text{Equ. 2.2}$$

$$D = \frac{RT}{Nf} \qquad \text{Equ. 2.3}$$

where *N* is Avogadro's number, and *R* is the universal gas constant. The molar masses are 205,721 Da and 209,429 Da for the 333 bp and 339 bp minicircles, respectively, as calculated from the DNA sequences [26] using http://molbiotools.com/dnacalculator. To aid comparisons of AUC experiments performed at different temperatures and buffer conditions, *s* and *D* are reported after correction to standard conditions (20 °C in water, see Equ. 2.4 and Equ. 2.5).

$$s_{20,W} = s_{TB}\frac{(1-\bar{\nu}\rho)_{20,W}\eta_{TB}}{(1-\bar{\nu}\rho)_{T,B}\eta_{20,W}} \qquad D_{20,W} = D_{T,B}\frac{293.15\ \eta_{T,B}}{T\ \eta_{20,W}}$$

Equ. 2.4 Equ. 2.5

$s_{20,W}$ correction requires knowledge of the PSV of DNA both at 20 °C in water, and at the experimental temperature in the buffer. Obtaining both values accurately requires instrumentation and sample amounts not available to us. For the purposes of this study, a constant PSV value was used for all temperatures, which is supported by earlier studies performed on DNA [27] that indicated a negligible variation of the PSV as a function of temperature. A PSV value for double-stranded DNA in water of 0.55 ml/g as reported earlier [27] was used here to correct to standard conditions. The Stokes-Einstein relationship relates the frictional coefficient, f_0, of a solid sphere to the radius R_0 of the sphere in a solvent with viscosity η (Equ. 2.6).

$$f_0 = 6\pi\ \eta R_0$$

Equ. 2.6

For any molecule, spherical or not, the hydrodynamic radius, R_h, is the radius of a sphere consistent with the measured frictional coefficient (Equ. 2.7).

$$R_h = \frac{f}{6\pi\eta}$$

Equ. 2.7

Substituting Equ. 2.3 into Equ. 2.7 yields the hydrodynamic radius, R_h, directly from the diffusion coefficient (Equ. 2.8).

$$R_h = \frac{RT}{6\pi\ \eta ND}$$

Equ. 2.8

2.6 Results

2.6.1 Rationale

The overall goal of our study was to evaluate DNA as a suitable molecular standard for AUC instruments. To explore the effect of temperature, shape, and mass on the hydrodynamics of DNA, we measured two different lengths (333 bp and 339 bp) of linearized, nicked (open circle) and negatively supercoiled DNA by AUC over the temperature range accessible by the instrument (4 ºC–37 ºC). Because AUC is sensitive to molar mass and conformation, the two sizes allow us to explore the impact of a small (six bp; ~0.6 helical turns) difference in length. Changes in topology and conformation will result in different frictional coefficients, which can also be measured by AUC.

Because AUC requires more material than other methods, such as electron cryotomography (cryoET), atomic force microscopy, or gel electrophoresis, we used negatively supercoiled minicircle DNA obtained from bacterial cells. Although it is possible to further modify the supercoiling post-extraction, and to purify individual topoisomers by preparative polyacrylamide gel electrophoresis, very few laboratories are equipped with the skills, experience, or equipment to do so, especially in the quantities required for AUC. If supercoiled DNA is to be used as an AUC molecular standard, it should be the same level of supercoiling as obtained from bacteria, without the need for additional manipulation.

The molecular mass of these minicircles (~206–209 kDa) falls in the middle of the range that can be interrogated using AUC, allowing meaningful data to be acquired over a wide range of rotor speeds, a benefit for a molecular standard. Previously, we performed AUC experiments at 20 ºC on minicircles of a similar length (336 bp) to validate theoretical models capable of

predicting their hydrodynamic properties [12]. Changing temperature changes DNA flexibility [13] and alters the helical repeat [28, 29], which will likely affect supercoiling-dependent conformations. Correcting for these effects may be non-trivial, possibly limiting the utility of supercoiled DNA as a molecular standard. Thus, we also tested linear and nicked (open circle) DNA samples that lacked supercoiling.

2.6.2 Composition analysis by polyacrylamide electrophoresis

Supercoiling in circular DNA molecules can be described by comparing the linking number (*Lk*), the total number of times the two single DNA strands coil about each other, to Lk_0, the *Lk* for a fully relaxed DNA molecule ($\Delta Lk = Lk - Lk_0$). Lk_0 is equal to the number of base pairs divided by the helical repeat. The helical repeat is ~10.45 bp/turn at room temperature (~23 ºC) under the conditions used for AUC (150 mM NaCl and 10 mM $CaCl_2$) (Fogg and Zechiedrich, unpublished results). This helical repeat value translates to $Lk_0 = 31.8$ for the 333 bp minicircle and $Lk_0 = 32.4$ for the 339 bp minicircle under these conditions. For simplicity, we hereby refer to ΔLk values rounded to the nearest integer. Actual (not rounded) values can be found in Supplementary Table 2.1.

Because of their short lengths, 333 and 339 bp minicircles adopt a narrow distribution of topoisomers reflecting the steady-state level of supercoiling in bacteria (Figure 2.2). The 333 bp topoisomer distribution was primarily (73 ± 1 %) $\Delta Lk = -3$, as well as 13 ± 1 % $\Delta Lk = -2$ and 13 ± 1 % $\Delta Lk = -1$ topoisomers. The negatively supercoiled 333 bp sample also contained a trace amount (1 %) of supercoiled 666 bp minicircle (double-length minicircle). The 339 bp supercoiled sample contained primarily (78 ± 3 %) $\Delta Lk = -2$, and also 10 ± 1 % $\Delta Lk = -1$, 8 ± 2 % $\Delta Lk = 0$, and 4 ± 1 % $\Delta Lk = -3$ topoisomers. Percentages reported for each species are the

mean values ± standard deviation from three separate gels. Because the linking number cannot be changed without breaking one or more of the DNA strands, the topoisomer distribution does not change with temperature or over time.

Although the linking number is unchangeable, the helical repeat of DNA depends on temperature [28, 29]. Lk_0, determined by the helical repeat, and consequently ΔLk, will both change with temperature. In particular, the negatively supercoiled samples become less supercoiled with increasing temperature. To allow comparison of different-sized molecules, ΔLk was scaled to the DNA length to give the superhelical density (σ), which is equal to $\Delta Lk/Lk_0$. The estimated values of σ at each temperature are listed in Supplementary Table 2.2. It should be emphasized that these values are estimates, not direct measurements. Over the full range of temperatures used in our study, we predict a small but significant temperature-dependent effect on σ, which should be taken into consideration when comparing data at different temperatures. For linear and nicked molecules, the DNA ends are not constrained and the Lk of these molecules can freely change without strand breakage. Thus, the effects of changes in helical repeat at different temperatures should be mitigated.

Electrophoresis experiments demonstrated that the linear and nicked samples consisted of single species with high purity (95–100 %) (Figure 2.2). Thus, a single enzymatic step (linearization or nicking) is sufficient to remove the heterogeneity and produce homogenous samples, which may be better as molecular standards for AUC.

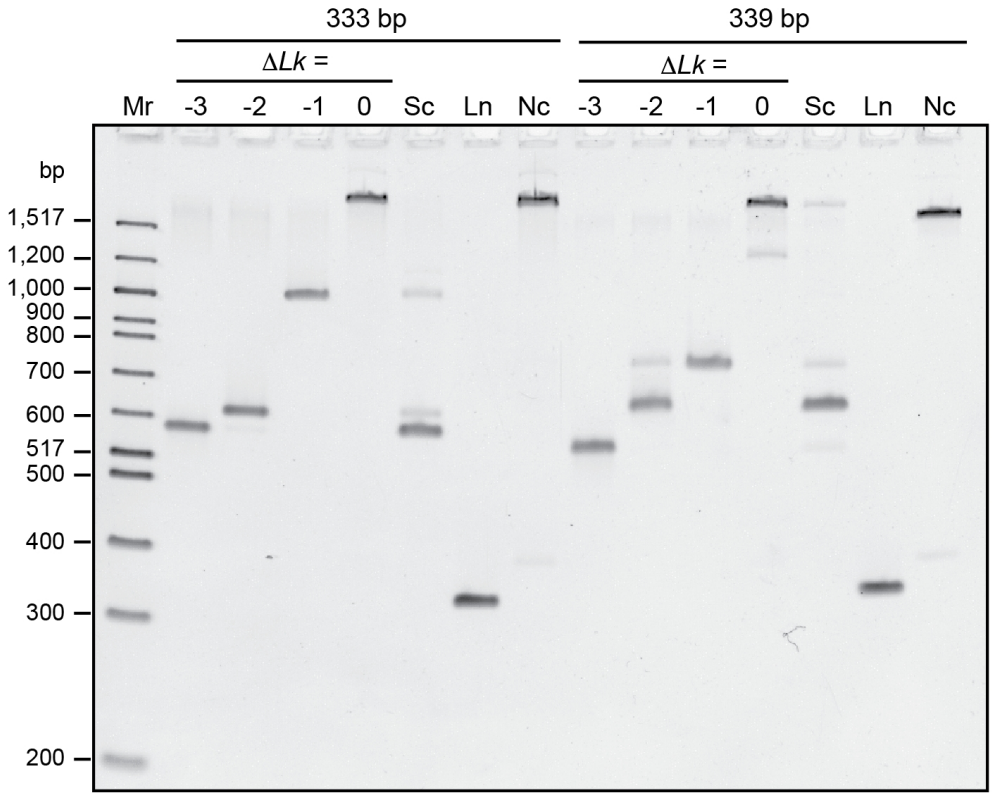


Figure 2.2 Composition of DNA samples determined by electrophoresis. DNA samples were analyzed by polyacrylamide gel electrophoresis (5% polyacrylamide) in 150 mM NaCl and 10 mM $CaCl_2$ (to replicate the conditions used in analytical ultracentrifugation). Mr: 100 bp DNA ladder (linear DNA, lengths in bp as indicated). 333 and 339 bp minicircle topoisomer markers (Δ*Lk* as indicated) were also loaded. Negatively supercoiled (Sc), linear (Ln), and nicked (Nc) 333 bp and 339 bp DNA samples were analyzed. The identity of the topoisomers in the supercoiled samples was determined by comparison to the topoisomer markers.

To be suitable as a molecular standard for AUC, the composition and conformation of a molecule in question should be consistent across a wide range of temperatures. To determine whether this was the case for the DNA samples, gel electrophoresis was also performed at 4 ºC and 37 ºC (Supplementary Figure 2.1). DNA samples migrated much more slowly at 4 ºC than at higher temperatures. The distances migrated by the linear samples were measured and plotted against temperature (Supp fig 2 (a)Supplementary Figure 2.2 (a)), revealing a linear relationship. This temperature dependence results from buffer viscosity decreasing with increasing

temperature [30], which allows for faster migration. To normalize for this generalized effect of temperature, distances migrated by the other samples were compared to the distance migrated by the linear samples. This normalization facilitated comparison among the gels run at different temperatures. Nicked samples migrated much more slowly than the linear and supercoiled samples at all temperatures. There were some subtle changes in the relative migration of the nicked minicircles at different temperatures (Supplementary Figure 2.2 (b)), but these differences are relatively small in comparison to the difference between nicked, linear, and supercoiled samples. For the supercoiled samples, there was little change in the relative migration of the major species within each sample (Supplementary Figure 2.2 (c) and (d)).

The electrophoresis results confirm that the composition of the DNA samples does not change with temperature. Any differences in relative electrophoretic migration were relatively subtle suggesting that the global conformation also does not significantly change, even for the supercoiled samples. The ability to use an orthogonal technique to verify composition and conformation meets another requirement for the suitability of DNA samples as a molecular standard for AUC.

2.6.3 Investigating the suitability of DNA as a molecular standard

As described above, a molecular standard for AUC must be stable for an extended period of time, produce meaningful data at a wide range of different rotor speeds, and exhibit predictable hydrodynamic behavior across a wide range of temperatures. We investigated whether DNA would meet these requirements by testing each DNA sample repeatedly at five different rotor speeds and five different temperatures. Each experiment was performed in triplicate and fitted globally over all replicates and rotor speeds. The DNA samples remained in the AUC cells

throughout the course of these experiments and were resuspended by shaking the cell before each subsequent AUC run. To ascertain that DNA did not aggregate, degrade, or become lost due to irreversible pelleting, or sticking to cell windows or centerpiece surfaces, we monitored the sample composition, DNA concentration (Figure 2.1 (a) and (b)), and sample volume (Figure 2.1 (c) and (d)) over time. Linear fit coefficients and the percent sample loss as a function of concentration and volume for each sample over the course of all experiments are shown in Table 2.1. The fits indicated a negligible drift in the total concentration, which was < 5 % for all samples over the course of the measurements, which included 60 runs with the same sample, collected over four months, five temperatures, four rotor speeds, and three replicates for each condition. Volume changes were negligible as well (Table 2.1). These data attest to the superior stability of double-stranded DNA of these lengths under these conditions in contrast to other biological macromolecules, such as RNA which is more susceptible to hydrolysis, or proteins prone to proteolysis, aggregation or denaturation at higher temperatures. It further suggests the suitability of DNA as a molecular standard for AUC experiments.

Table 2.1. DNA sample and volume loss over four months

		A_{260}		Sample volume	
Sample		Slope	Sample loss, %	Slope	Volume loss, %
333 bp	Linear	-0.0004	3.58	-0.00002	0.27
	Nicked	-0.0006	4.52	-0.00006	0.88
	Supercoiled	-0.0003	3.88	-0.00005	0.66
339 bp	Linear	-0.0002	1.53	-0.00005	0.70
	Nicked	-0.0001	0.84	-0.00003	0.42
	Supercoiled	-0.0002	1.59	-0.00004	0.57

2.6.4 Effect of temperature, circularity, and supercoiling on hydrodynamic parameters of DNA

The sedimentation coefficients, diffusion coefficients, and hydrodynamic radii of each linearized and nicked minicircle DNA sample are shown in Figure 2.3 as a function of temperature. With increasing temperature, the sedimentation and diffusion coefficients of all linearized and nicked samples increased while the hydrodynamic radii decreased. The changes in the sedimentation and diffusion coefficients as a function of temperature appeared to follow a linear relationship. Fitting to a straight line resulted in R^2 values close to 1 for most of the fits (Table 2.2). The PSV of DNA does not change significantly as a function of temperature [27] or topology [12]. Because the PSV and the molar mass of the DNA samples remain constant (across the different forms—linearized, nicked, and supercoiled), the change in sedimentation and diffusion coefficients of each sample, when corrected to standard conditions, must be a consequence of changes in the frictional coefficient caused by changes in the conformation.

At each temperature, a clear difference was observed between linearized and nicked minicircles with the linearized DNA samples always sedimenting slower than the nicked. This result is consistent with earlier studies showing circular DNA has a reduced radius of gyration relative to linear DNA [31]. These observations suggest that linear and nicked DNA are more rigid and extended at lower temperatures and become more compact or flexible at higher temperatures. Previous research showed that increasing temperature decreases the persistence length of linear duplex DNA [32], which is consistent with our observations. The persistence length is a measure of polymer stiffness, suggesting that the molecule becomes more flexible and less elongated with increasing temperature. Previous nuclear magnetic resonance experiments on

24 bp linear duplex DNA determined that the diffusion coefficient of short linear DNA increases with increasing temperature [33], which is also consistent with our results.

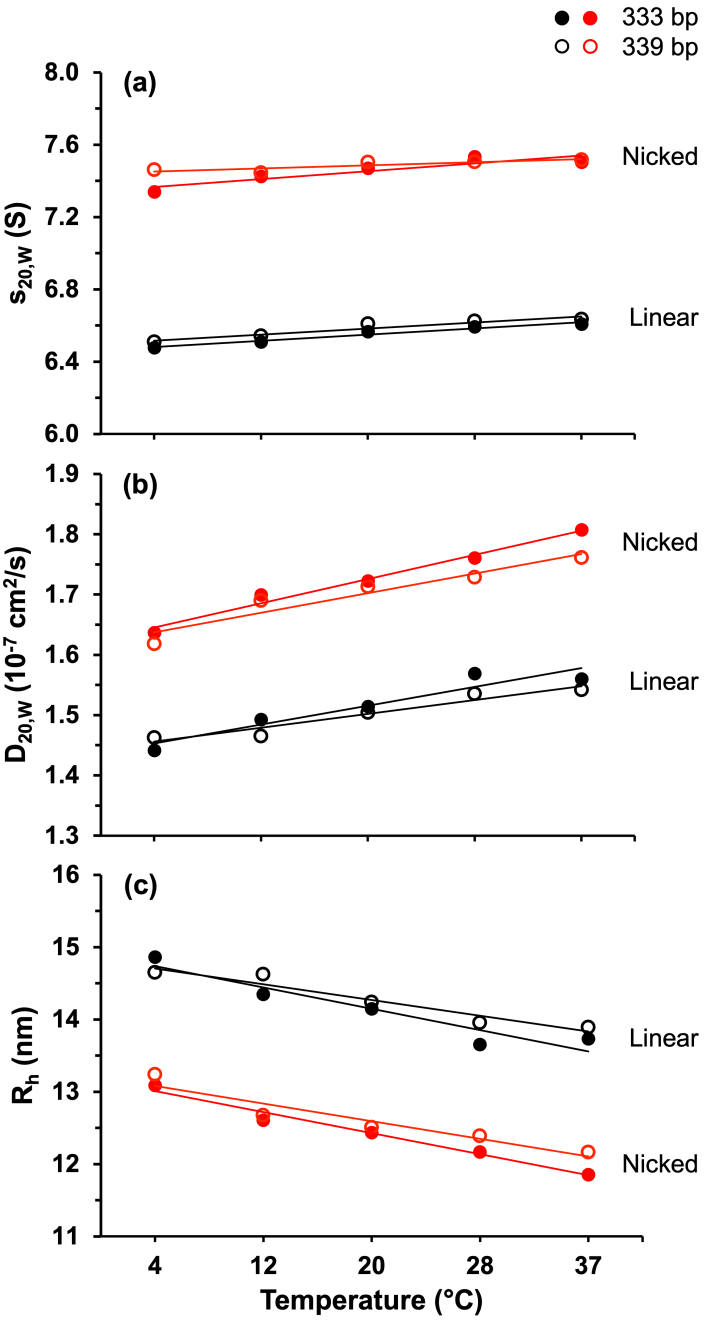


Figure 2.3. Hydrodynamic measurements of DNA. Hydrodynamic measurements of linearized (black) and nicked (red) minicircle DNA of 333 bp (filled circles) or 339 bp (open circles) length as a function of temperature for $s_{20,w}$ (top panel), $D_{20,w}$ (center panel), and R_h (bottom panel).

Table 2.2. Linear regression data for sedimentation and diffusion coefficients as a function of temperature.

	333 bp				339 bp			
	Linear		Nicked		Linear		Nicked	
	Slope	R^2	Slope	R^2	Slope	R^2	Slope	R^2
$s_{20,W}$	0.034 s/ºC	0.96	0.044 s/ºC	0.83	0.033 s/ºC	0.91	0.017 s/ºC	0.74
$D_{20,w}$	0.031 cm^2/(s ºC)	0.91	0.040 cm^2/(s ºC)	0.98	0.023 cm^2/(s ºC)	0.93	0.033 cm^2/(s ºC)	0.92
R_h	-0.295 nm/ºC	0.90	-0.291 nm/ºC	0.98	-0.218 nm/ºC	0.93	-0.243 nm/ºC	0.90

We observed that the rate of change of sedimentation and diffusion coefficients as a function of temperature for the linear and nicked DNAs were greater for the shorter DNA length (333 bp). While these changes were subtle, they were significant (see Figure 2.3). Although the molar masses of the two lengths of minicircles (333 and 339 bp) differ by less than 2 %, their differences in hydrodynamic properties demonstrate the sensitivity of AUC.

Although heterogenous, the supercoiled samples consisted primarily (> 70 %) of a single topoisomer (Figure 2.2). Moreover, we previously showed that small differences in *Lk* have only small effects on sedimentation [12]. At 4 ºC, the first temperature evaluated, we observed a primarily homogeneous behavior (see Figure 2.4). We therefore hoped that supercoiled DNA, as isolated out of bacteria, may have utility as a molecular standard. Overall, however, the sedimentation behavior of the supercoiled samples, and their response to temperature was much more complex than we had anticipated (Figure 2.4, temperatures above 4 ºC). The results reveal that supercoiled samples, when isolated directly from bacterial cells, are not suitable as a molecular standard. Nevertheless, the observations provide new insight into the behavior of supercoiled DNA.

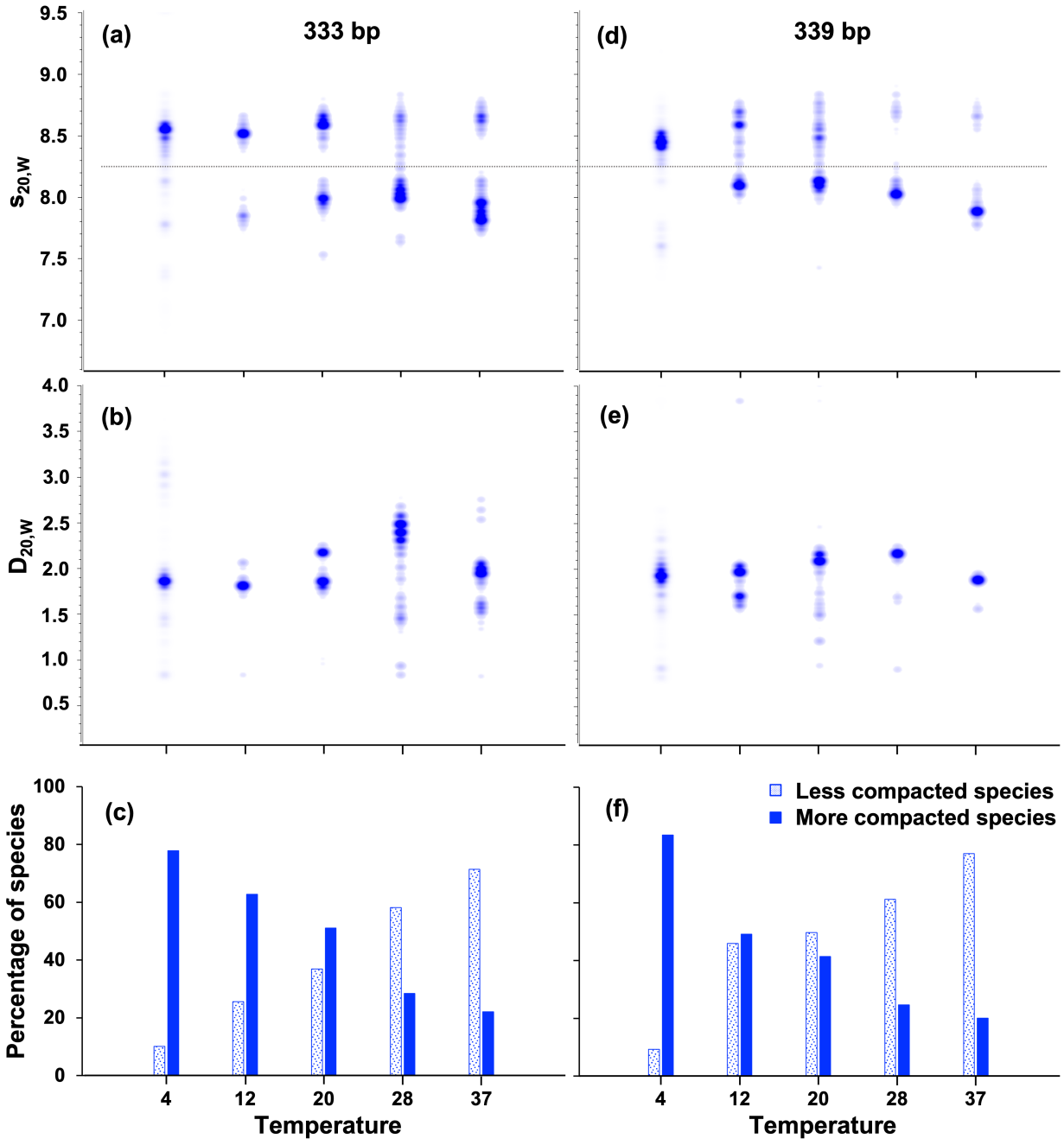


Figure 2.4. Sedimentation and diffusion coefficient distributions of DNA. Sedimentation coefficient distributions (top), diffusion coefficient distributions (center), and relative percentage of more compact (solid texture) and less compact (pattern texture) supercoiled minicircles. 333 bp minicircle results are shown in the left panels, and 339 bp supercoiled minicircle results are shown in the right panels. The horizontal line in panels (a) and (d) denotes the integration boundary for the integrated groups shown in (c) and (f).

Whereas the linearized and nicked minicircle DNA samples were sufficiently homogeneous to be analyzed by global GA analysis, the negatively supercoiled samples of multiple different topoisomers were best modeled by a global 2DSA because their heterogeneity was too great to lend themselves to parsimonious regularization. The global 2DSA analysis revealed that multiple species with differing sedimentation coefficients were present in the supercoiled samples (Figure 2.4). Two distinct peaks were observed in the sedimentation coefficient distributions (Figure 2.5). One of the peaks with a higher sedimentation coefficient was predominant at lower temperatures and presumably represents a more compact conformation. With increasing temperature, a second peak with lower sedimentation coefficient emerged. Although this peak presumably represents a less compact conformation, its sedimentation coefficient is still higher than that for nicked DNA, suggesting that supercoiled DNA may shift to more open conformations at higher temperatures, but never becomes as open as nicked DNA. These two groups and this temperature effect were observed for both the 333 bp and the 339 bp minicircles. To quantify the relative change in concentration for each major peak shown in the supercoiled sample of Figure 2.5, we integrated each peak ($s_{20,W}$ integration limits: 7.063–8.225 s and 8.268–9.134 s). This integration boundary is shown as a dashed line in Figure 2.4 (a) and (d). The relative concentration of each species changed from a predominantly fast sedimenting species at low temperatures to a significantly more populated species with a lower sedimentation rate at higher temperatures (Figure 2.4 (c) and (f)).

To facilitate comparison with the linear and nicked samples, weighted averages of the sedimentation coefficients of the supercoiled samples were plotted as a function of temperature (see Supplementary Figure 2.3). Whereas the sedimentation coefficients of the linear and nicked

samples increased with temperature, the opposite was the case for the supercoiled samples and the weighted average sedimentation coefficients decreased with temperature.

While the sedimentation coefficient results obtained for the negatively supercoiled minicircle samples provided a clear shifting pattern to slower sedimentation values as the temperature was increased, the diffusion coefficients (Figure 2.4 (b) and (e)) were difficult to distinguish, and individual species (resolvable by gel electrophoresis) could not be resolved. Diffusion coefficients are derived from the fitted slope of the sedimenting boundary, which is cross-correlated with the heterogeneity in the boundary. Boundary slopes from multiple species tend to merge into boundaries with average slopes, decreasing the resolution of the diffusion signal in heterogeneous samples.

A complete list of the sedimentation and diffusion coefficient values for the linear, nicked, and supercoiled samples at at each temperature can be found in Table 2.3. For the linear and nicked samples there was very high precision in these measured values. For the supercoiling samples the values reported with 95% confidence levels covered a broader range for each measurement.

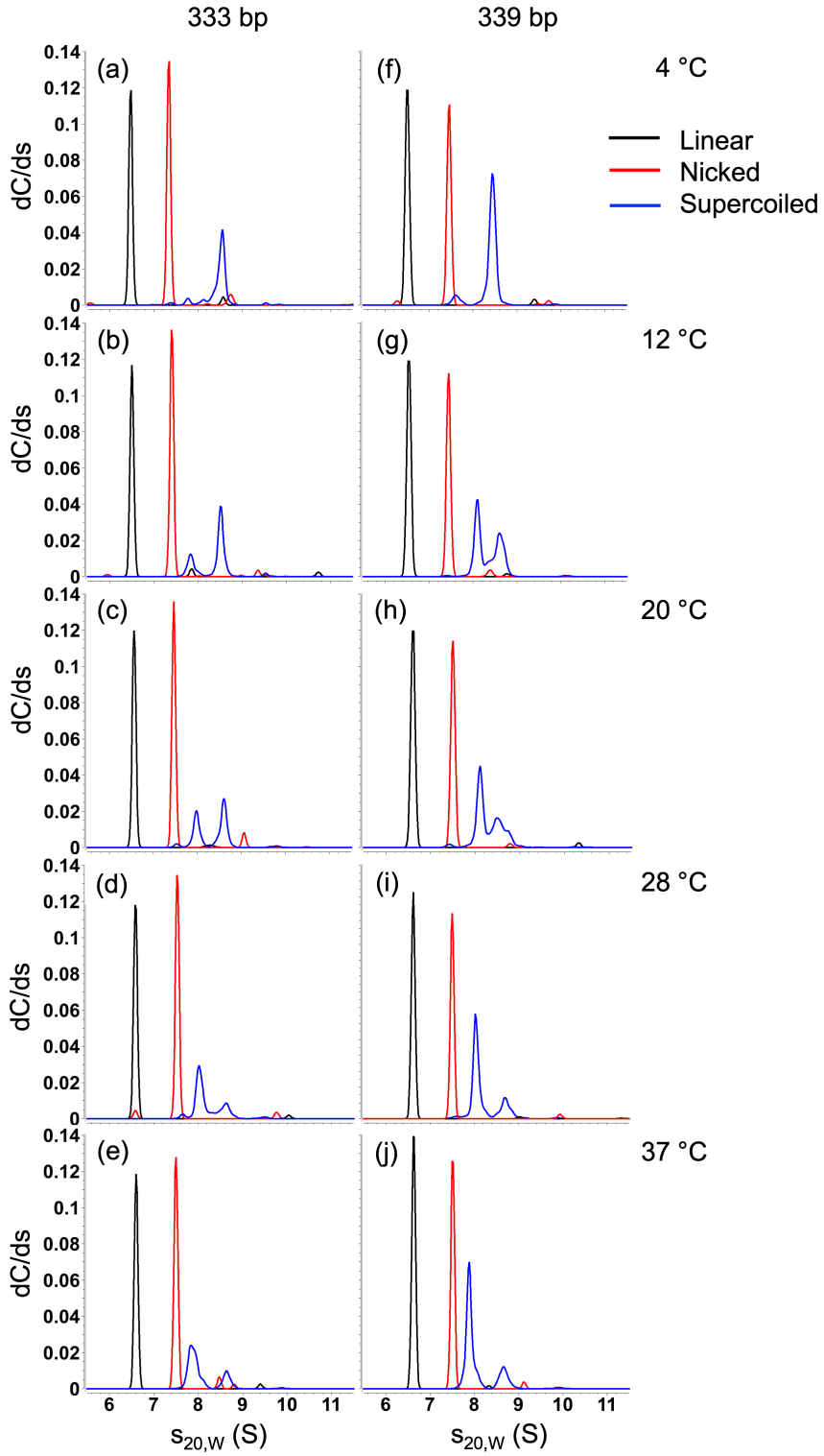


Figure 2.5. Effect on temperature on sedimentation coefficient distributions of DNA. Sedimentation coefficients of linearized (black), nicked (red) and supercoiled (blue) 333 bp (left) and 339 bp (right) DNA minicircles as a function of temperature.

