UC Irvine
UC Irvine Electronic Theses and Dissertations

Title

Simulation Aspects of the Mechanics of Biomolecular Filaments: Crackling in DNA unzipping and the Contraction of Bacteriophage Tails

Permalink

https://escholarship.org/uc/item/0r2330v7

Author

Chatterjee, Anupam

Publication Date

2019

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, availalbe at https://creativecommons.org/licenses/by-nc-sa/4.0/

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Simulation Aspects of the Mechanics of Biomolecular Filaments: Crackling in DNA unzipping and the Contraction of Bacteriophage Tails

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Anupam Chatterjee

Dissertation Committee:
Ioan Andricioaei, Chair
Douglas Tobias
David Mobley

2019

© 2019 Anupam Chatterjee

TABLE OF CONTENTS

Page

LIST OF FIGURES **iv**

LIST OF TABLES **vii**

ACKNOWLEDGMENTS **viii**

CURRICULUM VITAE **ix**

ABSTRACT OF THE DISSERTATION **xii**

1 Introduction **1**
- 1.1 Molecular Dynamics Simulations 2
- 1.2 Elastic Rod Model of Biofilaments 4
- 1.3 Thesis Summary 7

2 Elastic Continuum Stiffness of Contractile Tail Sheaths from Molecular Dynamics Simulations **9**
- 2.1 Introduction 9
- 2.2 Methods 15
  - 2.2.1 Building monomer models 15
  - 2.2.2 Building sheath sections 18
  - 2.2.3 Theory and simulation protocol 18
  - 2.2.4 Dynamic model of the injection process 20
- 2.3 Results and Discussion 21
  - 2.3.1 Convergence of stiffness constants and helical torsion/curvature fluctuations 21
  - 2.3.2 T4 vs. R2-pyocin strand stiffness in the same conformation (extended or contracted) 26
  - 2.3.3 Extended vs. contracted strand stiffness in the same sheath (T4 or R2-pyocin) 28
  - 2.3.4 Stiffness of T4 and R2-pyocin strands vs. actin filaments 29
- 2.4 Conclusion 30

3 **Energetics, dissipation and contraction pathway of the phage T4 injection machinery** **33**
3.1 Introduction 33
3.2 Incorporating Mechanisms that Dissipate Energy During Injection 35
3.2.1 Internal Dissipation of Sheath Strands 36
3.2.2 Sheath-Tail Tube Friction 40
3.3 Contribution of Dissipation Mechanisms to Injection Timescale 43
3.4 Dynamic Pathways for the Injection Process 46
3.5 Back-Calculating the Fully Atomic Pathways for Sheath Contraction 49
3.5.1 Sheath Contraction Intermediates from Continuum Data 50
3.5.2 Appending Tail-Tube, Capsid and Neck to Sheath Intermediates . . . 52

4 **DNA Homopolymer Unzipping Under Constant Force leads to Crackling Noise Type Avalanches** **54**
4.1 Introduction 54
4.2 Methods 60
4.2.1 Initial structure data and simulation protocol 60
4.2.2 Gathering unzipping velocity from trajectories 60
4.3 Results and Discussion 63
4.4 Conclusion 68

Bibliography **69**

LIST OF FIGURES

Page

1.1 Free body diagram of an infinitesimal element of elastic rod with variables related by the nonlinear Kirchoff rod model. 5

2.1 Schematic of (A) bacteriophage T4 and (B) R2-pyocin piercing a host cell membrane (not to scale). Components of the contractile tail assembly are illustrated. 11
2.2 All-atom structures of the front and top views with of full tail sheaths of (A) extended T4, (B) contracted T4, (C) extended R2-pyocin and (D) contracted R2-pyocin. The color coding is used to illustrate alternate 'helical' strands. In each figure, the boxed section depicts a closer view of 3 (out of a total of 6) adjacent strands with a surface representation used for the monomers to highlight inter-monomer contact. Non-starred sheath dimensions are reported experimental values from references [4] and [57] for T4 and from reference [31] for R2-pyocin. Starred dimensions are near estimates measured from the reported all-atom structures using VMD[44]. 12
2.3 Ribbon diagrams of (A) partial T4 (gp18M) and (B) R2-pyocin sheath monomers illustrating all domains. (C) Superposition of T4 (teal) and R2-pyocin (orange) sheath monomers oriented for maximum structural overlap. (D) Full gp18 structure modeled using the R2-pyocin structure superposition. 16
2.4 Various sections of the full tail sheath: (A) continuum rod with equivalent radius and elastic properties of an all-atom sheath strand, (B) free body diagram of an elastic rod with variables related by the nonlinear Kirchoff rod model, (C) a 4-disc section of the full sheath used in the actual MD simulations, with a best-fit helical strand passing through the centers of masses of monomers, (D) all-atom model of an individual sheath monomer. 21
2.5 Variation of (A) mean curvature and (B) mean torsion for extended and contracted sheath strands of T4 and R2-pyocin during the full 30ns unconstrained MD run. The highlighted sections are for the initial 15ns equilibration runs and were not considered for the stiffness constant calculations. 22
2.6 Bending and torsional stiffness constants for different sampling windows (Δt) obtained for extended and contracted sheath strands of phage T4 and R2-pyocin. 23

2.7 RMSDs relative to the unequilibrated structures for the core domains (Domain III for T4 and the N-terminal domain, see Fig. 2.5) for both extended and contracted conformations. The initial 15ns equilibration run is highlighted in yellow. 25

3.1 Arrangement of subunits in a partially contracted T4 sheath as proposed by Moody. Each circle represents a sheath monomer. Image adapted from Ref. [72]. 34
3.2 (a) Atomistic structure of the four-disc section of the T4 tail sheath with the middle ring surrounded by the superimposed rectangle. (b) The best-fit circle through the centers of masses (red dots) of the ring subunits of the middle ring has mean radius $r(t)$. The radial fluctuations of the filament from the circle are denoted by $u(t)$ 39
3.3 Autocorrelation of the transverse displacement of a middle ring of the sheath fragment from MD simulation for (a) the extended conformation and (b) the contracted conformation. The discrete data points indicate the MD-derived autocorrelation and the solid line represents the best fit per (20). Note log scale on vertical axis. 40
3.4 (a) Surface Coulomb potential distribution of (clockwise) four rings of the outer surface of the tail tube, the inner surface of the extended sheath, and the inner surface of the contracted sheath. Beneath are top views of the tube within the extended sheath and within the contracted sheath. (b) Kyte-Doolittle hydrophobicity of the same surfaces arranged in the same order as in (a). The outer surface of the tail tube and the inner surface of the sheath are largely hydrophilic (blue). 42
3.5 Components of the velocity profile of water within the nanoscale gap d between the sheath and the tail tube due to (a) translation v, and (b) rotation ω of the tail tube during injection. 43
3.6 The timescale of the injection process as a function of the sheath-tube gap viscosity coefficient η_w over the range $0.001 < \eta_w < 1000$ Pa-s and the internal friction coefficient $\eta^{'}$ over the range $0.005 < \eta' < 5$ Pa-s. Note log scales. Red region defines a likely injection time scale for T4 (approximately 5 ms) by drawing a comparison with that of the Type VI secretion system (T6SS) [10]. 46
3.7 Simulation reveals contraction dynamics of the sheath from the fully extended conformation to the fully contracted conformation, consistent with the experimental micrographs reported by Moody[72]. (a) Micrographs of sheath in extended (I) , partially contracted (II), and fully contracted (III) conformations[72]. The snapshots illustrate model simulated extended (IV), partially contracted (V), and fully contracted (VI) conformations. Intermediate conformation (V) captures contraction wave propagation from the (lower) baseplate towards the (upper) neck. (b) The helical strand angle θ as a function of location along the strand. As the contraction wave passes, θ decreases from that of the extended conformation (IV) with $\theta_{ext} = 59^o$ to that of the contracted conformation (VI) with $\theta_{cont} = 16^o$. 49

3.8 Steps in the contraction of the T4 tail sheath model as described in the mechanical contraption of Caspar (image reproduced from Fig. 13 of reference [15]). Note that the T4 sheath was believed to have 24 hexameric annuli when the work was published (more recent studies have concluded that the correct number is 23). The sheath subunits were represented by white knobs and were pinned to grooves in the central tailtube. The pins were sequentially removed starting from the baseplate to initiate contraction. 50
3.9 Complete model of the T4 injection machinery predicts the internal (elastic) energy of the contractile sheath that drives the injection process. (a) The sheath begins in the (high-energy) extended state where the tip of the tail tube remains 100 Åfrom the cell membrane. (b) The sheath contracts 100 Åso that the tip of the tail tube touches the outer cell membrane. (c) Further contraction of the sheath produces, in sequence, cell indentation, rupture of the outer cell membrane (after 60 Åindentation), and penetration into the (viscous) periplasmic space. (d) The fully contracted sheath with zero elastic energy. 51
3.10 Snapshots from the Back Calculated Atomic Pathway of T4 Contraction of (A) the initial extended tail, (B) a contraction structure and (C) the final contracted tail. 53

4.1 Distribution of neuronal avalanche sizes (S) in zebrafish larvae. Image from Fig. 2B of Ponce-Alvarez et al. [84] 56
4.2 Magnetization (J) or flux density (B) curve as a function of magnetic field intensity (H) in ferromagnetic material. The inset shows Barkhausen jumps. Image from [106]. 57
4.3 The DNA 18-mer homopolymer illustrating the constant applied force for unzipping $\vec{F}$ and the distance d used to identify the strand separation. . . . 61
4.4 Distance between pulled O3' atoms during unzipping with a smooth trend-line obtained after filtering the thermal noise contribution 62
4.5 Unzipping velocity with the red threshold line for defining avalanche clusters. 63
4.6 Distribution of avalanche duration with best-fit curve and straight line with slope equal to the critical exponent α. 65
4.7 Distribution of avalanche size with best-fit curve and straight line with slope equal to the critical exponent τ. 66
4.8 Avalanche duration v/s mean size with best-fit curve and straight line with slope equal to the critical exponent $1/a$. 67

LIST OF TABLES

Page

2.1 Bending (A) and torsional (C) stiffness constants and their ratios for extended T4 strand, contracted T4 strand, extended R2 (pyocin), contracted R2, and actin filaments for comparison. 24

2.2 Volume and solvent accessible surface areas (SASA) per monomer for the extended and contracted 4-disc sheaths, 4-monomer strands, and monomers for the extended and contracted conformations. R2 is short for the R2-pyocin sheath. 27

4.1 Critical exponents τ, α and a calculated for Barkhausen noise, crack growth and neuronal avalanches. PC refers to polycrystalline and AM refers to amorphous ferromagnetic films. 67

ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor Prof. Andricioaei for the extremely fruitful scientific discussions which have shaped almost all of the content of this thesis. Thank you Ioan for all of the guidance and mentoring you patiently provided during various points of (sometimes difficult) decision making throughout my tenure at UCI.

I would also like to thank our collaborators at the University of Michigan, Prof. Noel Perkins and Dr. Ameneh Maghsoodi, for the discussions on the continuum rod model and contributions to the contractile tail project.

In addition, I would thank all the current and previous members of the Andricioaei group: Jim, Gianmarc, Emel, Moises and Dhiman for their discussions and valuable suggestions through the past few years. I would also like to express my gratitude to Prof. Doug Tobias and Prof. David Mobley for their suggestions and being part of my dissertation and advancement committees. Thanks also to my P-Chem batchmates and UCI friends Saswata, Saleh, John, Andrew, Sree....Thank you Kanika for being supportive of me through the last year and helping and encouraging me through some difficult times. Thanks to my family and friends back home in India for being supportive and understanding. This would not have been possible without all of you.

Special acknowledgments to my Masters thesis advisor Prof. BL Tembe and my undergraduate research group at IIT Bombay who first introduced me to the exciting world of MD simulations and laid the platform for me to pursue graduate research in theoretical chemistry.

I would also like to acknowledge the research funding received during my graduate study from the National Science Foundation, the National Institutes of Health and the teaching assistantships received from the Department of Chemistry, UCI. Finally, I would also like to acknowledge the Greenplant computing facility at UCI and the San Diego supercomputing facility which enabled me to run my expensive simulations in good time.

CURRICULUM VITAE

Anupam Chatterjee

EDUCATION

Doctor of Philosophy in Chemistry **2019**
University of California, Irvine *Irvine, California*

Master of Science in Chemistry (5-Year Integrated) **2014**
Indian Institute of Technology, Bombay *Mumbai, India*

RESEARCH EXPERIENCE

Graduate Research Assistant **2009–2014**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2009–2014**
University of California, Irvine *Irvine, California*

HONORS AND AWARDS

Acceptance with Scholarship to CECAM School on Molecular Kinetics **2016**
EPFL, Lausanne

Kishore Vaigyanik Protsayan Yojana (KVPY) Fellowship **2009-14**
Department of Science and Technology, Govt. of India

Best Poster Award **2014**
DCCBS Conference, IIT Kanpur

Undergraduate Research Award **2013**
URA01, IIT Bombay

REFEREED JOURNAL PUBLICATIONS

Crackling Noise in DNA Constant Force Unzipping **2019**
In Preparation for The Journal of Chemical Physics

Elastic continuum stiffness of contractile tail sheaths from molecular dynamics simulations **2019**
In Preparation for The Journal of Chemical Physics

How the Phage T4 Injection Machinery Works: Energetics, Forces, and Dynamic Pathway **2019**
Submitted to Proceedings of the National Academy of Sciences of the USA

Dynamic Model Exposes The Energetics and Dynamics of the Injection Machinery for Bacteriophage T4 **2017**
Biophysical Journal

A First Model of the Dynamics of the Bacteriophage T4 Injection Machinery **2016**
Journal of Computational and Nonlinear Dynamics

Salting-Out of Methane in the Aqueous Solutions of Urea and Sarcosine **2016**
Journal of Chemical Sciences

Na^+ Cl^- Ion Pair Association in Water-DMSO Mixtures: Effect of Ion Pair Model Potentials **2016**
Journal of Chemical Sciences

Instantaneous, Parameter-Free Methods to Define a Solutes Hydration Shell **2015**
The Journal of Chemical Physics

Solvation Structures and Dynamics of the Magnesium Chloride Ion Pair in Water–Ethanol Mixtures **2013**
The Journal of Physical Chemistry A

CONFERENCE PRESENTATIONS

Elastic Properties of Contractile Tails in Pre- and Post-contraction States: Insight From Equilibrium MD Simulations **May 2017**
Southern California Theoretical Chemistry Symposium

Simulation Studies of Twist-Stretch Coupling in Nucleic Acids **March 2016**
Biophysical Society Annual Meeting

Parameter-Free Methods to Define a Solutes Hydration Shell from Molecular Dynamics Simulations **February 2014**
DCCBS Conference, IIT Kanpur

ABSTRACT OF THE DISSERTATION

Simulation Aspects of the Mechanics of Biomolecular Filaments: Crackling in DNA unzipping and the Contraction of Bacteriophage Tails

By

Anupam Chatterjee

Doctor of Philosophy in Chemistry

University of California, Irvine, 2019

Ioan Andricioaei, Chair

Both DNA and the contractile tail sheaths of bacteriophages are examples of biofilaments, whose monomer subunits consist of nucleotides and proteins respectively. The bending and torsional deformations of tail sheaths and strand separation of ds-DNA are important phenomena essential for their biological functions. Despite the great prevalence and biomedical importance of contractile delivery systems, many fundamental details of their injection machinery and dynamics are still unknown. On a similar note, a detailed theoretical understanding of the monomer-level dynamics of DNA unzipping under constant force is also lacking in literature. In the subsequent chapters of this thesis, I will describe how computer simulations can be used to perform an in-depth study of both of the above phenomena. I would begin by describing a method which uses molecular dynamics simuations to calculate the bending and torsional stiffness constants of two biologically relevant contractile tail sheaths: bacteriophage T4 and R2-pyocin. Next, I would describe how the stiffness constants can be incorporated in a continuum dynamic model to simulate the dynamics of contractile nano-injection machineries. Finally, I would describe how MD simuations can be used to study the unzipping dynamics of a long DNA homopolymer, which would to a fascinating discovery where the 'avalanches' in the unzipping velocity time series show a power law variation in avalanche size and time similar to crackling noise in other unrelated physical systems.

The studies of these phenomena are of great biological significance; studying contractile tail injection dynamics can open up new avenues in potential bio-nanotechnological applications like experimental phage therapy, and understanding of DNA unzipping at the monomer level is relevant to many essential genetic processes like replication, transcription, recombination, DNA repair, and -in biotechnology- to DNA sequencing.

Chapter 1

Introduction

Within the last four to five decades, with the advent of powerful computers, computer simulations have become a highly sophisticated tool to understand the structure, function and dynamics of biomolecules like protein assemblies and nucleic acids. The atomic-level interactions in these molecules can be explained to a high degree of accuracy using approximations based on quantum mechanical equations. However, these equations are not practical to be solved (at biologically relevant timescales) for systems beyond a few hundreds of atoms in size even with today's computing power. For considering larger biological systems in solvent, classical force-field based molecular dynamics (MD) simulations have been highly successful in studying the dynamical pathways of protein folding [26, 82] and DNA melting [87, 56], among others. Systems consisting of multiple large protein molecules like viruses, where interesting functional phenomena occur in the timescale of milliseconds, are practically beyond the size and time-scale of MD simulations and further coarse-grained representations like elastic rod models are necessary to study them. In the next two sections of this chapter, I will provide an overview of the theory MD simulations and an elastic rod modeling approach to biomolecular filaments. In the final section, I will summarize how MD simulations can be applied to study constant force DNA unzipping and the multi-protein filament of a contractile tail sheath. I will also summarize how the stiffness constants of the contractile tail sheaths can be incorporated in a hybrid modeling approach to describe the dynamics of the bacteriophage T4 contraction process.