Table 2.3. Sedimentation and diffusion coefficients of minicircle topoisomers with 95% confidence limits. Where confidence limits are not shown, the confidence limits were exactly zero as assessed by Monte Carlo analysis. The values of linear and nicked samples are obtained from global GA-MC models, whereas the values of supercoiled samples are obtained from global 2DSA-MC models.

Sample			(°C)	$s_{20,W}$ (S)	$D_{20,W}$ (10^{-7} cm^2 s^{-1})
333 bp	Linear		4	6.48	1.44
			12	6.51	1.49
			20	6.57	1.51
			28	6.59	1.57
			37	6.61	1.56
	Nicked		4	7.34	1.64
			12	7.42	1.70
			20	7.47	1.72
			28	7.53	1.76
			37	7.50	1.81
	Supercoiled	Less compact species	4	7.84 ± 0.63	1.61 ± 1.76
			12	7.86 ± 0.39	1.95 ± 0.38
			20	7.99 ± 0.29	2.17 ± 0.73
			28	8.03 ± 0.27	2.40 ± 0.75
			37	7.90 ± 0.41	1.97 ± 0.44
		More compact species	4	8.52 ± 0.30	2.00 ± 1.28
			12	8.52 ± 0.34	1.81 ± 0.90
			20	8.60 ± 0.32	1.85 ± 0.52
			28	8.56 ± 0.39	1.56 ± 0.61
			37	8.65 ± 0.30	1.55 ± 0.29
Sample			(°C)	$s_{20,W}$ (S)	$D_{20,W}$ (10^{-7} cm^2 s^{-1})
339 bp	Linear		4	6.51	1.46
			12	6.54	1.47
			20	6.61	1.50
			28	6.62	1.54
			37	6.63	1.54
	Nicked		4	7.46	1.62
			12	7.44	1.69
			20	7.50	1.71
			28	7.50	1.73
			37	7.52	1.76
	Supercoiled	Less compact species	4	7.71 ± 0.47	1.63 ± 0.50
			12	8.09 ± 0.34	1.99 ± 0.49
			20	8.10 ± 0.33	2.10 ± 0.51
			28	8.03 ± 0.31	2.16 ± 0.56
			37	7.90 ± 0.39	1.89 ± 0.31
		More compact species	4	8.45 ± 0.32	1.93 ± 0.93
			12	8.57 ± 0.39	1.72 ± 0.42
			20	8.57 ± 0.44	1.71 ± 0.60
			28	8.68 ± 0.44	1.69 ± 0.36
			37	8.68 ± 0.37	1.55 ± 0.34

2.7 Discussion

The goal of this study was to assess the suitability of different isomeric forms of double-stranded DNA as molecular standards for validating AUC instruments. Linearized and nicked circular samples demonstrated a very high degree of homogeneity, with highly reproducible sedimentation and diffusion coefficients. The effect of temperature on sedimentation and diffusion of these samples appeared to follow a linear relationship. Global analysis included triplicate measurements at five different speeds, reducing the 95 % confidence bands of each parameter value to near zero, producing very narrow sedimentation/diffusion coefficient distributions. Sedimentation and diffusion coefficients increased slightly with temperature, with distinct differences in both parameters between the two different minicircles. These parameter differences were greater at higher temperatures, and less noticeable at lower temperatures. Their stability in the AUC experiments, tested over four months on the same sample, resuspended by shaking the cell, and measured over 60 times at multiple rotor speeds, triplicate measurements, and temperatures ranging between 4 ºC and 37 ºC proved to be exceptional. No significant volume or concentration loss could be measured for any topoisomer, but linearized and nicked circles provided better homogeneity than supercoiled minicircles. Together, these results suggest that both linearized and open circular DNA molecules are quite suitable as a molecular standard, and satisfy all of the requirements.

Supercoiled samples isolated from bacterial cultures failed to satisfy the requirements for molecular standards. These heterogenous supercoiled samples yielded results that are complex and difficult to interpret. We had previously shown that purified supercoiled minicircle topoisomers with defined ΔLk adopt multiple conformations [11]. This result suggested that

individual negatively supercoiled minicircles may interconvert among the different conformations over time. The rate at which this possible conformational interchange takes place is not yet clear. Molecular dynamics revealed a possible pathway for interconversion, but could not provide accurate kinetics because of the way the solvent was simulated [11]. Future studies will be important in interpreting the species seen in AUC. The shift from a single major peak in the negatively supercoiled samples at lower temperatures to multiple peaks at higher temperatures suggests that the conformational equilibrium is shifted towards more open conformations. The fact that we observe differently sedimenting species suggests that the different conformations have sufficiently long lifetimes to be detected. If the interconversion is rapid, a bulk observation like AUC would report a weight averaged sedimentation profile for all interconverting conformations.

Electrophoresis and AUC are orthogonal approaches. The temperature dependent effects observed in the AUC data were not reflected in the electrophoresis data. Electrophoresis and AUC measure different parameters. The separation of supercoiled topoisomers by electrophoresis is poorly understood but is thought to be closely correlated to their average writhe in solution [34, 35]. Electrophoretic mobility is not a direct measure of hydrodynamic radius, especially when comparing different topologies. Sedimentation velocity, on the other hand, is inversely proportional to the hydrodynamic radius. The precise effect of temperature on supercoiled minicircles (and its reversibility) requires further study with purified supercoiled minicircles with a distinct ΔLk.

2.8 Conclusions

Our results demonstrate how AUC can be used to distinguish between the solution behavior of 333 bp and 339 bp minicircle DNA and differentiate these DNAs according to topology (linear, nicked, and negatively supercoiled). Global fitting across multiple replicates and multiple rotor speeds provided superior precision, significantly reducing the confidence limits for each measured parameter compared to single AUC experiments as assessed by Monte Carlo analysis. Temperature affects the sedimentation and diffusion coefficients, as well as the hydrodynamic radius of all the tested forms of minicircle DNA in a systematic fashion, and, while subtle, these changes were readily measured and distinguished by AUC. Both linear and nicked DNA samples displayed highly homogeneous and well defined sedimentation behavior across a wide range of temperatures and over time. There was a linear relationship among sedimentation, diffusion, and hydrodynamic radius as a function of temperature. A 2 % molar mass difference in DNA molecules could be clearly resolved. These results illustrate the suitability of these DNAs for molecular standards over the range of temperatures accessible in AUC instruments. In addition, AUC revealed a significant change in sedimentation and diffusion transport due to conformational changes as a function of temperature, but was unable to resolve the topoisomer composition of supercoiled species isolated directly from bacterial cells. Our results further show that under conditions used for these experiments, DNA of different topological states is remarkably stable during AUC and does not degrade over four months of replicate experiments performed at different speeds and temperatures. Our results suggest that linear and nicked circular DNA have superior properties compared to other biological macromolecules and may be suitable as a molecular standard for AUC. Furthermore, the results

revealed by this initial analysis have opened both a new method and a new line of investigation into the dynamics of supercoiled DNA at different temperatures. We suggest that the analysis for an eventually commercially available DNA molecular standard for AUC must be repeated on multiple, calibrated instruments to confirm the results.

Acknowledgements

This work was funded by National Institutes of Health grant R35 GM141793 and National Science Foundation grant MCB 2107527 (to LZ), and the Biomolecular Interaction Technology Center, University of Delaware, the Canada 150 Research Chairs program C150-2017-00015, the National Institutes of Health grant 1R01GM120600, and the Canadian Natural Science and Engineering Research Council Discovery Grant DG-RGPIN-2019-05637 (to BD). The Canadian Center for Hydrodynamics is funded by the Canada Foundation for Innovation grant CFI-37589 (BD). UltraScan supercomputer calculations were supported through NSF/XSEDE grant TG-MCB070039N, and University of Texas grant TG457201 (BD).

Data Availability Statement:

The UltraScan software used to analyze the AUC data is open source and freely available from the Github repository (https://github.com/ehb54/ultrascan3). The AUC data are available in openAUC format upon request from the authors, and is stored in the UltraScan LIMS server at the Canadian Center for Hydrodynamics. Electrophoresis quantitation data is available upon request from the authors.

Supplemental Information

Supplementary Table 2.1: Prediction of helical repeat, Δ*Lk*, and Lk_0 values as a function of temperature

Temp. (ºC)	Helical repeat* (bp/ turn)	333 bp					339 bp				
		Lk_0	Δ*Lk*				Lk_0	Δ*Lk*			
			Lk = 32	*Lk* = 31	*Lk* = 30	*Lk* = 29		*Lk* = 32	*Lk* = 31	*Lk* = 30	*Lk* = 29
4	10.39	32.05	-0.05	-1.05	-2.05	-3.05	32.63	-0.63	-1.63	-2.63	-3.63
12	10.42	31.97	+0.03	-0.97	-1.97	-2.97	32.55	-0.55	-1.55	-2.55	-3.55
20	10.44	31.90	+0.10	-0.90	-1.90	-2.90	32.47	-0.47	-1.47	-2.47	-3.47
23	10.45	31.87	+0.13	-0.87	-1.87	-2.87	32.44	-0.44	-1.44	-2.44	-3.44
28	10.47	31.82	+0.18	-0.82	-1.82	-2.82	32.39	-0.39	-1.39	-2.39	-3.39
37	10.49	31.73	+0.27	-0.73	-1.73	-2.73	32.30	-0.30	-1.30	-2.30	-3.30

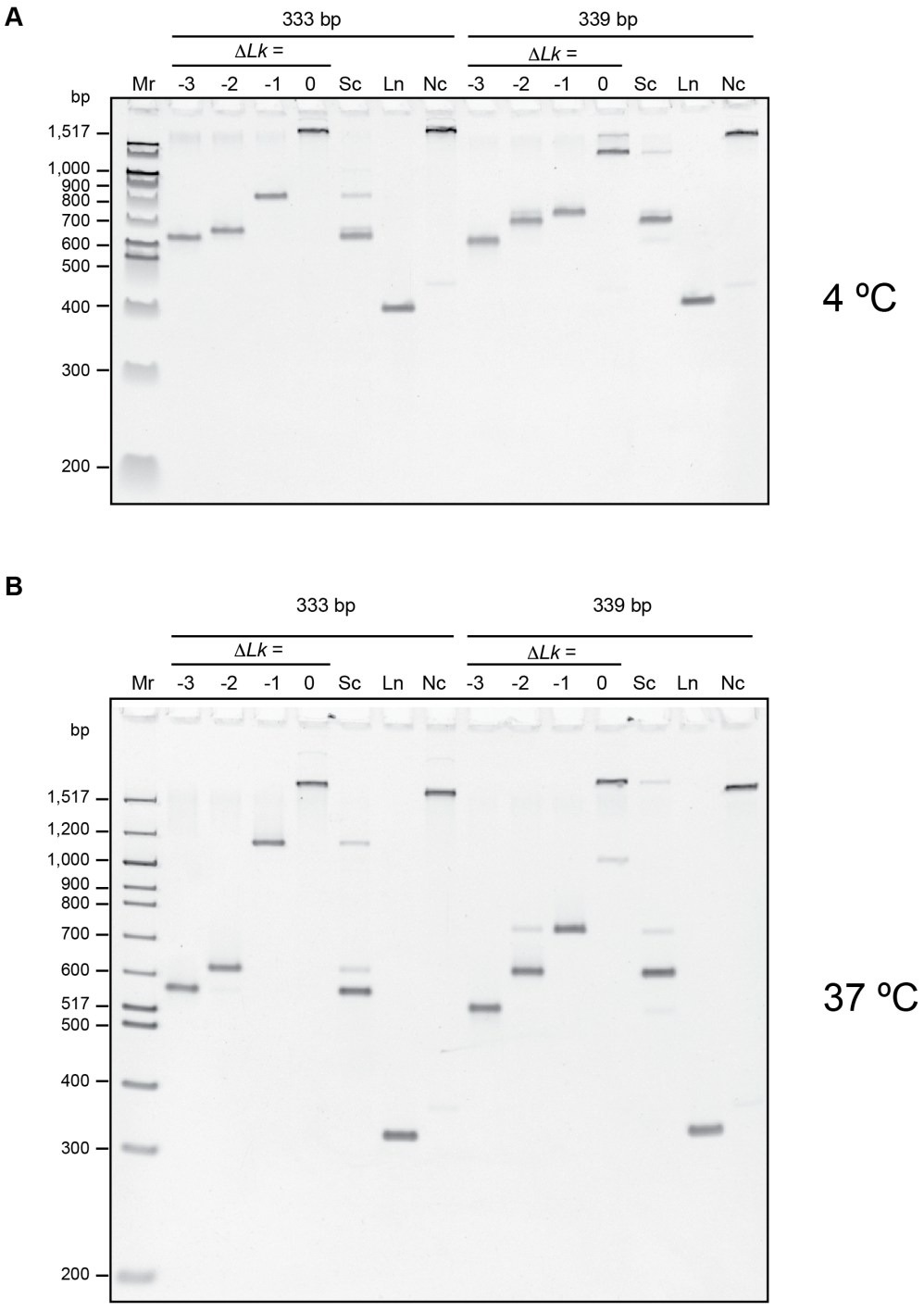
*****assuming 10.45 bp/turn at 23 ºC. The effect of temperature on helical repeat was estimated using the published value for change in winding angle per base pair with temperature of -0.0105º/K/bp[1]

Supplementary Table 2.2. Prediction of superhelical densities (σ) for each minicircle topoisomer as a function of temperature

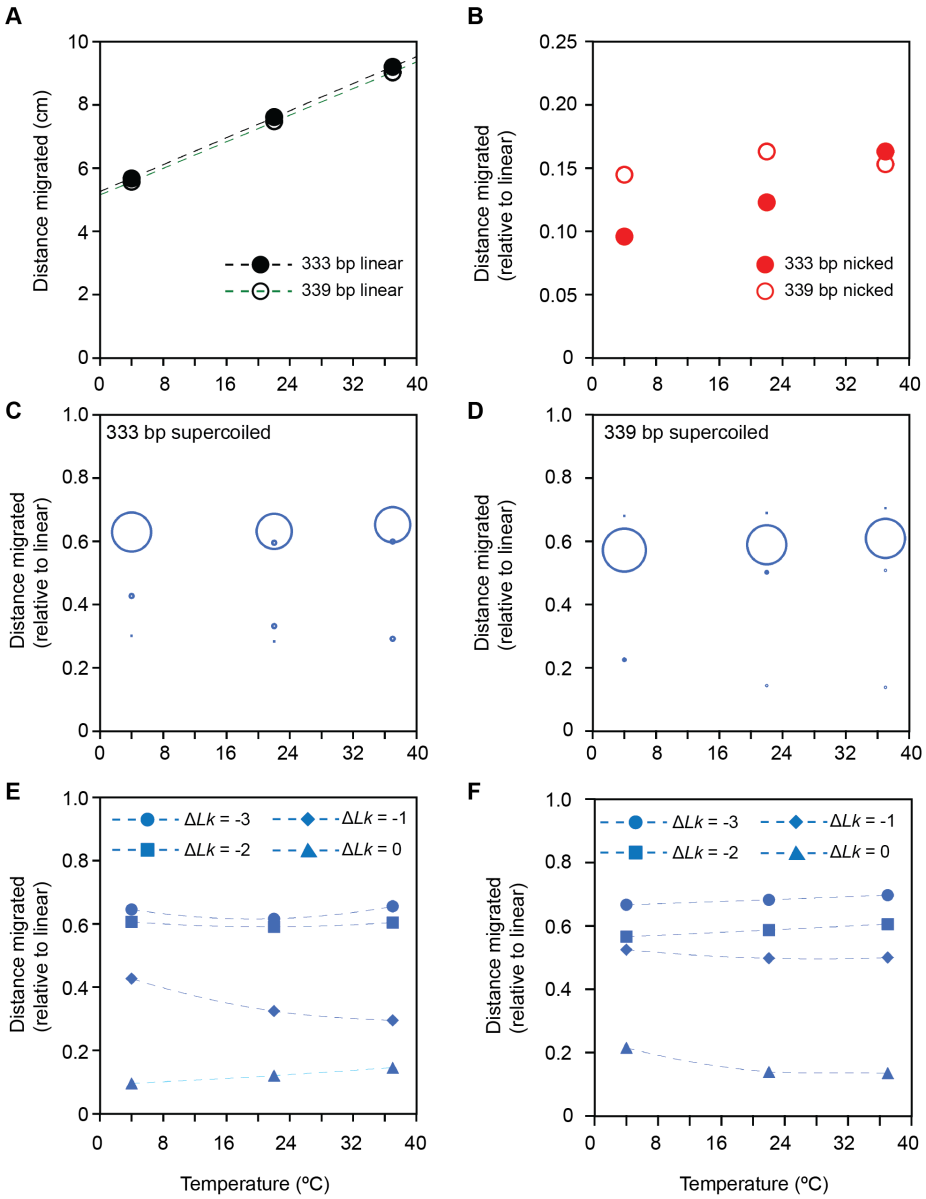
Temp. (ºC)	333 bp				339 bp			
	Δ*Lk* = 0	Δ*Lk* = -1	Δ*Lk* = -2	Δ*Lk* = -3	Δ*Lk* = 0	Δ*Lk* = -1	Δ*Lk* = -2	Δ*Lk* = -3
4	-0.002	-0.033	-0.064	-0.095	-0.019	-0.050	-0.081	-0.111
12	+0.001	-0.030	-0.062	-0.093	-0.017	-0.048	-0.078	-0.109
20	+0.003	-0.028	-0.059	-0.091	-0.014	-0.045	-0.076	-0.107
23	+0.004	-0.027	-0.059	-0.090	-0.014	-0.044	-0.075	-0.016
28	+0.006	-0.026	-0.057	-0.089	-0.012	-0.043	-0.074	-0.105
37	+0.009	-0.023	-0.055	-0.086	-0.009	-0.040	-0.071	-0.102

Obtained by scaling Δ*Lk* to the DNA length ($\sigma = \Delta Lk / Lk_0$). Actual (not rounded) values for Δ*Lk* were used for this calculation. See Supplementary Table 2.1 for Lk_0 and Δ*Lk* values used in the calculations.

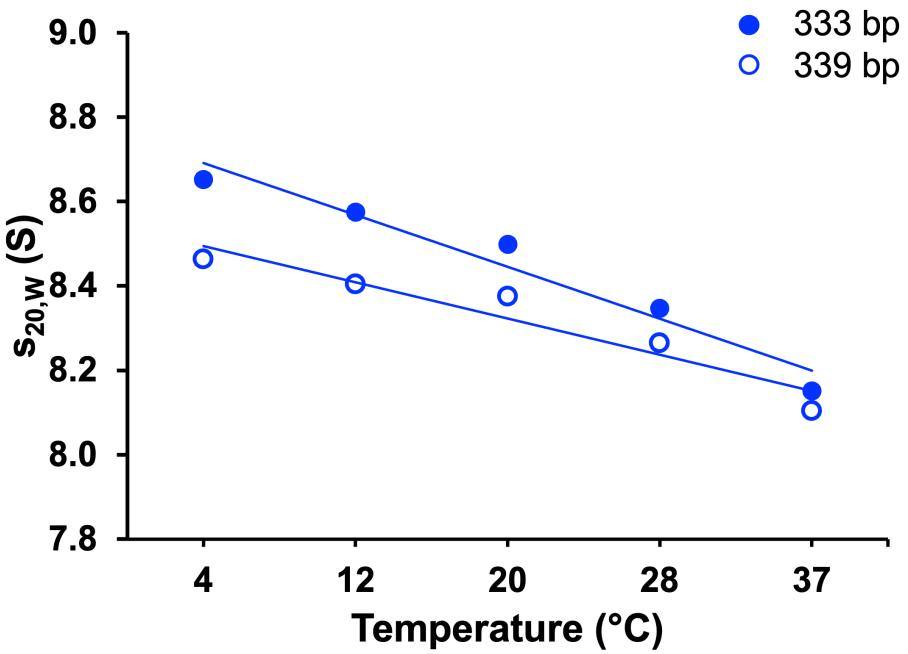
[1]Duguet M. (1993). The helical repeat of DNA at high temperature. Nucleic acids research, 21(3), 463–468. https://doi.org/10.1093/nar/21.3.463



Supplementary Figure 2.1. Gel electrophoresis at (**A**) 4 ºC and (**B**) 37 ºC. Samples loaded were the same as for Figure 2.2. Mr: 100 bp DNA ladder (linear DNA, lengths in bp as indicated). 333 and 339 bp minicircle topoisomer markers (Δ*Lk* as indicated) were also loaded. Negatively supercoiled, linear, and nicked 333 bp and 339 bp DNA samples were analyzed (Sc: supercoiled, Ln: linear, Nc: nicked).



Supplementary Figure 2.2. Comparison of electrophoretic migration at different temperatures. Distances migrated and band volumes were measured using image analysis software. (**A**) Distances migrated by linear samples as a function of temperature. (**B**) Relative migration of nicked samples at different temperatures. Distances migrated were measured and normalized by comparison to the linear samples. (**C**) Relative migration of supercoiled 333 bp samples at different temperatures. Each band from the gel is plotted separately. In the cases where the peaks for two bands overlapped they were plotted as a single datapoint. The size of the data point represents the relative fraction of the total DNA in the sample. (**D**) Relative migration of supercoiled 339 bp samples at different temperatures. Analyzes as for the 333 bp supercoiled samples. (**E**) Relative migration of 333 bp topoisomer markers at different temperatures. (**F**) Relative migration of 339 bp topoisomer markers at different temperatures.



Supplementary Figure 2.3. The weighted average sedimentation coefficients of the supercoiled minicircle DNA of 333 bp (filled circles) or 339 bp (open circles) length as a function of temperature.

2.9 References:

1 Hansen JC, Lebowitz J, Demeler B. Analytical ultracentrifugation of complex macromolecular systems. Biochemistry. 1994 Nov 15;33(45):13155-63. doi: 10.1021/bi00249a001. PMID: 7947722.

2 García de la Torre, J., & Hernández Cifre, J. G. (2020). Hydrodynamic Properties of Biomacromolecules and Macromolecular Complexes: Concepts and Methods. A Tutorial Mini-review. Journal of molecular biology, 432(9), 2930–2948. https://doi.org/10.1016/j.jmb.2019.12.027

3 Harding SE. The Svedberg Lecture 2017. From nano to micro: the huge dynamic range of the analytical ultracentrifuge for characterising the sizes, shapes and interactions of molecules and assemblies in Biochemistry and Polymer Science. Eur Biophys J. 2018 Oct;47(7):697-707. doi: 10.1007/s00249-018-1321-3.

4 Brookes E, Demeler B (2006) Genetic algorithm optimization for obtaining accurate molecular weight distributions from sedimentation velocity experiments. In: Wandrey C, Cölfen H (eds) Analytical Ultracentrifugation VIII. Prog Colloid Polym Sci 131:78–82

5 Cao W, Demeler B. Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution of the Lamm equation. Biophys J. 2005 Sep;89(3):1589-602. doi: 10.1529/biophysj.105.061135. Epub 2005 Jun 24. PMID: 15980162; PMCID: PMC1366663.

6 Cao W, Demeler B. Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution for multicomponent reacting systems. Biophys J. 2008 Jul;95(1):54-65. doi: 10.1529/biophysj.107.123950. Epub 2008 Apr 4. PMID: 18390609; PMCID: PMC2426643.

7 Stoutjesdyk M, Henrickson A, Minors G, Demeler B. A calibration disk for the correction of radial errors from chromatic aberration and rotor stretch in the Optima AUC™ analytical ultracentrifuge. Eur Biophys J. 2020 Dec;49(8):701-709. doi: 10.1007/s00249-020-01434-z. Epub 2020 May 9. PMID: 32388675.

8 Zhao H, Ghirlando R, Alfonso C, Arisaka F, Attali I, Bain DL, Bakhtina MM, Becker DF, Bedwell GJ, Bekdemir A, Besong TM, Birck C, Brautigam CA, Brennerman W, Byron O, Bzowska A, Chaires JB, Chaton CT, Cölfen H, Connaghan KD, Crowley KA, Curth U, Daviter T, Dean WL, Díez AI, Ebel C, Eckert DM, Eisele LE, Eisenstein E, England P, Escalante C, Fagan JA, Fairman R, Finn RM, Fischle W, de la Torre JG, Gor J, Gustafsson H, Hall D, Harding SE, Cifre JG, Herr AB, Howell EE, Isaac RS, Jao SC, Jose D, Kim SJ, Kokona B, Kornblatt JA, Kosek D, Krayukhina E, Krzizike D, Kusznir EA, Kwon H, Larson A, Laue TM, Le Roy A, Leech AP, Lilie H, Luger K, Luque-Ortega JR, Ma J, May CA, Maynard EL, Modrak-Wojcik A, Mok YF, Mücke N, Nagel-Steger L, Narlikar GJ, Noda M, Nourse A, Obsil T, Park CK, Park JK, Pawelek PD, Perdue EE, Perkins SJ, Perugini MA, Peterson CL, Peverelli MG, Piszczek G, Prag G, Prevelige PE, Raynal BD, Rezabkova L, Richter K, Ringel AE, Rosenberg R, Rowe AJ, Rufer AC, Scott DJ, Seravalli JG, Solovyova AS, Song R, Staunton D, Stoddard C, Stott K, Strauss HM, Streicher WW, Sumida JP, Swygert SG, Szczepanowski RH, Tessmer I, Toth RT 4th, Tripathy A, Uchiyama S, Uebel SF, Unzai S, Gruber AV, von Hippel PH, Wandrey C, Wang SH, Weitzel SE, Wielgus-Kutrowska B, Wolberger C, Wolff M, Wright E, Wu YS, Wubben JM, Schuck P. A multilaboratory comparison of calibration accuracy and the performance of external references in analytical

ultracentrifugation. PLoS One. 2015 May 21;10(5):e0126420. doi: 10.1371/journal.pone.0126420. PMID: 25997164; PMCID: PMC4440767.
9 Parker W, Song PS. Protein structures in SDS micelle-protein complexes. Biophys J. 1992 May;61(5):1435-9. doi: 10.1016/S0006-3495(92)81949-5.
10 Henrickson A, Montina T, Hazendonk P, Lomonte B, Neves-Ferreira AG, Demeler B. SDS-induced hexameric oligomerization of myotoxin-II from *Bothrops asper* assessed by sedimentation velocity and nuclear magnetic resonance. Eur. Biophys J (2023, in press).
11 Irobalieva, RN, Fogg, JM, Catanese, DJ, Sutthibutpong, T, Chen, M, Barker, AK, Ludtke, SJ, Harris, SA, Schmid, MF, Chiu, W, and Zechiedrich, L (2015) Structural diversity of supercoiled DNA. Nature Communications 6:8440
12 Waszkiewicz R, Ranasinghe M, Fogg JM, Catanese DJ, Ekiel-Jeżewska ML, Lisicki M, Demeler B, Zechiedrich L, Szymczak P. DNA supercoiling-induced shapes alter minicircle hydrodynamic properties. Nucleic Acids Res. 2023 Mar 27:gkad183. doi: 10.1093/nar/gkad183. Epub ahead of print. PMID: 36971110.
13 Driessen, R. P., Sitters, G., Laurens, N., Moolenaar, G. F., Wuite, G. J., Goosen, N., & Dame, R. T. (2014). Effect of temperature on the intrinsic flexibility of DNA and its interaction with architectural proteins. Biochemistry, 53(41), 6430–6438. https://doi.org/10.1021/bi500344j
14 Fogg, J. M., Kolmakova, N., Rees, I., Magonov, S., Hansma, H., Perona, J. J., & Zechiedrich, E. L. (2006). Exploring writhe in supercoiled minicircle DNA. Journal of physics. Condensed matter : an Institute of Physics journal, 18(14), S145–S159. https://doi.org/10.1088/0953-8984/18/14/S01
15 Fogg JM, Judge AK, Stricker E, Chan HL, Zechiedrich L. Supercoiling and looping promote DNA base accessibility and coordination among distant sites. Nat Commun. 2021 Sep 28;12(1):5683. doi: 10.1038/s41467-021-25936-2.
16 Demeler B., Gorbet G.E. (2016) Analytical Ultracentrifugation Data Analysis with UltraScan-III. In: Uchiyama S., Arisaka F., Stafford W., Laue T. (eds) Analytical Ultracentrifugation. Springer, Tokyo. https://doi.org/10.1007/978-4-431-55985-6_8
17 Savelyev A, Gorbet GE, Henrickson A, Demeler B. Moving analytical ultracentrifugation software to a good manufacturing practices (GMP) environment. PLoS Comput Biol. 2020 Jun 19;16(6):e1007942. doi: 10.1371/journal.pcbi.1007942. PMID: 32559250; PMCID: PMC7347214.
18 Cölfen, H., Laue, T. M., Wohlleben, W., Schilling, K., Karabudak, E., Langhorst, B. W., Brookes, E., Dubbs, B., Zollars, D., Rocco, M., & Demeler, B. (2010). The Open AUC Project. European biophysics journal : EBJ, 39(3), 347–359. https://doi.org/10.1007/s00249-009-0438-9
19 Pierce, M., S. Marru, B. Demeler, R. Singh, and G. Gorbet. (2014) The Apache Airavata Application Programming Interface: Overview and Evaluation with the UltraScan Science Gateway. In Proceedings of the 9th Gateway Computing Environments Workshop (GCE '14). IEEE Press, Piscataway, NJ, USA, 25-29. DOI=10.1109/GCE.2014.15
20 Brookes E, Cao W, Demeler B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. Eur Biophys J. 2010 Feb;39(3):405-14. doi: 10.1007/s00249-009-0413-5. Epub 2009 Feb 27. PMID: 19247646.
21 Brookes E, Demeler B. Parsimonious Regularization using Genetic Algorithms Applied to the Analysis of Analytical Ultracentrifugation Experiments. GECCO '07: Proceedings of the 9th

annual conference on Genetic and evolutionary computation, London, July 7-11, 2007, 361–368, https://doi.org/10.1145/1276958.1277035, ACM 978-1-59593-697-4/07/0007
22 Demeler B and E. Brookes. Monte Carlo analysis of sedimentation experiments. Colloid Polym Sci (2008) 286(2) 129-137
23 Gorbet, G.E., Mohapatra, S. & Demeler, B. Multi-speed sedimentation velocity implementation in UltraScan-III. Eur Biophys J 47, 825–835 (2018).
24 Williams TL, Gorbet GE, Demeler B. Multi-speed sedimentation velocity simulations with UltraScan-III. Eur Biophys J. 2018 Oct;47(7):815-823. doi: 10.1007/s00249-018-1308-0. Epub 2018 May 10. PMID: 29748855; PMCID: PMC6158090.
25 Lamm, O. 1929. Die Differentialgleichung der Ultrazentrifugierung. Ark. Mat. Astr. Fys. 21B:1-4.
26 Fogg JM, Judge AK, Stricker E, Chan HL, Zechiedrich L (2021) Supercoiling and looping promote DNA base accessibility and coordination among distant sites. Nat Commun 12(1):5683. https:// doi. org/ 10. 1038/ s41467- 021- 25936-2
27 Durchschlag, H. (1986). Specific Volumes of Biological Macromolecules and Some Other Molecules of Biological Interest. In: Hinz, HJ. (eds) Thermodynamic Data for Biochemistry and Biotechnology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-71114-5_3
28 Kriegel,F., Matek,C., Geršata,T., Kulenkampff,K., Tschirpke,S., Zacharias,M., Lankaš,F. and Lipfert,J. (2018) The temperature dependence of the helical twist of DNA. Nucleic Acids Res, 46, 7998–8009. https://doi.org/10.1093/nar/gky599
29 Duguet M. (1993). The helical repeat of DNA at high temperature. Nucleic acids research, 21(3), 463–468. https://doi.org/10.1093/nar/21.3.463
30 Seeton C.J. (2006). Viscosity temperature correlation for liquids. Tribology Letters 22, 67-78.
31 Robertson, R. M., Laib, S., & Smith, D. E. (2006). Diffusion of isolated DNA molecules: dependence on length and topology. Proceedings of the National Academy of Sciences of the United States of America, 103(19), 7310–7314. https://doi.org/10.1073/pnas.0601903103
32 Brunet, A., Salomé, L., Rousseau, P., Destainville, N., Manghi, M., & Tardin, C. (2018). How does temperature impact the conformation of single DNA molecules below melting temperature?. Nucleic acids research, 46(4), 2074–2081. https://doi.org/10.1093/nar/gkx1285
33 Lapham, J., Rife, J. P., Moore, P. B., & Crothers, D. M. (1997). Measurement of diffusion constants for nucleic acids by NMR. Journal of biomolecular NMR, 10(3), 255–262. https://doi.org/10.1023/a:1018310702909
34 Vetcher,A.A., McEwen,A.E., Abujarour,R., Hanke,A. and Levene,S.D. (2010) Gel mobilities of linking-number topoisomers and their dependence on DNA helical repeat and elasticity. Biophys. Chem., 148, 104–111.
35 Zivanovic,Y., Goulet,I. and Prunell,A. (1986) Properties of supercoiled DNA in gel electrophoresis. The V-like dependence of mobility on topological constraint. DNA-matrix interactions. J. Mol. Biol., 192, 645–660.

Chapter 3: Conclusions

3.1 Overview

By studying the hydrodynamic properties and solution behavior of DNA, we aimed to provide insights into its potential as a reliable standard material in AUC experiments. The DNA minicircles used in this study served as a model system to examine the properties of double-stranded DNA due to their well-defined topology and known molar mass [1].

The objectives outlined in this study were successfully addressed and accomplished. The first objective was to evaluate the hydrodynamic properties of DNA minicircles over time and assess the stability, reproducibility, and homogeneity of double-stranded DNA as a molecular standard. The DNA samples demonstrated exceptional stability throughout the study, maintaining consistent sedimentation behavior and preventing degradation over an extended period of time and under varying experimental conditions. The results confirm our hypothesis that double stranded DNA remains stable over an extended time in the AUC, survives many experiments over a large temperature range without any significant change, and can serve as a reliable standard for AUC experiments. The analysis further confirmed the consistent homogeneity of DNA samples, reducing variability in AUC data and enhancing the overall precision of measurements.

Objective 2 investigated the hydrodynamic properties of different topological states of circular and linearized DNA minicircles. Our study demonstrates the ability of AUC to distinguish between various topologies of double-stranded DNA, including linear, nicked, and negatively supercoiled topologies. Our findings strongly support the hypothesis that these distinct topological states of DNA exhibit distinct hydrodynamic properties.

In objective 3, the temperature-dependent effects on the hydrodynamic properties of DNA were evaluated. Sedimentation and diffusion coefficients, as well as the hydrodynamic radius of the tested minicircle DNA were examined, and the results revealed that these variations were relatively small, yet remained consistent for each topology. Despite the subtle nature of these changes, AUC has demonstrated exceptional sensitivity in accurately measuring these effects. Both linear and nicked DNA samples exhibited remarkable homogeneity and well-defined sedimentation behavior across a wide temperature range and extended time period. Negatively supercoiled DNA, as isolated from bacteria, demonstrated a higher heterogeneity, suggesting the presence of different supercoil populations. Furthermore, these topologies demonstrated conformational changes as a function of temperature. These findings suggest that temperature significantly influences the hydrodynamic properties of negatively supercoiled DNA.

Our results demonstrate that linear and nicked DNA have suitable properties to be used as molecular standards. However, supercoiled DNA did not meet the requirements for a molecular standard. By employing global fitting across multiple replicates and rotor speeds utilizing appropriate modeling approaches, such as global GA-MC analysis [2] for homogeneous linearized and nicked minicircle DNA samples and global 2DSA [3] for heterogeneous negatively supercoiled samples, we were able to achieve enhanced precision and reduce confidence limits for the measured parameters. This approach allows for more accurate and reliable characterization of macromolecules in AUC experiments, surpassing the limitations of single AUC measurements.

Ultimately, this research contributes to the validation of AUC experiments across different temperature conditions, enabling precise characterization of macromolecules in diverse

experimental settings. The use of DNA as a molecular standard also addresses the limitations associated with protein-based standards, such as potential aggregation, ensuring more accurate and reproducible AUC measurements.

In addition, the results indicate that DNA exhibits high hydrodynamic stability and possesses an extended shelf life when placed within an assembled centrifuge cell. It suggests the potential for using DNA as a long-lasting material within the AUC cell without the need for frequent refilling or replacement. It also has practical implications as it simplifies the experimental setup and reduces the chances of introducing variability due to frequent sample handling.

Furthermore, the insights gained from investigating the behavior of DNA as a function of temperature have broader implications. The changes in the conformation and dynamics of negatively supercoiled DNA may have biological function and play a role in interactions with other macromolecules.

3.2 Limitations and future directions

While this study has provided valuable insights into the potential use of DNA as a molecular standard in AUC experiments, the heterogeneity observed in negatively supercoiled samples demands additional work. Future research should focus on explaining the biological function of supercoiled DNA heterogeneity and determine if the variation in hydrodynamic properties of negatively supercoiled DNA as a function of temperature is reversible.

Development of molecular standards:

The findings provide a strong foundation for the development of a commercially available DNA molecular standard for AUC. It is important to establish standardized protocols for the

production of linear and nicked double-stranded DNA as molecular standards. These protocols should provide clear guidelines for sample preparation, handling, and storage conditions to help maintain consistency in the properties of DNA samples used as standards.

To address the potential variations across different batches of DNA samples, rigorous quality control measures and standardized protocols should be employed. Repeat experiments on multiple calibrated instruments with the same sample under the same conditions are required to ensure the robustness and reliability of the DNA molecular standard [4]. This validation process will confirm the consistency and reproducibility of the obtained results, informing confidence ranges.

According to the requirements for certification, it is essential to obtain a certified value assignment for the developed DNA standard from an authorized organization such as NIST (National Institute of Standards and Technology). This certification will officially classify the DNA material as a standard and further validate its use in AUC experiments. Additional work should be conducted to meet the criteria and standards set by such organizations, ensuring the credibility and widespread adoption of the DNA molecular standard in the scientific community. (NIST standards definitions: https://www.nist.gov/srm/srm-definitions)

Further research should focus on scaling up the production of linear DNAs with different lengths and conducting additional characterization to expand the range of available standard materials. By generating DNA samples of different lengths, it would be possible to assess the variations in sedimentation and diffusion coefficients observed across the range of DNA molecules. This approach will allow researchers to select DNA standards that are most suitable for their specific experimental needs and validate the performance of the instruments at different

rotor speeds, and for materials with variable anisotropies, enhancing the versatility and applicability of AUC measurements.

In addition to validating different DNA lengths, the mixing and matching of DNA samples at various concentrations can be employed to investigate other important parameters. These include determining the resolution achievable by the AUC instrument, as well as assessing the limits of detectability (LOD) and quantification (LOQ) [4]. By systematically varying the concentrations and compositions of DNA mixtures, the resolution and limitations of the AUC technique can be better understood.

Additionally, investigating the reversibility of DNA samples as a function of temperature would be a promising avenue of research. By subjecting DNA to temperature variations, the reversible transitions and conformational changes of the DNA molecule can be characterized, contributing to a comprehensive understanding of DNA behavior in solution.

Partial specific volume measurements from density meter: Due to limitations in available instrumentation and sample amounts, an orthogonal PSV measurement could not be obtained in the current study. For future research to enable the assignment of accurate sedimentation coefficients to DNA materials designated as molecular standards, it is recommended to use a density meter for experimental measurements of PSV.

Molecular mass determination: Considering the importance of accurate molecular mass determination in sedimentation and diffusion corrections for standard conditions, future directions of research include determining the molecular mass of the developed standard by employing a biophysical technique such as mass spectrometry. This approach will provide more

accurate values for the measured parameters of the molecular standard enhancing the reliability of the assigned values

Correlation between AUC observations and electrophoretic mobility: AUC directly measures thermodynamic effects of temperature variations that are inaccessible to gel electrophoresis or cryoEM. The disparity between the temperature-dependent effects observed in AUC data and the electrophoresis data can be a consequence of precise temperature control only being possible in AUC instruments. Further studies could employ complementary techniques such as molecular dynamics simulations or other biophysical methods such as light scattering or fluorescence correlation spectroscopy to obtain additional insights into the conformational changes and interconversion dynamics of supercoiled DNA. These approaches would also help elucidate the underlying mechanisms governing the separation of supercoiled topoisomers in electrophoresis and their hydrodynamic properties in AUC.

Optimizing AUC Run Conditions: This current study provides a foundation for future investigations into the sedimentation behavior of supercoiled DNA. With the observed significant variations across a broad range of supercoiling, the next step is to explore the more subtle differences between individual topoisomers and their ability to be distinguished using AUC. Achieving this resolution requires the optimization of AUC run conditions.

To effectively resolve different ΔLk topoisomers into separate discrete species in AUC, careful adjustment of experimental parameters is necessary. Optimizing parameters such as rotor speed, temperature, and buffer composition can enhance the resolution and accuracy in distinguishing between various topological states of DNA.

Application of DNA as a molecular standard in other techniques: The successful validation of DNA as a molecular standard in AUC opens up opportunities for its application in other analytical techniques, such as dynamic light scattering. Assessing the suitability of DNA as a standard across different techniques would enhance cross-validation and broaden its utility in the field of biomolecular characterization.

3.3 References

1 Fogg, J. M., Kolmakova, N., Rees, I., Magonov, S., Hansma, H., Perona, J. J., & Zechiedrich, E. L. (2006). Exploring writhe in supercoiled minicircle DNA. Journal of physics. Condensed matter: an Institute of Physics journal, 18(14), S145–S159. https://doi.org/10.1088/0953-8984/18/14/S01

2 Brookes E, Demeler B. Parsimonious Regularization using Genetic Algorithms Applied to the Analysis of Analytical Ultracentrifugation Experiments. GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation, London, July 7-11, 2007, 361–368, https://doi.org/10.1145/1276958.1277035, ACM 978-1-59593-697-4/07/0007

3 Brookes E, Cao W, Demeler B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. Eur Biophys J. 2010 Feb;39(3):405-14. doi: 10.1007/s00249-009-0413-5. Epub 2009 Feb 27. PMID: 19247646.

4 Savelyev, A., Gorbet, G. E., Henrickson, A., & Demeler, B. (2020). Moving analytical ultracentrifugation software to a good manufacturing practices (GMP) environment. PLoS computational biology, 16(6), e1007942. https://doi.org/10.1371/journal.pcbi.1007942

Appendices

Appendix A: DNA supercoiling-induced shapes alter minicircle hydrodynamic properties

Radost Waszkiewicz[1], Maduni Ranasinghe[2], Jonathan M. Fogg[3], Daniel J. Catanese, Jr.[4], Maria L. Ekiel-Je˙zewska[5,*], Maciej Lisicki[1,*], Borries Demeler[2,6,*], Lynn Zechiedrich[3,*], and Piotr Szymczak[1,*]

[1]Institute of Theoretical Physics, Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland

[2]University of Lethbridge, Dept. of Chemistry and Biochemistry, Alberta, T1K3M4, Canada
[3]Department of Molecular Virology and Microbiology, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Department of Pharmacology and Chemical Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
[4]Department of Biosciences, Rice University, 6100 Main St., Houston, TX 77005-1827, USA
[5]Institute of Fundamental Technological Research, Polish Academy of Sciences, A. Pawi´nskiego 5B, 02-106 Warsaw, Poland
[6]University of Montana, Dept. of Chemistry and Biochemistry, Missoula, MT 59812, USA

Abstract

DNA in cells is organized in negatively supercoiled loops. The resulting torsional and bending strain allows DNA to adopt a surprisingly wide variety of 3-D shapes. This interplay between negative supercoiling, looping, and shape influences how DNA is stored, replicated, transcribed, repaired, and likely every other aspect of DNA activity. To understand the consequences of negative supercoiling and curvature on the hydrodynamic properties of DNA, we submitted 336 bp and 672 bp DNA minicircles to analytical ultracentrifugation (AUC). We found that the diffusion coefficient, sedimentation coefficient, and the DNA hydrodynamic radius strongly depended on circularity, loop length, and degree of negative supercoiling. Because AUC cannot ascertain shape beyond degree of non-globularity, we applied linear elasticity theory to predict DNA shapes, and combined these with hydrodynamic calculations to interpret the AUC data, with reasonable agreement between theory and experiment. These complementary approaches, together with earlier electron cryotomography data, provide a framework for understanding and predicting the effects of supercoiling on the shape and hydrodynamic properties of DNA.

Introduction

Nearly seventy years after Rosalind Franklin's meticulous work that led to the first description of the structure of DNA [1] we are still working to understand how this remarkable molecule is organized, stored, activated, and segregated into daughter cells (2,3,4,5,6,7,8,9). It is becoming increasingly apparent that negative supercoiling (the underwinding of the DNA double helix) provides a secondary or 'hidden code' that contributes to the three-dimensional (3-D) organization of the genome (10) that can be used by cells as a 'molecular servomechanism' to detect and regulate gene expression (11).

We recently discovered that the degree of curvature, dictated by DNA loop length, additionally tunes supercoiling mediated effects and promotes mechanical crosstalk to expose DNA bases at specific distant sites (12). Exposed DNA bases drastically increase DNA flexibility to change the 3-D structure of DNA, which, conversely, influences the location and frequency of the disruptions to base pairing (12). Therefore, 3-D shape and base exposure are manifestations of supercoiling and looping (13). These findings underscore how supercoiling-dependent conformational changes may allow DNA to be an active participant in its transactions (13,14).

We previously used minicircles of a few hundred base pairs and defined supercoiling to determine how supercoiling and looping modulates the 3-D structure of DNA (13). DNA loops in this length range are found in nature, e.g., wrapped around the nucleosome, and are involved in gene regulation (15,16,17,18). Human topoisomerase II relaxes supercoiledDNA minicircles (13), demonstrating that they are biologically active. Thus, minicircles are biologically relevant and mimic DNA loops naturally occurring in cells.

We previously determined how supercoiling modulates the shapes of minicircles using electron cryotomography (cryoET) (13). Although informative, these assays are laborious and only limited DNA sequences and buffer conditions have been explored (12,13,19). CryoET provides 3-D information on individual DNA minicircles of defined supercoiling (DNA topoisomers) (13), but the approach requires skill in the art, is time-consuming and the resulting structures are of low resolution. Increasingly powerful, atomic force microscopy (AFM) can show the helical repeat ofDNA as well as areas of base pair disruption (19), but the sequence is unidentifiable and the technique requires that DNA is adsorbed onto a flat surface. Not only might this adsorption distort DNA conformations but it means that they are visualized as 2-D projections with limited 3-D information.

Computational modeling would be of great value in helping predict DNA negative supercoiling and looping behavior, but thus far fails to account for supercoiling- and looping-mediated site-specific base exposure or resulting conformational changes. For example, most efforts at understanding looping (20) or cyclizing of DNA (21) ignore supercoiling. Therefore, new modeling efforts are needed, including the parameters of supercoiling and degree of curvature (dictated by DNA loop length). Before this modeling can be improved, however, detailed parameters of supercoiled loops of DNA must be determined.

Whereas cryoET is impractical to use for multiple conditions, AUC and electrophoresis rapidly assess properties in solution, and can be used to test multiple conditions simultaneously. Toward the goal of understanding howDNA sequence and 3-D shape are affected by negative supercoiling and looping, here we combined state-of-the-art AUC (22) with mathematical modeling to determine hydrodynamic parameters of supercoiled DNA minicircles. We derived partial specific volume (PSV) and anisotropy, and measured the sedimentation and diffusion coefficient for minicircle DNA of different degrees of negative supercoiling and lengths. With these values, we determined the density of DNA. We discovered that DNA length and supercoiling strongly affect the sedimentation properties of minicircle DNA but have either no or only a minimal effect on the PSV.

We generalized the continuum elastic framework to accurately predict the previously observed DNA minicircle 3-D shapes (13). This generalization provides additional and complimentary

information that will allow us to interpret supercoiling- and curvature-dependent DNA structural alterations. Emboldened by this accomplishment, we then combined the measured elastic and hydrodynamic properties of DNA minicircles using bead models and considering force and hydrodynamic effects to compute the hydrodynamic sedimentation and diffusion coefficients. These modeling results compared favorably to AUC measurements.

Materials and methods

Chemicals and reagents

MseI, Nb.BbvCI, Proteinase K, T4 DNA Ligase, low molecular weight DNA ladder, and 100 bp DNA ladder were purchased from New England Biolabs (Ipswich, MA, USA). Adenosine triphosphate (ATP), antifoam 204, dithiothreitol (DTT), ethidium bromide, and RNase A were purchased from Sigma-Aldrich (St. Louis,MO, USA). Acrylamide, ampicillin, chloroform, and sodium chloride were purchased from Fisher Scientific (Pittsburgh, PA, USA). All other chemicals were purchased from VWR International (West Chester, PA, USA).

Generation and purification of minicircle DNA

Plasmid pMC336 (13) was used to generate both the 336 bp and 672 bp minicircles via λ-integrase-mediated site-specific recombination as described (23). Double-length 672 bp minicircles contain two copies of the 336 bp minicircle sequence in tandem orientation and are generated by the recombination used to generate 336 bp minicircle DNA.

Generation of different DNA topologies

The ‘supercoiled’ samples are the 336 bp or 672 bp minicircle products of the purification process. These were analyzed without further manipulation. To make nicked DNA, the minicircles were nicked at a single site using the nicking endonuclease Nb.BbvCI according to the manufacturer’s protocol. The 672 bp minicircle contains two copies of the BbvCI site and was thus nicked at both locations. Following nicking, the DNA was subsequently incubated at 80 °C for 20 minutes to inactivate Nb.BbvCI. Linear 336 bp was generated by incubating supercoiled 336 bp minicircle with MseI according to the manufacturer’s protocol. The linearized DNA was subsequently incubated at 65 ◦C for 20 minutes to inactivate the enzyme. ‘Relaxed’ 336 bp minicircle DNA was generated by incubating the nicked minicircles with T4 DNA ligase in 50 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, and 10 mM DTT overnight at room temperature. ‘Hypernegatively supercoiled’ 336 bp was generated in an identical manner as ‘relaxed’, except for the addition of ethidium bromide (6.5 g/ml) to the ligation reaction. Ligations were subsequently extracted with butanol (to both reduce the volume and to remove the ethidium bromide), extracted with chloroform, then precipitated with ethanol. The nicked, linearized, and supercoiled minicircle samples were also subjected to butanol and chloroform extraction, and ethanol precipitation in a similar manner to both remove the enzymes and to ensure that any differences observed could not be attributed to differences in how the samples were made. Following ethanol precipitation, DNA was resuspended in 50 mM Tris-Cl pH 8.0, 150 mM NaCl, and 10 mM $CaCl_2$. DNA samples were subsequently subjected to multiple rounds of buffer exchange in the same buffer using an Amicon 0.5 ml centrifugal filter to ensure that buffer conditions were equal across all samples. DNA concentrations were determined using a Nanodrop spectrophotometer.

Geometry and topology of DNA minicircles

DNA supercoiling is defined by the linking number (*Lk*), the total number of times the two single DNA strands coil about one another (24), thus it is an integer number by construction, if both strands are covalently closed. If one or both of the strands is not covalently closed, e.g., for nicked and linear DNA, *Lk* can adopt non-integer values. Another quantity determining the shape of DNA is the equilibrium helical repeat h defined as the number of base pairs between two locations where the backbones are aligned and is measured in base pairs per turn. The value of h is buffer dependent and ~10.42 bp/turn in 10 mM $CaCl_2$ (13). Using *h*, we can calculate the angle between terminal base pairs of a straight linear DNA segment of a given length *L*, which gives us the reference value $Lk_0 = L/h$. Since h can, in principle, take any value, Lk_0 is not restricted to integer values and usually has a fractional part. For relaxed 336 bp minicircles, we get $Lk_0 = 32.2$, while for 672 bp $Lk_0 = 64.4$. Therefore, the deviation from the most relaxed DNA structure is measured by the difference between Lk_0 and *Lk* denoted by $\Delta Lk = Lk - Lk_0$, which for the relaxed configuration of the 336 bp minicircle yields $\Delta Lk = -0.2$. ΔLk is typically scaled to the DNA length to give the superhelical density $\sigma = \Delta Lk/Lk_0$.

Since *Lk* is constrained to integer values it is sufficient to report ΔLk rounded to the nearest integer to uniquely identify experimental configurations, as used in Ref. (13). For the relaxed 336 bp minicircle, we round −0.2 to 0. For simplicity, we follow this convention when reporting experimental values in this work. At the same time, we keep track of the fractional parts to accurately compute the elastic properties of the minicircles.

Gel electrophoresis

DNA samples were analyzed by electrophoresis through 5 % (for 336 bp and 672 bp samples) or 4 % (for 672 bp samples) polyacrylamide gels (acrylamide:bis-acrylamide = 29:1) in Tris-acetate buffer (pH 8.2) containing either 150 mM NaCl and 10 mM CaCl2 (5 % gels) or 10 mM CaCl2 (4 % gels) at 125 V (~ 6 V/cm) for 8 hours. Buffer was continuously recirculated during electrophoresis. DNA samples were also analyzed by electrophoresis through 1.5 % and 3 % agarose gels (Seakem LE agarose, Lonza, Rockland, ME) in TAE (Tris-acetate + 1 mM EDTA) buffer at 100 V for 3 hours. Gels were subsequently stained with SYBR Gold (ThermoFisher Scientific, Waltham, MA), then visualized using a FOTO/ANALYST Investigator imaging system (Fotodyne, Hartland, WI, USA) with quantitation using ImageQuant TL, version 8.1 (GE Healthcare Life Sciences, Marlborough, MA, USA).

Analytical ultracentrifugation

Linear, nicked, relaxed, supercoiled, and hypernegatively supercoiled 336 bp minicircles, and supercoiled and nicked 672 bp samples were measured by sedimentation velocity using an AN50Ti rotor in a Beckman Coulter Optima AUC at the Canadian Center for Hydrodynamics at the University of Lethbridge in Alberta, Canada. For details of minicircles used, see Supplementary Table A.1.

All samples were measured in 50 mM Tris-Cl pH 8.0, 150 mM NaCl, and 10 mM $CaCl_2$. 460 µl of each sample at an absorbance (A) of 0.6 at 260 nm were loaded into cells fitted with sapphire windows and 12 mm double channel epon charcoal centerpieces (Beckman Coulter, Indianapolis, IN, USA). Data were collected in intensity mode at 260 nm, and at 20 °C at five different rotor

speeds of 10, 14, 25, 35, and 45 krpm. After data collection at each speed was completed, AUC cells were thoroughly shaken to redistribute the minicircle DNA uniformly. Depending on minicircle topology and length, at 10 krpm, pelleting occurred between 559–770 scans, requiring 50–70 hours. At 45 krpm, pelleting occurred after 88–159 scans, requiring 2–4 hours. The density and viscosity of the buffer, estimated with UltraScan, was 1.00682 g/ml and 1.02667 cP, respectively.

AUC data analysis

All data were analyzed with UltraScan-III, version 4.0 (6345) (22), using the UltraScan data acquisition module (25). UltraScan fits experimental data to finite element solutions of the Lamm equation, deriving distributions for sedimentation and diffusion coefficients (26, 27). Optimization is achieved by parallel distributed data analysis, which was performed on the UltraScan Science Gateway using XSEDE resources (Expanse, Bridges 2, Stampede), and high-performance computing clusters at the University of Montana and University of Lethbridge. The optimization process proceeds through a series of model refinement steps, which employs the two-dimensional spectrum analysis (2DSA) (28). This refinement process removes systematic noise contributions contained in the raw data and obtains exact boundary conditions (the radial positions at the meniscus and the bottom of the cells) as described in (29). The final 2DSA refinement result is used to initialize a genetic algorithm analysis (GA) (30), which is followed by a Monte Carlo GA analysis (31). The total concentration determined from each speed between identical samples was also compared to ensure no material was lost due to aggregation or degradation, and samples were comparable across all speeds for a global analysis. The Monte Carlo GA results from identical samples and different speeds were combined to initialize a global GA analysis over all speeds. UltraScan supports simultaneous fitting to datasets from multiple experiments performed at different speeds. A global analysis benefits from the enhanced signal of the diffusion coefficient at low speeds and the improved sedimentation signal at higher speeds (32, 33). This feature also enhances signal-to-noise ratios and improves the confidence limits for the determined hydrodynamic parameters. The global fitting algorithm in UltraScan is further explained in (22).

Hydrodynamic properties

Translational diffusion coefficient The translational diffusion coefficient D is inversely proportional to the translational frictional coefficient f,

$$D = \frac{k_B T}{f}$$

Equ. A.1

where k_B is the Boltzmann constant, and T the absolute temperature. In this work, diffusion occurs at very low concentrations of solute, which allows the analysis of transport coefficients in terms of single-particle properties only. For microscopic solid spheres of radius R suspended in a liquid of temperature T and viscosity η the Stokes-Einstein relationship reads

$$D = \frac{k_B T}{6\pi\eta R}$$

Equ. A.2

This relationship can be generalized to non-spherical molecules by introducing the effective hydrodynamic radius R_h, defined as

$$R_h = \frac{k_B T}{6\pi\eta D}$$

Equ. A.3

In aqueous solvents, macromolecules are typically hydrated, which adds to their apparent size and friction. The value of R_h derived from measured D includes these effects.

The hydrodynamic anisotropy of a given macromolecule was characterized by the frictional ratio f / f_0, being the ratio of the measured frictional coefficient f and the frictional coefficient f_0 of a spherical particle of the same volume. The anisotropy equals 1.0 for a spherical molecule and exceeds 1.0 for non-spherical molecular shapes.

Sedimentation coefficient The sedimentation coefficient s depends on the molar mass M, the translational frictional coefficient f and the buoyancy of the particle, which is a function of its PSV, $\bar{v}$, and solvent density ρ,

$$s = \frac{M(1-\bar{v}\rho)}{Nf}$$

Equ. A.4

where N is Avogadro's number. The Svedberg equation describes the ratio of the two parameters measured in a sedimentation velocity experiment, s and D, and provides a way to estimate the molar mass M, if the PSV is known:

$$s = \frac{M(1-\bar{v}\rho)}{N k_B T}$$

Equ. A.5

Apparent PSV Equ. A.5 considers a two-component system–an analyte with anhydrous molar mass M and a solvent with density ρ. However, our experimental solution also contains buffer components and ions that may be bound to the analytes. The degree of counterions bound to the analyte is dependent on solvent conditions and the ionic strength of the solvent, particularly for charged molecules (34). PSV is defined as the change in volume when one gram of analyte is added to the solvent, and is typically reported in units of ml/g. Because we do not know the precise amount of counterions bound to the analyte, we consider an apparent partial specific volume $\bar{v}'$, which can only be considered constant for a single solvent at a constant temperature and pressure. Rearranging the Svedberg equation allows the determination of the apparent PSV, provided the molar mass and the solvent density are known and the sedimentation and diffusion coefficients have been determined experimentally from a sedimentation velocity experiment

$$\bar{v}' = \frac{1}{\rho}\left(1 - \frac{N s k_B T}{MD}\right)$$

Equ. A.6

In our case, the molar masses are 207.576 kDa for the 336 bp minicircle and 415.152 kDa for the 672 bp minicircle, as calculated from the sequence (12) using molbiotools.com/dnacalculator. UltraScan automatically estimates the solvent density and viscosity from the buffer composition, adjusts the experimental $s_{T,B}$ and $D_{T,B}$ values to standard conditions (water at 20 °C) using the density and viscosity estimates from the buffer components, see (35, p. 117)

$$s_{20,W} = s_{T,B} \frac{(1 - \bar{v}'\rho)_{20,W}}{(1 - \bar{v}'\rho)_{T,B}} \frac{\eta_{T,B}}{\eta_{20,W}}$$

Equ. A.7

$$D_{20,W} = D_{T,B} \frac{T_{20}}{T} \frac{\eta_{T,B}}{\eta_{20,W}}$$

Equ. A.8

where $s_{T,B}$ is the observed sedimentation coefficient at experimental conditions (temperature T = 293.15 K and buffer B). However, for the $s_{20,w}$ corrections, the partial specific volume of DNA at standard conditions is required, but it is not known to us and impossible for us to measure. While literature values are reported for NaDNA (0.54 – 0.55 ml/g) (36, 37), topoisomers here were studied in 10 mM calcium, which has a higher binding affinity to DNA than Na (38). Hence, we report here the experimentally measured values of s and D for all topoisomers, and the apparent partial specific volume under experimental conditions, Embedded Image, calculated by Equ. A.6, and assuming a two component system.

Finding equilibrium shapes of loops

To model the shapes of DNA minicircles with a given *Lk*, a variant of the Kirchhoff beam theory for inextensible rods (39) was used, which describes the twisting and bending of a uniform elastic filament of constant steric thickness d_s, which was set to 20 Å in all computations. The helical repeat of the DNA yields a reference value of $Lk_0 = L / h$ for a given length *L*.

To model a DNA minicircle with a given *Lk* (and ΔLk), we use an elastic beam representation in which a (closed) beam is characterized by two constants: bending rigidity A and geometric torsional stiffness ω (describing the cross-sectional shape, equal to 2/3 for circular cross sections). The energy density has quadratic contributions from the residual excess twist density Ω and local curvature κ. The total energy is thus given by

$$E_{rod} = \frac{1}{2} \int A \left(\kappa^2 + \omega \Omega^2\right) ds$$

Equ. A.9

where Ω is computed from *Lk* and the shape of the filament centerline with the help of the Călugăreanu theorem (40)

$$Lk = T\omega + \omega r$$

Equ. A.10

where twist ($T\omega$) is defined as

$$T\omega = Lk_0 + \frac{1}{2\pi}\int_0^L \Omega \, ds$$

Equ. A.11

and writhe (Wr) is defined in the standard way (41). Scaling by kBT, the energy function can be made dimensionless, leaving the (width-to-length) aspect ratio ds/L and ω as the only parameters of the model. This approach was used by Coleman and Swigon (42) to categorize equilibrium shapes of looped filaments for a single aspect ratio d_s/L = 8.2 ×10−3 (corresponding to a DNA minicircle of length 718 bp and d_s = 20 Å), which constitutes a benchmark for our computations. Coleman & Swigon began by solving the problem of a free beam segment subject to boundary conditions at each end, and two beams in contact along a contact line. Such solution fragments can be glued together at contact points to form a looped solution, subject to gluing conditions that ensure the continuity of the first two derivatives and appropriate jump conditions to account for beam-beam steric forces. This approach uses the same expression for the beam energy but addresses the energy minimization in a different way—either by solving an ordinary differential equation subject to appropriate boundary conditions when no contact forces are needed, or by direct minimization subject to no-overlap constraint when contact forces are present.