1.1 Molecular Dynamics Simulations

In molecular dynamics simulations, a macromolecular system is represented by a collection of atoms, which are considered as point charged particles without any internal degrees of freedom. Interaction between the atoms is defined by an effective potential field. The effective potential is a sum of bonded terms and non-bonded terms. Bonded terms are interactions that consider covalent bond stretching, angle bending, proper and improper dihedrals. Non-bonded terms are for long-range interactions like the van der Waals and electrostatic Coulomb interactions. The general form of an MD potential function with the atomic Cartesian coordinate vectors $\vec{r}_i$ can be written as:

$$\begin{aligned} V(r) = & \sum_{bonds} k_b(b-b_0)^2 + \sum_{angles} k_\theta(\theta-\theta_0)^2 + \sum_{dihedrals} k_\phi(1+cos(n\phi-\delta)) + \sum_{improper} k_\psi(\psi-\psi_0)^2 \\ & + \sum_{UreyBradley} k_{UB}(S-S_0)^2 + \sum_{nonbonded} 4\epsilon_{i,j}[\left(\frac{\sigma}{r_{i,j}}\right)^{12} - \left(\frac{\sigma}{r_{i,j}}\right)^{6}] + \sum_{nonbonded} \frac{q_i q_j}{\epsilon + dr_{i,j}} \end{aligned} \tag{1.1}$$

The first five terms are the bonded terms and the last two non-bonded interaction terms. The bonded terms for bonds, angles, improper dihedrals and the Urey-Bradley component are all harmonic approximations. The nonbonded interactions are the more computationally expensive parts of the full potential, and are evaluated with a user-defined cut-off for long-range interactions (for details on the set of parameters in the functional form of the potential see, for example, ref. [5]).

An MD simulation run usually starts with an initial configuration of the system which is most oftenly derived experimentally. For biomolecules like proteins and nucleic acids, the initial configuration can be derived from high-resolution X-ray crystallography, NMR and/or cryo-EM experiments. The molecule is then either solvated using models for explicit water molecules or using an implicit continuum solvent model. During the MD simulation, the particles in the system starting from the initial configuration experience forces $\vec{F_i}$ which can be determined by differentiating the potential function:

$$\vec{F_i} = -\frac{\delta V(\vec{r_i})}{\delta \vec{r_i}} \tag{1.2}$$

From the forces, a deterministic time series of the coordinates of all particles of the system can be generated by a time integration of Newton's equations of motion. The particle initial velocities are assigned from the knowledge of the temperature of the system. For a detailed description of the technical aspects of MD simulations, I would refer the readers to one of the numerous standard texts available on the subject [5].

The trajectories generated from MD simulations can be used to investigate various structural properties and conformational dynamics of macromolecules [49]. Using the ergodic hypothesis of statistical mechanics, a sufficiently long MD trajectory can be used to approximate an equilibrium ensemble of structures. The ensemble averages can be used to calculate useful bulk thermodynamic properties of the systems like enthalpy, entropy, heat capacity, X-ray temperature factors, reaction rates and equilibrium constants. Equilibrium fluctuations in MD trajectories can also be used to calculate elastic properties of biomolecular filaments like their persistence lengths, Young's moduli and stiffness constants [55, 46, 68]. MD simulations serve as a useful (and sometimes vital) link between experimentally observed three-dimensional structures of macromolecules and their dynamics. Their ability as a relatively inexpensive tool to augment experimental work has made them extremely popular among theoretical scientists. Multiple software packages are available to run MD simulations, and they are continuously updated to make them faster and more parallelizable [14, 81, 100].

Although all-atom MD simulations provide an accurate representation of a molecular system, there are limitations in their utility as a tool for biophysical analysis. A major bottleneck is the timescale currently achievable by MD simulations. The calculations of non-bonded pair interactions are computationally expensive and the CPU cost rises roughly as $\sim N^2$, with N the number of atoms in the system. The presence of significantly fast bond vibrational frequencies limits the integration timestep to one or two femtoseconds. As a result, the timescales currently achievable by standard MD simulations for systems with tens of thousands of atoms is in the range of microseconds or less. This introduces severe limits when the slow conformational changes involved in the phenomenon to be studied are greater than the microsecond scale and systems sizes are in the range of millions of atoms. It is therefore useful to exploit coarse grained representations of molecular systems to probe long-timescale phenomenon. One such representation, the elastic rod model, is discussed in the next section.

1.2 Elastic Rod Model of Biofilaments

Biofilaments like DNA, microtubules, actin filaments, and sheaths of contractile tails are essential to functions in living cells and viruses. An elastic rod model is essentially a coarse-grained continuum approximation which sacrifices atomic detail to describe the long length

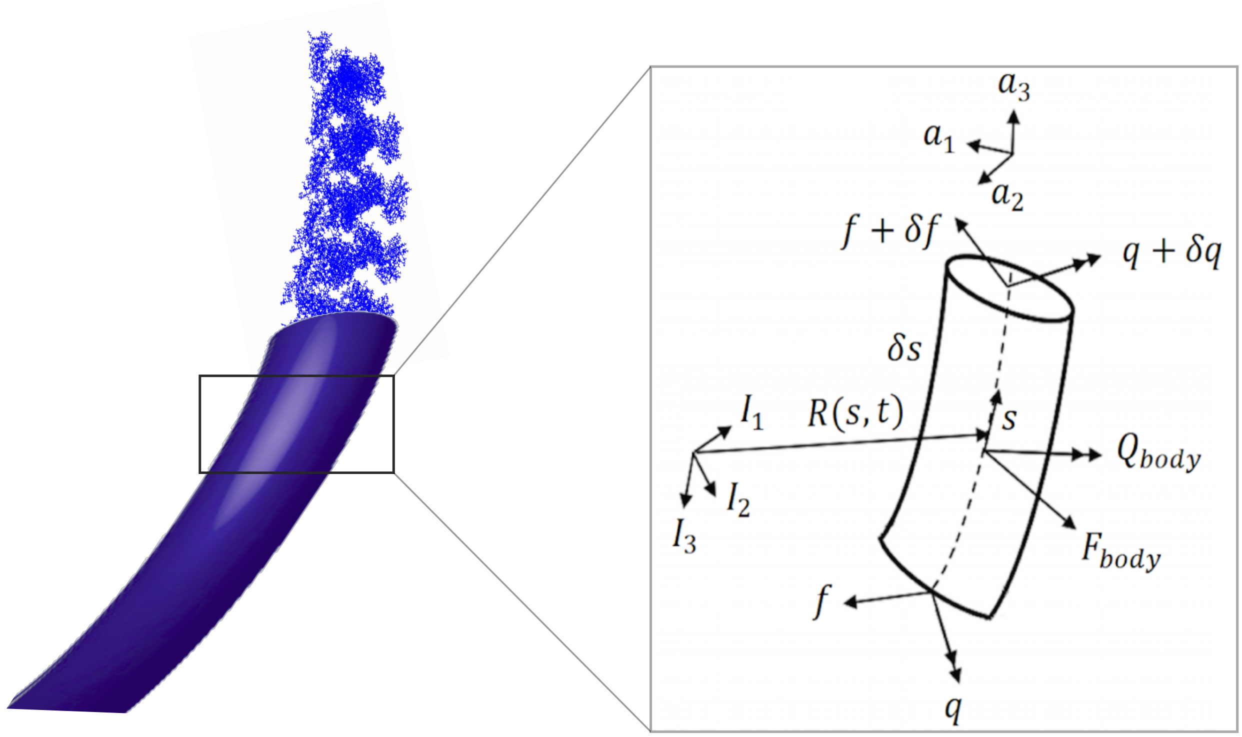

Figure 1.1: Free body diagram of an infinitesimal element of elastic rod with variables related by the nonlinear Kirchoff rod model.

and time scale functions of biofilaments. The rod model formulation has been previously employed to study protein-mediated DNA looping [36], DNA supercoil formation and relaxation [38, 60, 59], and DNA buckling in viral cavities [42]. The advantages of using rod theory to approximate biopolymers like DNA include the ability to resolve large-scale structures like DNA loops and supercoils. With the exception of the formulation originating in [37], the prior rod formulations above exclusively employ equilibrium rod theory and thereby preclude from consideration the dynamic contraction of a biofilament like a phage sheath strand, which is the focus of study in Chapters two and three of this thesis. I will now outline a dynamical rod theory as applicable to the work in this thesis, where we consider the simultaneous conformational change of six helical sheath strands of a six-fold symmetric contractile tail sheath. In the dynamic model, each helical strand of the sheath is modeled as a nonlinear Kirchhoff rod [63] (Fig. 1.1). The governing dynamical equations of the ith strand of sheath (i=(1,2,...6)) are the following,

$$\left\{\frac{\partial f}{\partial s} + \kappa \times f = m_s\left(\frac{\partial v}{\partial t} + \omega \times v\right) - F_{body}\right\}^i \tag{1.3}$$

$$\left\{\frac{\partial q}{\partial s} + \kappa \times q = I_s\frac{\partial \omega}{\partial t} + \omega \times I_s\omega + f \times a_3 - Q_{body}\right\}^i \tag{1.4}$$

$$\left\{\frac{\partial v}{\partial s} + \kappa \times v = \omega \times a_3\right\}^i \tag{1.5}$$

$$\left\{\frac{\partial v}{\partial s} + \kappa \times v = \frac{\partial k}{\partial t}\right\}^i \tag{1.6}$$

Equations (1.3) and (1.4) describe the balance laws for linear and angular momentum, respectively, and equations (1.5) and (1.6) describe constraints on rod inextensibility and rotation, respectively. The variables $\omega^i(s,t)$ and $v^i(s,t)$ are the angular velocity and translational velocity of ith strand cross-section, respectively, $\kappa^i(s,t)$ is the curvature/twist vector, $m_s^i(s)$ is the mass of the ith strand per unit contour length, and $I_s^i(s)$ denotes the diagonal 3×3 tensor of principal mass moments of inertia per unit contour lengths. The quantities $f^i(s,t)$ and $q^i(s,t)$ are the internal force vector and internal moment vector, respectively, and a_3^i is the unit tangent vector at each cross-section. Finally, $F_{body}^i(s,t)$ and $Q_{body}^i(s,t)$ denote the sum of all distributed external body forces and moments per unit contour length, respectively, including the drag and moment on the sheath strands from water. In eq. (1.4), the internal moment $q^i(s,t)$ is related to the curvature/twist vector through an assumed linear elastic constitutive law,

$$q^i(s,t) = B^i(\kappa^i - \kappa^i_0) \tag{1.7}$$

where κ^i_0 is the known intrinsic curvature/twist vector of the ith helical strand in the contracted state of the sheath. $B^i(s,t)$ is a diagonal 3×3 stiffness tensor of ith strand defined as follows,

$$B^i = \begin{bmatrix} A^i & 0 & 0 \\ 0 & A^i & 0 \\ 0 & 0 & C^i \end{bmatrix} \tag{1.8}$$

where A^i and C^i are the bending and torsional stiffness constants of the ith sheath strand. Further details of this rod model formulations for a contractile sheaths are described in [63, 64].

1.3 Thesis Summary

The approach followed in this thesis is to investigate how MD simulations can be used on their own and in conjunction with highly coarse-grained continuum elastic rod models to describe interesting dynamical phenomenon in biofilaments like DNA and bacteriophage tails. In the next (second) chapter, I describe how classical MD simulations can be used to determine the elastic stiffness constants of the tail sheaths of two contractile systems, Bacteriophage T4 and R2-Pyocin. In the third chapter I describe how, in collaboration with the group of Dr. Noel Perkins in the University of Michigan, we employed the stiffness constants calculated for the bacteriophage T4 in a companion continuum model of the entire phage sheath that also couples to a model of the capsid and tail tube assembly. The resulting model of the entire T4 injection machine predicts the nonlinear and rapid dynamic conformational changes induced during the injection process. Doing so exposes the energetics, timescale, and pathway of these dynamical changes. In the fourth and final chapter, I describe how ‘crackling’ noise type dynamics can be observed when MD simulations are used to unzip a double stranded DNA homopolymer strand using a constant external force applied at its two ends perpendicular to the helical axis. The specific applications of the simulation approaches presented in this thesis can be expanded to better understand other challenging biological phenomenon which occur across a wide range of length and time scales.

Chapter 2

Elastic Continuum Stiffness of Contractile Tail Sheaths from Molecular Dynamics Simulations

2.1 Introduction

A number of related biological nano-machines have specialized organelles in the form of tails that contract, whose essential function is to pierce the membrane of a host cell and deliver DNA, protein, or ions through a conduit involving a tail tube-sheath assembly. Two intriguing classes of such systems –whose atomic-level structures are known– are contractile-tailed bacteriophages and R-type pyocins. Bacteriophages infect the bacterial host cells by injecting genomic DNA into the host [57, 58]. R-type pyocins, produced by *P. aeruginosa*, attack competing bacteria by channeling protons and cations through their tail structure inside the bacteria and thereby dissipating their membrane potential [71, 32, 31]. Significant research interest has been devoted over the past decades towards understanding the function of contractile nanomachines, with potential bio-nanotechnological applications ranging from detection and control of pathogens [40] and peptide display [78], to experimental phage therapy [77] and generating novel bactericidal protein complexes [107, 89].

Despite differences between the two systems in terms of the nature of the material injected (DNA vs. protons and cations), in intriguing illustrations of "form follows function," the overall structures of bacteriophages and R-type pyocins are remarkably similar. Phage T4, a virus of the *Myoviridae* family of tailed phages that infects the *E-coli* bacterium, has been extensively studied, and its structure has been characterized in detail using X-ray crystallography and cryo-electron microscopy (cryo-EM) methods [4, 110, 57, 29, 48]. It consists of a long contractile tail assembly which connects to a multi-protein capsid containing the genomic DNA (172 kbp)[4, 24, 90]. The tail assembly which is responsible for transferring DNA from capsid to the host consists of a long rigid tube surrounded by a contractile sheath. The tail sheath attaches to the capsid by a neck (collar) at the upper end and to the baseplate at the lower end (Fig. 2.1A). In comparison, R2-pyocin is a headless and DNA-free injection machinery [31, 111], but its long contractile tail assembly is closely similar to that of phage T4. It consists of a tail tube surrounded by a sheath, a collar at the upper end, and a baseplate at the lower end of the tail (Fig. 2.1B).

From the cryo-EM images and the crystal structure fits, it is clear that the sheaths of both bacteriophage T4 and R2-pyocin are constructed from the backbone of six interacting helical 'strands' of a sheath protein monomer. During contraction, the sheaths undergo a large conformational change from an extended (high-energy) state to a contracted (low-energy) state, reducing to about half of their initial length and significantly increasing their diameter (Fig.2.2). The sheath monomer subunits largely translate and rotate as rigid bodies without significant changes in monomer structure from the extended to contracted conformation for both T4 and R2-pyocins [4, 31].

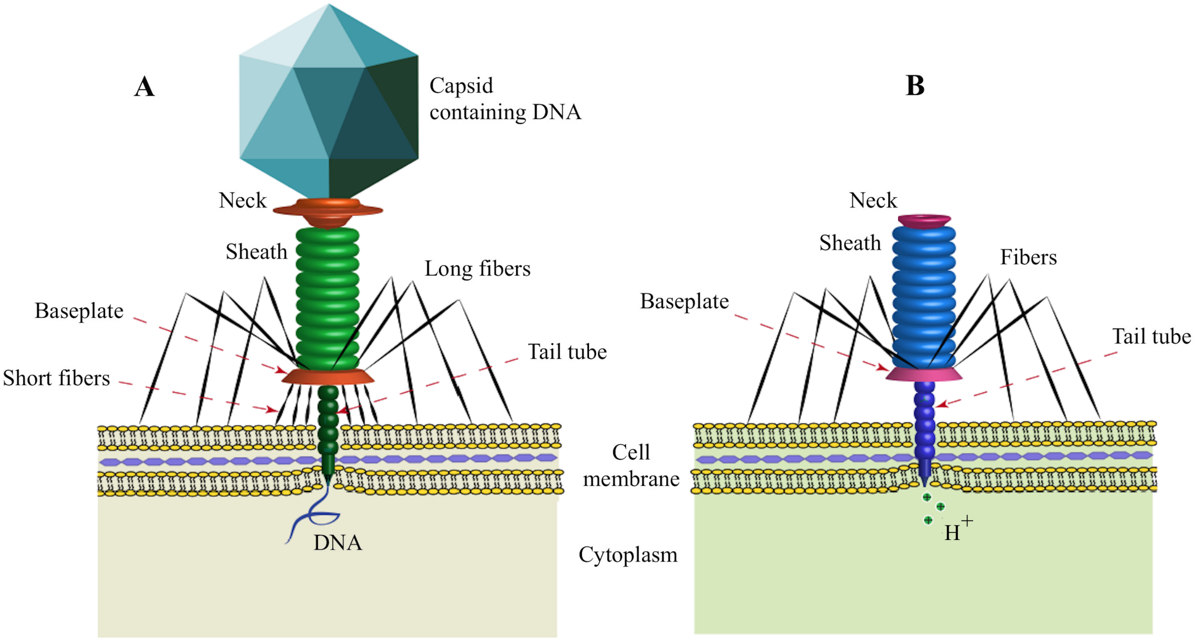
A
Capsid
containing DNA
Neck
Sheath
Long fibers
Baseplate
Tail tube
Short fibers
DNA
B
Neck
Sheath
Fibers
Baseplate
Tail tube
Cell
membrane
Cytoplasm
H^+

Figure 2.1: Schematic of (A) bacteriophage T4 and (B) R2-pyocin piercing a host cell membrane (not to scale). Components of the contractile tail assembly are illustrated.

Despite rich structural data known about both T4 and R2-pyocin in the extended and contracted states, not much is known about the intermediate structures or dynamical details of their injection processes. As all known contractile tails are essentially similar in assembly and function [58], the contraction pathway is expected to be similar for both T4 and R2-pyocin. On the experimental side, only one known microcalorimetric study in literature has probed the enthalpy of phage T4 contraction process, where contraction was induced *in vitro* by both heat and urea [7]. Moody in 1973 predicted a possible 'contraction wave' pathway for sheath contraction from observation of electron micrographs of partially contracted sheaths *in vitro* [72]. On the theoretical side, Falk and James used elasticity theory to describe an approximate coarse grained free energy surface for the extended to contracted conformational change of the T4 sheath, but the theory fails to account for any dynamics

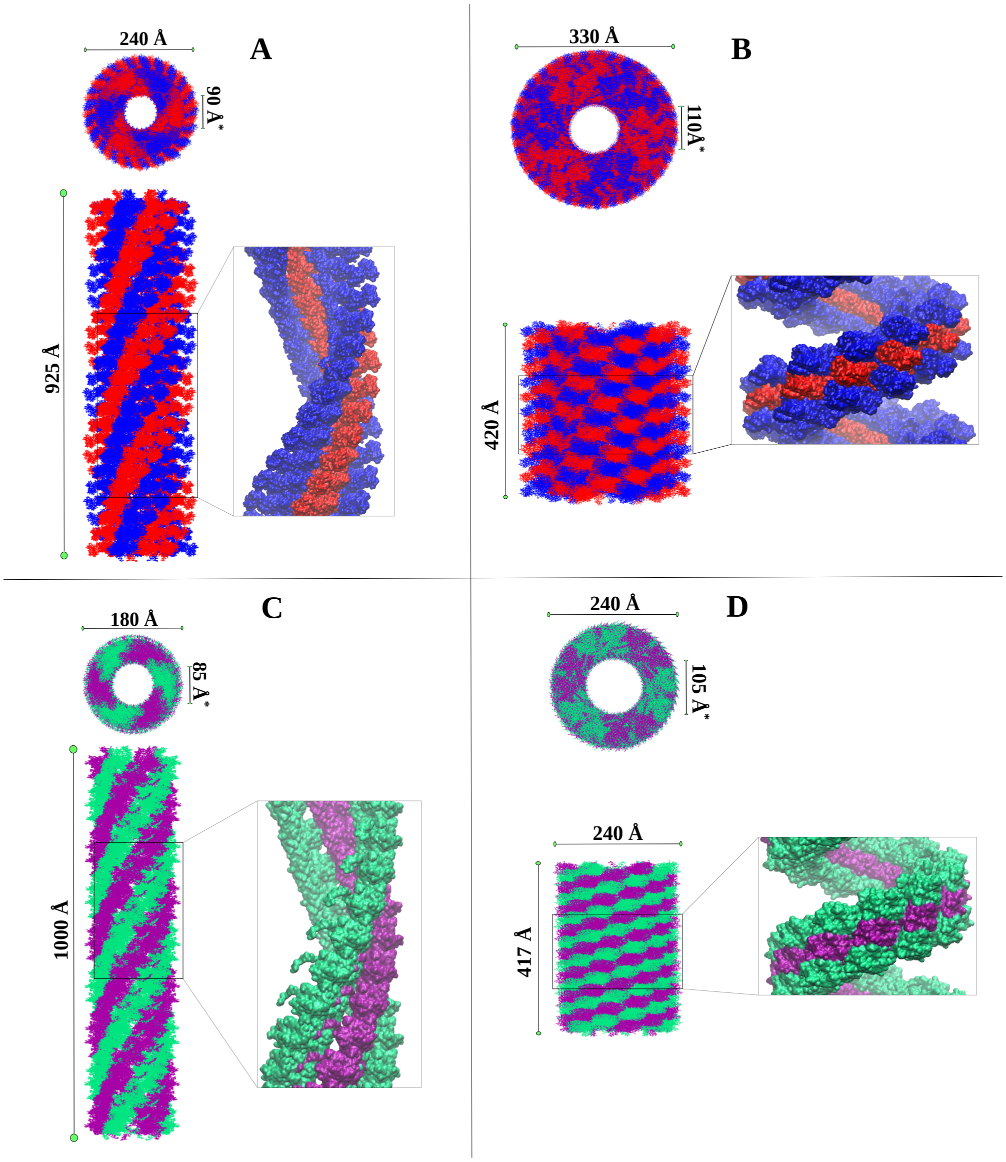

Figure 2.2: All-atom structures of the front and top views with of full tail sheaths of (A) extended T4, (B) contracted T4, (C) extended R2-pyocin and (D) contracted R2-pyocin. The color coding is used to illustrate alternate 'helical' strands. In each figure, the boxed section depicts a closer view of 3 (out of a total of 6) adjacent strands with a surface representation used for the monomers to highlight inter-monomer contact. Non-starred sheath dimensions are reported experimental values from references [4] and [57] for T4 and from reference [31] for R2-pyocin. Starred dimensions are near estimates measured from the reported all-atom structures using VMD[44].

of the process [27]. Ge et al. calculate the predicted theoretical free energy of R2-pyocin contraction using PISA [53], but experimental corroboration of the predicted energies are lacking.

In principle, the entire dynamics of sheath contraction could be simulated with atomic detail using molecular dynamics (MD) simulations [50], and results could be used to make models of the contraction mechanism and compare them to existing models. For example, the pathways from atomically detailed trajectories could be compared with intermediates along the contraction wave of T4 observed by Moody [72]. However, the systems are extremely large, counting multiple million atoms (approximately a million atoms for just the T4 sheath alone, plus surrounding water and the newly resolved tail tube and base-plate [97]). In addition, the time scales currently achievable by straightforward molecular dynamics simulations for such large systems fall short (by at least six orders of magnitude) of the expected millisecond-second or longer time scales needed for a single transition from the extended to the contracted state. To model the complete dynamics of sheath contraction, one clearly needs to exploit coarse grained representations for the tail sheath in the entourage of the entire capsid, tail tube, and host membrane.

Various avenues have been explored to describe the dynamics of large macromolecules like multi-domain proteins and biofilaments by reducing the dimensionality of the modeled system via multi-scale approaches [28, 8]. Continuum models, which dispense completely with atomic detail have been successful in describing mechanical properties of large proteins and assemblies [11, 2]. These models involve atomic structural parameters which need to be derived from either experiments or smaller-scale simulations so that the results are physically interpretable. Hicks et al. report an approach to derive the dynamics of large protein assemblies like viral capsid proteins by treating each individual protein as one or few rigid coarse grained units [41]. Interaction between the rigid units is modeled using a simplified potential, for example a generalized harmonic potential, and parameters for the potential can be derived using all-atom MD simulations of a small number of subunits over manageable time-scales. During contraction, the sheath helical strands function very much like springs that undergo changes in their elastic parameters from the extended state to a contracted state during injection (Fig. 2.2). The entire contraction process can be modeled as six continuum elastic springs dynamically contracting according to defined equations of motions and appropriate boundary conditions. We believe that a continuum model will be most appropriate for this system because, (1) the large sheath conformational deformation involved is on a length scale that is comparable to that of the component monomer subunits, and continuum models have been shown to describe such large changes, like the icosahedral to dodecahedral "buckling transition" of a viral capsid, very well [2], and (2) a continuum model enables us to see direct kinematics of sheath contraction over relevant timescales.

To model sheath contraction in a way such that we can make accurate quantitative predictions, we require a continuum model with elastic parameters that have a microscopic basis in the actual sheath structure. Elastic parameters of biofilaments such as microtubules and actin have been calculated experimentally using micromanipulation[51, 99] or direct observation of thermal fluctuations [33]. In these experiments the filaments used were of the μm length scale and fluctuations were observed over the μs - ms timescale. Similar experimental data is unavailable for phage T4 and pyocin sheaths, so an alternative is to use all-atom MD simulations for calculating the elastic parameters. In principle, an MD simulation on the μs - ms timescale should capture all possible spatial fluctuations and highly accurate elastic constants can be derived, but simulations on that timescale for large systems are not feasible. However, MD simulations even on the ns scale have been used extensively to get good estimates of elastic properties of microtubules[105], actin filaments[80, 93], FtsZ filaments [35] and viral capsids[68]. Matshushita et al. calculated the extensional and torsional stiffness of actin filaments from MD simulations and in an analysis of sampling time vs. stiffness constants, showed that using a sampling window of 16 ns duration resulted in constants which were in good agreement with experimental constants derived from μs scale experiments[67].

In the present study, we derive bending and torsional elastic stiffness constants of a helical sheath strand of phage T4 and R2-pyocin, in both the pre- (extended) and post-contraction (contracted) states, from large-scale *ns* length MD simulations of about one-sixth of the complete sheaths. As a main application of this work, it is described how the derived stiffness constants can be incorporated into an elastic-continuum dynamic model to explore unknown parameters including the energetics, dynamics, time scale and pathways of the injection process for contractile nano-injection machineries including phage T4 and R2-pyocin. As an example, the constants calculated for phage T4 have been employed in a continuum model of the T4 sheath composed of six interacting helical strands by representing each strand as a homogenous, isotropic elastic rod with time-varying elastic constants. Our recently published work describing the results of the model [64] predicts an energy of contraction for T4 which is close to the experimentally calculated enthalpy of contraction. Another work describing a similar continuum model for R2-pyocin is in progress. In future, as their all-atom structures are experimentally reported, our methodology can be extended to other known bio-systems featuring contractile tails, for example, type-VI secretion systems (T6SS) [9] which transport macromolecules such as protein across target cell membranes, and phage ϕ812[76], which injects DNA into the host bacterium *Staphylococcus aureus*.

2.2 Methods

2.2.1 Building monomer models

All-atom structures of the complete R2-pyocin monomers in their pre- and post-contraction states were reported by cryo-EM to a resolution of 3.5 Å and 3.9 Å respectively (PDB codes 3J9Q and 3J9R[31]). The R2-pyocin monomer (Fig. 2.3B) in the extended conformation consists of a prominent N terminal domain with a six-stranded β-sheet, a C domain with a

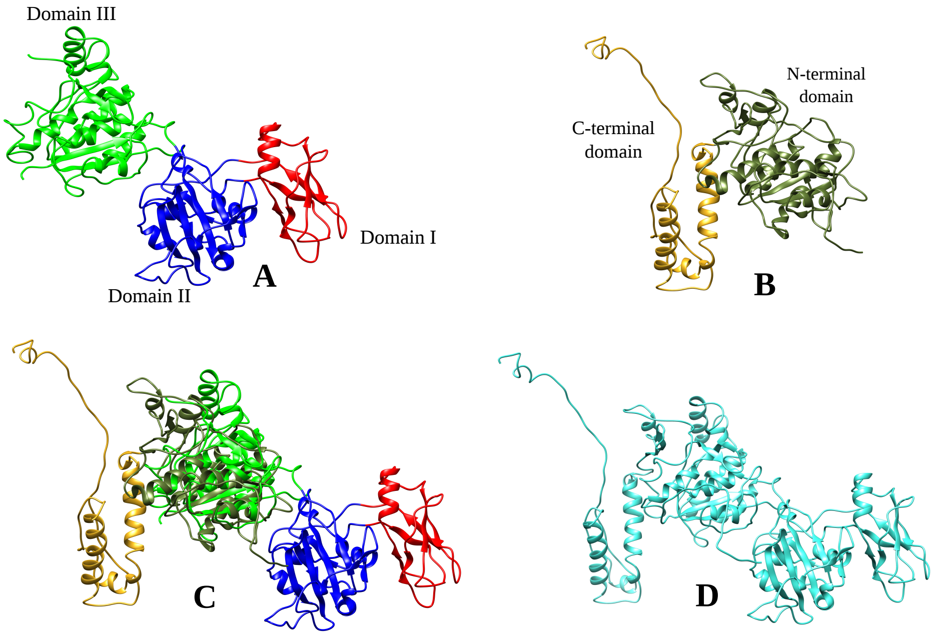

Figure 2.3: Ribbon diagrams of (A) partial T4 (gp18M) and (B) R2-pyocin sheath monomers illustrating all domains. (C) Superposition of T4 (teal) and R2-pyocin (orange) sheath monomers oriented for maximum structural overlap. (D) Full gp18 structure modeled using the R2-pyocin structure superposition.

four stranded β-sheet and two α-helices, and long extension arms at both N and C termini of the protein, which extend into neighboring subunits of the sheath [31].

An all-atom crystal structure of about three-fourths of the T4 sheath monomer gp18 (gp18M) is available (PDB code 3FOA)[4]. The partial T4 monomer consists of three domains (Fig. 2.3A). Domain I has a six stranded β sheet and an α-helix, domain II consists of two β-sheets and four small α-helices, and domain III (homologous to the N-terminal domain of the R2-pyocin monomer, see (Fig. 2.3C) consists of a six stranded β-sheet plus six α-helices. The available structure is missing the N-terminal disordered linker arms and the entire fourth inner domain, which includes the C-terminal linker arms. It has been established beyond reasonable doubt that the missing fourth domain of the T4 monomer is homologous to the corresponding regions of other *Myoviridae* phage tail sheath proteins [3], which in turn is homologous to the C-terminal domain of the R2-pyocin monomer[86]. Recent studies of the

T6SS sheath structures show that the inner domain in the T6SS sheath forms a mesh identical to the pyocin [9, 86]. Thus, the missing inner domains of the T4 sheath are expected to interact with each other within each strand and across the neighboring strands, and the missing linker arms act as hinges to keep the sheath structure intact during contraction. It is necessary to model the missing structure of the T4 monomer before MD simulations are performed.

To model the missing domain of the T4 monomer, an initial superposition of the gp18M and the R2-pyocin monomers was created for maximum structural overlap, using the CLICK method[75]. The structural overlay clearly shows that the N-terminal domain of the pyocin monomer is homologous to Domain III of the T4 monomer. Within the overlapping region of 166 C_α residues, the two monomers have a sequence identity of 14%, a structural overlap of 40.8%, and an RMS deviation of 2.20 Å (Fig. 2.3C). This low sequence identity at the amino acid level, but significant similarity at the structural level of sheath monomers of distinct contractile tail systems is consistent with earlier comparisons of phage gp18 and T6SSs [61, 86]. Next, the program MODELER [103] was used along with knowledge of the full T4 sequence to generate homology models of the entire missing C-terminal domain and the N-terminal linker arm of the gp18. Superposing the homology models on the partial T4 structure (in the same orientation as in the previous structural overlay) gave rise to a near-complete atomic structure of the sheath monomer gp18 (Fig. 2.3D). The resultant molecule was then fitted into the cryo-EM reconstructions of the extended and contracted T4 tails using the program UCSF Chimera [79] to get the correct orientation of the monomer with respect to the complete phage in both conformations.

2.2.2 Building sheath sections

The all-atom structure of one "disc" of six T4 sheath monomers in both extended and contracted states was obtained by aligning the complete gp18 monomer in the same orientations as that of the incomplete gp18 monomers in the available single disc pdbs (PDB ids: 3FOH, 3FOI) [4]. Single disc structures of six monomers are also available for R2-pyocin sheaths (PDB ids: 3J9Q, 3J9R)[31]. All-atom structures of the complete sheaths were created by repeatedly applying symmetry transformations to the single disc structure using published helical parameters [57, 31]. The full T4 sheath is made of 23 discs and the R2-pyocin has 27 discs (Fig. 2.2). An input structure of 4 adjacent discs (Fig. 2.4C), about one-sixth of the entire sheath's length for both T4 and R2-pyocin, was used for all MD simulations and volume/surface area analysis. We chose 4 repeats of the disc because, (1) we believe this length is sufficient to deduce the equivalent bending and torsional stiffness properties of the entire sheath strand, and (2) the size of the system (about 239,200 atoms for T4 and 138,400 atoms for R2-pyocin) is small enough such that an all-atom MD simulation is feasible over the required nanosecond time-scale.

2.2.3 Theory and simulation protocol

Langevin dynamics simulations with a friction coefficient of 5 ps^{-1} were performed with the NAMD package[81] using the CHARMM 36 all-atom force field[14]. A generalized-Born implicit solvent model, as implemented in NAMD [96], was used to represent the solvent. The systems were minimized and gradually heated to ambient temperature over a 1ns MD run with harmonic constraints in place for the heavy atoms, followed by a short equilibration of 1ns during which the harmonic constraints were gradually removed. This was followed by an unconstrained long equilibration of 15ns, followed by a final 15ns production run.

From the MD trajectories, the center of mass positions for each of the twenty four monomer subunits (four subunits for each strand and a total of six such strands) were recorded as time series variables. For each strand, the pitch and radius of a best-fit helix parametrically given by $(r\cos t,\ r\sin t,\ ct\ ;\ t \in [0, 2\pi))$ and passing through the center of masses of the four subunits was determined. The curvature and torsion of each best-fit helix was then determined from the pitch and torsion using the Frenet-Serret formulas,

$$\gamma = \frac{r}{r^2 + c^2}\,,\ \tau = \frac{c}{r^2 + c^2} \tag{2.1}$$

where γ and τ are the curvature and torsion of the helix, r is radius of the helix, and $2\pi c$ is the helical pitch, i.e. the vertical separation of one helical turn. In this way, for each frame in the MD trajectory, we calculated the torsion and curvature for each of the six helical strands.

To get the stiffness constants, we measure equilibrium fluctuations in γ and τ from the 15 ns production run windows for all conformations (Fig. 2.5). The windows were chosen such that the torsion and curvature fluctuations are within an equilibrated mean. Given that the strain energy per unit length of the strand depends quadratically on torsion and curvature (see Eq. 2.3), we use the equipartition theorem of classical statistical mechanics [69] to arrive at the stiffness constants,

$$A = \frac{k_B T}{\langle (\gamma - \gamma_0)^2 \rangle}\,,\ C = \frac{k_B T}{\langle (\tau - \tau_0)^2 \rangle} \tag{2.2}$$

where A is bending stiffness, C is torsional stiffness, γ_0 is the mean curvature and τ_0 is the mean torsion. Angular brackets denote thermal averaging.