Determination of critical Δ*Lk*

The stability of a computed minicircle shape depends on its Δ*Lk* (42, 43). For sufficiently small | ΔLk|, a flat circular configuration is the only equilibrium solution. Upon increasing |Δ*Lk*|, at a thickness-dependent threshold value of critical (minimal) Lk_{crit}, a figure-8 solution becomes admissible and the flat circular and figure-8 shapes coexist. Supercoiling further, above the thickness-independent threshold of $Lk_{max} = \sqrt{3}/\omega$, the flat circular shape is no longer a solution and only writhed configurations exist (43). For the prediction of Lk_{crit} for initial writhing of a minicircle, an approach based on solving an ordinary differential equation for centerline shape was used. The minimal value of |Δ*Lk*| required for writhing can be characterized by the existence of a configuration with a single contact point but with zero contact force. An ordinary differential equation was written for the beam centerline with minimum energy in a Cartesian parametrization subject to the boundary condition of a single contact point and no contact force and solved numerically using Mathematica, where boundary value problems are solved by the shooting method with conjugate gradient descent. For any given separation of the centerline at the self-contact location, one value of residual twist density was found. The relationship between the two was used to derive $Lk_{crit}(d_s/L)$.

Energy minimization of DNA minicircle shapes

Having determined the range of |Δ*Lk*| for which writhed configurations can be stable, the space of admissible configurations was examined and those configurations that minimized the elastic

energy were investigated. Here, because of the presence of contact forces, a different numerical method was used. Representing the centerline shape with periodic cubic splines, direct energy minimization was performed over all possible shapes without self-intersections. Curves with 16 nodal points with enforced dihedral symmetry were subjected to a Monte Carlo minimization procedure. The bending energy was calculated directly from curvature using the adaptive Simpson's algorithm and taking advantage of the twice continuously-differentiable nature of the cubic splines. The precise estimation of *Wr*, required to compute *Tw*, was performed by approximating the curve by 200 linear segments and using an algorithm proposed by Levitt (44) to deal with the singularities of the Gauss formulation. Steric interaction was introduced by tracking self-intersections through a large number (20 L/d_s) of sample points along the curve and a suitable steric energy penalty. Length constraint was imposed by computing the apparent length at each optimization step using the adaptive Simpson's method and by imposing an energy penalty for the deviation from the prescribed length. Multiple sets of different control parameters for numerical optimization were tested to ensure both fast convergence and satisfactory precision. All final computations were done with identical discretization and penalty characteristics. Final values of penalty parameters, Monte Carlo procedure parameters and initial conditions are available on Github.

Models for hydrodynamic radius

Solutions for Stokes flow around a slender toroidal object were developed by Johnson (45) that provide an asymptotic approximation in terms of slender-body theory. A fully analytical approach based on toroidal harmonics used by Goren & O'Neill (46) allows exact computations of all elements of the mobility matrix for a torus with an arbitrary aspect ratio. For a rigid axially symmetric particle of a given length L and hydrodynamic thickness dh, the mobility coefficients for translation along the symmetry axis and perpendicularly to it, μ_z (L, d_h) and μ_x (L, d_h), respectively, are sufficient to compute the hydrodynamic radius, R_h, by taking the inverse of their arithmetic mean

$$R_h = \frac{1}{6\pi\eta}\left(\frac{3}{2\mu_x(L,d_h)+\mu_z(L,d_h)}\right)$$

Equ. A.12

To theoretically determine R_h for an arbitrarily shaped molecule, a rigid bead model of its structure was constructed and its hydrodynamic radius was calculated using the ZENO software package (47, 48). In our case, the configuration of a minicircle was represented by 400 spherical and overlapping beads placed on the shape centerline, with diameters corresponding to the hydrodynamic thickness of the DNA molecule and the distance between overlapping beads summing up to the length of the molecule. This structure was then used to evaluate Rh for the composite particle. The diffusion coefficient at a given T and η is calculated from the definition of R_h in Equ. A.2. The sedimentation coefficient is obtained from the Svedberg relation, Equ. A.5.

Results

Rationale

The study of DNA supercoiling and curvature has benefited from multiple complementary theoretical and experimental approaches. The combination of multiple approaches takes advantage of the knowledge gained from each approach while helping shore up their individual limitations.

We previously [13] used 336 and 672 bp minicircles to study the effect of supercoiling and looping on DNA structure and have extensive 3-D structural data. AUC requires more material than cryoET (13), AFM (19), or other biochemical and biophysical analyses (12). A wide range of supercoiling was explored by testing relaxed, nicked, supercoiled, and hyper-negatively supercoiled minicircle samples. Supercoiled minicircle DNA was obtained from bacterial cells and, therefore, approximates the steady-state level of supercoiling in bacteria. Hypernegatively supercoiled samples have been further manipulated to increase the level of supercoiling and allowed us to determine whether AUC can distinguish different topoisomers from each other. To determine the effect of circularity, linearized DNA samples were also analyzed.

We characterized the minicircle samples by gel electrophoresis, which allows the topoisomer distribution of each sample to be precisely determined. Polyacrylamide gel electrophoresis effectively separates minicircle topoisomers and provides some insight into the conformational differences, although the theory underlying the differential migration is not fully understood.

We then applied advanced theoretical modeling to see whether it can explain the previously observed 3-D conformations of these minicircles [13]. It was reasonably successful, and these advanced mathematical models could then be used to help analyze and interpret AUC data. Algorithms used elsewhere to interpret AUC data approximate shape as a sphere to a line, but because of the diversity of shapes of DNA, in this paper a theoretical model was constructed to determine specific shapes and use them to evaluate diffusion and sedimentation coefficients.

Electrophoretic characterization of DNA minicircles

DNA minicircles were analyzed by polyacrylamide gel electrophoresis. Both the helical repeat and conformation of DNA are sensitive to solution conditions (49, 50). Here we used the same conditions used in analytical ultracentrifugation experiments (150 mM NaCl and 10 mM $CaCl_2$). Supercoiled topoisomers migrated much more rapidly on the polyacrylamide gel than relaxed topoisomers (Figure A.1). In comparison, the different topologies had relatively similar mobilities on an agarose gel (Supplementary Figure A.1). This difference in migration on polyacrylamide gels can be at least partially explained by the relative compactness of supercoiled minicircle conformations (13). The nicked and relaxed topoisomers had near-identical migration, suggesting that the single-strand break in the nicked minicircle does not significantly affect the global conformation. The lack of difference is explained by the number of helical turns in the 336 bp minicircle studied being close to a perfect integer value of 32 (under these conditions), resulting in the base pairs flanking the nick site being in close rotational alignment, allowing for favorable base stacking across the nick (12). We previously showed that when the rotational alignment is out of phase (i.e., when the number of helical turns deviates from a perfect integer value), the effect of a nick on polyacrylamide gel migration is much more pronounced (12).

The topoisomer distribution for the samples taken through to AUC analyzes was measured from quantification of digital images of the fluorescently stained gels using image analysis software. The 'supercoiled' 336 bp sample contained primarily ΔLk=−3 (48 %), ΔLk=−2 (41 %), and ΔLk=−1 (7 %) topoisomers. The sample also contained trace amounts of nicked 336 bp (1 %) and supercoiled 672 bp (3 %) minicircle DNA. The topoisomer distribution obtained (mean $\sigma \sim -0.08$) reflects the supercoiling level in the bacterial strain used to generate the minicircles. The 'hypernegatively supercoiled' sample contained primarily ΔLk=−6 (61 %) and ΔLk=−5 (33 %) topoisomers, with trace amounts of nicked 336 bp (4 %) and supercoiled 672 bp (3 %) minicircle DNA. This sample (with mean $\sigma \sim -0.15$) is representative of the very high levels of dynamic supercoiling generated transiently during transcription.

The supercoiled 672 bp sample contained primarily ΔLk=−4 (63 %), ΔLk=−5 (24 %), ΔLk=−6 (6 %) topoisomers, and trace amounts of nicked 672 bp (2 %), ΔLk=−2 (4 %) and ΔLk=−3 (2 %) topoisomers. The identity of the topoisomers present in the supercoiled 672 bp sample was determined on a separate gel with 672 bp topoisomer markers. The topoisomer distributions of each sample are compiled in Supplementary Table A.1.

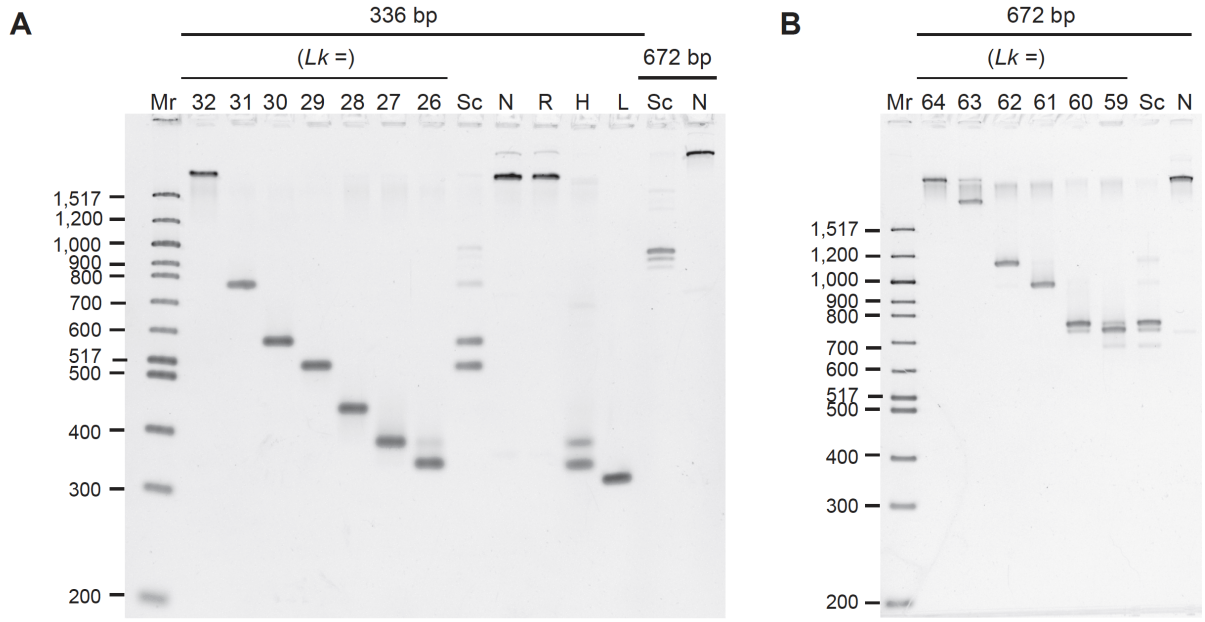


Figure A.1. Electrophoretic mobility of minicircle DNA. **(A)** DNA samples were analyzed by polyacrylamide gel electrophoresis (5 % polyacrylamide) in 150 mM NaCl and 10 mM $CaCl_2$ (the same conditions used in analytical ultracentrifugation). Mr: 100 bp DNA ladder, lanes 2–8: 336 bp minicircle topoisomer markers (*Lk* as indicated), lanes 9–13: 336 bp minicircle DNA samples (Sc: "supercoiled", N: nicked, R: relaxed, H: "hypernegatively supercoiled," L: linear), lanes 14–15: 672 bp DNA samples (Sc: "supercoiled," N: nicked). **(B)** Determination of topoisomer identity in 672 bp samples. DNA samples were analyzed by electrophoresis on a 4 % polyacrylamide gel in the presence of 10 mM CaCl2. Mr: 100 bp DNA ladder, lanes 2–8: 672 bp minicircle topoisomer markers (*Lk* as indicated), lanes 9–10: 672 bp DNA samples (Sc: "supercoiled" as isolated from the bacteria, N: nicked).

Table A.1. Apparent partial specific volume for DNA minicircle topoisomers in the buffer.

	Partial specific volume[a] in ml/g	
Sample	333 bp	672 bp
Linear	0.479	ND
Relaxed	0.470	ND
Nicked	0.469	0.495
Supercoiled	0.488	0.494
Hypernegatively supercoiled	0.479	ND
Average value[b]	0.482 ± 0.011	

[a]Determined by global sedimentation velocity analysis using the known molar masses. [b]For all the 336 bp and 672 bp minicircle DNA species. ND, not determined.

Analytical ultracentrifugation of DNA minicircles

Apparent PSVs obtained from a global multispeed genetic algorithm-Monte Carlo analysis for each minicircle sample are summarized in Table 2.1. The derived PSV values did not show any apparent pattern that would indicate a dependence of the PSV on topoisomer conformation, and resulted in a near constant value of 0.482 ± 0.011 ml/g over all tested minicircles (see Table A.1). The average PSV for 336 bp minicircles was 0.477 ml/g while the average PSV for 672 bp minicircles was 0.494 ml/g.

Using the determined average PSV value of 0.482 ml/g, a frictional ratio f/f_0 was derived from the sedimentation and diffusion coefficients obtained in the global analysis. Plots of the frictional ratio as a function of sedimentation coefficient and the diffusion coefficient as a function of the sedimentation coefficient are shown in Figure A.2. DNA topology had a significant effect on sedimentation and diffusion coefficients. In contrast to polyacrylamide gel electrophoresis, for which linear migrates fastest (Figure A.1), circular molecules (nicked, relaxed, supercoiled, and hypernegatively supercoiled) all sedimented faster than linear.

AUC was additionally able to differentiate relaxed, supercoiled, and hypernegatively supercoiled samples. The supercoiled or hypernegatively supercoiled samples containing a mixture of topoisomers behaved as single species in AUC. This result can be rationalized by noting that the topoisomers in experiments differ in ΔLk by 1 only. While the difference between nicked and supercoiled minicircles is clearly seen, it may be difficult to resolve by AUC minicircle topoisomers differing by only one or a few ΔLk.

Predicted shapes of DNA minicircles

Before employing our mathematical models to predict the effect of supercoiling and curvature on the hydrodynamic properties of DNA, we needed to first test how well these models predicted known equilibrium 3-D minicircle shapes. We adopted the strategy of building a coarse-grained representation of the equilibrium shapes of the minicircles obtained using our energy minimization codes. These models are reduced representations of macromolecules still capable of

retaining key physical aspects (51). This approach has been highly successful in calculations of biomolecule properties in solution (52). Having found the shapes of DNA minicircles, we calculated the hydrodynamic radius for each conformation. The hydrodynamic radius was used to calculate the sedimentation and diffusion coefficients. We also tested our models against the 3-D structures of these minicircles, previously observed experimentally (13). Our aim was to develop a practical predictive theoretical framework to determine the measured transport coefficients.

Electrostatic screening and Brownian contributions DNA molecules have substantial charge, which can be exploited, e.g., in electrophoretic measurements of DNA of different lengths. A qualitative comparison of the elastic and electrostatic forces is possible by considering scaling arguments. The Debye-H¨uckel equation is a well-established model of electrostatic interaction in a buffer containing counterions (53). In this approach, the interaction potential decays exponentially with separation due to screening. The decay rate is quantified by a characteristic distance, the Debye length (RD). Comparing RD with the typical distances between base pairs gives a crude estimation of the influence of electrostatic forces. We estimated RD for our setup to be 1.45 Å from the ionic strength of 230 mM using an ionic strength-based estimate Embedded Image, where Cs is the molar salt concentration in moles per litre and $\kappa_0 = 0.329$ $Å^{-1}l^{1/2}mol^{-1/2}$, as reported by Lim et al. (54). Thus the Debye length is much smaller than an average distance between different segments of the DNA molecule. Notably, this value of RD was also much lower than that of earlier work, such as 30 Å in Ref. (55) for different buffer conditions.

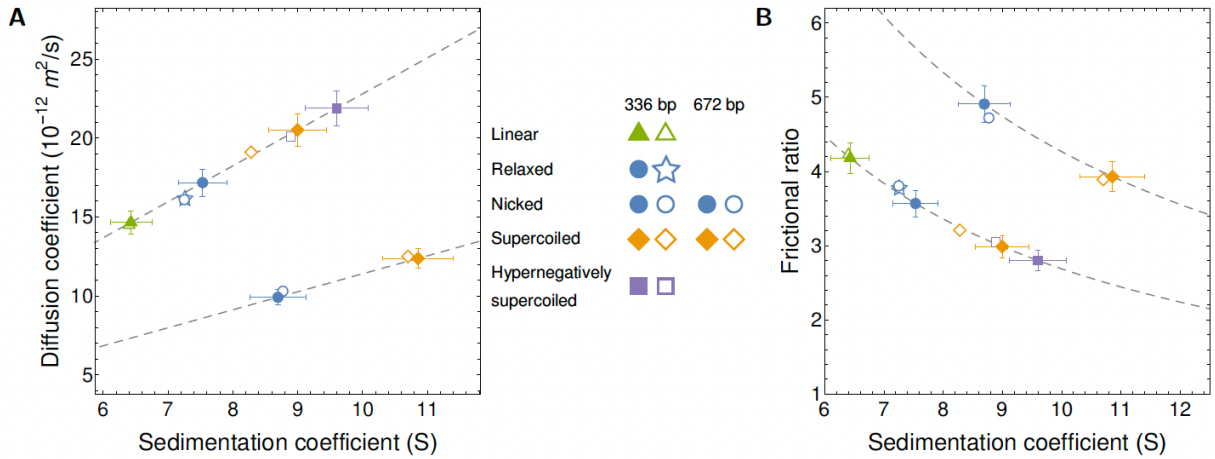


Figure A.2. Measured and predicted diffusion and sedimentation coefficients for DNA minicircles. AUC measurements using global Monte Carlo-Genetic Algorithm analysis are marked as empty symbols. Theoretical predictions are presented with filled symbols. (A) Diffusion coefficient as a function of the sedimentation coefficient for topoisomers of 336 bp (upper branch) and 672 bp (lower branch) minicircle DNA. Dashed lines represent the constant mean value of PSV determined from AUC experiments on which theoretical predictions of the sedimentation coefficient are based. Experimental and theoretical data for 336 bp and 672 bp relaxed and nicked minicircles overlay almost completely and thus are impossible to discern in the plots. (B) Frictional ratio as a function of the sedimentation coefficient. Sedimentation coefficients s are measured in svedberg units (S), with 1 S=10−13 s.

The persistence length *P* of polyelectrolytes is the sum of two contributions,

$$P = P_{in} + P_{el}$$

Equ. A.13

where P_{in} is an intrinsic persistence length due to the rigidity of the backbone, and P_{el} is an electrostatic persistence length, which accounts for buffer-dependent repulsion between neighboring ionic sites (56). The latter can be related to the Debye length as $P_{el} = (4\kappa^2 l_B)^{-1}$, with lB being the Bjerrum length, according to the Odijk-Skolnick-Fixman theory (57, 58, 59). For linear or relaxed double-stranded DNA at room temperature, the persistence length in 0.1 M NaCl is approximately 500 Å (150 bp) (60). Based on persistence length, one can define effective bending and twisting energies for a circular shape as

$$E_{bend} = 4\pi^2 k_B T \frac{P}{L} = 4\pi^2 \frac{A_{eff}}{L} \quad \text{and}$$

Equ. A.14

$$E_{twist} = 6\pi^2 k_B T (\Delta Lk)^2 \frac{P}{L} = 6\pi^2 (\Delta Lk)^2 \frac{A_{eff}}{L} \quad ,$$

Equ. A.15

which, via Equ. A.13, include both elastic and electrostatic contributions. Here, $A_{eff} = k_B TP$ is the effective bending rigidity of the DNA. The remaining long-ranged electrostatic contribution E_{lr} can be estimated as the interaction energy between N equal charges *q* at a typical distance comparable to the radius of the loop, $L/2\pi$, which amounts to

$$E_{lr} \approx \frac{q^2 N}{2\varepsilon L} \exp\left(-\frac{L}{2\pi R_D}\right) \quad ,$$

Equ. A.16

where ε is the dielectric permittivity of the buffer. Note that these estimates do not take into account short-range electrostatic interactions between distant parts of the DNA that may come close together when supercoiled. For a circular loop with $|\Delta Lk| = 2$ and equivalent length of 336 bp, we found $E_{bend} \approx 17 k_B T$, $E_{twist} \approx 105 k_B T$ and a negligible value of E_{lr}. The effective bending and twisting energies considerably exceed the typical energy of thermal fluctuations, k_{BT}, which is somewhat surprising given that the minicircles are longer than P. This result may be a consequence of the relatively high energy stored in the form of bending energy Ebend, which renders the circular configuration less prone to Brownian shape disturbances than a torsional-stress-free linear DNA. We therefore use the stiff beam approximation to describe minicircle shapes.

Critical ΔLk for writhing The shape of a DNA minicircle depends primarily on $|\Delta Lk|$ and the width to length ratio d_s/L. To predict 336 bp and 672 bp minicircle conformations for $\Delta Lk = 0$ to $|\Delta Lk| = 5$, we developed a theoretical framework that assumed that DNA can be modelled by a continuous elastic beam of ΔLk, with the steric thickness d_s and stiffness A as determined by the inter-phosphate distance and persistence length, respectively. The previous theoretical study of

Coleman & Swigon (42) focused on a circularised 718 bp DNA fragment and determined the structure of possible stable configurations, describing the contact diagrams in great detail. Our results were consistent with their findings for a particular thickness to length ratio (d_s/L = 0.0082) but were applied to an arbitrary aspect ratio.

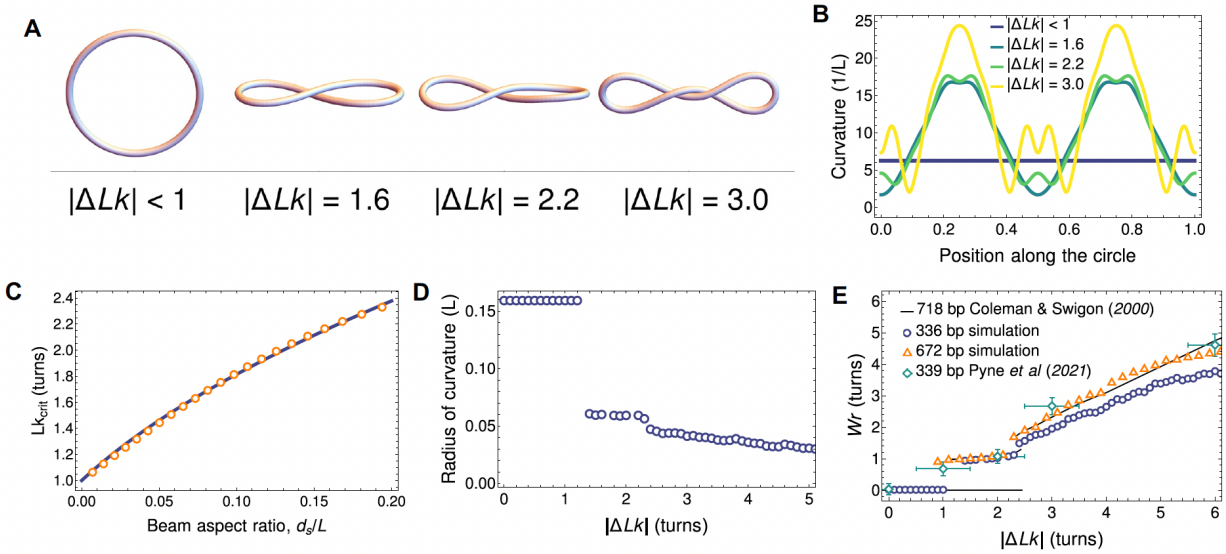


Figure A.3. Elastic equilibrium shapes of model DNA minicircles. **(A)** Energy minimizing shapes of 336 bp minicircle of various ΔLk (with d_s/L=0.018). For small $|\Delta Lk|$ values (<1.6), loops adopt open circle configurations. For intermediate values (ΔLk = 1.6 or 2.2), a single point of polymer contact is observed; for larger values of ΔLk (>2.2), continuous contact is observed. **(B)** Curvature distribution along the supercoiled loop centerline in shapes corresponding to panel (A). The position along the loop is measured from the point of contact (or the center of symmetry in multiply touching configurations). **(C)** Lk_{crit} above which a writhed configuration can be stable, plotted as a function of the beam aspect ratio (steric diameter to length, d_s/L). Solid line shows the approximation of eq. (17). **(D)** Minimal radius of curvature along the loop as a function of ΔLk for the 336 bp minicircle. This radius decreases monotonically with ΔLk, leading to increasing bending stresses. (E) Writhe of energy-minimizing shapes as a function of the aspect ratios d_s/L = 0.018, 0.0090, and 0.0082 for 336 bp, 672 bp, and 718 bp DNA minicircles respectively. For small values of $|\Delta Lk|$, only flat (open circle) configurations are permitted but above Lk_{crit}, one or more twists is relaxed by writhing. For large values of $|\Delta Lk|$, around 90 % of torsional energy is relaxed by shape deformation. Our results are shown next to the continuum model predictions of Coleman and Swigon (42) and atomistic MD simulations of Pyne et al. (19).

In Figure A.3A we present the shapes obtained by our minimization procedure. Close to ΔLk = 0, circular configurations were stable. At a critical value of $|\Delta Lk|$, a figure-8 shape became energetically favorable, and the loop writhed to relax excess twist. For higher $|\Delta Lk|$, the number of selfcontacts increased, leading to a more writhed configuration. In this process, regions of high and low curvature emerged along the loop, as sketched in Figure A.3B, with higher curvature

present for higher values of $|\Delta Lk|$, as expected intuitively. For open configurations, the curvature was nearly constant and equal to $2\pi/L$. The position along the loop changed from 0 to 1 and was measured from the point of contact or (in the case of multiple contacts) from the center of symmetry. With increasing $|\Delta Lk|$, regions of increased curvature appear far away from self-contact points.

For larger $|\Delta Lk|$, increased curvature was present even at points close to the contact line. The variation of curvature along the loop centerline, shown in Figure A.3C, might hint at sites of potential base pair instability or other configurational changes. The effect of localized curvature regions would be further compounded by the transmission of mechanical stress along the DNA backbone to promote DNA kinking and base pair disruption at distant sites (13, 19, 61, 62) as reported by Fogg et al. (12).

The value of Lk_{crit} for the open circle-figure-8 transition depends on the aspect ratio ds/L. In Figure A.3C, we show the dependence of Lk_{crit} on the aspect ratio numerically, showing that for thicker beams a transition to writhed configurations required an increased negative supercoiling. A very thin filament can be stably writhed for almost any value of $|\Delta Lk|$ greater than 1. However, for beams with a larger thickness, more torsional stress was required to stabilize writhed configurations and prevent them from unwrithing to a circular, open configuration.

For increased values of d_s/L, we also empirically found a convenient approximate expression for

$$Lk_{crit} = \sqrt{1+23.1\frac{d_s}{L}} \quad ,$$

Equ. A.17

which we show in Figure A.3C. To further characterize the writhed shapes, in Figure A.3D we plotted the minimal radius of curvature of a twisted beam as a function of $|\Delta Lk|$. For low values of $|\Delta Lk|$, the constant value reflected the purely circular equilibrium. Above Lk_{crit}, the loop became writhed, with an increased curvature at the apices. The monotonic decrease of radius of curvature with $|\Delta Lk|$ showed that increasingly supercoiled beams tended to have tighter bends and, therefore, stored larger bending energy. Finally, in Figure A.3E, we present the writhe of the resulting configuration, calculated with our method as a function of $|\Delta Lk|$. The values were calculated for both lengths–336 and 672 bp. For comparison, we also plotted the results of computations of Coleman and Swigon (42) for L = 718 bp as a solid line, and results of base pair-resolution MD simulations of Pyne et al. (19) for a 339 bp minicircle, confirming the observed trend. The total Wr as a function of $|\Delta Lk|$ seemed to be weakly dependent on the minicircle length. Writhing, therefore, emerged as a universal mechanism of stress release for twisted loops.

Shape and stability of supercoiled configurations. Although purely elastic considerations suggest a plethora of possible writhed configurations, we mostly observed the simple minicircle conformations in the cryoET measurements of Irobalieva et al. (13), rather than the more intricate shapes. The latter have higher energies and are thus less frequently realized (42). For a given ΔLk and at low temperatures, the shapes associated with higher energies were unfavorable compared to the ground state solution determined by our energy minimization procedure.

Equilibrium shapes measured in cryoET experiments in Ref. (13) were divided into eight groups termed open circle, open figure-8, figure-8, racquet, handcuffs, needle, rod, and other. For a quantitative comparison, we reduced the complexity by distinguishing only between contact-free (corresponding to open minicircle configurations) and selftouching (corresponding to writhed minicircle configurations) solutions. Using the elastic beam framework, we determined the regions of stability of these solutions in terms of ΔLk for a given DNA loop length. We present them in Figure A.4A. The orange-shaded region close to ΔLk=0 favors an open circle as the stable configuration. For undertwisted configurations, we identified a region of multistability, where we found both open circular and figure-8 shapes. Finally, when the inherent ΔLk was large enough, marked by the blue shading in the figure, the self-touching shapes became the only stable energy minimum. The stable circular shape region narrowed down from $|\Delta Lk| < 1.1$ for 336 bp minicircles to $|\Delta Lk| < 1.2$ for 672 bp and 718 bp minicircles.

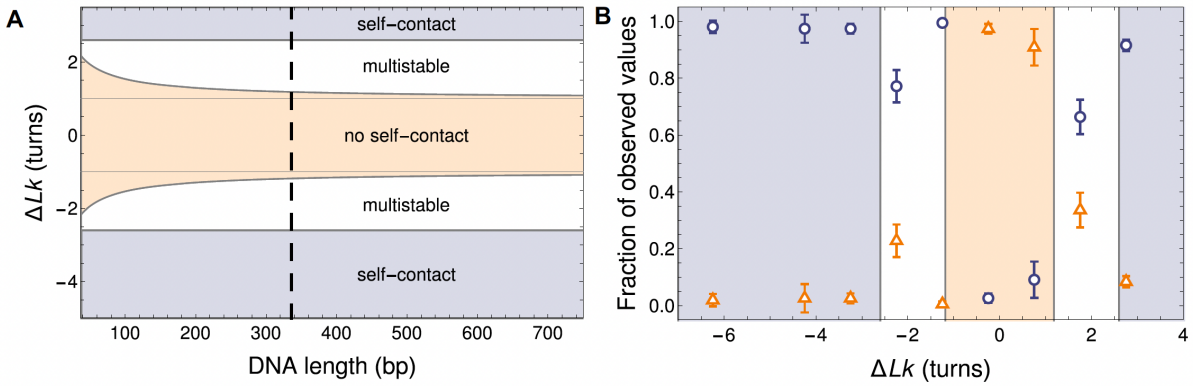


Figure A.4. Regimes of shape stability for model DNA minicircles. (A) Phase diagram of different regimes of stability of supercoiled DNA as a function of ΔLk and DNA length. The phase space is divided into three regions: only open circular configurations permissible (tan-shaded area), both circular and writhed configurations permissible (multistability; white area), and only writhed configurations permissible (blue-shaded area). The dashed line marks the 336 bp minicircles, for which experimental shape data are available. (B) Plot of the relative occurrence rate of configurations with (blue circles) and without (orange triangles) points of self contact, based on cryoET shapes of 336 bp minicircles measured by Irobalieva et al. (13) and overlayed on our theoretical stability predictions with the same color coding as in panel (A). Error bars represent 2 standard deviations.