Volumes, mass densities and solvent accessible surface areas of the sheath monomers and the 4 disc sheath sections were calculated using the 3V volume assessor web server [102], with a grid resolution of 0.5 Å and a probe radius of 1.5 Å.

2.2.4 Dynamic model of the injection process

The stiffness constants calculated from MD simulations can be used as input variables to simulate the dynamics of entire injection machineries of phage T4 and R2-pyocin. The dynamic model is the same as described in the introductory Chapter 1 under the section *Elastic Rod Model of Biofilaments*. The bending and torsional stiffness constants of the ith strand derived as required as inputs to the model were derived from the MD simulations of the four-disc sheath strands described earlier. The stiffness constants for yet structurally unknown intermediate sheath conformations are those calculated by a linear interpolation of the extended and contracted conformations. To simulate the remaining components of the injection machinery, the capsid/DNA/neck/tail tube assembly for phage T4 and the collar/tail tube assembly for R2-pyocin are added to the sheath model in the upper boundary conditions of the sheath strands. The baseplate defines the lower boundary of the sheath strands for both phage T4 and R2-pyocin. Upon considering all six strands, Eqs. (1.3-1.6) yield a system 24 nonlinear partial differential equations. Solution of these equations under specified initial and boundary conditions yields solutions for 24 unknown field variables $\kappa^i(s,t)$, $\omega^i(s,t)$, $v^i(s,t)$, and $f^i(s,t)$, where i=(1,2,...6). The internal energy of the sheath driving the injection process is defined as the internal strain energy of the six interacting strands,

$$U(t) = \sum_{i=1}^{6} \int_0^1 \frac{1}{2}(\kappa^i - \kappa_0^i)^{tr} B^i (\kappa^i - \kappa_0^i) ds \tag{2.3}$$

where κ, κ_0 and B follow from eqn. (1.7).

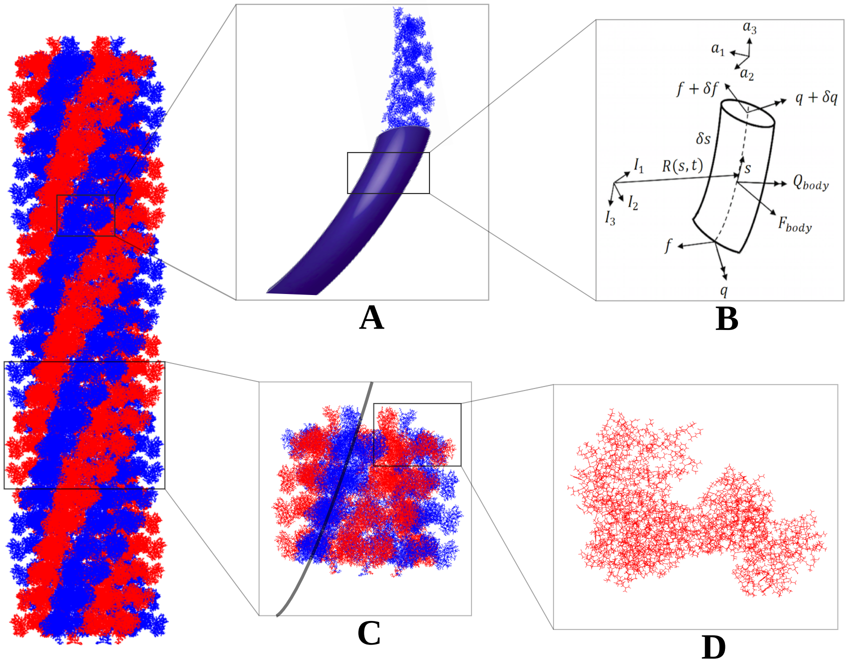

Figure 2.4: Various sections of the full tail sheath: (A) continuum rod with equivalent radius and elastic properties of an all-atom sheath strand, (B) free body diagram of an elastic rod with variables related by the nonlinear Kirchoff rod model, (C) a 4-disc section of the full sheath used in the actual MD simulations, with a best-fit helical strand passing through the centers of masses of monomers, (D) all-atom model of an individual sheath monomer.

2.3 Results and Discussion

2.3.1 Convergence of stiffness constants and helical torsion/curvature fluctuations

The stiffness constants of the helical strands in the continuum model (reported as an average over all six strands in (Table 2.1) are inversely proportional to the mean squared torsion and curvature fluctuations of a helical curve passing through the center of mass of the monomers. In 2.5, curvature and torsion of the best-fit helices averaged over all six strands are plotted for both the extended and contracted sheaths over the full 30 ns unconstrained MD run. It can be noticed that after some initial structural reorganization during the first 15 ns of equilibration runs (highlighted in yellow), the fluctuations are broadly about an equilibrated

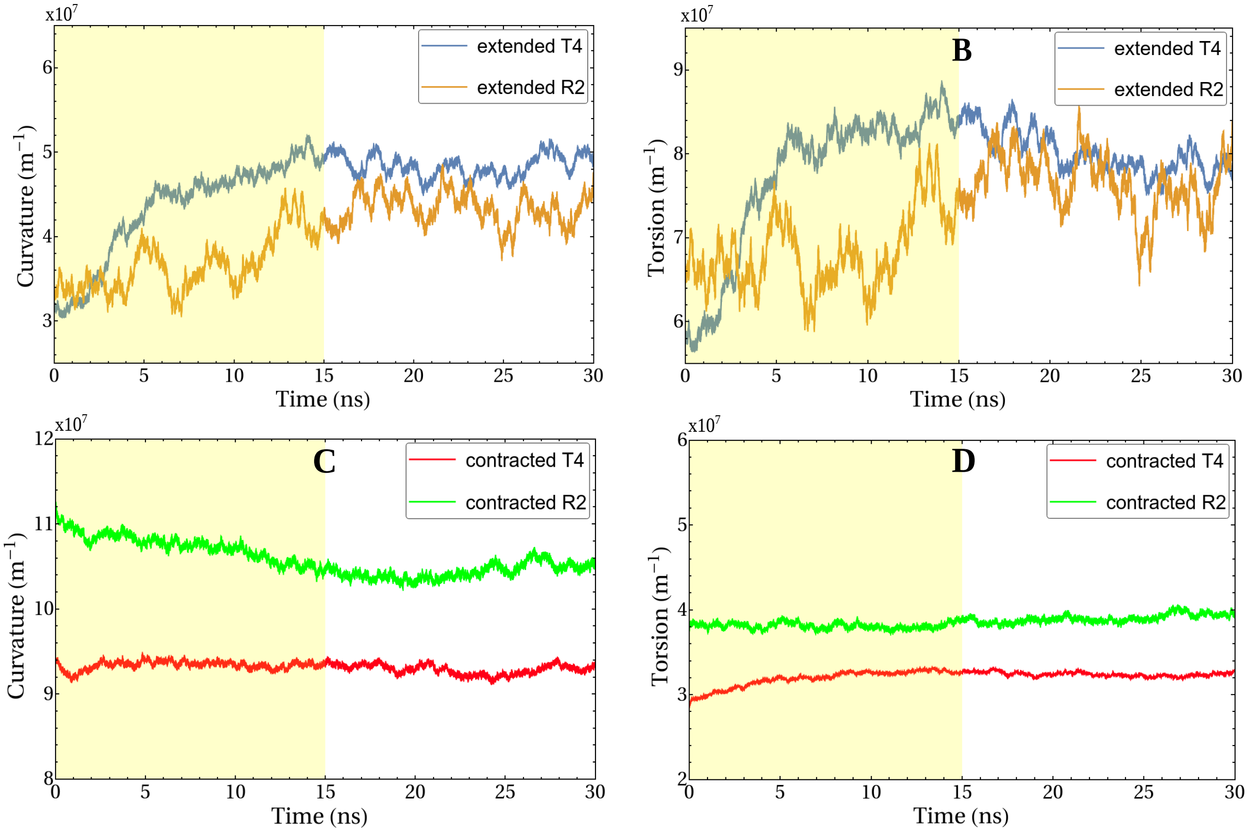

Figure 2.5: Variation of (A) mean curvature and (B) mean torsion for extended and contracted sheath strands of T4 and R2-pyocin during the full 30ns unconstrained MD run. The highlighted sections are for the initial 15ns equilibration runs and were not considered for the stiffness constant calculations.

mean for the 15 ns production runs. It is also apparent that structural reorganization during the initial equilibration are more pronounced for the extended sheaths, where the mean torsion and curvatures change by about 2 x 10^{7} m^{-1} from the initial unequilibrated structure, compared to negligible corresponding changes in the contracted sheaths. This is most likely due to small movements of the extended sheath monomers relative to each other, in the absence of the tail tube which acts as an inner boundary for the extended sheath. Boundary conditions due to the collar/tail tube assembly are incorporated explicitly in the dynamic continuum model as described previously in the methods section.Fig. 2.5 also shows that fluctuations about the mean of the extended sheaths strands have a larger amplitude overall than the contracted strands, which is consistent with earlier calculations which report the extended sheath to be in a higher energy meta-stable potential well and the contracted

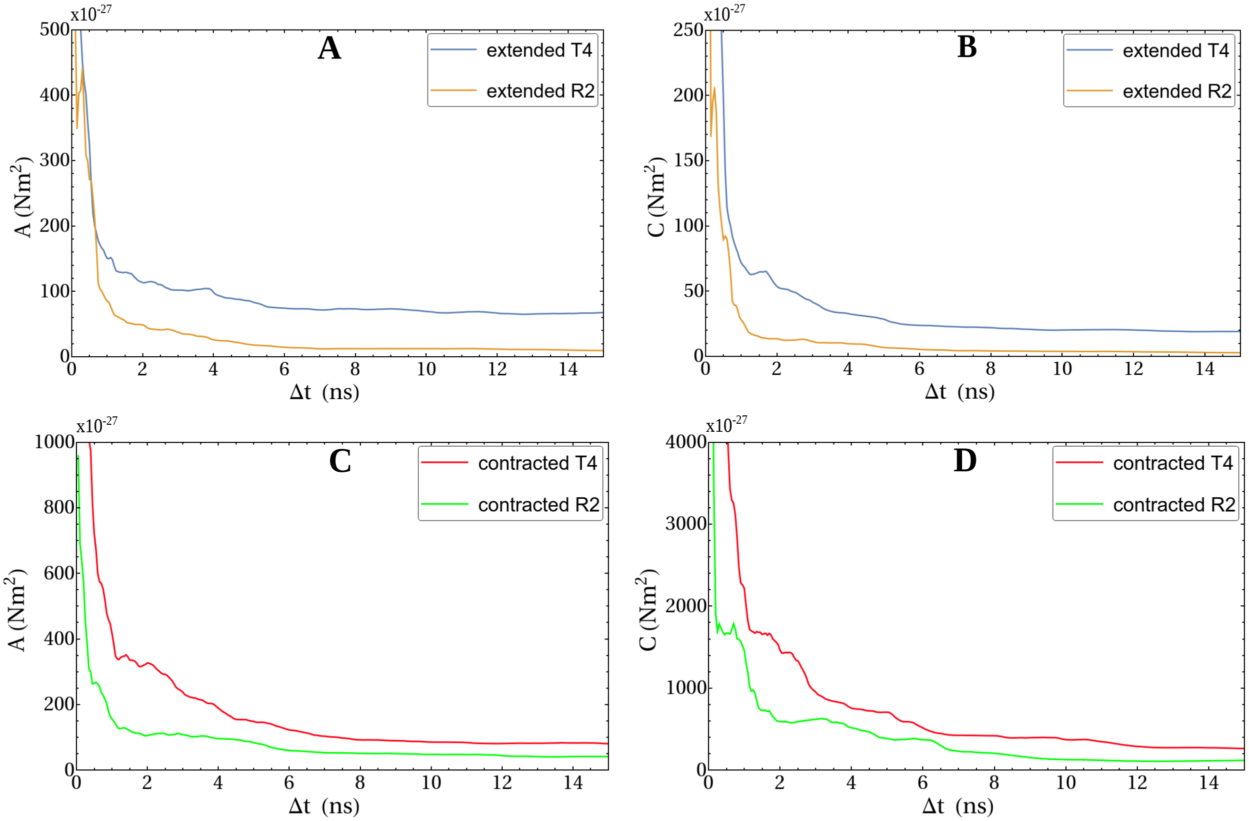

Figure 2.6: Bending and torsional stiffness constants for different sampling windows (Δt) obtained for extended and contracted sheath strands of phage T4 and R2-pyocin.

Table 2.1: Bending (A) and torsional (C) stiffness constants and their ratios for extended T4 strand, contracted T4 strand, extended R2 (pyocin), contracted R2, and actin filaments for comparison.

Strand	A (10^{-27}N-m^2)	C (10^{-27}N-m^2)	A/C (ratio)
Extended T4	67.7	18.9	3.58
Contracted T4	81.0	263	0.31
Extended R2	14.1	4.27	3.30
Contracted R2	41.3	119	0.34
Actin (exp)	73.0[33]	80.0[99]	0.91

sheath to be in a lower potential energy well. The difference in amplitude of fluctuations results in significant differences between stiffness constants of the extended and contracted conformations of the same sheath (T4 or R2-pyocin). In addition, although both T4 and R2-pyocin sheaths have very similar overall architecture, assembly and function, there are significant differences between their strand stiffness constants in the same conformation (extended or contracted). We report the resultant stiffness constants in Table 2.1, and explain the above differences by analyzing structural features of the all-atom sheath structures at the monomer level.

To verify if the calculated stiffness constants converge within the 15 ns sampling window, we plot the strand stiffness constants calculated for different sampling windows ranging from 50 ps to 15 ns (Fig. 2.6). It is clear from the plots that as the sampling window is increased, the stiffness constants become smaller and converge reasonably within the 15ns window. Similar plots reported earlier for MD simulations of actin filaments found that increasing the sampling windows led to stiffness constants which converged close to values which were

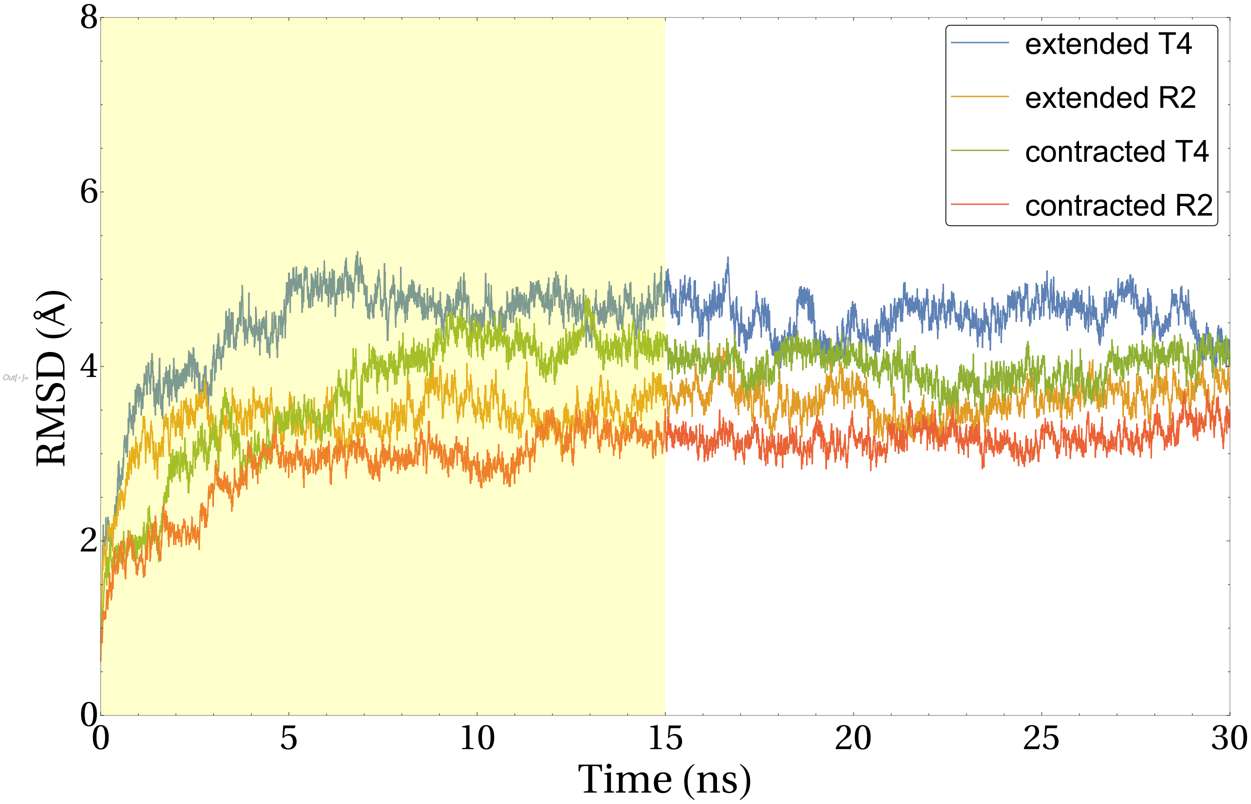

Figure 2.7: RMSDs relative to the unequilibrated structures for the core domains (Domain III for T4 and the N-terminal domain, see Fig. 2.5) for both extended and contracted conformations. The initial 15ns equilibration run is highlighted in yellow.

comparable to available experimental data [67, 93].

To probe convergence at the monomer level, we plot Root Mean Square Deviation (RMSD) curves of the backbone atoms of the monomer core domains for the 30ns MD run 2.7. The convergence of RMSD curves shows that the sheath structures are broadly equilibrated at the monomer level during the 15 ns production runs. Higher RMSDs of the extended monomers compared to the contracted monomers are due to the highly packed structures of the contracted sheaths compared to the extended sheaths.

2.3.2 T4 vs. R2-pyocin strand stiffness in the same conformation (extended or contracted)

There are differences in size, architecture and inter-subunit interactions of the T4 and R2-pyocin subunits within the sheath. The all-atom structure of the full T4 monomer model has 9979 atoms with an occupied volume of 91300 Å^3, and a surface area of 26694 Å^2. The all-atom structure of the R2-pyocin monomer has 5755 atoms with a volume of 54583 Å^3 and a surface area of 16734 Å^2. The mass densities of both T4 and R2-pyocin monomers are similar at about 1.24 x 10^3 kg m^{-3}, as would be expected for the density of folded proteins with a hydrophobic core, but the R2-pyocin monomer is about 40% smaller in volume than the T4 monomer (Table 2.2). Smaller monomer volume results in the R2-pyocin sheath being about 37% thinner in the extended state and about 39% thinner in the contracted state, across the tail cross-section, compared to the corresponding T4 sheaths (Fig. 2.2 top views). The helical twist and rise are almost the same for the extended T4 (17.2°, 40.6 Å) and the extended R2-pyocin (18.3°, 38.4 Å) strands, and also for the contracted T4 (32.9°, 16.4 Å) and the contracted R2-pyocin (33.1°, 16.2 Å) strands. As the strand architecture is essentially identical for both sheaths in the same conformation, the reduction in sheath thickness will also translate to a reduction in the thickness of a single strand. For two isolated cylindrical elastic rods of the same density, the rod with a smaller diameter will have lower mass per unit length, leading to higher thermal fluctuations and lower stiffness constants at the same temperature. However, the sheath strands of T4 and R2-pyocin are not isolated as there are significant inter-strand contacts with adjacent strands. These contacts will result in additional interactions and change the strand stiffness compared to isolated strands. Such inter-strand interactions are complex, and while their influence is present in the computed fluctuations, their contribution cannot be easily separated from that of the intra-strand contributions that are also present. Thus, the resulting stiffness parameters reflect both inter- and intra-strand contributions.

Table 2.2: Volume and solvent accessible surface areas (SASA) per monomer for the extended and contracted 4-disc sheaths, 4-monomer strands, and monomers for the extended and contracted conformations. R2 is short for the R2-pyocin sheath.

Quantity	Extended sheath — strand	Contracted sheath — strand	Isolated Monomer
Volume T4 (10^4 Å^3)	9.23 — 9.19	9.23 — 9.23	9.13
Volume R2 (10^4 Å^3)	5.53 — 5.51	5.62 — 5.55	5.46
SASA T4 (10^4 Å^2)	2.47 — 2.54	2.22 — 2.45	2.67
SASA R2 (10^4 Å^2)	1.24 — 1.41	1.09 — 1.40	1.67

Referring to Fig. 2.2, the increase in inter-strand contact area (equal to a decrease in solvent accessible surface area (SASA)) for a monomer in the extended T4 sheath is 0.07 x 10^4Å^2 from the isolated, extended T4 strand. This is only about 2.3% of the isolated monomer's surface area. For the extended R2-pyocin strand, the increase in contact area per monomer is 0.17 x 10^4 Å^2, which is about 10% of the isolated monomer surface area, so inter-strand interaction is significant and greater than that in T4 (Fig. 2.2C boxed section). From the stiffness constants in Table 2.1, the extended conformation of the T4 strand is about 4 times stiffer than the R2-pyocin strand for both bending and torsion. Greater stiffness of the extended T4 strand would be expected even for isolated strands, so inclusion of inter-strand interactions does not skew the relative stiffness constants for the extended strands.

For the contracted T4 strand, the increase in inter-strand contact area per monomer is 0.23 x 10^4Å^2, which is about 8.6% of the isolated monomer surface area; and for the contracted R2-pyocin strand, the increase per monomer is 0.31 Å^2 which is about 18.6% of the isolated monomer surface area (see Fig. 2.2B, 2.2D boxed sections). In the contracted conformation, the T4 strand is about twice as stiff as the R2-pyocin strand for both bending and torsion. Clearly, for the contracted sheaths, inter-strand interactions are much more significant and skew the relative values of the stiffness constants compared to isolated strands. Ge et al.[31] mention complex and spread out charged interactions in the contracted conformation of the R2-pyocin sheath compared to the extended conformation, which is probably also true for the corresponding T4 sheath conformations. This is corroborated by a closer view of the sheath strands. Boxed sections in Fig. 2.2 reveal that in the contracted conformations, inter-strand monomer contact is greater than intra-strand monomer contact, so inter-strand cooperativity is likely to affect strand stiffness to a much larger extent than that in the extended conformations.

2.3.3 Extended vs. contracted strand stiffness in the same sheath (T4 or R2-pyocin)

Within the experimental resolutions, the monomer structure for R2-pyocin changes minimally on contraction, with only the N-terminal extension arm consisting of the first 15 residues showing any significant conformational change[86]. It was reported that within the limits of 3.55 Å resolution, the partial gp18 monomer structure fits into the lower resolution cryo-EM density maps of both the extended and contracted T4 sheaths as a rigid body [4, 57, 52], indicating that, excluding possible small changes in domain-domain positions, the monomer structure is the same for both extended and contracted sheaths. That the sheath monomer moves essentially as a rigid body on contraction, with negligible change in its internal structure, has also been hypothesized in recent cryo-EM studies of the T6SS sheath at 3.3 Å resolution[86].

While the monomer subunits of the sheaths probably do not undergo any major structural change during the contraction process, significantly different strand architecture in both conformations results in the contracted strands being slightly thinner than the corresponding extended strands (see ref. [64] table 2, radius of rods). This is a direct result of the contracted strands having a longer helical arc-length than the extended strands for both T4 and R2-pyocin. In isolation, the thinner contracted strands will have lower stiffness constants than the extended strands. However, there is a significant increase in inter-strand interactions for the contracted sheaths compared to the extended sheaths, as the inter-strand contact area per monomer increases by 6.3% and 8.6% respectively, for T4 and R-2 pyocin, from the respective extended strands (Fig. 2.2 and Fig. 2.2B, 2.2D boxed sections). Increase in inter-strand interaction in this case leads to greater bending and torsional stiffness constants for the contracted sheaths than the extended sheaths, for both T4 and R2-pyocin. This significant increase in inter-strand contact area on contraction was mentioned previously in cryo-EM reconstructions of both T4 and R2-pyocin tail sheaths [4, 31].

2.3.4 Stiffness of T4 and R2-pyocin strands vs. actin filaments

For comparison with T4 and R2-pyocin sheath strands, stiffness constants of actin filaments measured experimentally [33, 99] are also tabulated in 2.1. The range between the stiffness constants of contracted and extended sheath strands includes the values for actin. The change in directional flexibility (bending vs. torsion) of all strands is reflected in the ratio A/C for the extended and contracted conformations (Table 2.1, final column). The ratio changes from 3.58 for the extended T4 strand to 0.31 for the contracted T4 strand, and from 3.30 to 0.34 for the corresponding conformations of R2-pyocin strands. These changes indicate that as the sheaths contract, the stiffness constants of the strands change from being bending-dominated to torsion-dominated. For actin filaments, the ratio is 0.91, which indicates slightly torsion-dominated stiffness constants. The reversal in elastic behavior of T4 and R2-pyocin strands is possibly related to changes in intra-strand structure on contraction, where monomers from neighboring strands are inserted into the intra-strand gaps created by monomer rotation. A detailed analysis into the reversal of this elastic behavior is beyond the scope of the present study and can be considered in future studies.

Finally, the bending and torsional stiffness constants for T4 are used as input to a coarse-grained dynamic model of the sheath starting from eqs. (1.3-1.6) and as developed in [64]. The resulting model estimates that the T4 injection process is powered by approximately 14500 kT of internal (strain) energy stored in the extended conformation of the sheath, which is in reasonable agreement with the experimentally estimated 10000 kT (5860 kcal/mol) enthalpy of heat-induced sheath contraction [7]. In addition, the maximum cell rupture force computed from eqs. (1.3-1.7) is 2260 pN, which is greater than the minimum cell rupture force of 103 pN estimated earlier by an analytical method [27]. Details of the dynamical model of phage T4 have been described in our previous works [63, 64]. For the capsid-less R2-pyocin, the model predicts about 6000kT of total strain energy, which is of the order of the analytically estimated free energy of sheath contraction, 3400 kT (2000 kcal/mol) calculated by Ge et al. using PISA [31, 54]. The maximum cell rupture force for R2-pyocin is estimated to be about 840 pN. Further details on the energetics and dynamics of the R2-pyocin injection process will be described in future works.

2.4 Conclusion

We derive the elastic bending and torsional stiffness constants of the helical strands of bacteriophage T4 and R2-pyocin tail sheaths from equilibrium MD simulations by considering a small section of the sheath. We further relate the stiffness constants to structural differences at the monomer level. The estimated stiffness constants suggest that the extended strand is more flexible for both bending and torsion than the contracted strand in both systems, primarily due to greater inter-strand interactions in the contracted conformation. Both extended and contracted R2-pyocin strands are more flexible than the corresponding T4 strands, primarily due to a reduction in strand thickness from T4 to R2-pyocin. The ratio of bending and torsional stiffness constants suggest that in the extended conformation, the strands are more flexible for torsion, but on contraction the flexibility is greater for bending, for both T4 and R2-pyocin. The ratio of bending and torsional stiffness constants for actin filaments is greater than that of the contracted strands, but lesser than that of the extended strands. Importantly, the calculated elastic constants can be passed to a continuum dynamic model to predict the energetics and timescales of the injection processes of phage T4 and R2-pyocin. The resulting dynamic model for phage T4 predicts that the injection process is driven by approximately 14500 kT of internal energy stored in the extended conformation of the sheath, and generates a maximum force of 2260 pN to pierce the cell membrane. A similar model for R2-pyocin reveals that the total elastic energy in the extended R2-pyocin sheath is 6000 kT and the maximum cell rupture force is 840 pN.

The modeling approach presented here can be expanded in the future to explore the dynamics and energetics of other contractile tail injection machineries having helically-stranded sheaths. Two other well-known structures are type VI secretion system and phage ϕ812. The bacterial type VI secretion system is a headless contractile injection system used to inject toxic effectors into competing bacterial and eukaryotic cells [9]. Bacteriophage ϕ812, from family *Myoviridae*, infects the *Staphylococcus aureus* cell by injecting its genomic DNA from the capsid into the host [76]. Despite many differences in biology, protein composition and targets among these injection machineries, they share the same essential feature: a contractile sheath composed of six helical strands that drives the injection process. As the all-atom structures of these contractile tail systems are resolved in future experiments, insights can be gained into their function by an immediate application of the modeling methods described in this chapter. Studying the structure, function, and dynamics of these nano-injection machineries has important bio-nanotechnological implications. Understanding contractile tail injection could provide examples for synthetic biology or bio-mimicry, which may further lead to the development of useful nano-technological devices with the ability to fight bacterial infections resistant to common antibiotics.

Chapter 3

Energetics, dissipation and contraction pathway of the phage T4 injection machinery

3.1 Introduction

Not much is known about the intermediate structures or dynamical details of the phage T4 injection process. On the experimental side, only one known microcalorimetric study in literature has probed the enthalpy of phage T4 contraction process, where contraction was induced *in vitro* by both heat and urea [7]. They estimated the free energy enthalpy of sheath contraction to be 3400 kcal/mol gp18 for urea-induced contraction and 6000 kcal/mol gp18 for heat-induced contraction. The authors considered the latter number to be more accurate. Moody in 1973 predicted a possible 'contraction wave' pathway for sheath contraction from observation of electron micrographs of partially contracted sheaths *in vitro* [72]. He proposed that the sheath contraction mechanism is likely *displacive*, with contraction initiating at the baseplate and propagating towards the neck (Fig. 3.1). Later, Casper [15] devised a mechanical model to visualize the sheath contraction mechanism predicted by Moody. The mechanical model employs temporary links between all subunits of the sheath and the tail tube that are released sequentially starting from the baseplate to the neck, following a conformational change of the baseplate. On the theoretical side, Falk and James used elasticity theory to describe an approximate coarse grained free energy surface for the extended to contracted conformational change of the T4 sheath, but the theory fails to account for any dynamics of the process [27].

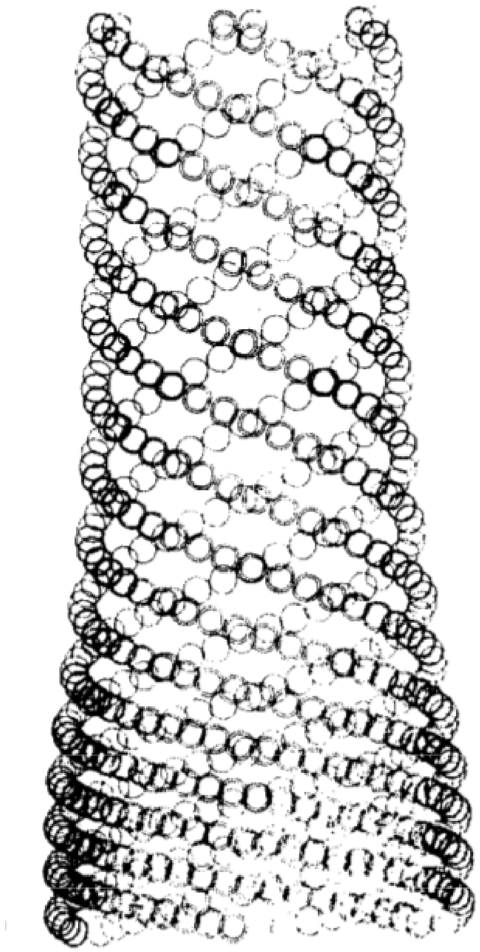

Figure 3.1: Arrangement of subunits in a partially contracted T4 sheath as proposed by Moody. Each circle represents a sheath monomer. Image adapted from Ref. [72].

Fluorescence imaging experiments on contractile tails of T6SS report that the timescale of sheath contraction is approximately 5 ms. The continuum model described in Chapter 1, while providing a good estimate of the sheath contraction energy, had a few crucial shortcomings. The first one being that it greatly underestimated the timescale of sheath contraction and found it to be $\sim 6\mu s$. The second shortcoming was the lack of observation of a 'contraction wave' as hypothesized by Moody [72]: The continuum model contraction intermediates implied that contraction proceeded uniformly through the sheath and not in a wave like manner. We concluded that underestimating of the timescale was due to ignoring the possible dynamic friction between the tail tube and the sheath as well as friction between the tail tube and the cell membrane during the injection process. These friction forces would dissipate energy and also increase the time scale of the injection process relative to the results reported in Chapter 1. We also concluded that the absence of a contraction wave was due to the assumption of the sheath bending and twist stiffness constants being homogeneous throughout the contraction process. To be consistent with the contraction wave mechanism, for a partially contracted sheath, one would expect the sheath stiffness to be non-homogeneous, with the region closest to the baseplate possessing stiffness constants close to those of the contracted conformation and the region closest to the neck possessing stiffness constants close to those of the extended conformation. The continuum rod model was thereby modified by our collaborator Dr. Amaneh Maghsoodi to include these two factors. I will subsequently discuss in this chapter sections relevant to my contribution to the above improved model.

3.2 Incorporating Mechanisms that Dissipate Energy During Injection

We identified four possible mechanisms that dissipate energy during contraction. These include 1) the hydrodynamic dissipation on the capsid and sheath from the surrounding extra-cellular environment, 2) the dissipation from the host cell membrane interacting with the tip of the tail tube, 3) dissipation due to internal friction of the sheath strands during the large conformational change, and 4) the hydrodynamic interactions between the flexible sheath and the tail tube during contraction. I will discuss the last two mechanisms in the larger context of overall dissipation during the injection process.

3.2.1 Internal Dissipation of Sheath Strands

During sheath contraction, the helical protein strands undergo a nonlinear conformational change from the extended state to the contracted state. Cryo-EM maps reveal that, during this change, the gp18 subunits within each strand rotate and translate in forming new contacts without significant change to their atomic structure [4]. The resulting motion produces internal dissipation due to forming new contacts.

Poirier and Marko examined the effect of dissipation due to the effect of internal conformational rearrangements of biofilaments [83]. Their model, based on a modified Langevin formulation of a thermally fluctuating elastic rod, describes the bending dynamics of a biofilament as,

$$B\frac{\delta^4\mathbf{u}(s,t)}{\delta s^4} + \eta\frac{\delta\mathbf{u}(s,t)}{\delta t} + \eta_b^{'} I\frac{\delta}{\delta t}\frac{\delta^4\mathbf{u}(s,t)}{\delta s^4} = \mathbf{n}(s,t) \tag{3.1}$$

in which $\mathbf{u}(s,t)$ denotes the fluctuating transverse displacement of the filament as a function of the countour length s and time t and $\mathbf{n}(s,t)$ is a random thermal excitation with delta function correlations, η is the external hydrodynamic drag coefficient (Stoke's regime) and $\eta_b^{'}$ is the internal dissipation coefficient (refer to equation (3) in reference [83] for more details of the formulation and subsequent derivations). This model forms the basis of understanding the relative importance of the external drag (dependence on η) and internal friction (dependence on $\eta_b^{'}$) for thermally fluctuating filaments. To calculate the internal friction coefficient, the expressions for the autocorrelation $R(t) = \langle \mathbf{u}(s)\mathbf{u}(s,t) \rangle$ and its relaxation time τ are deduced from 3.1 (via time-inverse Fourier transform of Eqn. (8) in[83]) as

$$R(T) = Rexp(\frac{-T}{\tau}) \tag{3.2}$$

$$\tau = \frac{\eta + \eta_b^{'} I q^4}{B q^4} \tag{3.3}$$

where B is the bending stiffness constant. In the large wavenumber ($q \to \infty$) limit, the relaxation time reduces to,

$$\tau \approx \frac{\eta_b^{'}}{B} \tag{3.4}$$

Building on Marko's above method, our collaborators at Michigan proposed a Timoshenko beam theory for thermally fluctuating biofilaments that captures shear deformation effects in addition to bending deformation effects to internal friction [65]. An equally easier and computationally faster alternative is to approximate the internal dissipation of sheath as an equivalent hydrodynamic drag from the surrounding fluid defined by a new drag coefficient $\eta^{'}$ which yields the same net energy dissipation as the internal dissipation. The relaxation time of the autocorrelation defined in (3.2) is described in terms of $\eta^{'}$ as,

$$\tau = \eta^{'} \frac{\kappa S + Bq^2}{\kappa S B q^4} \tag{3.5}$$

where κ denotes the shear correction factor for Timoshenko beam theory which is the product of the filament shear modulus G and cross sectional area A , and where $B = EI$ is the filament bending stiffness which is the product of the filament Youngs modulus E and area moment of inertia I.