For the particular length of L=336 bp, in Figure A.4B we compared the theoretical predictions to the cryoET data (13). We measured the fractions of open and self touching configurations in the population of minicircle shapes. For ΔLk close to zero, we saw that most of the shapes were open loops, with only a fraction of about 10 % showing self-contact. This situation changed with increased ΔLk. In the predicted region of multistability, we saw that the fractions of writhed configurations increased, but there was still a pronounced population of open circular shapes that vanished almost completely when increasing $|\Delta Lk|$. This observation is in excellent agreement

with our prediction for the new stability region at $|\Delta Lk| > 2.6$. In particular, the most relaxed state (closest to $\Delta Lk = 0$) was slightly undertwisted ($\Delta Lk<0$). For configurations with $\Delta Lk= -0.2$ and $\Delta Lk= 0.8$, only open configurations were predicted while for $\Delta Lk= -2.2$, -1.2 and 1.8, we observed large conformational variability corresponding to two solutions in the uniform elasticity model. For $\Delta Lk = -4.2$, -3.2, and 2.8, open configurations were no longer permitted by the uniform elasticity theory and these open configurations were largely absent in the prior cryoET measurements for these topoisomers. Interestingly, an outlier at $\Delta Lk = -1.2$ was observed in the cryoET data, where we saw a surprising lack of open circular configurations, which could not be simply explained by our coarse-grained model. One potential explanation for this puzzling observation has been postulated that involves a coupling of limited base pair disruption with writhing (12).

Predictions of hydrodynamic radius of DNA minicircles with different ΔLk

Non-writhed configurations. Configurations with small values of ΔLk adopt toroidal conformations; for those configurations, we have found a convenient expression for the hydrodynamic radius, consistent with asymptotic solutions by Johnson and Wu (45) and Johnson (63)

$$R_h = \left(\frac{L}{2\pi}+\frac{d_h}{2}\right)\frac{72\pi\left(\frac{1}{16x^2}+1\right)x}{25x+6(11x+8)\log 8x+16} \quad ,$$

Equ. A.18

with $x=L/\pi d_h$. Here, log denotes the natural logarithm. Notably, Equ. A.18 agrees with the numerical results of Goren and O'Neill (46) even for non-slender tori. For more slender tori, when $L/d_h >30$, a simpler expression can be fitted without loss of accuracy, given by

$$R_h = L\left(\frac{11}{6}\log\left(\frac{2L}{d_h}\right)+1.13\right)^{-1}$$

Equ. A.19

as obtained by Adamczyk et al. (64) from the numerical results of the bead model shape approximations. Qualitatively, R_h is of the same order of magnitude as the experimental DNA lengths and thus it approximately scales with the mass of the DNA molecule. This scaling for approximating DNA is in contrast with globular models used for some proteins that scale with the cubic root of their mass instead.

Writhed configurations. Configurations with intermediate values of $|\Delta Lk|$ adopted either toroidal or writhed shapes. To generate hydrodynamic predictions forDNA minicircles with larger values of $|\Delta Lk|$,we combined elastic energy minimization with hydrodynamic bead models. To this end, we used equilibrium shapes obtained for each $|\Delta Lk|$ within the elastic beam model described before, and produced their representation as a collection of 400 rigidly attached and overlapping spherical beads arranged such that the length and thickness of the DNA molecule were left

unchanged. The bead model was then used to calculate the hydrodynamic mobility of the conglomerate within the ZENO package (47). We estimated the error of our predictions of the hydrodynamic radius to be about 5 % by comparing the results of bead-model calculations to known analytical solutions for highly symmetric shapes of model molecules.

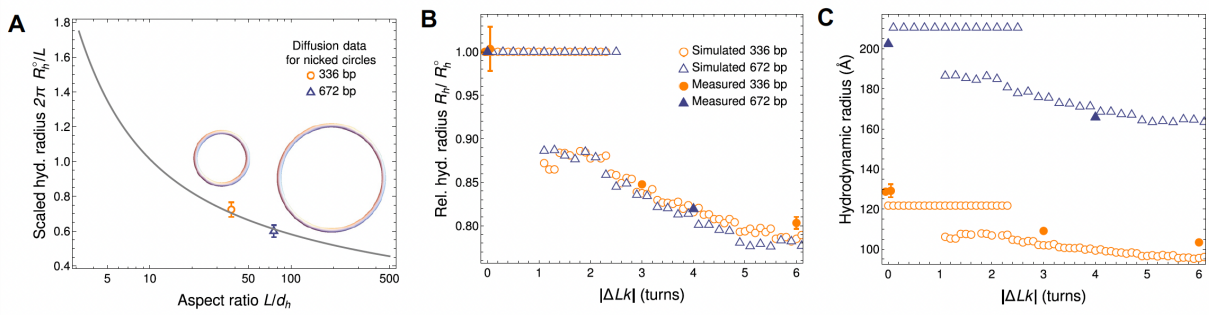


Figure A.5. Hydrodynamic radius of DNA minicircles. (A) Hydrodynamic (hyd.) radius R_h° of open-circular DNA, scaled by the geometric radius $L/2$ of a torus, plotted for a range of DNA aspect ratios L/d_h. Comparing the ZENO results for toroidal particles (solid line) to diffusion measurements of minicircles with 336 bp and 672 bp yields the fitted (common) hydrodynamic thickness d_h = 29.4 Å. We note that ZENO approximation yields high precision results for toroidal particles (48). This value was used in all subsequent computations. Circular sketches representing molecules preserve both the relative scale and thickness. Note the logarithmic scale on the horizontal axis. (B) R_h of minicircle shapes, relative (rel.) to the hydrodynamic radius R_h° of the relaxed, open-circular shape for a range of $|\Delta Lk|$. We present results of simulations (open circles and triangles) along with experimental data for 336 bp (filled diamonds) and 672 bp (filled squares) minicircles. For 336 bp (nicked), 336 bp (relaxed), and 672 bp (nicked), the simulations show a region of constant Rh where the shape is independent of $|\Delta Lk|$. For supercoiled 336 bp containing a mixture of $\Delta Lk = -1$, -2 and -3, the Rh is about 15 % smaller than that of an open circle. For larger values of $|\Delta Lk|$, the theoretical approach seems to correctly grasp R_h of the resulting highly compact conformers. (C) Absolute values of the R_h for minicircles from (B), with the same symbols.

Our modeling approach reduced the problem of finding diffusion and sedimentation coefficients to the computation of the hydrodynamic radius R_h. Once calculated theoretically, R_h was used to predict the diffusion coefficient from the Stokes-Einstein relation, Equ. A.2), provided that the viscosity of the environment η is known. The presented results were adjusted for the buffer viscosity. The prediction of R_h for DNA of any shape required additionally the knowledge of the hydrodynamic thickness of the loop d_h. To determine the effective thickness of the molecules (related to the existence of a hydration layer), we assumed that, as a local property, it does not depend on the shape. Then, we used the measured AUC data for nicked and relaxed (ΔLk=0) 336 bp and 672 bp minicircles, which we assumed to be toroidal. Knowing the dependence of the DNA Rh on the molecule aspect ratio from the ZENO software tool, we compared it to the AUC results for the Rh of the 336 bp and 672 bp DNA minicircles, and we fitted the same value of dh to both minicircles to reproduce their experimental Rh. The dependence of R_h on the aspect ratio

is presented in Figure 5A, together with the two measured values. Fitting the theoretical curve to these two data points yielded the hydrodynamic diameter of d_h = 29.4 Å (corresponding to d_h/L = 2.6 × 10^{-2} for the 336 bp minicircles and d_h/L = 1.3 × 10−2 for the 672 bp minicircles).We used this value for all subsequent calculations. We note, however, that the estimation of dh from diffusion measurements cannot be precise due to the logarithmic dependence of hydrodynamic parameters on this value, so the fitted value should be treated as more approximate than the number of digits provided.

The value of excess thickness over inter-phosphate distance was significantly larger than that reported by Fernandes et al. (65) (22.8 Å)—without the details of the solvent ionic strength; however, as argued by Penkova et al. (66), the hydration shell can be as thick as 16 Å, corresponding to diameters as large as 40 Å. This value is sensitive not only to the ionic strength of the solvent, but also to the details of ion composition (67). Moreover, small deviations from the toroidal shape of the nicked and relaxed DNA shapes are expected, as caused by Brownian motion.

Figure A.5B demonstrates that the difference of shapes with different |Δ*Lk*| is significant and similar in theory and experiments. We present therein the hydrodynamic radius for writhed configurations calculated for the two investigated DNA lengths with the same thickness (d_h = 29.4 Å). The plot shows the radius relative to that of a toroidal particle of the same d_h, denoted by R_h°. We normalized theoretical results by calculations at Δ*Lk*=0, while experimental data are rescaled by respective results for a nicked/relaxed configuration of a minicircle of a given length. Open circles and triangles in the graph mark theoretical results. For the 336 bp minicircle, two experimental values were available, namely that of a nicked and relaxed configuration. In these cases, we rescaled the experimental data by the mean of the two radii. The values of the relative radii are equal to unity for Δ*Lk* = 0 by definition, but even the non-zero Δ*Lk* simulations predicted a region of stable circular configurations with unchanged R_h.

Increasing |Δ*Lk*| led to the emergence of highly writhed configurations, which tended to be more compact and therefore had a smaller R_h than the open circular configurations seen at low |Δ*Lk*|.We predict that topological writhing could reduce R_h by about 15 % in experimentally relevant conditions as shown in Figure A.5B.

We additionally plotted available experimental results for relaxed and supercoiled minicircles, which compared favorably with theoretical predictions. The theory seems to correctly determine R_h obtained from AUC measurements, which opens an efficient route to calculate the hydrodynamic transport coefficients also for highly writhed conformations. In Figure A.5C we also presented the same results in absolute terms. Because the fitted value of hydrodynamic thickness in Figure A.5A lies between the estimates based on 336 bp and 672 bp only, we see the value of R_h for Δ*Lk* = 0 to be slightly overestimated by the simulation for 672 bp and underestimated for 336 bp. We emphasize that this difference is an effect of the thickness fitting procedure. This systematic difference can be caused by slightly non-toroidal DNA shapes caused by Brownian motion.

Our AUC data, together with the theoretical modeling shown in Figure A.5, yield a length-invariant observation—the ratio of R_h of open circular (nicked or relaxed) minicircle topoisomers to the Rh of compact (supercoiled) topoisomers is approximately 5:4. This value holds true for

both 336 bp and 672 bp minicircles, even though the length L is longer than the persistence length P for 672 bp minicircles, with L≈4P. This length-invariant ratio of 5:4 holds true as long as thermal effects are negligible. This theoretical result is particularly robust because it is independent of the viscosity of the buffer or the minicircle length. We attributed this robustness to the logarithmic dependence of the hydrodynamic models on the thickness of slender bodies. Additionally, the ratio of R_h of 336 bp linear to R_h of 336 bp nicked is approximately 7:6. These results can also be viewed in absolute rather than relative terms (Figure A.3), where the theoretical predictions for the values of the sedimentation and diffusion coefficients are displayed together with the experimental results. We obtained a good agreement for the larger of the two minicircles (672 bp) and for linear, relaxed, and nicked samples of the smaller minicircle (336 bp). For highly writhed configurations of the smaller minicircle, however, theoretical predictions suggest a more compact conformation than observed. This result could be attributed to the limits of applicability of the linear elasticity theory to the very tight bend radii necessary for energy minimization in these configurations, or to a rough estimate of shapes and width of nicked and relaxed minicircles.

Table A.2. Comparison of predicted and measured diffusion and sedimentation coefficients in the buffer

		D (μm^2/s)		s(S)	
Sample	$\|\Delta Lk\|$	predicted	measured	predicted	measured
336 linear	not applicable	14.7 ± 0.7	14.5	6.4 ± 0.3	6.40
336 relaxed	0	17.2 ± 0.9	16.2	7.5 ± 0.4	7.26
336 nicked	0	17.2 ± 0.9	16.1 ± 0.4	7.5 ± 0.4	7.25 ± 0.02
336 supercoiled[a]	**3**, 2, 1	20.5 ± 1.0	19.0	9.0 ± 0.5	8.28
336 hypernegatively supercoiled[b]	**6**, 5	21.9 ± 1.1	20.1 ± 0.2	9.6 ± 0.5	8.89
672 nicked	0	9.9 ± 0.5	10.3	8.7 ± 0.4	8.77
672 supercoiled[c]	**4**, 5, 6	12.4 ± 0.6	12.5	10.9 ± 0.5	10.72

[a]Predictions done for the dominant species in the mix (ΔLk=−3) highlighted in bold. [b]Predictions done for the dominant species in the mix (ΔLk=−6). [c]Predictions done for the dominant species in the mix (ΔLk=−4).Where experimental errors are not shown, the confidence limits from the global Monte Carlo analysis of the AUC data were exactly 0. We roughly estimated the errors of theoretical values as 5 %. Sedimentation coefficients are measured in svedbergs, with 1 S = 10^{-13} s. Experimental confidence limits are reported only when the last digit of the result would be affected.

To test the effectiveness of our modeling approach, in Table A.2 we compared the measured and predicted values of D and s. To compute D, data from nicked samples were used as a calibration of the only fitting parameter: the effective hydrodynamic diameter of the DNA molecule. With

the use of d_h, all the D values were purely theoretical predictions based on the DNA length and the value determined in separate experiments.

Because supercoiled and hyper-supercoiled samples contain mixtures of minicircles with different ΔLk, theoretical predictions are given for the most common species in a mixture. These configurations are sketched in Figure A.6. To calculate s, we used the approximately constant value of the PSV determined from the AUC experiments. In all cases, the deviation between theory and experiment was below 8 %. The simplified hydroelastic model provides practical estimates of the experimentally accessible quantities, and thus may be used to discern populations of minicircles differing in ΔLk.

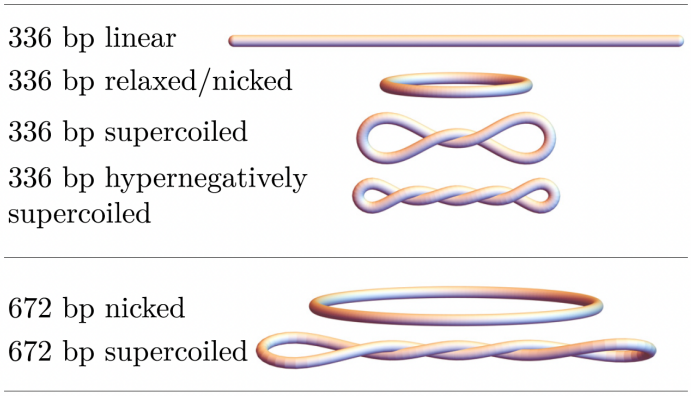


Figure A.6. Sketches of model shapes in a givenminicircle configuration used for hydrodynamic simulations with Lk specified in the caption of Table 2. The sketches have realistic aspect ratios (dh/L) and preserve the relative size.

Our modeling experiments reveal that the change of ΔLk by one turn (for example from $\Delta Lk = -4$ to $\Delta Lk = -5$) changes the frictional ratio by ~2 %—a value similar to theory-experiment deviations and similar to the noise level coming from theMonte Carlo optimizations of energy minimizing shapes. Such small gradients could be the reason why AUC measurements cannot distinguish topoisomers that differ in ΔLk by only a single turn, especially if the concentrations of the different topoisomers in the mix differ. The samples analyzed by AUC contained a mixture of topoisomers that behaved as a single species because the major species in each sample were within one ΔLk of each other.

Discussion

Methods that measure the hydrodynamic radius R_h, such as gel electrophoresis and diffusion-sedimentation AUC experiments, can be qualitatively described as sorting molecules by their size. Unlike for gel electrophoresis, however, when assessing diffusion via AUC, size contributions can be separated from viscosity, temperature, and PSV by the appropriate scaling, which yields R_h.

In electrophoresis, molecules squeeze their way through pores in the gel matrix. One of the major differences between agarose and polyacrylamide is the average pore size, with agarose gels typically having larger sized pores (68). In typical electrophoresis experiments, the primary determinant of the migration speed is the molecular weight and charge of the molecule. Here, the DNA molecule of a given length has a fixed charge and molecular weight, but the speed of migration through the matrix of small pores like the polyacrylamide gel varies greatly when the shape of the molecule is changed (although charge, and thus cations and hydration shell may be altered with negative supercoiling (Randall 2006;Randall 2009)). Comparing gel results to the AUC data allows us to confidently say that knowing the R_h is insufficient to predict its electrophoretic mobility as the gel electrophoresis and AUC separated the minicircle topoisomers differently. Supercoiling influences minicircle electrophoretic mobility much more than hydrodynamic mobility (as measured by AUC); the same change in supercoiling increases electrophoretic mobility by 400 % while increasing mobility only by 20%. Differences between the methods can be even more dramatic. The linearized 336 bp minicircle showed 600%increase of electrophoretic mobility, while at the same time (in the AUC measurements) having a 14 % smaller hydrodynamic mobility as compared to nicked or relaxed form.

Introducing circularization presents a significant obstacle to extending quantitative electrophoresis methods, such as one proposed by Ziraldo et al. (69). Calibrating gel electrophoresis measurements with a ladder of relaxed or nicked circular DNA of different lengths would be insufficient because the contribution of the degree of supercoiling strongly affects the apparent DNA mass derived from that method (at least by a factor of two). These difficulties are further compounded by the dependence on the absolute value of applied electrophoretic-mediated force, as reported by Iubini et al. (70). For large electrophoreticmediated forces, linear DNA is expected to migrate faster than the circular form, which is opposite to what occurred in AUC, while for small forces, the circular form migrates faster than the linear, which is the same as in AUC. One possible explanation for the differences between electrophoresis and AUC is that as the DNA is pulled through the pores of the gel matrix, it has to change its shape (this would be most significant for the ΔLk=0 samples) whereas there is no sieving in AUC. AUC, therefore, better reflects the solution properties of DNA, while gel electrophoresis offers higher resolution for separation by means we do not yet fully understand.

Diffusion-sedimentation measurements provided by AUC allow accurate hydrodynamic modeling. To elucidate the shapes and properties of DNA minicircles, we proposed a coarse-grained model that represents the DNA as an elongated, uniform elastic beam. By minimizing the elastic energy of a beam with a given superhelical density, we were able to find equilibrium shapes of the minicircles. We note that our coarse-grained models of minicircles are oblivious to their sequence and do not exploit information on sequence-dependent elastic properties. However, the uniform beam model of DNA elasticity presented here predicts shapes of DNA minicircles (shown in Figure A.3A and Figure A.6) that compare favorably with the direct observations of 336 bp minicircles (13) (shown in Figure A.4B).

Our theoretical model predicts a very weak dependence of the shapes of DNA minicircles with a few hundred base pairs on the aspect ratio d_s/L, as shown in Figure A.4A. According to the model, the change in ΔLk, alone, gives sufficient information to describe basic features of the

minicircle configuration: $\Delta Lk = 0$ is an open circle; $|\Delta Lk| = \pm 1$ is at the transition between open circle and writhed conformations; $|\Delta Lk| = \pm 2$ is multi-stable; and $|\Delta Lk| > 3$ always exhibits self-contact (compare with the corresponding shapes in Figure A.3A and Figure A.6). This result means that no matter whether the minicircle was 336 or 672 bp, the loss of three helical turns was enough to disallow the open circle conformation as an energyminimal solution. These results suggest that it is the absolute value of ΔLk and not the superhelical density $\sigma = \Delta Lk/Lk_0$ that governs the conformational landscape of small DNA minicircles (with the length having a small influence via the d_s/L ratio). This finding provides an important input to future models of circularized polymers.

Incorporating torsional interactions is more difficult than just bending interactions and is often neglected when constructing models of circularized molecules (71, 72, 73, 74). Because the torsional forces play a role even when ΔLk is close to zero, care must be taken when generalizing such models for the context of supercoiled DNA. We have shown here that regardless of the length of the DNA molecule the torsional forces are of the same order of magnitude as pure bending forces.

While both bending and shearing forces remain comparable in magnitude even for longer minicircles, the relative importance of thermal fluctuations increases proportionally to the length. For the short minicircles considered here, thermal effects are not strong enough to push them out of their global energy minimum corresponding to the presented shapes. The results of Coleman and Swigon (42), however, show that mechanically stable branched plectonemic configurations exist that correspond to local energy minima of higher elastic energy. These high energy states can be accessed provided that the forces from thermal motion exceed the elastic forces, which is the case for minicircles of lengths much longer than the persistence length. For the minicircles studied here, we assumed that the minimal elastic energy shape was rigidly maintained throughout an experiment. To predict hydrodynamic properties of longer DNA circles, prone to conformational flexibility due to Brownian effects, one has to use different modeling techniques, e.g., such as those inspired by models of intrinsically disordered proteins (75).

The elastic-energy minimizing centerline shapes were found for minicircles with different ΔLk with hydrodynamic thickness determined by calibration based on AUC experimental data. These shapes were then used to construct hydrodynamic models to compute Rh, which, in turn, was used to compare the theoretical diffusion and sedimentation coefficients with the AUC measurement. The comparison (Table A.2) shows general agreement, which is satisfactory given the simplicity of the underlying coarsegrained model. These results confirm the predictive capabilities of uniform elasticity models, combined with a hydrodynamic calculations package (ZENO) to interpret and guide AUC measurements. We expect that the proposed modeling strategy could be beneficially employed to similar problems in the dynamics of DNA and perhaps extended further to account for sequence-specific effects and modified intramolecular interactions. One way forward would be to introduce local, sequence-dependent elasticity by allowing the bending stiffness of the elastic beam model to vary along the centerline of the minicircle. This extension should aid the studies of localized kinking, seen in earlier experiments (61), and the formation of bubbles in DNA minicircles under negative supercoiling (12).

This work represents significant progress in understanding and modeling the sedimentation of a biological molecule with a complex and dynamic conformation. Sedimentation of roughly

spherical molecules (e.g., many proteins) is fairly well understood, but linear and minicircle DNA do not adopt a spherical conformation. Using 336 bp minicircles as a model system we are now able, for the first time, to fully test the theoretical models. We anticipate that this testing will allow us to further improve these models and expand the use of AUC to include a larger repertoire of important and complex biological molecules. We have successfully applied two stage modeling (combination of energy minimization to find the shape then applying hydrodynamic modeling for rigid configurations) to find hydrodynamic properties of a DNA molecule. Our approach developed here yielding the prediction of shape and hydrodynamic properties is applicable to other complex biomolecules.

Conclusion

The work presented here is a step toward understanding how supercoiling and curvature affect DNA shape and hydrodynamic properties, which in turn affect important DNA activities. The DNA solvation shell and counterions as well as DNA shape all affect DNA structure to influence how the DNA code is protected, accessed, modified, and activated. It will take a combination of approaches to fully understand this remarkable molecule.

Data Availability

The UltraScan software used to analyze the AUC data is open source and freely available from Github repository. The AUC data itself is available upon request from the UltraScan LIMS server at the Canadian Center for Hydrodynamics. The algorithm for finding minimal energy shapes, initial conditions, and final configurations can be found in the Zenodo repository 7501673 (76). Lk-Wr pairs computed from energy minimal shapes as well as hydrodynamic radii computed using the ZENO software can be found in the Zenodo repository Zenodo repository 7501675 (77).

Acknowledgements

The Authors thank Prof. De Witt Sumners and Dr. David Swigon for helpful remarks.

Funding

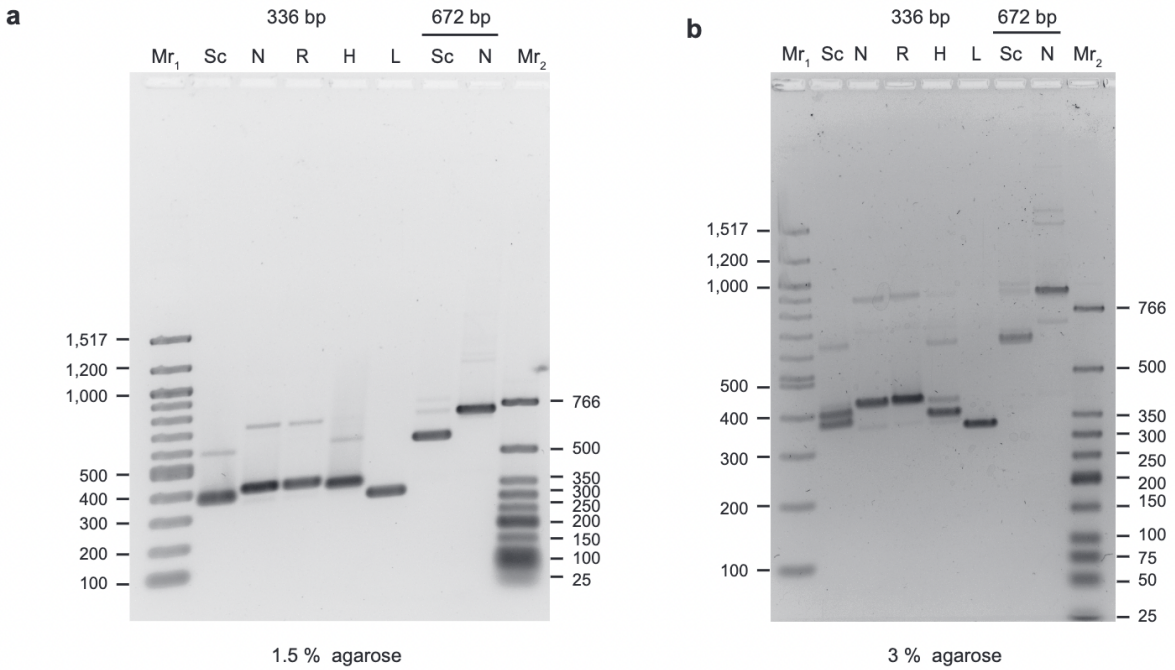
The work of RW and ML was supported by the National Science Center of Poland grant Sonata to ML no. 2018/31/D/ST3/02408. This work was also supported by the Canada 150 Research Chairs program (C150-2017-00015, BD), the National Institutes of Health (1R01GM120600, BD) and the Canadian Natural Science and Engineering Research Council (DG-RGPIN-2019-05637, BD). The Canadian Center for Hydrodynamics is funded by the Canada Foundation for Innovation (CFI- 37589, BD). UltraScan supercomputer calculations were supported through NSF/XSEDE grant TG-MCB070039N (BD), and University of Texas grant TG457201 (BD). MLEJ was supported in part by National Science Center of Poland under grant UMO-2018/31/B/ST8/03640. The authors acknowledge funding (to LZ) from National Institutes ofHealth grant R35 GM141793 andNational Science Foundation grant MCB 2107527.

Conflict of interest statement. Daniel J. Catanese, Jr., Jonathan M. Fogg, and Lynn Zechiedrich are co-inventors on issued and pending patents covering the supercoiled minicircle technology and uses and furthermore hold equity stake in Twister Biotech, Inc.

Supplemental Information

Supplementary Table A.1. Topological composition of the DNA samples used. The composition of the samples was determined by quantification of fluorescently stained gels using image analysis software. N.D.: not detected.

Minicircle length, bp	Sample	Topoisomers present (% of total)	Trace species (% of total)
336	"Supercoiled"	$\Delta Lk = -3$ (48%) $\Delta Lk = -2$ (41%)	$\Delta Lk = -1$ (7 %) Nicked 336 bp (1 %) Supercoiled 672 bp (3 %)
	Nicked	Nicked (93 %)	Nicked 672 bp (7 %)
	Relaxed	Relaxed (95 %)	Relaxed 672 bp (5 %)
	"Hypernegatively supercoiled"	$\Delta Lk = -6$ (61 %) $\Delta Lk = -5$ (33 %)	Nicked 336 bp (4 %) Supercoiled 672 bp (3 %)
	Linear	Linear (100 %)	N.D.
672	"Supercoiled"	$\Delta Lk = -4$ (63 %) $\Delta Lk = -5$ (24 %)	$\Delta Lk = -6$ (6 %) $\Delta Lk = -2$ (4 %) $\Delta Lk = -3$ (2 %) Nicked (2 %)
	Nicked	Nicked (100 %)	N.D.



Supplementary Figure A.1. Electrophoretic mobility of minicircle DNA on agarose gels. DNA samples were analyzed by agarose gel electrophoresis in either a) 1.5 % agarose or b) 3 % agarose in the presence of 1 mM EDTA. Mr_1: 100 bp DNA ladder, lanes 2–6: 336 bp minicircle DNA samples (Sc: "supercoiled", N: nicked, R: relaxed, H: "hypernegatively supercoiled", L: linear), lanes 7–8: 672 bp DNA samples (Sc: "supercoiled", N: nicked), Mr_1: low molecular weight DNA ladder.

References

1 Franklin,R.E. and Gosling,R.G. (1953) Molecular configuration in sodium thymonucleate. Nature, 171, 740–741.

2 Lieberman-Aiden,E., van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,P.J., Dorschner,M.O. et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science, 326, 289–293.

3 Dorman,C.J. (2013) Genome architecture and global gene regulation in bacteria: making progress towards a unified model?. Nat. Rev. Microbiol., 11, 349–355.

4 Fisher,J.K., Bourniquel,A., Witz,G., Weiner,B., Prentiss,M. And Kleckner,N. (2013) Four-dimensional imaging of E. coli nucleoid organization and dynamics in living cells. Cell, 153, 882–895.

5 Kleckner,N., Fisher,J.K., Stouf,M., White,M.A., Bates,D. And Witz,G. (2014) The bacterial nucleoid: nature, dynamics and sister segregation. Curr. Opin. Microbiol., 22, 127–137

6 Dorman,C.J. (2019) DNA supercoiling and transcription in bacteria: a two-way street. BMC Mol. Cell Biol., 20, 26.

7 Baranello,L., Levens,D., Gupta,A. and Kouzine,F. (2012) The importance of being supercoiled: How DNA mechanics regulate dynamic processes. Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms, 1819, 632–638.