In phage T4, thermal fluctuations in one ring of the tail sheath (having relatively large radius compared to the amplitude of the thermal fluctuations) mimic transverse thermal fluctuations of a (nominally straight) filament[39]. The radial displacement of the ring is analogous to the transverse displacement of the filament. Accordingly, for thermal fluctuations of one ring of the sheath, the autocorrelation of the radial displacement $u(t)$ and its relaxation time are given by eqs. 3.2 and 3.4. We utilize these results to analyze the thermal fluctuations of a ring of the tail sheath obtained from atomistic (MD) simulations. We employ a similar MD procedure described above to simulate a four-ring portion of the sheath, and then calculate the transverse fluctuations $u(t)$ for one of the middle rings as a representative (circular) filament for the T4 sheath structure; refer to Fig. 3.2(a). A middle ring is chosen to minimize solvent surface effects on the radial fluctuations. Transverse fluctuations $u(t)$ are defined as the radial fluctuations from the best-fit circle passing through the center of masses of the six subunits of the ring (Fig. 3.2(b)). For each frame of the trajectory, the center and radius $r(t)$ of the best-fit circle was calculated with the radial deflection $u(t)$ being the deviation $u(t) = r(t) - \langle r(t) \rangle$. The autocorrelation function $R(T) = \langle u(t), u(t+T) \rangle$ is then calculated from the trajectories for both the extended and the contracted conformations. Fitting a single exponential function (eqn. 3.2) for the extended and contracted rings yields the relaxation time and then the equivalent dissipation coefficient for each conformation as in (eqn. 3.4). The equivalent dissipation coefficient is estimated as $\eta^{'}= 0.003$ Pa-s in the extended state and $\eta^{'}= 0.008$ Pa-s in the contracted state. As a further illustration of this procedure, Figure 3.3 shows the autocorrelation of the transverse displacement of a middle ring of the sheath fragment in both the extended and contracted conformations using MD simulation. Superimposed on this data are the best-fit curves employing 3.2 from which we calculate the relaxation time τ.

Finally, the (Stokes) drag model is again employed to model the net internal dissipation as an equivalent (but now significantly increased) external hydrodynamic drag using the friction coefficient $\eta^{'} \approx 0.005$ Pa-s which is the average across the extended and contracted conformations.

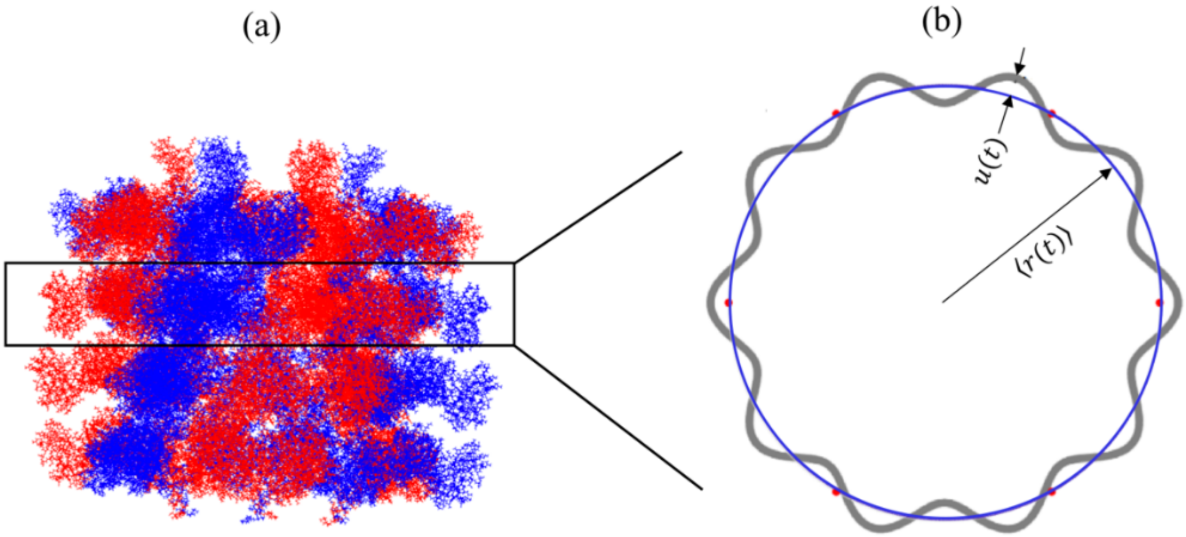

Figure 3.2: (a) Atomistic structure of the four-disc section of the T4 tail sheath with the middle ring surrounded by the superimposed rectangle. (b) The best-fit circle through the centers of masses (red dots) of the ring subunits of the middle ring has mean radius $r(t)$. The radial fluctuations of the filament from the circle are denoted by $u(t)$

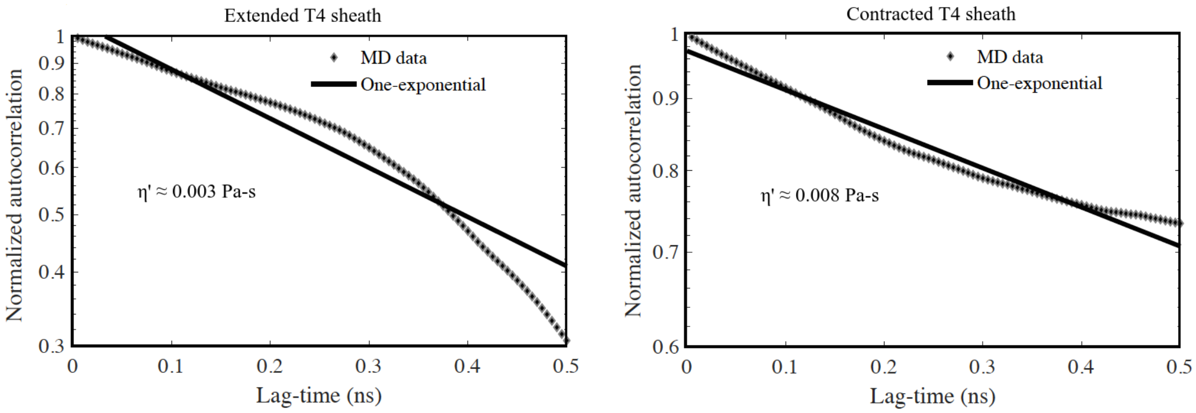

Figure 3.3: Autocorrelation of the transverse displacement of a middle ring of the sheath fragment from MD simulation for (a) the extended conformation and (b) the contracted conformation. The discrete data points indicate the MD-derived autocorrelation and the solid line represents the best fit per (20). Note log scale on vertical axis.

3.2.2 Sheath-Tail Tube Friction

Despite the wealth of information on the atomic structure of the sheath and the tail tube, possible interactions between them remain largely unknown. Potential sheath-tube interactions may arise from electrostatic and non-bounded forces and from viscosity in the nano-scale gap (interstitial water) between the tail tube and the surrounding sheath. Importantly, these interactions likely retard the injection process thereby affecting the injection time-scale.

To understand the sheath-tail tube interactions, I would begin by reporting the surface Coulomb potential distribution and the Kyte-Doolittle hydrophobicity for the tail tube, the extended sheath, and the contracted sheath. The atomic structures of the sheath are determined following the same procedure used to study the thermal fluctuations of the sheath; refer to the previous subsection titled *Internal Dissipation of Sheath Strands*. The atomic structure of the lower two rings of the tail tube are extracted from the cryo-EM structure of the entire T4 baseplate (pdb id 5IV5), and two additional rings were constructed using the published tail tube helical parameters[97]. The surfaces were generated using UCSF Chimera[79].

As illustrated in Fig. 3.4(a), the inner surface of the extended sheath is largely positively charged at the edges (blue) where it forms complementary charged interactions with the mostly negatively charged (red) outer tail tube surface. However, apart from these edges, the inner sheath surface is equally neutrally, positively and negatively charged. In the contracted conformation, the inner sheath edge is negatively charged, implying that during contraction there is a redistribution of charges along the inner sheath surface. Importantly, these charge distributions along the inner sheath surface for the both extended and contracted conformations and along the outer tube surface are relatively uniform. This implies that the net electrostatic and nonbonded forces remain largely perpendicular to the tube axis and thus contribute insignificant work as the tube translocates through the sheath. By contrast, significant work and interaction may arise during translocation from the viscosity of the interstitial nano-scale gap.

As illustrated in Fig. 3.4(b), the outer surface of the tail tube and the inner surface of the sheath are largely hydrophilic (blue), indicating that the nano-channel between the tube and sheath is essentially hydrophilic.

Previous experiments using Atomic Force Microscopy (AFM)[34, 66], where a layer of water was confined within two hydrophilic surfaces separated by a few nanometres, estimated that nanoconfined water has a viscosity which is about six to seven orders of magnitude higher than bulk water viscosity at room temperature (η_{bulk}=8.6x10^{-4} Pa-s). In the extended conformations of phage T4, the gap between the inner surface of the tail sheath and outer surface of the tube is negligible, but it increases to $\sim$ 10 Åin the contracted state. The size of the gap is such that water molecules from surrounding extracelluar environment can percolate inside the gaps. Accordingly, the viscosity of this interstitial nano-scale gap is expected to be far greater than that of bulk water. To simulate the friction (hydrodynamic drag) between the contracted portion of the sheath and the tube, we employ a classic model of fluid motion between two parallel surfaces with linearly varying velocity profile due to

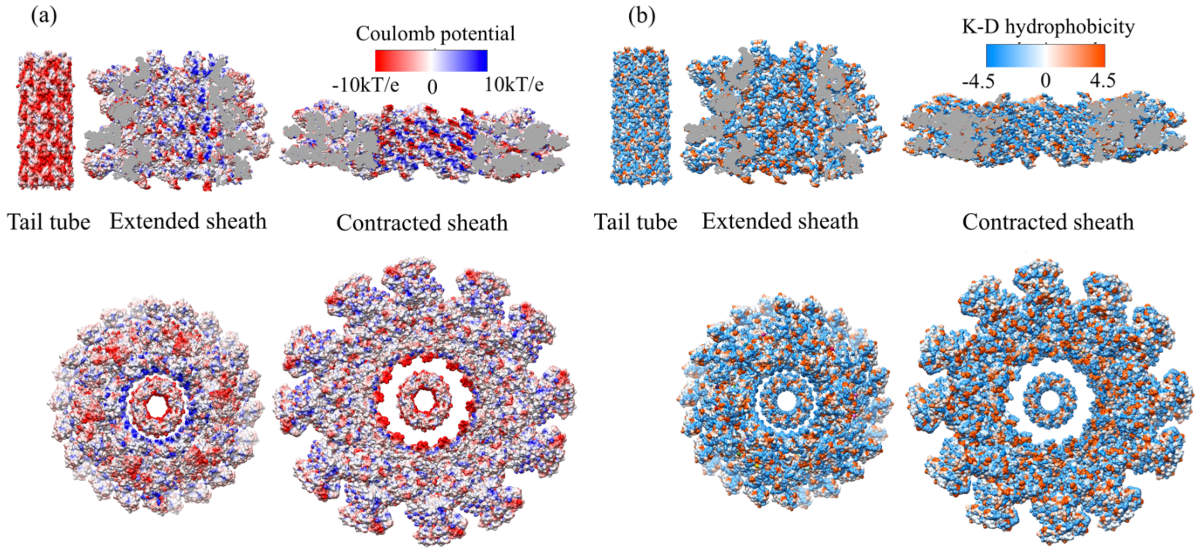
(a)
Coulomb potential
-10kT/e
0
10kT/e
Tail tube
Extended sheath
Contracted sheath
(b)
K-D hydrophobicity
-4.5
0
4.5
Tail tube
Extended sheath
Contracted sheath

Figure 3.4: (a) Surface Coulomb potential distribution of (clockwise) four rings of the outer surface of the tail tube, the inner surface of the extended sheath, and the inner surface of the contracted sheath. Beneath are top views of the tube within the extended sheath and within the contracted sheath. (b) Kyte-Doolittle hydrophobicity of the same surfaces arranged in the same order as in (a). The outer surface of the tail tube and the inner surface of the sheath are largely hydrophilic (blue).

shear both parallel and perpendicular to the translating and rotating tail tube; refer to Fig. 3.5. Since the tail tube attaches to the upper end of the sheath strands, the tail tube rotates and translates with the same linear $v(t)$ and angular $\omega(t)$ velocities of the upper end of the sheath, i.e., $v_t(t) = v(L,t)$ and $\omega_t(t) = \omega(L,t)$. The linear and angular velocities of the contracted portion of the sheath are almost zero. The resulting friction (viscous) forces and moments from the water molecules on the inner surface of the sheath and the outer surface of tube are given by,

$$F_{tube} = \eta_w \frac{v(L,t)}{d} A_t(t) \tag{3.6}$$

$$Q_{tube} = \eta_w \frac{r_t \omega(L,t)}{d} A_t(t) \tag{3.7}$$

where η_w is the effective water viscosity in the nano-scale gap, $d = 10\mathring{A}$, A_t is the wetted area of the tube interacting with the contracted portion of the sheath, r_t is the outer radius of the tube. The frictional force F_{tube} and moment Q_{tube} are respectively applied to the upper boundary condition of sheath strands in the continuum model.

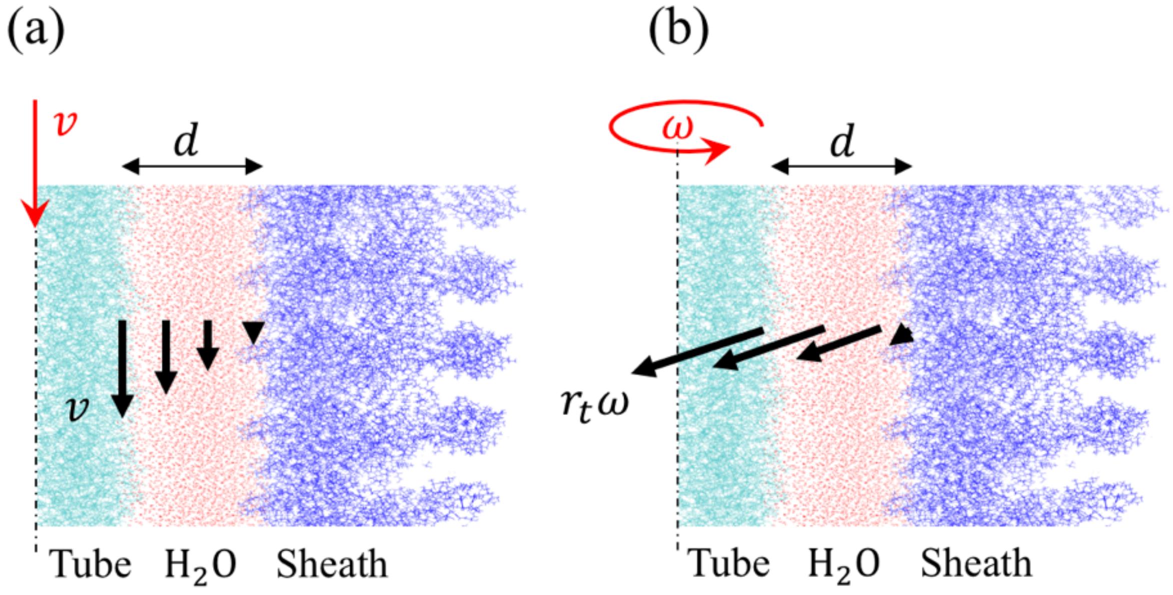

Figure 3.5: Components of the velocity profile of water within the nanoscale gap d between the sheath and the tail tube due to (a) translation v, and (b) rotation ω of the tail tube during injection.

3.3 Contribution of Dissipation Mechanisms to Injection Timescale

During injection, the elastic energy stored in the extended sheath is dissipated through the four energy dissipation mechanisms mentioned earlier in the chapter introduction. Of these mechanisms, the parameters describing the hydrodynamic dissipation on the capsid/sheath and the cell-tail tube dissipation are reasonably well understood and have been discussed in our recent collaborative work submitted for publication [Maghsoodi et al. 2019]. However, we must consider a wider range of model parameters for the remaining mechanisms that are lesser understood; namely, the internal dissipation of the sheath strands (IDSS) and the dissipation due to sheath-tail tube friction (STTF).