8 Magnan,D. and Bates,D. (2015) Regulation of DNA replication initiation by chromosome structure. J. Bacteriol., 197, 3370–3377.
9 Drew,H.R., Weeks,J.R. and Travers,A.A. (1985) Negative supercoiling induces spontaneous unwinding of a bacterial promoter. EMBO J., 4, 1025–1032.
10 Kim,S.H., Ganji,M., Kim,E., van der Torre,J., Abbondanzieri,E. And Dekker,C. (2018) DNA sequence encodes the position of DNA supercoils. eLife, 7, e36557
11 Levens,D. (2008) How the c-myc promoter works and why it sometimes does not. J. Natl. Cancer Ins. Monograp., 39, 41–43
12 Fogg,J.M., Judge,A.K., Stricker,E., Chan,H.L. and Zechiedrich,L. (2021) Supercoiling and looping promote DNA base accessibility and coordination among distant sites. Nat. Commun., 12, 5683
13 Irobalieva,R.N., Fogg,J.M., Catanese,D.J., Sutthibutpong,T., Chen,M., Barker,A.K., Ludtke,S.J., Harris,S.A., Schmid,M.F., Chiu,W. et al. (2015) Structural diversity of supercoiled DNA. Nat. Commun., 6, 8440.
14 Fogg,J.M., Randall,G.L., Pettitt,B.M., Sumners,D. W.L., Harris,S.A. and Zechiedrich,L. (2012) Bullied no more: when and how DNA shoves proteins around. Quart. Rev. Biophys., 45, 257–299.
15 Bond,L.M., Peters,J.P., Becker,N.A., Kahn,J.D. and Maher,L.J. (2010) Gene repression by minimal lac loops in vivo. Nucleic Acids Res., 38, 8072–8082.
16 Czapla,L., Grosner,M.A., Swigon,D. and Olson,W.K. (2013) Interplay of protein and DNA structure revealed in simulations of the lac operon. PLoS ONE, 8, e56548.
17 Garcia,H.G., Grayson,P., Han,L., Inamdar,M., Kondev,J., Nelson,P.C., Phillips,R., Widom,J. and Wiggins,P.A. (2007) Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity. Biopolymers, 85, 115–130.
18 Luger,K., M¨ader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 ˚A resolution. Nature, 389, 251–260.
19 Pyne,A.L.B., Noy,A., Main,K.H.S., Velasco-Berrelleza,V., Piperakis,M.M., Mitchenall,L.A., Cugliandolo,F.M., Beton,J.G., Stevenson,C. E.M., Hoogenboom,B.W. et al. (2021) Base-pair resolution analysis of the effect of supercoiling on DNA flexibility and major groove recognition by triplex-forming oligonucleotides. Nat. Commun., 12, 1053.
20 Basu,A., Bobrovnikov,D.G., Qureshi,Z., Kayikcioglu,T., Ngo,T.T.M., Ranjan,A., Eustermann,S., Cieza,B., Morgan,M.T., Hejna,M. et al. (2021) Measuring DNA mechanics on the genome scale. Nature, 589, 462–467.
21 Li,K., Carroll,M., Vafabakhsh,R., Wang,X.A. and Wang,J.-P. (2022) DNAcycP: a deep learning tool for DNA cyclizability prediction. Nucleic Acids Res., 50, 3142–3154.
22 Demeler,B. and Gorbet,G.E. (2016) Analytical ultracentrifugation data analysis with UltraScan-III. In: Uchiyama,S., Arisaka,F., Stafford,W.F. and Laue,T. (eds.) Anal. Ultracentr. SpringerJapan, Tokyo, pp. 119–143.
23 Fogg,J.M., Kolmakova,N., Rees,I., Magonov,S., Hansma,H., Perona,J.J. and Zechiedrich,E.L. (2006) Exploring writhe in supercoiled minicircle DNA. J. Phys. Cond. Matt., 18, S145–S159.
24 Fuller, F. B. (1978) Decomposition of the linking number of a closed ribbon: A problem from molecular biology. Proc. Natl Acad. Sci. USA, 75(8), 3557–3561.
25 Savelyev, A., Gorbet, G. E., Henrickson, A., and Demeler, B. (6, 2020) Moving analytical ultracentrifugation software to a good manufacturing practices (GMP) environment. PLOS

Computational Biology, 16(6), e1007942.
26 Cao, W. and Demeler, B. (9, 2005) Modeling Analytical Ultracentrifugation Experiments with an Adaptive Space-Time Finite Element Solution of the Lamm Equation. Biophysical Journal, 89(3), 1589–1602.
27 Cao, W. and Demeler, B. (7, 2008) Modeling Analytical Ultracentrifugation Experiments with an Adaptive Space-Time Finite Element Solution for Multicomponent Reacting Systems. Biophysical Journal, 95(1), 54–65.
28 Brookes, E., Cao, W., and Demeler, B. (2, 2010) A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. European Biophysics Journal, 39(3), 405–414.
29 Demeler, B. (4, 2010) Methods for the Design and Analysis of Sedimentation Velocity and Sedimentation Equilibrium Experiments with Proteins. Current Protocols in Protein Science, 60(1).
30 Brookes, E. H. and Demeler, B. (2007) Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments. In GECCO '07 London, Englandthe 9th annual conference: ACM Press [Online; accessed 2022-01-18] p. 361.
31 Demeler, B. and Brookes, E. (2, 2008) Monte Carlo analysis of sedimentation experiments. Colloid and Polymer Science, 286(2), 129–137.
32 Williams, T. L., Gorbet, G. E., and Demeler, B. (10, 2018) Multi-speed sedimentation velocity simulations with UltraScan-III. European Biophysics Journal, 47(7), 815–823.
33 Gorbet, G. E., Mohapatra, S., and Demeler, B. (10, 2018) Multi-speed sedimentation velocity implementation in UltraScan-III. European Biophysics Journal, 47(7), 825–835.
34 Eisenberg, H. (11, 2003) Modern analytical ultracentrifugation in protein science: Look forward, not back. Protein Science, 12(11), 2647–2649.
35 an Holde, K. E. (1985) Solutions of macromolecules. Physical Biochemistry, pp. 24–50.
36 Durchschlag, H. (1986) Specific Volumes of Biological Macromolecules and Some Other Molecules of Biological Interest. In Hinz, H.-J., (ed.), Thermodynamic data for Biochemistry and Biotechnology, Springer Verlag Berlin, Heidelberg.
37 Durchschlag, H. (December, 1989) Determination of the partial specific volume of conjugated proteins. Colloid Polym. Sci., 267(12), 1139–1150.
38 Korolev, N., Lyubartsev, A. P., Rupprecht, A., and Nordenski ¨old, L. (November, 1999) Competitive Binding of Mg^{2+}, Ca^{2+}, Na^{+}, and K^{+} Ions to DNA in Oriented DNA Fibers: Experimental and Monte Carlo Simulation Results. Biophysical Journal, 77(5), 2736–2749.
39 Landau, L. D. and Lifshitz, E. M. (2009) Theory of elasticity, Course of theoretical physicsElsevier, Butterworth-Heinemann, Amsterdam Heidelberg 3. engl. ed., rev. and enlarged edition.
40 C˘alug˘areanu, G. (1959) L'int´egrale de Gauss et l'Analyse des noeuds tridimensionnels. Revue Roumaine de Math´ematiques Pures et Appliqu´ees, 4, 5–20.
41 Bates, A. D. and Maxwell, A. (2005) DNA topology, Oxford University Press, 2nd edition.
42 Coleman, B. D. and Swigon, D. (2000) Theory of Supercoiled Elastic Rings with Self-Contact and Its Application to DNA Plasmids. Journal of Elasticity, 60(3), 173–221.
43 Zajac, E. E. (1962) Stability of Two Planar Loop Elasticas. Journal of Applied Mechanics, 29(1), 136–142.
44 Levitt, M. (1983) Protein folding by restrained energy minimization and molecular dynamics. Journal of Molecular Biology, 170(3), 723–764.

45 Johnson, R. E. and Wu, T. Y. (1979) Hydromechanics of low- Reynolds-number flow. Part 5. Motion of a slender torus. Journal of Fluid Mechanics, 95(2), 263–277.
46 Goren, S. L. and O'Neill, M. E. (1980) Asymmetric creeping motion of an open torus. Journal of Fluid Mechanics, 101(1), 97–110.
47 Juba, D., Audus, D. J., Mascagni, M., Douglas, J. F., and Keyrouz, W. (2017) ZENO: Software for calculating hydrodynamic, electrical, and shape properties of polymer and particle suspensions. Journal of Research of National Institute of Standards and Technology, 122(20).
48 Hubbard, J. B. and Douglas, J. F. (1993) Hydrodynamic friction of arbitrarily shaped Brownian particles. Physical Review E, 47(5), R2983–R2986.
49 Rybenkov,V.V., Vologodskii,A.V. and Cozzarelli,N.R. (1997) The effect of ionic conditions on DNA helical repeat, effective diameter and free energy of supercoiling. Nucleic Acids Res., 25, 1412–1418.
50 Xu,Y.C. and Bremer,H. (1997) Winding of the DNA helix by divalent metal ions. Nucleic Acids Res., 25, 4067–4071.
51 Pak, A. J. and Voth, G. A. (2018) Advances in coarse-grained modeling of macromolecular complexes. Current Opinion in Structural Biology, 52, 119–126.
52 de la Torre, J. G. and Bloomfield, V. A. (1981) Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. Quarterly Reviews of Biophysics, 14(1), 81– 139.
53 H¨uckel, E. and Debye, P. (1923) The theory of electrolytes: I. lowering of freezing point and related phenomena. Physikalische Zeitschrift, 24, 185–206.
54 Lim, S., Kim, Y., and Swigon, D. (2011) Dynamics of an electrostatically charged elastic rod in fluid. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 467(2126), 569–590.
55 Robertson, R. M., Laib, S., and Smith, D. E. (2006) Diffusion of isolated DNA molecules: Dependence on length and topology. Proceedings of the National Academy of Sciences, 103(19), 7310–7314.
56 Mantelli,S., Muller,P., Harlepp,S. and Maaloum,M. (2011) Conformational analysis and estimation of the persistence length of DNA using atomic force microscopy in solution. Soft Matt., 7, 3412.
57 Odijk,T. (1977) Polyelectrolytes near the rod limit. J. Poly. Sci.: Poly. Phys. Edn., 15, 477–483.
58 Skolnick,J. and Fixman,M. (1977) Electrostatic persistence length of a wormlike polyelectrolyte. Macromolecules, 10, 944–948.
59 Manning,G.S. (2001) Counterion condensation on a helical charge lattice. Macromolecules, 34, 4650–4655
60 Manning,G.S. (2006) The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force. Biophys. J., 91, 3607–3616.
61 Lionberger,T.A., Demurtas,D., Witz,G., Dorier,J., Lillian,T., Meyh¨ ofer,E. and Stasiak,A. (2011) Cooperative kinking at distant sites in mechanically stressed DNA. Nucleic Acids Res., 39, 9820–9832.
62 Wang,Q., Irobalieva,R.N., Chiu,W., Schmid,M.F., Fogg,J.M., Zechiedrich,L. and Pettitt,B.M. (2017) Influence of DNA sequence on the structure of minicircles under torsional stress. Nucleic Acids Res., 45, 7633–7642.

63 Johnson,R.E. (1980) An improved slender-body theory for Stokes flow. J. Fluid Mech., 99, 411–431.
64 Adamczyk,Z., Sadlej,K., Wajnryb,E., Ekiel-Je˙zewska,M.L. and Warszy ´nski,P. (2010) Hydrodynamic radii and diffusion coefficients of particle aggregates derived from the bead model. J. Colloid Interface Sci., 347, 192–201.
65 Fernandes,M.X. (2002) Calculation of hydrodynamic properties of small nucleic acids from their atomic structure. Nucleic Acids Res., 30, 1782–1788.
66 Penkova,N.A., Sharapov,M.G. and Penkov,N.V. (2021) Hydration shells of DNA from the point of view of terahertz time-domain spectroscopy. Int. J. Mol. Sci., 22, 11089.
67 Laage,D., Elsaesser,T. and Hynes,J.T. (2017) Water dynamics in the hydration shells of biomolecules. Chem. Rev., 117, 10694–10725.
68 Stellwagen,N.C. (2009) Electrophoresis of DNA in agarose gels, polyacrylamide gels and in free solution. Electrophoresis, 30, S188–S195.
69 Ziraldo,R., Shoura,M.J., Fire,A.Z. and Levene,S.D. (2019) Deconvolution of nucleic-acid length distributions: a gel electrophoresis analysis tool and applications. Nucleic Acids Res., 47, e92–e92.
70 Iubini,S., Orlandini,E., Michieletto,D. and Baiesi,M. (2018) Topological sieving of rings according to their rigidity. ACS Macro Lett., 7, 1408–1412.
71 Waszkiewicz,R., Szymczak,P. and Lisicki,M. (2021) Stability of sedimenting flexible loops. J. Fluid Mech., 919, A14.
72 Gruziel,M., Thyagarajan,K., Dietler,G., Stasiak,A., Ekiel-Je˙zewska,M.L. and Szymczak,P. (2018) Periodic motion of sedimenting flexible knots. Phys. Rev. Lett., 121, 127801.
73 Gruziel-Słomka,M., Kondratiuk,P., Szymczak,P. and Ekiel-Je˙zewska,M.L. (2019) Stokesian dynamics of sedimenting elastic rings. Soft Matter, 15, 7262–7274.
74 Gruziel-Słomka,M., Kondratiuk,P., Szymczak,P. and Ekiel-Je˙zewska,M.L. (2022) Correction: Stokesian dynamics of sedimenting elastic rings. Soft Matter, 18, 4811–4811.
75 Cichocki,B., Rubin,M., Niedzwiecka,A. and Szymczak,P. (2019) Diffusion coefficients of elastic macromolecules. J. Fluid Mech., 878, R3.
76 Waszkiewicz,R. (2023) RadostW/twisted-loop: Manuscript version. Zenodo, https://doi.org/10.5281/zenodo.7501673.
77 Waszkiewicz,R. (2023) RadostW/twisted-dna-results: Manuscript version. Zenodo, https://doi.org/10.5281/zenodo.7501675.

Appendix B: Photocatalytic hydrogen evolution by a de novo designed metalloprotein that undergoes Ni-mediated oligomerization shift

Dr. Pallavi Prasad[1], Dr. Leigh AnnaHunt[1], Ashley E. Pall[2], Maduni Ranasinghe[3], Dr. Ashley E. Williams[1], Prof. Timothy L. Stemmler[2], Prof. Borries Demeler[3], Prof. Nathan I. Hammer[1], Prof. Saumen Chakraborty[1*]

[1]Dr. P. Prasad, Dr. L. A. Hunt, Dr. A. E. Williams, Prof. N. I. Hammer and Prof. S. Chakraborty Department of Chemistry and Biochemistry The University of Mississippi University, MS, 38677 (USA) E-mail: saumenc@oleimss.edu Twitter: @ChakrabortyS16
[2]A. E. Pall and Prof. T. L. Stemmler Department of Pharmaceutical Sciences Wayne State University 259 Mack Avenue Detroit, MI, 48201-2417 (USA)
[3]M. Ranasinghe and Prof. B. Demeler Department of Chemistry and Biochemistry University of Lethbridge 4401, University Dr W Lethbridge, AB T1K 6T5 (CA)

Abstract

De novo metalloprotein design involves the construction of proteins guided by specific repeat patterns of polar and apolar residues, which, upon self-assembly, provides a suitable environment to bind metals and produce artificial metalloenzymes. While a wide range of functionalities have been realized in de novo designed metalloproteins, the functional repertoire of such constructs towards alternative energy-relevant catalysis is currently limited. Here we show the application of de novo approach to design a functional H_2 evolving protein. The design involved the assembly of an amphiphilic peptide featuring cysteines at tandem *a/d* sites of each helix. Intriguingly, upon Ni^{II} addition, the oligomers shift from a major trimeric assembly to a mix of dimers and trimers. The metalloprotein produced H_2 photocatalytically with a bell-shape pH dependence, having a maximum activity at pH 5.5. Transient absorption spectroscopy is used to determine the timescales of electron transfer as a function of pH. Selective outer sphere mutations are made to probe how the local environment tunes activity. A preferential enhancement of activity is observed via steric modulation above the Ni^{II} site, towards the N-termini, compared to below the Ni^{II} site towards the C-termini.

Keywords

de novo proteins; self-assembly; photocatalysis; solar H_2; transient absorption kinetics

Introduction

The relationship between primary sequence and 3D structure and function of metalloproteins is highly complex. De novo protein design approach enables the creation of artificial metalloproteins as structural and functional analogs of native proteins, providing a deep understanding of how the sequence determines the structure and function [1]. Over the last decade, advances have been made in three important areas of de novo protein design [2]. These include rational design by inspection, computational protein design, and database-driven construction of functional proteins. In this context, de novo designed self-assembled peptide

amphiphiles have produced artificial metalloproteins with defined oligomeric states and tunable coordination environment for various metals [1b, 3]. De novo designed metalloproteins help bridge the gap between molecular catalysts and native enzymes, as they harbor native-like environments for the metal site, providing aqueous conditions to perform the chemistry [3a]. Specific folding pattern of de novo proteins help stabilize unique metal coordination environments and avoids undesired polymeric metal-ligand complexes. Outer sphere interactions provide a handle to further tune the properties of the metal site.

A relatively unexplored area in de novo protein design is to design metalloenzymes for energy-relevant processes. The use of fossil fuels as the primary energy source is not sustainable. In addition, burning of fossil fuels releases CO_2, which causes environmental pollution [4]. Developing sustainable and alternative energy resources while mitigating the carbon footprint in the atmosphere is a major challenge. H_2 gas is an attractive clean fuel source that can be combusted to produce energy along with anon-toxic by product, H_2O [5]. However, H_2 is available only at low levels from natural resources, which necessitates that H_2 must be produced in a sustainable manner from a natural resource such as the sunlight and employing earth-abundant transition metal catalysts. Photosynthetic water splitting is one approach where solar energy is stored chemically in the form of H_2 [6]. In the reductive quenching mechanism, photo-induced charge separation occurs when a photosensitizer (PS) absorbs sunlight. A sacrificial electron donor (SED) quenches the PS^{n+}*, which then transfers electrons to the catalytically active metal center for subsequent reduction of H^+ to H_2. Alternatively, in the oxidative quenching pathway the PS^{n+*}directly transfers an electron to the catalyst, which then enables H^+ reduction to H_2. The oxidized $PS^{(n+1)}$ is then reduced by a SED for regeneration.

Hydrogenase enzymes found in archaea, bacteria and eukarya, hold an excellent promise in alternative energy because they are natural producers of H_2 [7]. Depending on the location of the enzyme and the energy demand of the organism, hydrogenases can either remove reducing equivalents to make H_2 from H^+ or oxidize H_2 to generate energy for cellular processes. [NiFe] hydrogenases harbor a bimetallic active site where the Ni and Fe share two bridged Cys ligands. Ni is coordinated to two additional Cys plus a fifth bridging ligand, which is H^- in the active Ni-R and Ni-C forms, and OH^- in the inactive Ni-A and Ni-B forms.

Water-soluble biomolecular catalysts for hydrogen evolution reaction (HER) have been reported over the last few years, which involve incorporation of synthetic catalysts or transition metal ions into biomolecular scaffolds [8]. Here, we report the de novo design of a peptide oligomer (4SCC), which, intriguingly, stabilizes a lower oligomeric statein the presence of Ni^{II}. In the metal-bound form, 4SCC demonstrates photocatalytic H_2 production activity. The role of pH on H_2 production activity and the kinetics of electron transfer (ET) during photocatalysis has been determined from pH titration and transient absorption spectroscopic (TAS) studies. Outer sphere variants are employed to assess how steric modulation alters the activity.

Results and discussion

Peptide design: In conceiving the 4SCC design, we hypothesized that if two Ni^{II} sites are introduced within a de novo scaffold, synergistic interactions between the metals would improve the H^+ reduction activity. To incorporate the intended dual $Ni^{II}(Cys)_4$ binding site, we chose a

sequence that is conducive to produce tetrameric coiled coils, with the expectation that Cys residues in the -Cys-Xxx-Xxx-Cys-motif will produce the dual $Ni^{II}(Cys)_4$ sites. The tetramer assembly is governed by the knob-into-holes (KIH) packing of the ***a*** and ***d*** site Leu and Ile residues, respectively [9]. Oppositely charged Lys and Glu at ***e*** and ***g*** sites induce the formation of inter-helical ion pairs, which further stabilizes the assembly and dictate the parallel orientation. Ala at ***b*** and ***c*** sites induce α-helicity while polar residues at external ***f*** positions increase water solubility of the constructs. As the ***a*** and ***d*** residues form the hydrophobic core and point to the interior, we chose the 9th and 12th layers to introduce Cys residues via the $-Cys_{\alpha}^{9}$-Xxx-Xxx-Cys_{d}^{12-} chelate motif. The sequences of the peptides used in this study are shown in Table B.1.

Table B.1. Peptide sequences used here.

Peptide	Sequence[a]
4SCC	Ac-GLAAIEQKCEACEKKLEAIEKKLAAIEWKG-CONH2
4SCC-1	Ac-GLAAIEEKCAACKQELAAIEKKLAAIEWKG-CONH2
C12I-4SCC	Ac-GLAAIEQKCEAIEKKLEAIEKKLAAIEWKG-CONH2
C9L-4SCC	Ac-GLAAIEQKLEACEKKLEAIEKKLAAIEWKG-CONH2
I5A-4SCC	Ac-GLAAAEQKCEACEKKLEAIEKKLAAIEWKG-CONH2
L16A-4SCC	Ac-GLAAIEQKCEACEKKAEAIEKKLAAIEWKG-CONH2
4SCC-Cys	Ac-GLAAIEQKLEAIEKKLEAIEKKLAAIEWKG-CONH2

Oligomerization behavior of apo and Ni^{II}-bound 4SCC: To study the oligomerization behavior of 4SCC, we performed sedimentation velocity (SV) experiments in apo and metal bound forms as a function of concentration. SV experiments produce sedimentation coefficient profiles, which provide a distinction between homogeneous preparations and reversible self-association in response to mass action [10]. A concentration series of the apo form demonstrates the reversible oligomerization behavior of 4SCC, as the sedimentation coefficient distribution shifts to higher s values with a sedimentation profile consistent with reversible mass action, as the loading concentration is increased from 3.7 μM to 185 μM (Figure B.1A, black, blue, red). The addition of Ni^{II} to higher concentration samples shifts the equilibrium from anoligomeric species with a molar mass consistent with a trimer to a species consistent with a dimer, while Ni^{II} addition at lower concentration has no obvious effect. In the apo form, approximately 83% of the total concentration of 4SCC sediments as a trimer (red line in Figure B.1B). After addition of Ni^{II} some of the trimer dissociates into dimer, and only about 42% of the trimer species exists in the metal-bound form (blue line in Figure B.1B), the remaining peptide is dimeric. This result suggests that in the presence of Ni^{II}, the dimeric assembly is favored over the trimer. This is intriguing, since metal binding is expected to stabilize the overall assembly. The following scenario is consistent with the addition of Ni^{II}: when Ni^{II} is added to the apo peptide, the Ni^{II} ions stabilize the dimeric conformation preferentially, causing the trimer to dissociate into dimer and monomers. The monomers are then captured by free Ni^{II} ions, forming more dimer, until all Ni^{II} is consumed (Scheme 1). The coordination environment of Ni^{II} imposed by the intact

assembly may not be optimal. Hence, Ni^{II} dictates its geometric preference, which requires the assembly to dissociate.

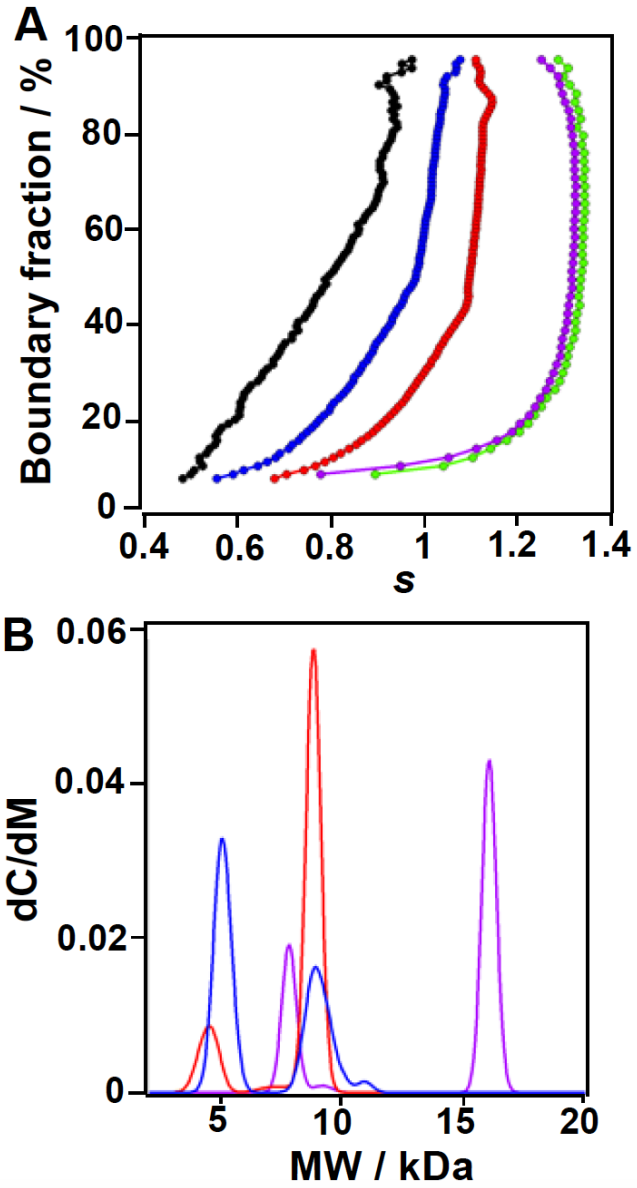
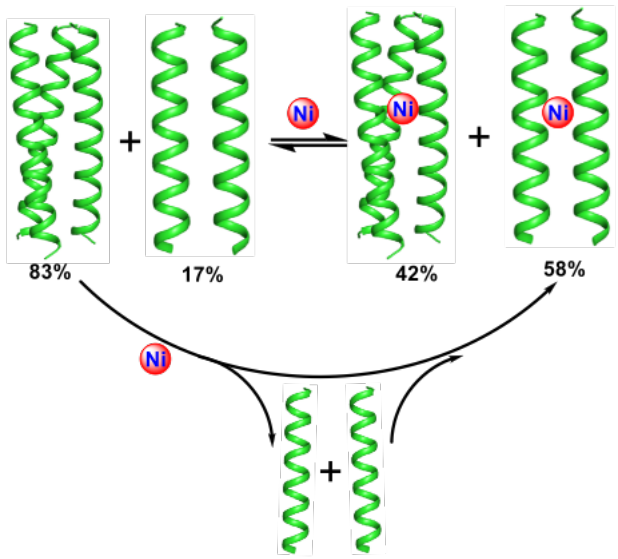


Figure B.1. A) van Holde-Weischet diffusion-corrected integral sedimentation coefficient distributions for apo 4SCC at increasing concentration (black: 3.7 μM, blue: 51 μM, red: 185 μM), demonstrating reversible self-association as sedimentation distributions shift to the right with higher concentration. The corresponding plots for 4SCC-Cys (purple, green) do not show mass action. B) Molar mass distributions of apo-4SCC (red) and Ni^{II}-4SCC (blue) at high concentration. Addition of Ni^{II} causes a shift in the equilibrium distribution from trimeric to dimeric oligomers. The major species (~80%) in 4SCC-Cys is tetramer (purple).

We believe this process leads to more dimer production upon Ni^{II} addition to the apo mixture. This is consistent with our previous observation that the $Ni^{II}(Cys)_4$ coordination is also stabilized when Ni^{II} is added to the 2SCC peptide that is designed to exist as a dimer under similar experimental conditions [3c]. To test how the oligomerization behavior is affected by the introduction of Cys residues, we prepared the Cys-free peptide (4SCC-Cys: Table B.1) where the 9th and 12th layer contained Leu and Ile, respectively, instead of Cys.



Scheme 1. Scheme showing Ni^{II}-induced shift in oligomerization behaviour of 4SCC from a primarily trimer species in the apo form to a major dimer species in the presence of Ni^{II}. The bottom part shows that the trimer dissociates into dimer and monomer. The monomer combines with more monomer to form dimer, adding to the dimer pool.

The 4SCC-Cys peptide does not undergo concentration-dependent change in sedimentation and molar mass distribution (Figure B.1A green, purple). Approximately 80% of 4SCC-Cys exists as a tetramer, and the remaining 20%as dimer. While the desired tetramer was not detected in either apo or Ni^{II} bound forms of 4SCC, the presence of ~80%tetramer in the unmodified peptide 4SCC-Cys suggests that replacing Leu and Ile residues with Cys lowers the driving force necessary to producethe tetrameric self-assembly. This can be attributed to the fact that favorable Van der Waals packing of the hydrophobes is compromised in preparing the 4SCC construct which involves replacing Leu and Ile with Cys.

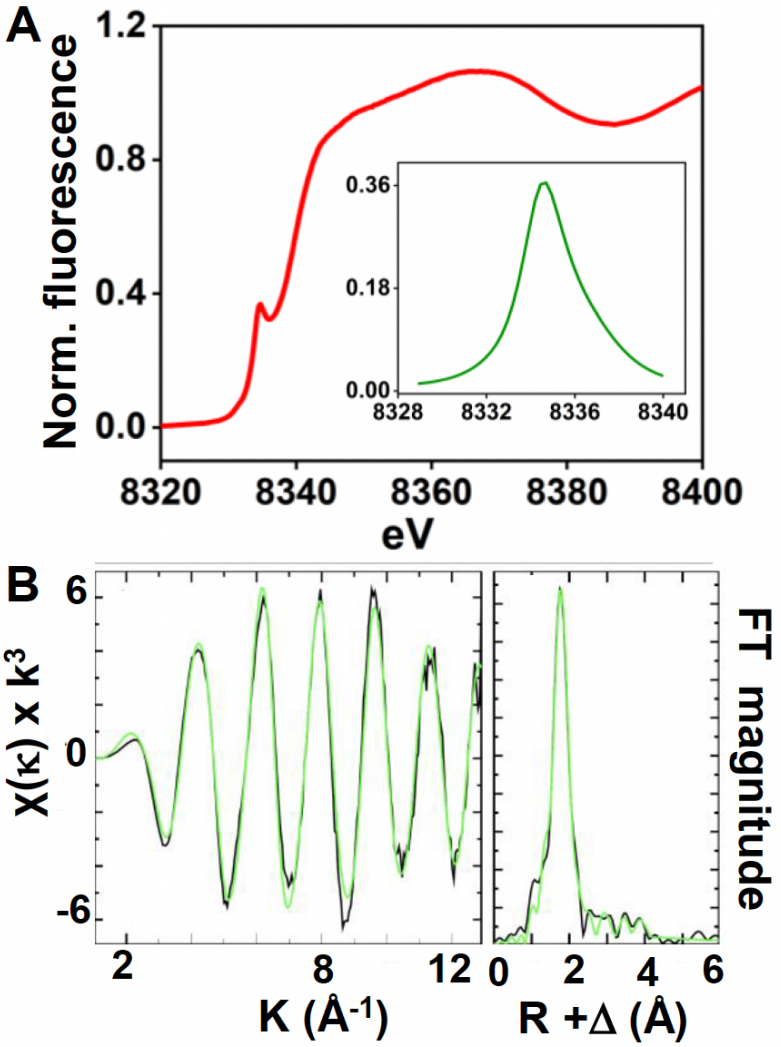


Figure B.2. A) Normalized k-edge Ni XANES spectra of Ni^{II}-4SCC with baseline subtracted pre-edge data in the inset. B) Raw unfiltered EXAFS data and the Fourier transforms of the raw EXAFS (black) and simulated traces (green) for Ni environments. Samples contained 1 mM 4SCC plus 500 μM Ni^{II} at pH 8.5.