The friction coefficient $\eta^{'}$ governing the internal dissipation of the sheath strands is estimated to be 0.005 Pa-s from MD-derived thermal fluctuations (as discussed earlier), which is on the order of the viscosity of bulk water. However, the amplitude of the fluctuations in the MD simulations of the extended and contracted conformations are on atomistic length scales, whereas the sheath undergoes an overall conformational change that is several orders of magnitude larger. Consequently, the friction coefficient $\eta^{'}$ may actually be far greater and perhaps similar to that of thermally fluctuating actin filaments (1 Pa-s) which is about three orders of magnitude greater than the viscosity of bulk water[83]. Similarly, the sheath-tail tube friction coefficient η_w may vary considerably with the properties of the interstitial nano-scale gap between the sheath and tail tube. Experimental estimates for nano-confined water viscosity between hydrophilic surfaces ranges from bulk-like viscosity at a separation of 2-3 nm [85] to a drastic increase in viscosity below 1 nm[34, 66, 6]. There is also a lack on consensus on the variation of viscosity from the subnanometer range to the bulk, with some studies predicting an exponential decay to bulk viscosity [88], while others reporting an oscillatory decay from the subnanometer scale to the bulk [73].

The main question is what is the likely range of the time scale given the acknowledged uncertainties in the parameters and for the two dominant dissipation mechanisms described above? To address this, we first explore how the injection time scale varies over the broad ranges $0.001 < \eta_w < 1000$ Pa-s and $0.005 < \eta^{'} < 5$ Pa-s. Over these ranges, the computed injection time varies from the microsecond scale to a hundred milliseconds. In particular, for $\eta_w < 10$ Pa-s, the injection time remains approximately on the order of microseconds for all values of $\eta^{'}$. In contrast, for $\eta_w > 10$ Pa-s, the injection time rapidly grows to the millisecond time scale (see Fig. 3.6). In this range ($\eta_w > 10$ Pa-s), the dominant energy dissipation mechanism derives from sheath-tail tube friction (STTF) as discussed above.

Next, we focus on the parameter range that yields likely injection times by taking a cue from experimental results. While there are presently no experimental results that resolve the injection time for T4, there are for the Type VI secretion system (T6SS) which is a headless, contractile tail system used to inject toxic effectors into competing bacterial cells[9]. Recent experiments on T6SS report that the timescale of sheath contraction is 5 ms[10]. Despite obvious morphological differences between phage T4 and T6SS (head versus headless) and the injected material (DNA versus toxic effectors), their actual injection machineries possess remarkable similarities. Both injection machineries possess a long rigid tail tube surrounded by an elastic sheath that is formed by six interacting helical protein strands. Furthermore, both sheaths power the injection process by contracting from high-energy extended states to low-energy contracted states[9]. Given these significant structural similarities of the injection machineries, one might also anticipate similar injection times. Proceeding under that assumption, we identify the region within Fig. 3.6 that yields similar injection times; refer to region in red (defined by an injection time of 5 ms). If we assume that the timescales of sheath contraction for phage T4 and T6SS are approximately the same, then the corresponding value for η_w would be 60 Pa-s (see Fig. 3.6) which is approximately four orders of magnitude greater than the viscosity of bulk water (0.001 Pa-s), yet substantially below the extreme value (1000 Pa-s) for nano-scale layers with highly hydrophilic surfaces[34]. In this region, the internal energy dissipation exerts only weak influence on the injection time (see the red line in Fig. 3.6 which remains largely insensitive across the range $0.005 < \eta^{'} < 5$ Pa-s). Thus, the energy dissipation from the sheath-tail tube friction (STTF) emerges as the likely dominant dissipation mechanism controlling the injection time scale.

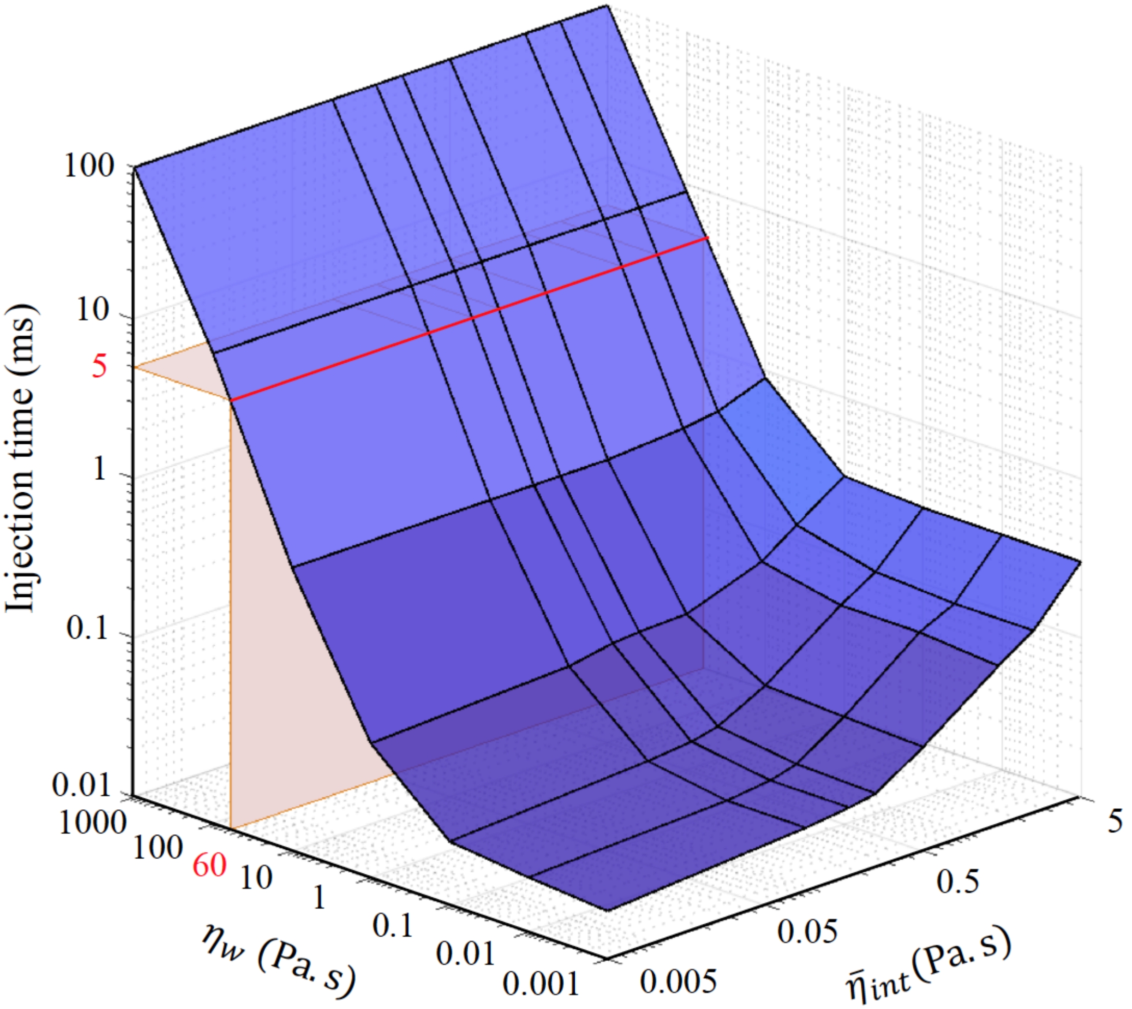

Figure 3.6: The timescale of the injection process as a function of the sheath-tube gap viscosity coefficient η_w over the range $0.001 < \eta_w < 1000$ Pa-s and the internal friction coefficient η' over the range $0.005 < \eta' < 5$ Pa-s. Note log scales. Red region defines a likely injection time scale for T4 (approximately 5 ms) by drawing a comparison with that of the Type VI secretion system (T6SS) [10].

3.4 Dynamic Pathways for the Injection Process

On incorporating appropriate values of dissipation coefficients such that they can replicate the overall expected timescale of contraction, the dynamic model of the phage T4 injection machinery reveals the conformational change of the sheath as a nonlinear contraction wave and the companion dynamics of the attached capsid/neck/tail tube assembly.

During sheath contraction, the sheath subunits translate and rotate during the large conformational change from the extended state to the contracted state. From the micrographs of partially contracted sheaths[72], Moody hypothesized that the sheath contraction is displacive whereby the subunits forming a single ring displace identically and in unison to produce a contraction wave that propagates upwards from the baseplate (adjacent to first ring) to the neck (adjacent to last ring); see Fig. 3.7a(I-III). In fact, Caspers mechanical contraption[15] utilized removable connections (links) between the mechanical elements (i.e., sheath subunits) and a central column (i.e., tail tube) to release each (elastically strained) ring of mechanical elements in upward succession 3.8. Our MD simulations described in the previous chapter also reveal that, due to changes in the sheath subunit contacts, the elastic stiffness constants of the sheath strands in the contracted state are larger than those in the extended state (Table 2.1). Consistent with Moodys contraction wave hypothesis and our MD-derived stiffness constants, we propose the following sheath contraction model. Prior to injection, the sheath remains in the extended conformation wherein interactions between the sheath and tail tube subunits retain the sheath in the high energy state. Sheath contraction is triggered by a large conformational change of the baseplate that breaks the interactions between the local sheath-tail tube subunits enabling displacive contraction starting at the first ring of sheath subunits. The sheath-tube subunit interactions are then broken sequentially upwards in each ring from the baseplate towards the neck enabling the sheath subunits to rotate and translate in forming new contacts and thus new local stiffness properties. As a consequence, an intermediate conformation of the sheath would consist of a partially contracted region extending upwards from the baseplate possessing larger elastic stiffness constants (similar to the fully contracted sheath) and a partially extended sheath (extending downwards from the neck) possessing smaller elastic stiffness constants. Hence, the resulting intermediate would possess non-homogenous stiffness constants at any point of time between the beginning and end of contraction. This non-homogeneity of stiffness constants has been incorporated in the continuum contraction model, details of which have been discussed in our recently submitted manuscript [Maghsoodi et al. 2019].

Model simulation captures this dynamic contraction wave as illustrated by the results of Fig. 3.7. In particular, snapshots of the simulated injection dynamics in Fig. 3.7(a) illustrate the fully extended sheath (IV), a partially contracted sheath (V), and the fully contracted sheath (VI) (We might want to add references to the atomistic movie snapshots here as well). These images from the model are consistent with experimental micrographs images for the extended (I), a partially contracted (II), and the fully contracted (III) sheath conformations reported by Moody [72]. The model simulations follow from the computed dynamic conformations of the six interacting helical gp18 protein strands that form the sheath. As the contraction wave passes, the helical strand angle decreases and the helical strand radius increases. The resulting wave front of the propagating contraction wave is best visualized in Fig. 3.7b that shows the helical strand angle as a function of location along the strand for the fully extended, a partially contracted, and the fully contracted sheath strands. Note that the helical strand angle is reduced by approximately 45 degrees during contraction and that this rotation begins at the baseplate and propagates upwards along each strand to the neck. This finding is consistent with the cryo-EM data that reveals that the sheath subunits rotate 45 degrees about the radial axis passing through the subunits[4] following contraction.

The continuum model also predicts how the elastic energy stored in the extended sheath is released as a function of time during the entire injection process 3.9.

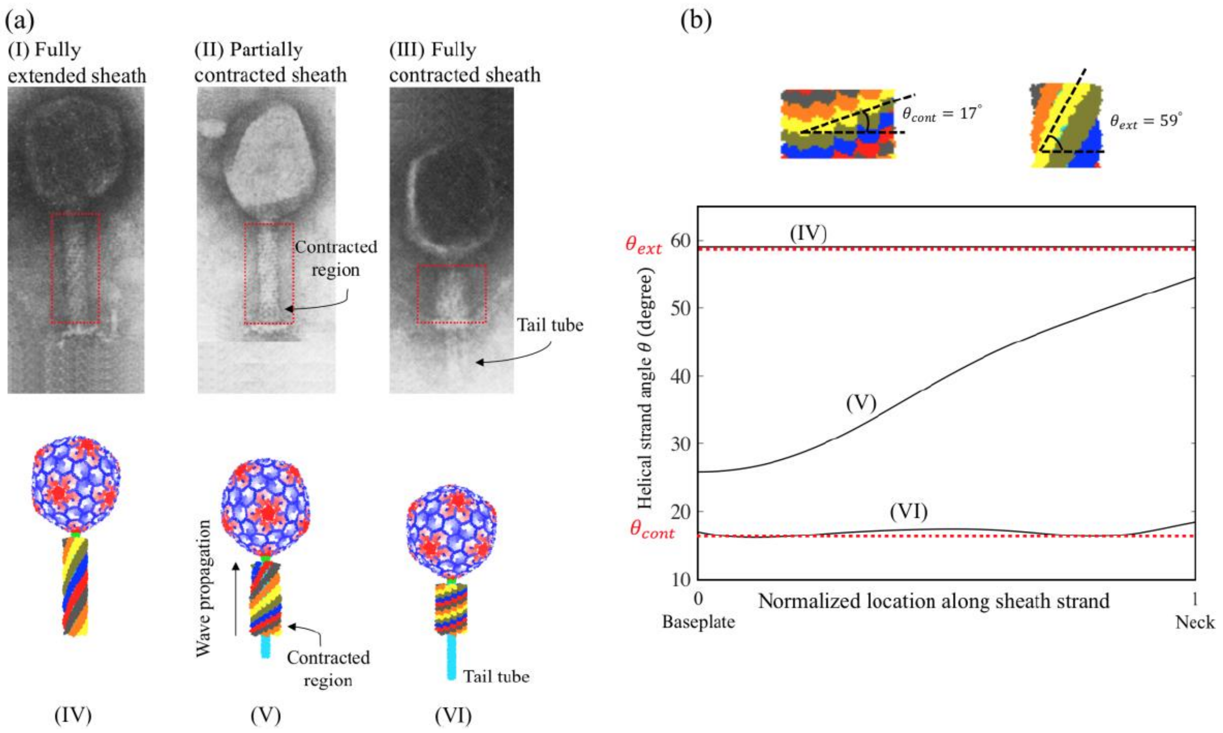

Figure 3.7: Simulation reveals contraction dynamics of the sheath from the fully extended conformation to the fully contracted conformation, consistent with the experimental micrographs reported by Moody[72]. (a) Micrographs of sheath in extended (I) , partially contracted (II), and fully contracted (III) conformations[72]. The snapshots illustrate model simulated extended (IV), partially contracted (V), and fully contracted (VI) conformations. Intermediate conformation (V) captures contraction wave propagation from the (lower) baseplate towards the (upper) neck. (b) The helical strand angle θ as a function of location along the strand. As the contraction wave passes, θ decreases from that of the extended conformation (IV) with $\theta_{ext} = 59^o$ to that of the contracted conformation (VI) with $\theta_{cont} = 16^o$.

3.5 Back-Calculating the Fully Atomic Pathways for Sheath Contraction

To illustrate the complete sheath contraction process at the atomistic level, we created a pathway of contraction using reported atomic structures of the T4 sheath, capsid, tail tube, and a section of the neck. The atomistic pathway was guided by the time series of continuum contraction dynamics of the helical sheath strands.

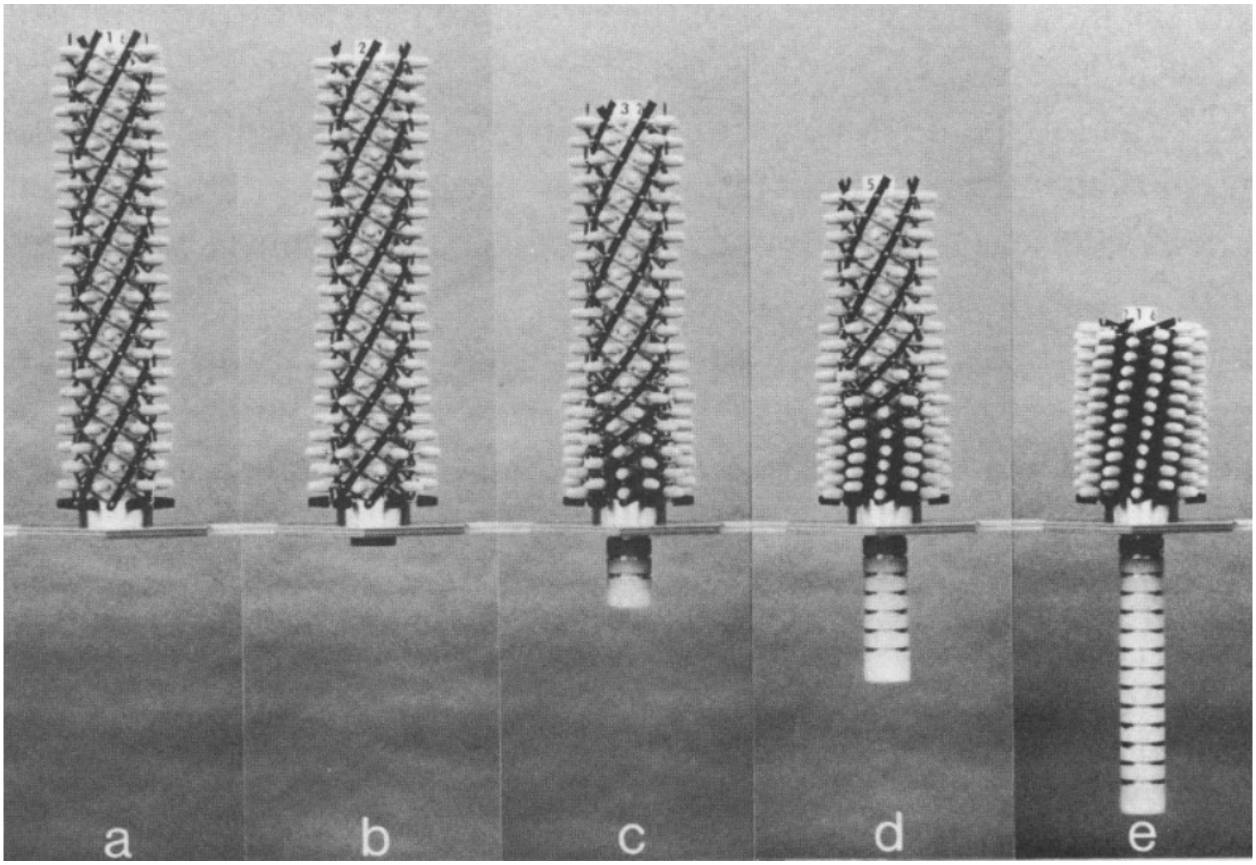

Figure 3.8: Steps in the contraction of the T4 tail sheath model as described in the mechanical contraption of Caspar (image reproduced from Fig. 13 of reference [15]). Note that the T4 sheath was believed to have 24 hexameric annuli when the work was published (more recent studies have concluded that the correct number is 23). The sheath subunits were represented by white knobs and were pinned to grooves in the central tailtube. The pins were sequentially removed starting from the baseplate to initiate contraction.