Ni^{II} binds 4SCC via a distorted tetrahedral geometry: To characterize the oxidation state and the metal-ligand geometry, Ni K-edge X-ray absorption spectroscopy (XAS) is used. The Xray absorption near edge (XANES) pre-edge feature analysis provides a direct insight into the metal-ligand coordination while the first inflection energy for the excitation edge provides information about the redox state of the metal. The pre-edge region of XANES (Figure B.2A) data shows a low 1s→3d transition at ca. 8331 eV, indicative of a square planer Ni^{II} environment, while an additional feature at 8334.9 eV is ascribed to the 1s→4p transition.[11] The corresponding 1s→4p peak area is 1.37×10^{-2} eV^2, while that of the 1s→3d transition is not statistically measurable. The edge energies in 4SCC are shifted by ca. -3 eV relative to the distorted Ni^{II}-

$S_2(O/N)_2$ tetrahedral sites in NiSOD models [12]. The shift to lower energy reflects a complete sulfur ligation of Ni in 4SCC [11].

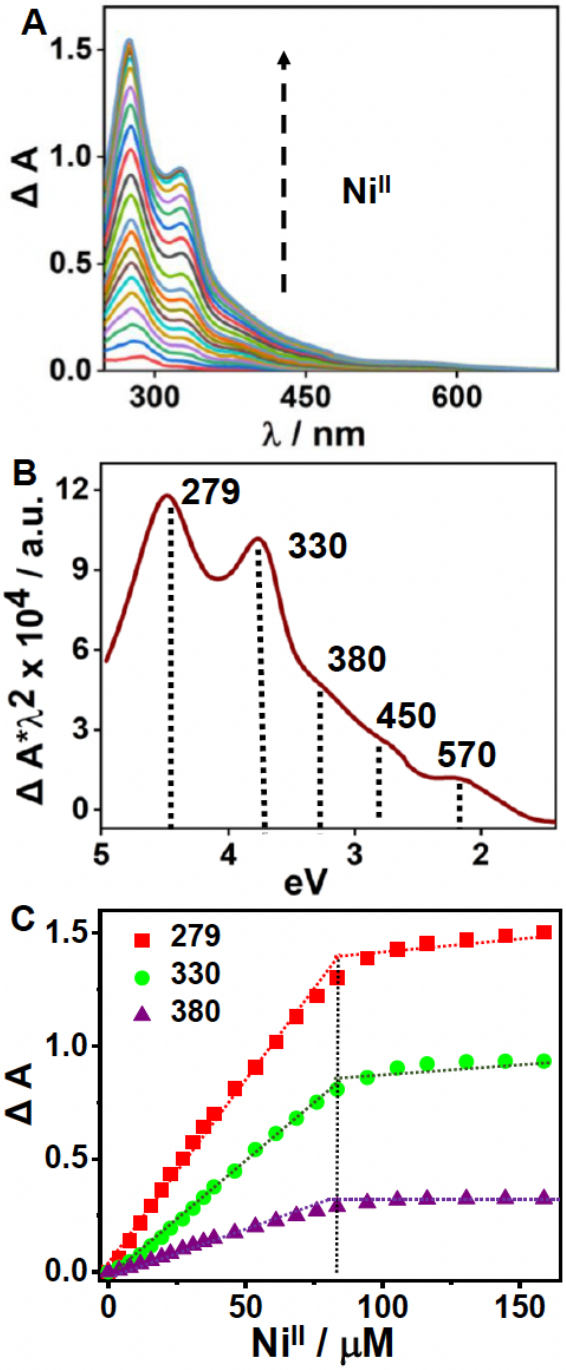


Figure B.3. A) Spectral changes upon NiII addition to 160 μM 4SCC. B) An energy plot highlighting the resolved absorption maxima for each peak. C) Differential absorption vs [Ni^{II}] plot at different wavelengths showing saturation at ~80 μM NiII, i.e. 2:1 monomer: Ni^{II}.

Table B.2. UV-vis and XAS parameters of 4SCC.

Peptide	λ_{max} (nm)	$\Delta\varepsilon$ ($M^{-1}cm^{-1}$)	Ni-S[a] (Å)	CN[b]	σ^{2}[c] Å x 10^{3}
Ni-4SCC	279	15620	2.22	3.5 ± 0.5	4.71
	330	9690	2.44	0.5 ± 0.5	2.41
	380	3460			
	450	970			
	570	395			

[a]Average bond distance and [b]coordination number from EXAFS fitting of 2 scans. [c]Debye Waller factor.

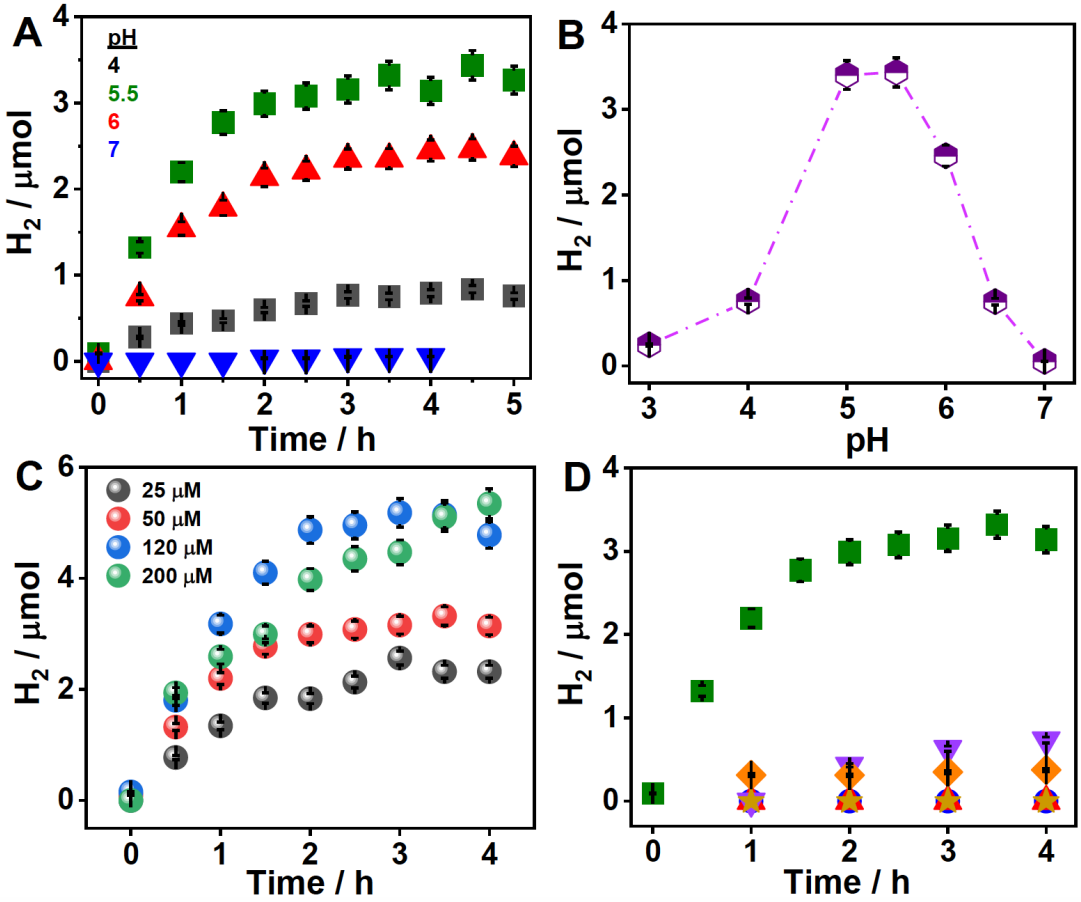


Figure B.4. A) pH-dependent H_2 production time courses of 50 µM 4SCC plus 25 µM Ni^{II}. B) Bell-shaped [H_2] vs pH plot. C) Concentration-dependent (monomer) H_2 production of 4SCC at pH 5.5 with 2:1 peptide:Ni^{II} ratio. D) H_2 production by the single-site variants C12I-4SCC (purple) and C9L-4SCC (orange), and 4SCC-Cys (yellow), all at 50 µM peptide concentrations in the presence of 25 µM Ni^{II} at pH 5.5. Apo peptide (red) and 25 µM $NiSO_4$ (blue) are also shown at pH 5.5. The olive trace in D for Ni-4SCC is shown for comparison against the controls in the same plot. Light source is a white LED (λ>400 nm).

The sinusoidal amplitude pattern in the extended X-ray absorption fine structure (EXAFS) indicates a strong homogeneity in Ni^{II} ligand scattering, suggesting a single nearest neighbor ligand environment. The EXAFS simulations indicate a Ni^{II}-nearest neighbor ligand environment constructed by a four-coordinate ligand system with 3.5 ± 0.5 sulfur ligands at 2.22 Å and 0.5 ± 0.5 sulfur ligands at 2.44 Å (Figure B.2B, Table B.2). There is no evidence for oxygen or nitrogen scattering in the Ni^{II}-nearest neighbor environment. Since the peptide exists as a mixture, the bond lengths and coordination numbers we obtained are likely from both the 3- and 4-coordinate species, since the geometrical parameters obtained in the fits are an average. Long

range scattering (R > 2.6 Å) is also observed in the Fourier transform of the Ni EXAFS (Supplementary Table B.1). These signals can be fit with Ni^{II}•••carbon scattering at independent but resolvable distances of 3.4, 3.95 and 4.37 Å, likely due to Cys carbon side chain atoms. A distinct Ni•••Ni interaction is not detected; this data supports the SV experiments. If the intact tetramer was formed, then Ni•••Ni scattering would be expected to be observed in the EXAFS from the two Ni^{II} bound to the peptide.

Ni^{II} binding to 4SCC: The peptide to metal stoichiometry was determined by Ni^{II} titrations (Fig. 3) to 160 μM apo peptide. With each NiII addition, bands appearing at 279 nm, 330 nm, 380 nm, 450 nm, and 570 nm (Fig. 3A-B) gradually increased in intensity, before reaching a saturation at ~80 μM Ni^{II} (Fig. 3C). The overall spectral features are similar to Ni^{II}-2SCC,[3c] so are the molar absorptivity (Δε: 15620, 9690, 3460, 970, and 395 $M^{-1}cm^{-1}$) values (Table 2). The 4SCC to Ni^{II} stoichiometry is further established from Ni^{II}-dependent CD spectra (Supplementary Figure B.1, Supplementary Table B.2), which show that Ni^{II} addition increases the helical character of the peptide up to a peptide:Ni^{II} ratio of 2:1, similar to what is observed in UV-vis titrations as well. More Ni^{II} addition does not affect the helicity. An overall stabilizing effect of Ni^{II} is also observed in the thermal melt profiles (Supplementary Figure B.2) where the Ni^{II} bound 4SCC has ~8°C higher tm of 45°C, compared to the apo peptide. To ensure that the observed UV-vis features are arising due to Ni^{II}-thiolate interactions and not from adventitiously bound Ni^{II}, the control peptide without Cys (4SCC-Cys) was used (Table B.1). A lack of spectral signatures upon Ni^{II} addition (Supplementary Figure B.3 A-B) confirms that the UV-vis features observed in the 4SCC construct are due to Ni^{II}-thiolate interactions. Furthermore, control experiments with the corresponding single Cys site peptides that contain Cys in either the a site (C12I-4SCC) or d site (C9L-4SCC) demonstrate (Supplementary Table Figure B.3 C-D) that the -Cys-Xxx-Xxx-Cys- chelate motif is a critical requirement to produce the Ni^{II}-thiolate chromophore observed in Figure B.3.

4SCC shows pH-dependent H_2 evolution: Light-induced H_2 evolution was studied using $[Ru(bpy)_3]^{2+}$ (Ru) as the PS and ascorbic acid as a SED. We investigated the pH-dependence for H_2 production, since solution acidity is a critical determinant of light-induced H_2 production in tri-component systems.[3c, 8k, 13] 4SCC demonstrates a bell-shaped activity profile with a maximum H_2 production of ~3.6 μmol [14] from 50 μM monomer plus 25 μM Ni^{II} at an optimal pH of 5-5.5 (Figure B.4A-B). In comparison, our previously reported 2SCC system produced 2.6 μmol H_2 from 60 μM monomer peptide plus 30 μM Ni^{II}.[3c] Thus, the current 4SCC system has a higher activity than the 2SCC construct. 4SCC shows a linear concentration dependence for H2 production up to ~120 μM (Figure B.4C). No significant increase in H_2 production was observed with further increase in peptide concentration. The final H_2 yield from 200 μM is comparable to that obtained with 120 μM 4SCC. Control experiments (Figure B.4D) with apo 4SCC (red), the Cys-free peptide 4SCC-Cys (yellow), and free Ni^{II} (blue) produce no detectable H_2. The single Cys-site variants C12I-4SCC (Figure B.4D purple) and C9L-4SCC (Figure B.4D orange) produce 1.3 μmol and 0.5 μmol H_2, respectively.

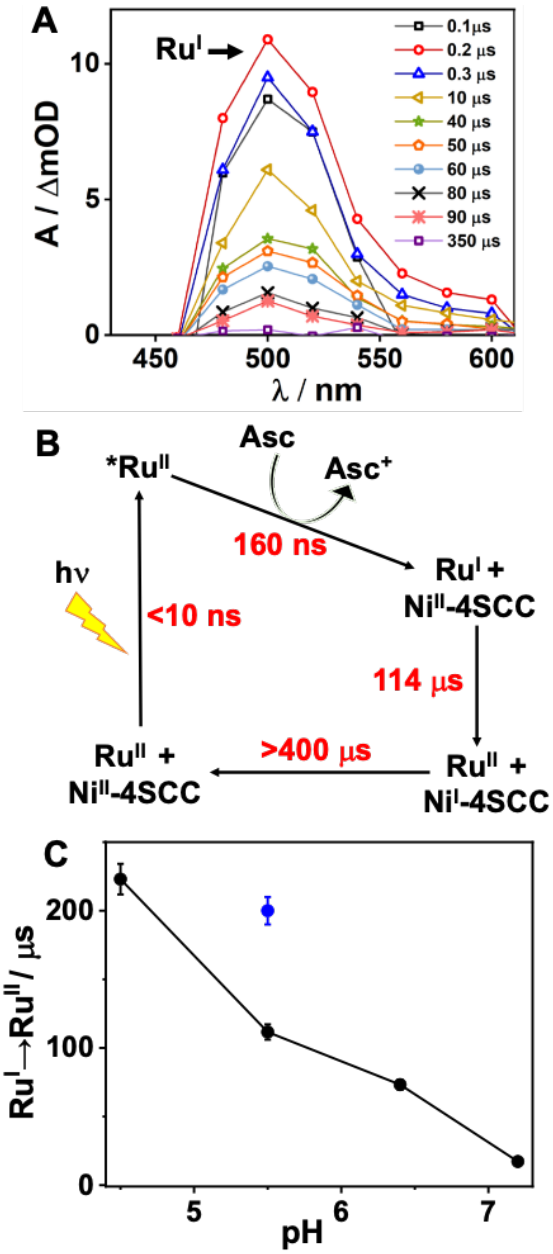


Figure B.5. A) TA spectra of Ni-4SCC at pH 5.5 showing the formation and decay kinetics of RuI signal at ~510 nm. B) Schematic representation of the time constants for different steps of the photochemical process derived from the TA spectra. C) Decay timescales of Ru^{I} to Ru^{II} obtained from pH-dependent TA kinetics. Blue data point represents the blank PS/SED sample in the absence of the peptide.

As the Cys in the **a** site orient towards the center of the peptide assembly and the **d** site side chains point towards the interfaces [15], the Cys in the **a** site (C12I-4SCC) produce a favorable Ni^{II} binding environment compared to the peptide where the Cys is in the **d** site (C9L-4SCC). This differential metal binding ability of the two sites is reflected in H_2 production of the corresponding single-site

peptides where C12I- 4SCC likely produces a well-defined Ni^{II}-thiolate site than C9L- 4SCC. However, the fact that the single site variants produce a significantly less H_2 than 4SCC, bolster the notion that the presence of Cys at both **a** and **d** sites is needed to form a suitable Ni^{II} coordination motif via chelation. Photostability studies show that no degradation of the photosystem occurs in the studied time frame of 5h (Supplementary Figure B.4). Bren and co-workers have reported that stable Co-porphyrin catalysts based on cyt c_{552} and Comimochrome VI*a can sustain prolonged H_2 production during electrolysis,[8j, 8n] emphasizing that a robust scaffold is beneficial for activity. The photocatalytic H_2 production by [NiFe] hydrogenases vary depending on the type of modifications introduced to facilitate ET to the active site. The Armstrong group has employed silver nanoclusters as PS with Hyd-1 and Hyd-2, producing H_2 with TOFs of 40 s^{-1} and 193 s^{-1}, respectively [16]. A Ru-linked [NiFe] hydrogenase showed a TOF of 0.16 s^{-1} [17]. The Ni-rubredoxin system has a TOF of 0.00833 s^{-1} with RuPS.[8i] Incorporation of the DuBois Ni catalyst into photosystem I produced H_2 with a TOF of 0.73 s^{-1}, while its encapsulation in flavodoxin increased the TOF to 1.25 s^{-1}.[8e]

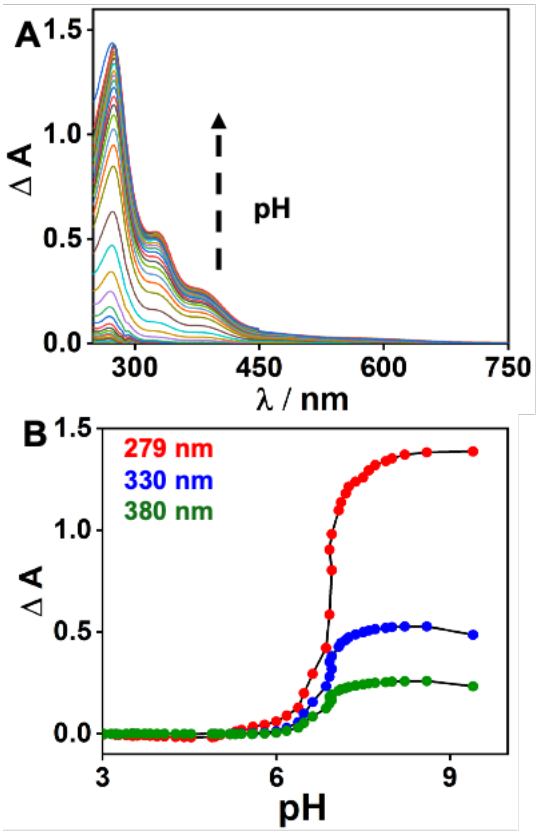


Figure B.6. A) pH-dependent spectral changes during titration of 160 µM 4SCC plus 80 µM Ni^{II} solutions with KOH, showing the appearance of absorption bands with an increase in pH. B) ΔA vs pH plot for different wavelengths extracted from the corresponding traces.

Timescales of ET: It is known that the H_2 production activity in low acidic pH is controlled by the acidity of the SED [3c, 8k, 13, 18]. The protonated form of ascorbic acid (pKa ~ 4.1) is a sluggish electron donor to photoexcited $*Ru^{II}$ compared to its conjugate base.[3c, 13, 19] Thus, Ru^{I} production by

reductive quenching via ascorbate anion and subsequent reduction of Ni^{II} to Ni^{I} occurs only above the pKa of ascorbic acid. We have employed TAS to determine the timescale of ET by following the formation and decay kinetics of Ru^{I} at 510 nm at pH 5.5 (Figure B.5A), where the maximal activity is observed. The initial photoexcitation of Ru^{II} to $*Ru^{II}$ is fast, <10 ns (Figure B.5B). Utschig and coworkers have observed and assigned a broad absorption feature at ~550-750 nm to this excited $*Ru^{II}$ species [8k]. The photoexcited $*Ru^{II}$ is quenched to Ru^{I} by the SED within ~160 ns (τ_1) via reductive quenching. The reduced Ru^{I} then transfers an electron to the Ni^{II}-peptide leading to the formation of Ni^{I}-peptide. This process is accompanied by a decay of the Ru^{I} signal at ~114 μs (τ_2). In the absence of Ni^{II}-peptide, the decay of Ru^{I} occurs at ~200 μs (Figure B.5C). The faster decay kinetics in the presence of peptide compared to its absence suggests that ET from Ru^{I} to Ni^{II}-peptide is expedited in the presence of the metallopeptide because it serves as the electron acceptor. In the absence of an electron acceptor, Ru^{I} decays at a slower rate. A weak Ni bleach signal has been observed in PSI and synthetic Ru systems [11]. However, no Ni bleach signal is detected for our system, which precluded the direct determination of Ni^{I} formation timescales. The photosystem relaxes at longer timescales (>400 μs; Figure B.5B). Although no H_2 evolution occurs at higher pH, ET from Ru^{I} still occurs (Figure B.5C), as evidenced by a fast decay kinetics of Ru^{I} signal in the TAS. This implies that the formation of Ni^{I} is not limited at higher pH that would account for the low H_2 production. To determine how solution acidity and the protonation state of active site Cys affects the yield of H_2, a pH titration experiment was performed.

Role of Cys pKa: The pH dependence of Ni^{II} binding to 4SCC is probed by monitoring the changes in the absorption bands as a function of pH. In these experiments, pH is adjusted by the addition of small aliquots of KOH to unbuffered solutions containing 160 μM peptide plus 80 μM Ni^{II}. In the pH range of 3- 5.5 no bands are observed (Figure B.6A). A further increase in pH leads to the appearance of characteristic absorption features due to the formation of thiolates and their interaction with Ni^{II}. The spectral features fully develop with an approximate pKa of ~6.5 (Figure B.6B), which is ascribed to the deprotonation of active site Cys [3c]. In 4SCC, the titration curve is complex, due to the presence of multiple Cys residues in the peptide and that the peptide exists as a mixture of trimer and dimer in solution. Nonetheless, the pH profile is consistent with the deprotonation of metal-bound Cys thiols [20]. We propose that during catalysis H^+ binds to Ni^{I} center, producing a Ni^{III}-H− intermediate, analogous to the Ni-C species of [NiFe] H2ases [7, 21], and the 2SCC system [3c]. This species is then reduced to Ni^{II}-H− (Ni-R analog of [NiFe] H_2ases). Next, the protonated Cys thiol(s) in 4SCC facilitates H_2 production after combining with the Ni^{II}-H− intermediate. The lack of protonated Cys at higher pH limits H_2 production.[3c, 21]

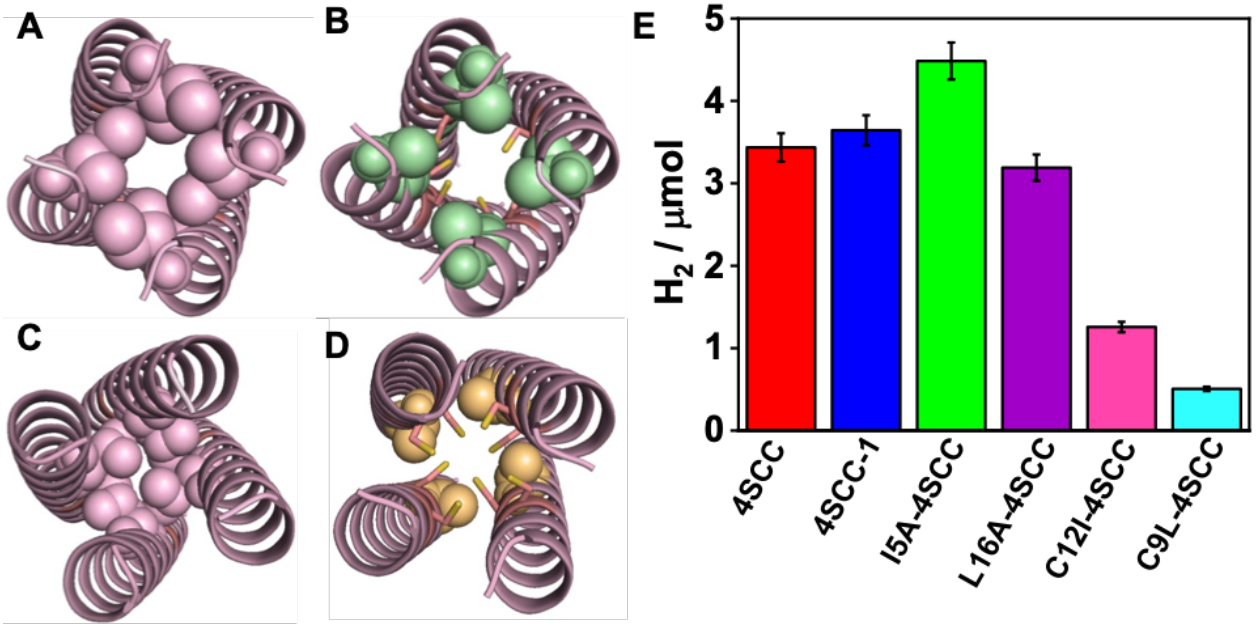


Figure B.7. Space filling model of the 5th layer Ile (A) in 4SCC and the Ala layer in I5A-4SCC (B) as viewed from the N-termini. The corresponding views of the 16th layer in 4SCC (C) and L16A-4SCC (D) from the C-termini. While trimer is predominant in the apo form experimentally, the tetramer model, as generated from the sequence, is shown to represent the space availability at the top and bottom Cys layers. The space filling models showing outer sphere steric changes are valid irrespective of the oligomeric state of the peptide. The bar graph for H_2 production of 4SCC and its variants is shown in E.

Activity tuning by altering local environment: To probe how the local environment influences H_2 production, several mutants are designed. The first mutant (4SCC-1) is designed to probe the contribution of Glu residues near the Cys in the -Cys9-Glu-Ala- Cys12-Glu- motif. Glu residues near the coordinating Cys sites have been proposed to bind CdII in a dimeric coiled coil system. [22] To eliminate the possibility of Glu coordination, a new motif, - Cys^9-Ala-Ala-Cys12-Lys- was designed (Table B.1) that contained Ala instead of Glu at the 10th position and a Lys at the 13th position. Ala was chosen as a non-metal coordinating residue, which also stabilizes the α-helix, while Lys13 in the new motif was introduced to maintain the charge neutrality of the peptide. A similar 4SCC:Ni^{II} stoichiometry of 2:1 and a similar H_2 production amount of ~3.6 µmol is observed in 4SCC-1 (Figure B.7E, blue), which suggests that the Glu residues are not involved in Ni^{II} binding or these residues do not have any favorable contribution to H_2 production activity. Next, we probed whether decreasing outer sphere steric around Ni^{II} would alter the reactivity. Two mutants are constructed by replacing the hydrophobic residues at the 5th Ile and 16th layer Leu to Ala, one at a time. Space filling models in the tetramer conformation show that the introduction of Ala generates space above (Figure B.7 A-B) and below (Figure B.7 C-D) the Ni^{II} site. While the models are generated with tetramers, space availability in Ala mutants is representative of all assemblies, irrespective of the exact oligomeric state of 4SCC in solution, since the Ile and Leu residues directly line the top and bottom layers of the coordinating Cys ligands. In addition to determining how an increased space above and below the active site alters

the reactivity, these variants also help address if there is any preference of H^+ binding from above or below the Ni^{II} site. The L16A-4SCC mutant does not show any increase in activity compared to the parent 4SCC, rather it has a slightly lower H_2 production of 3.2 μmol (Figure B.7E, purple). However, the top layer Ala variant, I5A-4SCC, produces ~4.5 μmol H_2 (Figure B.7E, green) compared to ~3.6 μmol produced in 4SCC. This increase in H_2 production also correlates with an increase in the initial rate by ~1.5-fold. Ghirlanda and coworkers have shown that axial M7A mutation of Co-substituted cyt b562 increased the TON for H_2 production by ~2.5-fold, compared to the WT protein, which was attributed to the direct access of H^+ to Co [23]. However, in the case of I5A-4SCC the increase is observed via an outer sphere mutation instead of direct mutation of a coordinating residue, which highlights that outer sphere steric can control the reactivity for even a small substrate such as H^+. It is interesting to note that the L16A-4SCC mutant does not enhance the activity even though more space is created below the Ni^{II} site. This observation likely suggests that H^+ preferentially binds to the top side of Ni^{II} towards the N-termini. Combined, these results highlight how steric changes in the outer sphere metal coordination environment tunes the reactivity of de novo designed metalloproteins in the context of photocatalytic H_2 production.

Conclusions

Application of de novo designed proteins into energy-relevant catalysis is currently limited. Here we aimed to expand the function of self-assembling peptide amphiphiles into photocatalytic HER by mimicking the reductive half of artificial photosynthesis. A concentration-dependent mass action is observed where the trimeric species is produced as a major component at higher concentrations. Intriguingly, Ni^{II} addition caused the trimer to dissociate into dimers and monomers, a process which is governed by the coordination and geometrical preference of Ni^{II}. We believe the quaternary structure of the intact assembly does not produce optimal coordination parameters for Ni^{II}. The maximal photocatalytic H_2 production activity of the metalloprotein is observed ~pH 5.5, which is governed by a tradeoff between the acidity of the SED and active site Cys pKa that shifts the balance between the availability of reduced PS and protonated Cys thiol, which combines with the metal hydride to produce H_2. TAS studies led to the determination of ET timescales for the photocatalytic process. Tweaking the hydrophobic packing near the active site leads to selective enhancement of activity when space is created towards the N-termini above the metal site, as opposed to below the Ni center. This observation suggests that the peptide pocket shows preferential activation of even the smallest and achiral substrate such as H^+. We demonstrate a significant expansion of de novo design approach to advance the construction artificial metalloproteins catalyzing an energyrelevant process. A current limitation is that the introduction of metal binding residues severely perturbs the Van der Waals packing of hydrophobic residues that is necessary to produce specific oligomeric species. Increasing the length of the peptide by 1-2 additional heptads is envisioned to counter this effect of lost oligomer specificity.

Experimental Section

All chemicals and reagents are of analytical grade and used as received, without further purification. All the glassware and plasticware used for the experiments were soaked in a 10 mM

ethylenediaminetetraacetic acid (EDTA) bath overnight, followed by overnight soaking in 10% and 1% nitric acid baths, respectively. These are then thoroughly washed with deionized water. All buffers are chelexed (Sigma-Aldrich) overnight followed by pH adjustment and filtration.

Peptide design: Energy minimized models for the peptide sequence of the tetrameric coiled coil peptide were obtained by using CC Builder 2.0 [24]. The model with the lowest BUDE energy has a pitch of 174.8 Å, radius of 6.9 Å and an interface angle of 19.65°.

Peptide synthesis and purification: The peptides were synthesized and purified as previously described [3c]. The purified peptide was lyophilized and kept at -80°C until further use. Concentrations of purified peptides were calculated using an extinction coefficient ε_{280} = 5500 M^{-1} cm^{-1} for one Trp per peptide monomer.

Mass spectrometry: The purity and identify of the peptides were determined using MALDI-MS (Bruker Voyager). Sinapinic acid (10 mg/ml in 50% acetonitrile, 50% H_2O, 0.1% TFA) was used as the matrix. 1:1 (matrix: peptide) mixture was spotted on 100-SS stainless steel plate and analyzed.

Preparation of the metal bound peptide samples: The solid peptides were dissolved in degassed buffer solutions and then incubated with 50-fold excess Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) for 15 minutes. The reduced peptides were then passed through a PD10 column (GE Healthcare). The eluted peptides were degassed under nitrogen using Schlenk techniques for 10 minutes, followed by sequential addition of degassed $NiSO_4$ to a peptide:Ni ratio of 2:1. Ni^{II} binding was confirmed by UV-vis.

Ni^{II} titration: Titrations were performed anaerobically in a septacapped quartz cuvette (Starna Cells) by adding $NiSO_4$ to 160 μM (monomer) of reduced peptides in 100 mM Tris pH 8.5. A gastight syringe (Hamilton) was used for metal addition. After each addition, the sample was stirred and incubated for 10 min prior to recording the spectra in a Cary 5000 spectrophotometer (Agilent). The absorbance of the apo peptide is subtracted from the metal bound forms to calculate the differential absorption, which was then plotted against [Ni^{II}] to determine the stoichiometric ratio of peptide to metal. Dilution corrections were done to eliminate volume error.

pH titration: pH titrations were done by adding small aliquots of a KOH solution to 160 μM peptide in the presence of 80 μM Ni^{II} in unbuffered solutions. After each addition of KOH, the solution was equilibrated for 10 minutes, after which the pH was recorded and the spectra were collected at each pH.

CD spectroscopy: CD spectra were collected using a JASCO spectrometer. 30 μM of apo and Ni^{II}-bound peptide solutions were prepared under inert conditions in 5 mM Tris pH 8.5. CD scans were collected in a 1 cm path length quartz cuvette. Raw data are converted to mean residue ellipticity (MRE).

Analytical ultra centrifugation: All samples were measured in a buffer containing 20 mM MES and 50 mM NaCl, pH 6.0. Sample concentrations of apo and NiII-bound peptides were varied to

establish the existence or non-existence of concentration dependent mass action. Different wavelengths were chosen to keep the sample absorbance within the dynamic range of the detector. All samples were measured at 60 krpm by UV intensity detection in a Beckman Optima AUC analytical ultracentrifuge using an An60Ti rotor and standard 2-channel epon centerpieces (Beckman-Coulter) at the Canadian Center for Hydrodynamics at the University of Lethbridge in Alberta, Canada (CCH). All data were analyzed with UltraScan-III, version 4.0 (6348), using the UltraScan data acquisition module. The data were first imported and converted into UltraScan-III OpenAUC format and edited. Parallel distributed data analysis was performed on the UltraScan Science Gateway using XSEDE resources (Expanse, Bridges-2), and high-performance computing clusters at the University of Lethbridge. Subsequently, two-dimensional spectrum analysis (2DSA) [25] was performed with a sedimentation coefficient range of 0.1-3 s and a frictional ratio range of 1-4, to remove time and radially invariant noise, and to fit the meniscus and bottom position. A final 2DSA iterative refinement was obtained for each dataset. The models resulting from the final 2DSA fit were used to initialize a genetic algorithm analysis (GA) [26], which was followed by a Monte Carlo GA analysis [27] to obtain 95% confidence intervals for the sedimentation and diffusion coefficients, and the partial concentration of each solute in a dataset, and to quantify the effect of experimental stochastic noise on the obtained parameter values. Noise-corrected data were analyzed by the enhanced van Holde–Weischet analysis [28] (vHW) to generate diffusion corrected sedimentation coefficient distributions.

X-ray absorption spectroscopy (XAS): XAS samples were prepared anaerobically inside an MBraun glove box. Samples were prepared in 100 mM Tris pH 8.5 containing 1 mM 4SCC (monomer) plus 500 μM Ni^{II} prepared by adding 125 μM Ni^{II} at a time and stirring for 15 min in-between. The metallated peptide was then mixed with glycerol as the glassing agent to a final concentration of 20%. The sample was then transferred to XAS cells sealed with Kapton tape and flash frozen in liquid nitrogen. XAS data were collected at the Stanford Synchrotron Radiation Light Source (SSRL) on beamline 7-3. Beamline 7-3 utilizes a Si[220] double crystal monochromator with an in-line mirror for Xray focusing and for harmonic rejection. During data collection, samples were maintained at 12 K using a liquid He continuous flow cryostat. Fluorescence XAS spectra were collected using a 30 element Ge-Detector (Canberra) in a path tangential to using ion chamber detectors filled with nitrogen gas. Spectra were measured in 5 eV increments in the pre-edge region (8,265-8,325 eV), 0.25 eV increments in the edge region (8,325-8,405 eV) and 0.05 $Å^{-1}$ increments in the extended X-ray absorption fine structure (EXAFS) region (out to k = 13.3 $Å^{-1}$), integrating from 1 to 25 s in a k^3-weighted manner for a total scan length of approximately 50 min. X-ray energies were individually calibrated by collecting a Ni-foil absorption spectrum simultaneously with the compound; the first inflection point of the Ni-foil spectrum was assigned to 8,333 eV. Data presented here represent the average from 8 scans. XAS spectra were processed using the Macintosh OS X version of the EXAFSPAK program suite integrated with the Feff v8 software for theoretical model generation [29]. Data reduction utilized a Gaussian spline for background removal in the pre-edge region and a three-region cubic spline throughout the EXAFS. Data were converted to k-space using a Ni E0 value of 8,333 eV. The k^3 weighted EXAFS was truncated between 1.0 and 13.0 $Å^{-1}$ for filtering purposes. This k-range corresponds to a spectral resolution of ca. 0.13 Å for all nickel-ligand interactions; therefore, only independent scattering environments outside 0.13 Å were considered resolvable in the EXAFS fitting analysis [30]. EXAFS fitting analysis was performed first on

filtered data and then verified on the raw unfiltered data. EXAFS data were fit using both single and multiple scattering amplitude and phase functions calculated with the program Feff v8. Single scattering theoretical models were calculated for carbon, nitrogen, oxygen, and sulfur coordination to simulate nickel nearest-neighbor ligand environments. Scale factor (Sc) and E0 values, used in a static manner during the simulations, were calibrated by fitting crystallographically characterized Ni models; specific values include a Scale Factor of 0.9 and E0 values of -10.55 eV for O/N/C, -12.22 eV for S and -14 for Ni. Criteria for judging the bestfit simulation utilized both the lowest mean square deviation between data and fit (F'), corrected for the number of degrees of freedom, and a reasonable Debye-Waller factor [31]. Pre-edge analysis was done using EDG_FIT software. A spline function was best-fit between 8,328 eV and 8,342 eV and a 3-peak model was applied to accommodate the pre-edge features.

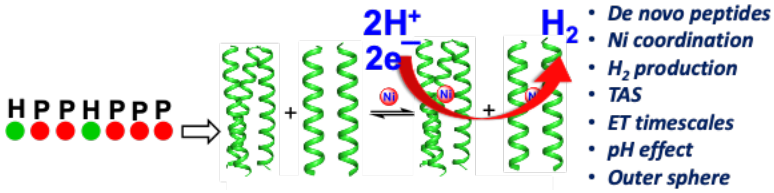
Photogenerated hydrogen evolution: Photocatalytic H_2 production experiments were performed in a 3-component system using tris(2,2'-bipyridine)dichlororuthenium(II)hexahydrate, ($[Ru(bpy)_3Cl_2].6H_2O$, Alfa Aesar) as the photosensitizer, 100 mM ascorbic acid (VWR) as the sacrificial electron donor and peptides in 10 mM Tris buffer. 2 mL samples were prepared anaerobically in a setpa-capped pyrex tube and irradiated with a 180 mW white light source ($\lambda >$ 400nm; Thor Laboratories) for 4-6 hours with constant stirring. In order to study pH dependence, separate samples at each pH were prepared and pH was adjusted in the final samples using small volumes of NaOH. At the end of each experiment, the pH was again recorded which showed a +/- 0.5 unit variability before and after the experiment. The peptide concertation dependence was performed by using variable concentrations of the peptide while keeping the other reagents constant. Every 30 min, 250 μL of head space gas was sampled using gas-tight syringes (VICI) and injected into a 2014-Shimadzu instrument operating with N_2 as the carrier gas. A thermal conductivity detector was used to detect the H_2 and the area under the peak was converted to ppm using calibration standards.

Transient absorption spectroscopy: TAS data were collected on an Edinburgh LP980 optical system (Edinburgh Instruments, U.K.) using same parameters as has been described before.[3c] Briefly, 300 μL anaerobically prepared 700 μM peptides plus 350 μM Ni^{II} in 10 mM Tris buffer with 0.04 mM of Ru^{II} and 100 mM of ascorbic acid were used. Solutions were stirred and purged with N_2 for 15 min prior to measurements. Data were recorded in a 1 mm path length cuvette positioned at a 45° angle relative to the pump and probe sources to maximize overlap and direct scattered light away from the entrance slit to the monochromator. A 150 W CW xenon arc lamp is used as the probe source while a photomultiplier tube (R928, Edinburgh Instruments, U.K.) is used to detect the absorption. The processing of the outputs from the detector are done using Tekrtronix MDO3022 Mixed Domain Oscillosope (200 MHz, 2.5GS s^{-1}) interfaced to a PC and Edinburgh's L900 (version 8.2.3) software package. Timeresolved absorption is monitored at 510 nm and averaged over 200 laser shots. Kinetic traces are fit with a single exponential function $\Delta OD(t) = \Delta OD_0\, e^{-k1\tau}$, where ΔOD_0 is the initial ΔOD value, k_1 is the first order rate constant and τ is the lifetime.

Acknowledgements
This work was supported by funds from the National Institutes of Health for T.L.S. (R01 DK068139). Portions of this research were carried out at the Stanford Synchrotron Radiation

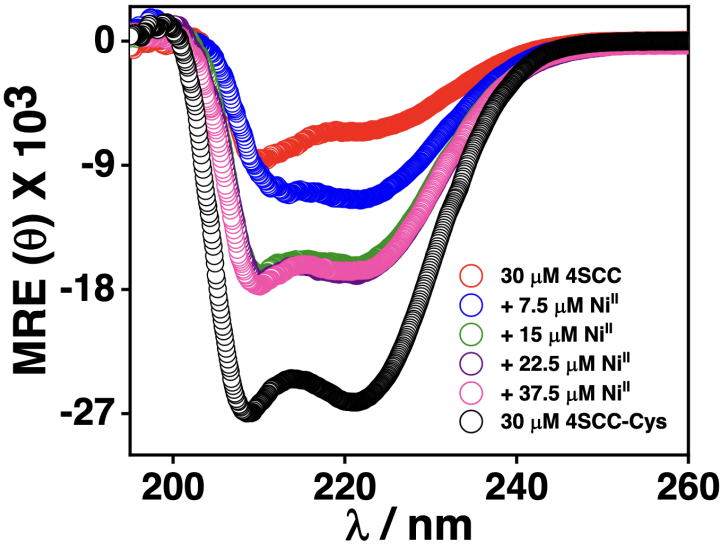
Light source (SSRL). SSRL is a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the NIH, National Center for Research Resources, Biomedical Technology Program. B.D. is supported by the Canada 150 Research Chairs program (C150-2017-00015) and the Canadian Natural Science and Engineering Research Council (DG-RGPIN-2019-05637). UltraScan supercomputer calculations were supported through NSF/XSEDE grant TG-MCB070039N, and University of Texas grant TG457201 (both to B.D.). The development of the UltraScan software is supported by NIH through grant GM120600 (B.D.). N.I.H thanks the National Science Foundation (Grant OIA- 1757220). S.C. thanks the National Institutes of Health (GM131260) for support.

Table of contents graphic

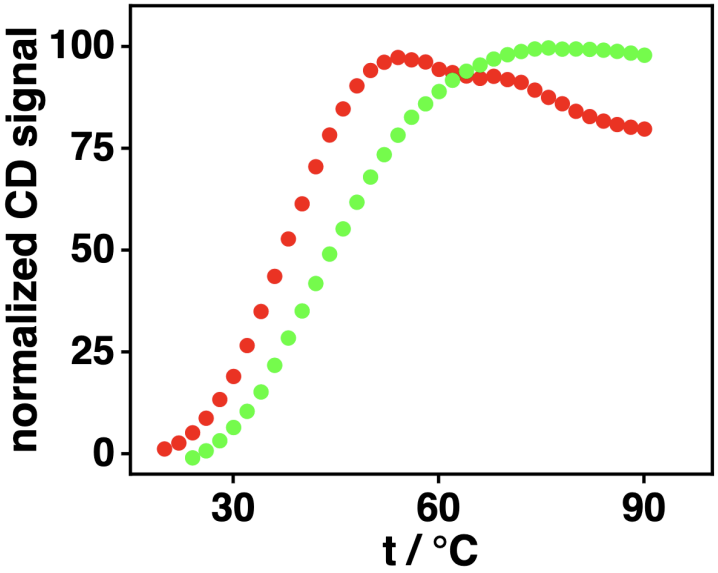


A de novo H_2 evolving metallopeptide is constructed from first principles, which undergoes oligomerization change in the presence of Ni^{II}. Under white light, the metallopeptide produces H_2 via a reductive quenching pathway. The timescales of ET and the role of thiol pKa is determined from TAS kinetics and pH titration experiments, respectively. Outer sphere steric modulation shows preferential binding of H^+ to Ni from N-termini.

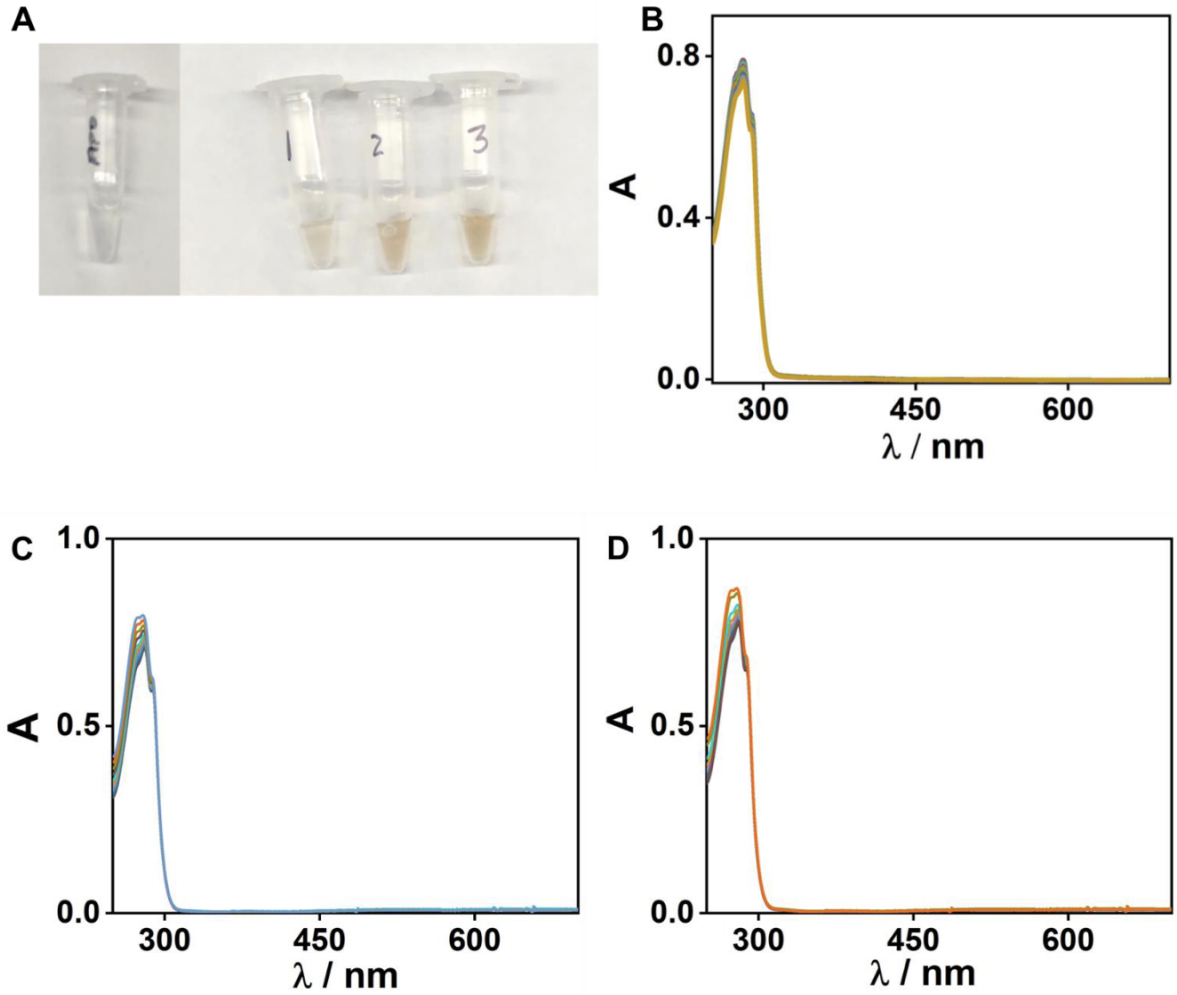
Supplemental Information



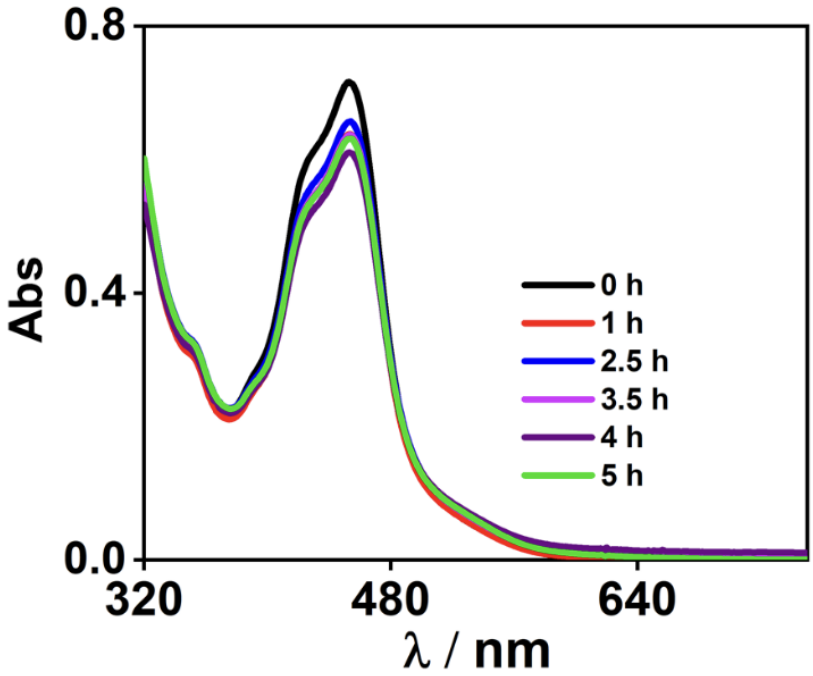
Supplementary Figure B.1. CD data of 30 μM 4SCC in apo (red) and in the presence of different Ni^{II} concentrations (blue, green, purple, and pink). The 4SCC-Cys peptide (black) that lacks Cys is the most helical due to packing of unmodified hydrophobes. Data collected in 5 mM Tris pH 8.5.



Supplementary Figure B.2. Thermal melt profiles of 30 μM apo 4SCC (red) and in the presence of 15 μM Ni^{II} (green) monitored as CD signal at 222 nm in 5 mM Tris pH 8.5.



Supplementary Figure B.3. A) Colour comparison of apo and Ni^{II}-bound 4SCC. B, C and D are Ni^{II} titration of the Cys-free variant 4SCC-Cys, and the single site variants C12I-4SCC and C9L-4SCC, respectively, all at 160 μM peptide concentration. All data collected in 100 mM Tris pH 8.5.



Supplementary Figure B.4. Absorbance of RuPS from photocatalysis experiments monitored at different time intervals over the 5h of experimental timeframe. No significant degradation of the dye is observed during the assay.

Supplementary Table B.1. Summary of best fit simulation for k-edge EXAFS, showing long-range ligand environment. Data were fit over a k range of 1 to 12.94 $Å^{-1}$, using a scale factor of 0.9 and E_0 values of -10.55 eV for Ni-O/N/C, -12.22 eV for Ni-S and -14 eV for Ni•••Ni interaction to calibrate the simulations

Atom[a]	R(Å)[b]	CN[c]	σ^{2d}	F'[e]
S	2.22	3.5	4.71	0.28
S	2.44	0.5	2.41	
C	3.40	1	2.76	
C	3.95	1	1.38	
C	4.37	2	2.37	

[a] Scattering atoms: S (Sulfur) and C (carbon).
[b] Average metal-ligand bond length from two scans.
[c] Average metal-ligand coordination number from two scans.

[d] Average Debye-Waller factor in $Å^2$ $x10^3$ from two scans.
[e] Number of degrees of freedom weighted mean square deviation between data and fit.

Supplementary Table B.2. Summary of CD results.

Sample	% α-helix
30 μM apo 4SCC	20
+ 7.5 μM Ni^{II}	35
+ 15 μM Ni^{II}	50
+ 22.5 μM Ni^{II}	50
+ 37.5 eq. Ni^{II}	52
30 μM 4SCC-Cys	81

References

1 a) I. V. Korendovych, W. F. DeGrado, Q. Rev. Biophys. 2020, 53; b) F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum, V. L. Pecoraro, Chem. Rev. 2014, 114, 3495-3578; c) C. S. Mocny, V. L. Pecoraro, Acc. Chem. Res. 2015, 48, 2388-2396; d) C. M. Rufo, Y. S. Moroz, O. V. Moroz, J. Stöhr, T. A. Smith, X. Hu, W. F. DeGrado, I. V. Korendovych, Nat. Chem. 2014, 6, 303-309; e) A. Lombardi, F. Pirro, O. Maglio, M. Chino, W. F. DeGrado, Acc. Chem. Res. 2019, 52, 1148-1159.

2 a) P.-S. Huang, S. E. Boyken, D. Baker, Nature 2016, 537, 320-327; b) D. N. Woolfson, J. Mol. Biol. 2021, 167160.

3 a) K. J. Koebke, T. B. Pinter, W. C. Pitts, V. L. Pecoraro, Chem. Rev. 2022; b) S. Mitra, D. Prakash, K. Rajabimoghadam, Z. Wawrzak, P. Prasad, T. Wu, S. K. Misra, J. S. Sharp, I. Garcia-Bosch, S. Chakraborty, ACS Catal. 2021, 11, 10267-10278; c) S. Malayam Parambath, A. E. Williams, L. A. Hunt, D. Selvan, N. I. Hammer, S. Chakraborty, ChemSusChem 2021, 14, 2237-2246.

4 G. P. Peters, C. Le Quéré, R. M. Andrew, J. G. Canadell, P. Friedlingstein, T. Ilyina, R. B. Jackson, F. Joos, J. I. Korsbakken, G. A. McKinley, S. Sitch, P. Tans, Nat. Clim. Change 2017, 7, 848-850.

5 a) G. W. Crabtree, M. S. Dresselhaus, MRS Bull. 2008, 33, 421-428; b) A. Sartbaeva, V. Kuznetsov, S. Wells, P. Edwards, Energy Environ. Sci. 2008, 1, 79-85.

6 a) G. Ciamician, Science1912, 36, 385-394; b) N. S. Lewis, Science 2007, 315, 798-801; c) N. S. Lewis, D. G. Nocera, Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 15729.

7 W. Lubitz, H. Ogata, O. Ruediger, E. Reijerse, Chem. Rev. 2014, 114, 4081-4148.

8 a) A. K. Jones, B. R. Lichtenstein, A. Dutta, G. Gordon, P. L. Dutton, J. Am. Chem. Soc. 2007, 129, 14844–14845; b) Y. Sano, A. Onoda, T. Hayashi, Chem. Commun. 2011, 47, 8229-8231; c) L. M. Utschig, S. C. Silver, K. L. Mulfort, D. M. Tiede, J. Am. Chem. Soc. 2011, 133, 16334-16337; d) A. Roy, C. Madden, G. Ghirlanda, Chem. Commun. 2012, 48, 9816-9818; e) S. C. Silver, J. Niklas, P. Du, O. G. Poluektov, D. M. Tiede, L. M. Utschig, J. Am. Chem. Soc. 2013, 135, 13246-13249; f) M. L. Reback, G. W. Buchko, B. L. Kier, B. Ginovska-Pangovska, Y. Xiong, S. Lense, J. Hou, J. A. S. Roberts, C. M. Sorensen, S. Raugei, T. C. Squier, W. J. Shaw, Chem. Eur. J. 2014, 20, 1510-1514; g) D. J. Sommer, M. D. Vaughn, G. Ghirlanda, Chem. Commun. 2014, 50, 15852-15855; h) J. G. Kleingardner, B. Kandemir, K. L. Bren, J. Am. Chem. Soc. 2014, 136, 4-7; i) J. W. Slater, H. S. Shafaat, J. Phys. Chem. Lett. 2015, 6, 3731-3736; j) B. Kandemir, S. Chakraborty, Y. Guo, K. L. Bren, Inorg. Chem. 2016, 55, 467-477; k)S. R. Soltau, J. Niklas, P. D. Dahlberg, K. L. Mulfort, O. G. Poluektov, L. M. Utschig, ACS Energy Lett. 2016, 2, 230-237; l) B. Kandemir, L. Kubie, Y. Guo, B. Sheldon, K. L. Bren, Inorg. Chem. 2016, 55, 1355-1357; m) J. W. Slater, S. C. Marguet, H. A. Monaco, H. S. Shafaat, J. Am. Chem. Soc. 2018, 140, 10250-10262; n) V. Firpo, J. M. Le, V. Pavone, A. Lombardi, K. L. Bren, Chem. Sci. 2018, 9, 8582-8589; o) A. Call, C. Casadevall, A. Romero-Rivera, V. Martin-Diaconescu, D. J. Sommer, S. Osuna, G. Ghirlanda, J. Lloret-Fillol, ACS Catal. 2019, 9, 5837-5846; p) P. Prasad, D. Selvan, S. Chakraborty, Chem. Eur. J. 2020, 26, 12494-12509; q) D. Selvan, P. Prasad, E. R. Farquhar, Y. Shi, S. Crane, Y. Zhang, S. Chakraborty, ACS Catal. 2019, 9, 5847-5859.

9 a) J. Walshaw, D. N. Woolfson, Biochem. Soc. Trans. 2000, 28, A423-A423; b) J. M. Fletcher, A. L. Boyle, M. Bruning, G. J. Bartlett, T. L. Vincent, N. R. Zaccai, C. T. Armstrong, E. H. C. Bromley, P. J. Booth, R. L. Brady, A. R. Thomson, D. N. Woolfson, ACS Synth. Biol. 2012, 1, 240-250.

10 a) B. Demeler, E. Brookes, R. Wang, V. Schirf, C. A. Kim, Macromol. Biosci. 2010, 10, 775-782; b) M. Łuczkowski, M. Stachura, V. Schirf, B. Demeler, L. Hemmingsen, V. L. Pecoraro, Inorg. Chem. 2008, 47, 10875-10888; c) P. Anzini, C. Xu, S. Hughes, E. Magnotti, T. Jiang, L. Hemmingsen, B. Demeler, V. P. Conticello, J. Am. Chem. Soc. 2013, 135, 10278-10281.

11 G. J. Colpas, M. J. Maroney, C. Bagyinka, M. Kumar, W. S. Willis, S. L. Suib, P. K. Mascharak, N. Baidya, Inorg. Chem. 1991, 30, 920-928.

12 R. A. Steiner, S. P. Dzul, T. L. Stemmler, T. C. Harrop, Inorg. Chem. 2017, 56, 2849-2862.

13 A. Reynal, E. Pastor, M. A. Gross, S. Selim, E. Reisner, J. R. Durrant, Chem. Sci. 2015, 6, 4855-4859.

14 Turn over number (TON) has not been reported because it is not possible to distinguish individual contributions of the dimer and trimer to the resultant H2 production.

15 S. Chakraborty, D. S. Touw, A. F. A. Peacock, J. Stuckey, V. L. Pecoraro, J. Am. Chem. Soc. 2010, 132, 13240-13250.

16 a) L. Zhang, S. E. Beaton, S. B. Carr, F. A. Armstrong, Energy Environ. Sci. 2018, 11, 3342-3348; b) L. Zhang, G. Morello, S. B. Carr, F. A. Armstrong, J. Am. Chem. Soc. 2020, 142, 12699-12707.

17 O. A. Zadvornyy, J. E. Lucon, R. Gerlach, N. A. Zorin, T. Douglas, T. E. Elgren, J. W. Peters, J. Inorg. Biochem. 2012, 106, 151-155.

18 P. Gotico, D. Moonshiram, C. Liu, X. Zhang, R. Guillot, A. Quaranta, Z. Halime, W. Leibl, A. Aukauloo, Chem. Eur. J. 2020, 26, 2859-2868.

19 a) Z. Han, W. R. McNamara, M.-S. Eum, P. L. Holland, R. Eisenberg, Angew. Chem. Intl. Ed. 2012, 51, 1667-1670; b) M. Natali, ACS Catal. 2017, 7, 1330-1339.
20 O. Iranzo, S. Chakraborty, L. Hemmingsen, V. L. Pecoraro, J. Am. Chem. Soc. 2011, 133, 239-251.
21 H. Ogata, K. Nishikawa, W. Lubitz, Nature 2015, 520, 571-574.
22 a) S. Chakraborty, O. Iranzo, E. R. P. Zuiderweg, V. L. Pecoraro, J. Am. Chem. Soc. 2012, 134, 6191-6203; b) D. V. Zaytsev, V. A. Morozov, J. Fan, X. Zhu, M. Mukherjee, S. Ni, M. A. Kennedy, M. Y. Ogawa, J. Inorg. Biochem. 2013, 119, 1-9.
23 D. J. Sommer, M. D. Vaughn, B. C. Clark, J. Tomlin, A. Roy, G. Ghirlanda, Biochim. Biophys. Acta 2016, 1857, 598-603.
24 C. W. Wood, D. N. Woolfson, Protein Sci. 2018, 27, 103-111.
25 E. Brookes, W. Cao, B. Demeler, Eur. Biophys. J. 2010, 39, 405-414.
26 E. H. Brookes, B. Demeler, in Proceedings of the 9th annual conference on Genetic and evolutionary computation, Association for Computing Machinery, London, England, 2007, pp. 361–368.
27 B. Demeler, E. Brookes, Colloid Polym. Sci. 2008, 286, 129-137.
28 B. Demeler, K. E. van Holde, Anal. Biochem. 2004, 335, 279-288.
29 J. Rehr, A. Ankudinov, J. Synchrotron Radiat. 2001, 8, 61-65.
30 P. J. Riggs-Gelasco, T. L. Stemmler, J. E. Penner-Hahn, Coord. Chem. Rev. 1995, 144, 245-286.
31 J. J. Cotelesage, M. J. Pushie, P. Grochulski, I. J. Pickering, G. N. George, J. Inorg. Biochem. 2012, 115, 127-137.