3.5.1 Sheath Contraction Intermediates from Continuum Data

From the continuum contraction trajectory, 200 equally spaced time frames were extracted to create corresponding time frames for the atomic trajectory. For each of the 200 time frames, radius, height and cylindrical angle data were calculated for 23 equidistant points along the arc length of a single helical sheath strand. Each of these points would correspond approximately to the center of mass of a disc monomer. To generate the intermediate sheath structures, we first created a database of 100 intermediate structures of the previously modeled single disc hexamer of the full gp18 molecule as it transitions from the extended

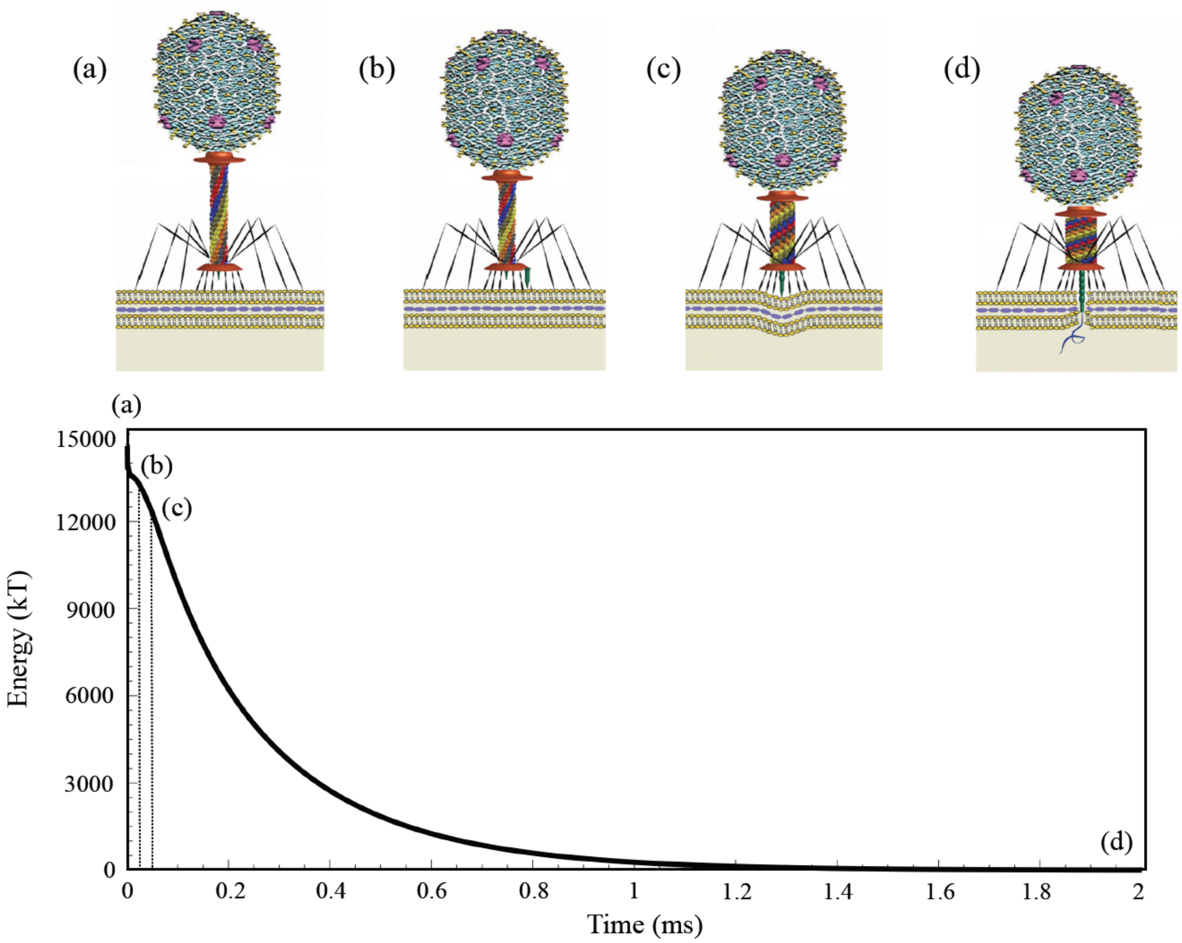

Figure 3.9: Complete model of the T4 injection machinery predicts the internal (elastic) energy of the contractile sheath that drives the injection process. (a) The sheath begins in the (high-energy) extended state where the tip of the tail tube remains 100 Åfrom the cell membrane. (b) The sheath contracts 100 Åso that the tip of the tail tube touches the outer cell membrane. (c) Further contraction of the sheath produces, in sequence, cell indentation, rupture of the outer cell membrane (after 60 Åindentation), and penetration into the (viscous) periplasmic space. (d) The fully contracted sheath with zero elastic energy.

to the contracted conformation. The intermediate disc structures were created by linearly interpolating the atomistic cartesian coordinates of the six sheath monomers between the extended and contracted states. For each time frame, we create a stack of 23 sheath discs, resulting in the full sheath contraction intermediate. Each disc in the stack is chosen from the database of the intermediate structures by selecting a structure whose radius is closest to the radius of the corresponding point in the continuum contraction data. Once the correct 23 disc intermediates are chosen, they are aligned together to create the full sheath using helical translation and rotation parameters from the continuum data points.

3.5.2 Appending Tail-Tube, Capsid and Neck to Sheath Intermediates

The atomic structure of a single ring hexamer of the tail tube was extracted from the cryo-EM structure of the entire T4 baseplate (pdb id 5IV5) [97], and the full structure of the tail tube was constructed by repeatedly applying the accompanying tail tube helical parameters to generate the required 23 copies of the tail tube (Fig. 3.10A) [109]. Twenty five such rings were required to cover the entire inner surface of the T4 sheath . The tail tube was capped at the top by the tail terminator protein gp15 hexamer, it structure being extracted from the gp15-gp18 hexamer complex (pdb id 3J2M) [29], while preserving its orientation relative to the top disc of the sheath. The atomic structure of the full icosahedral T4 capsid was obtained from the protein data bank (pdb id 5FV3) [18]. To the 200 sheath contraction intermediates, atomic structures of the tail tube, capsid and a section of the neck (gp15 hexamer) were appended to create contraction intermediates of the entire phage minus the baseplate, tail fibres and missing sections of the neck. The exact relative orientation of the capsid relative to the gp15 hexamer and the rest of the tail is unknown, so an appropriate orientation was chosen by visual inspection using the VMD program [44]. Within known cryo-EM resolutions, no rearrangement of subunits of the capsid, tail tube and the gp15 neck has been observed on contraction, and the orientation of the gp15 hexamer relative to the top of the sheath remains preserved [29]. Therefore, in the contraction intermediates, the translation and rotation of the capsid-tail tube-gp15 hexamer complex is identical to the top disc of the sheath. The contraction intermediates were combined to create a single trajectory, and a movie of the contraction process was created as a result. To enable efficient rendering of the multi-million atom trajectory, the capsid is represented by only its C_α atoms in the movie. All other components of the phage are shown in all-atom representations.

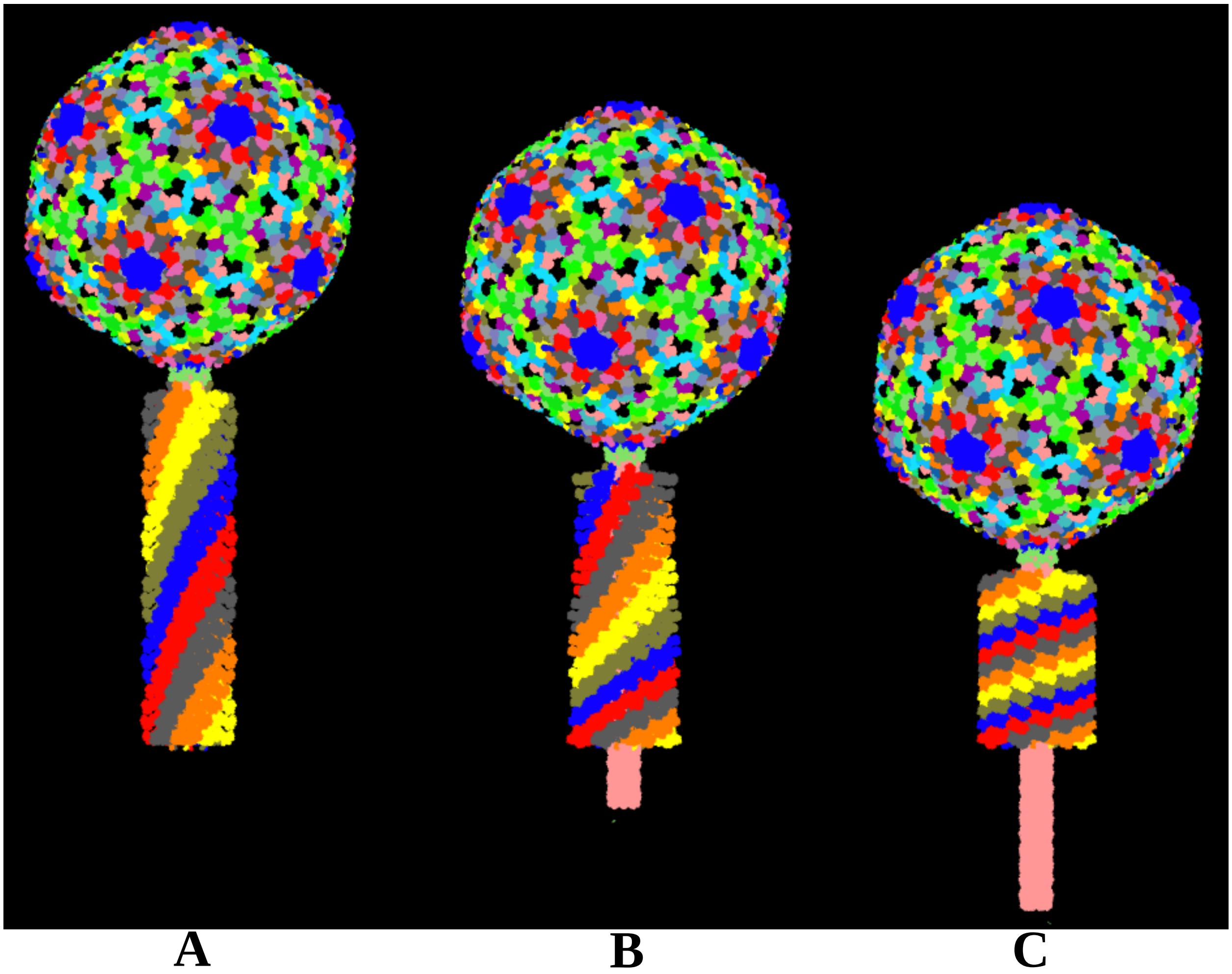

Figure 3.10: Snapshots from the Back Calculated Atomic Pathway of T4 Contraction of (A) the initial extended tail, (B) a contraction structure and (C) the final contracted tail.

Chapter 4

DNA Homopolymer Unzipping Under Constant Force leads to Crackling Noise Type Avalanches

4.1 Introduction

Crackling noise is a characteristic of varied systems with multiple degrees of freedom whose response to being slowly driven by an external force is through avalanche-like events which are power-law distributed in duration and sizes [92, 91, 22, 17]. Examples of such systems span a wide range of fundamentally unrelated systems including Barkhausen noise in ferromagnets [92], noise emitted from paper crumpling [43], crack propagation in materials [13], earthquake magnitudes over time [19], light flux in stars[94], fluctuations in financial market prices[62] and recently neuronal activity in the brain [84].

Most systems respond to external forces by moving between two stable states via multiple metastable states. This movement usually happens in two major ways: either with many small similar-sized random events (like the *popping* of popcorn when heated), or with one single event (like the *snapping* of chalk when broken). When systems *crackle*, their response to external forces is between these two extremes: the interaction between the system's subunits is stronger than that in popcorn but weaker than that in chalk. In crackling, the transition events span multiple size and time scales. For example, a recent study on neuronal avalanches in zebrafish larvae found that the avalanche sizes span up to six decades of power law scaling (Fig. 4.1).

The most well known example of crackling noise (and also the one studied in most detail) is Barkhausen noise of ferromagnetic materials, where on an application of a slowly varying external magnetic field, the magnetic material's magnetization increases in a series of jerky steps. Zooming in on a seemingly smooth curve of magnetization or flux density as a function of magnetic field reveals that the graph is in fact not microscopically smooth: it consists of multiple small bursts or avalanches (Fig. 4.2). The avalanches are of different sizes, contrary to what would be expected if the microscopic domains flipped continuously one after the other. A natural question pops up: What is the distribution of the Barkhausen avalanche sizes? Bohn et al. studied the Barkhausen noise in polycrystalline and amorphous ferromagnetic films with varied thicknesses, and studied the distribution of avalanche sizes and times [12]. They found signatures of critical dynamics in the distributions: avalanches were power-law distributed with scale-invariant properties and universal features characteristic to different unrelated systems.

Physical systems which crackle and show the same power law exponents for crackling noise belong to the same universality class [92]. It is widely believed (but not proven) that all systems belonging to the same universality class can be dynamically described close to a phase transition using the same simple models (like the 2-D Ising model, or percolation model [98],

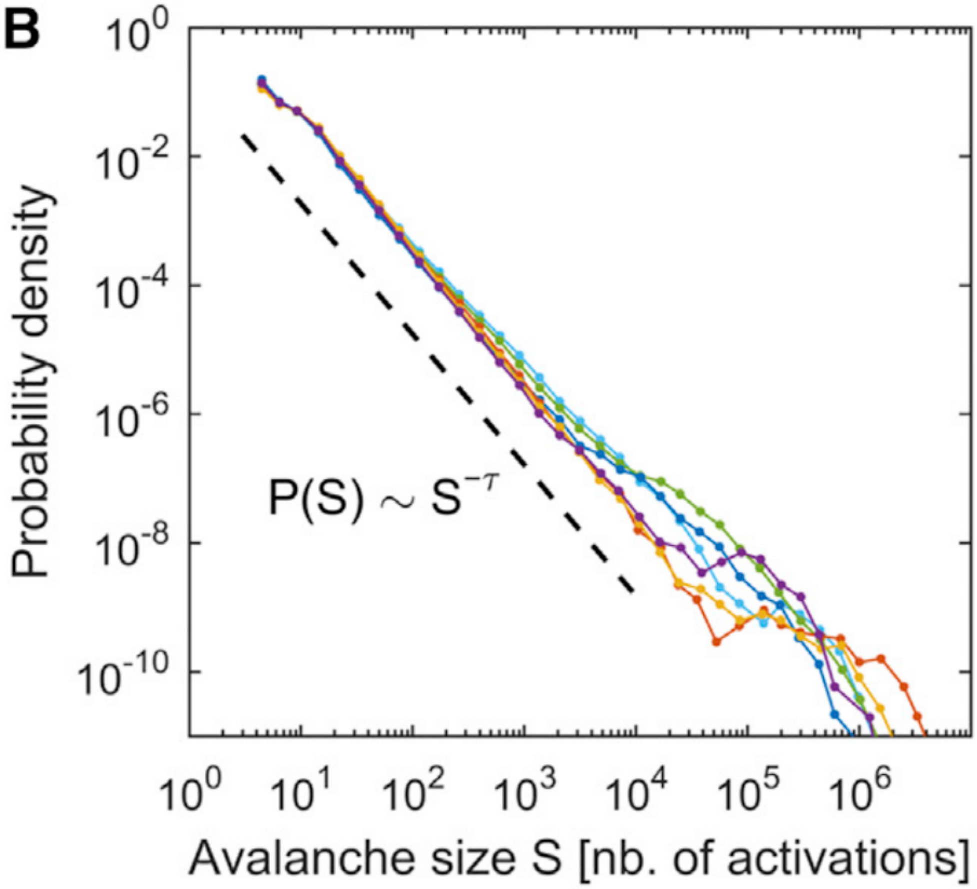

Figure 4.1: Distribution of neuronal avalanche sizes (S) in zebrafish larvae. Image from Fig. 2B of Ponce-Alvarez et al. [84]

irrespective of their microscopic complexity. It is therefore of interest to discover new systems which potentially display crackling noise, and if they do, then find out if they belong to an existing universality class. For example, in the case of Barkhausen noise in ferromagnetic materials, experimental exponents showed that there are two distinct universality classes of polycrstalline and amorphous magnetic materials respectively [25, 12].

In biological systems, crackling noise has only been observed in neuronal avalanches [108, 84]. However, there are multiple other driven systems known in biology: For example, it is well known that protein unfolding can be accelerated when it is driven by an external force at its ends [45]. Jose and Andricioaei, in an elegant study relating two otherwise microscopically unrelated systems, showed that force and temperature induced protein unfolding shows

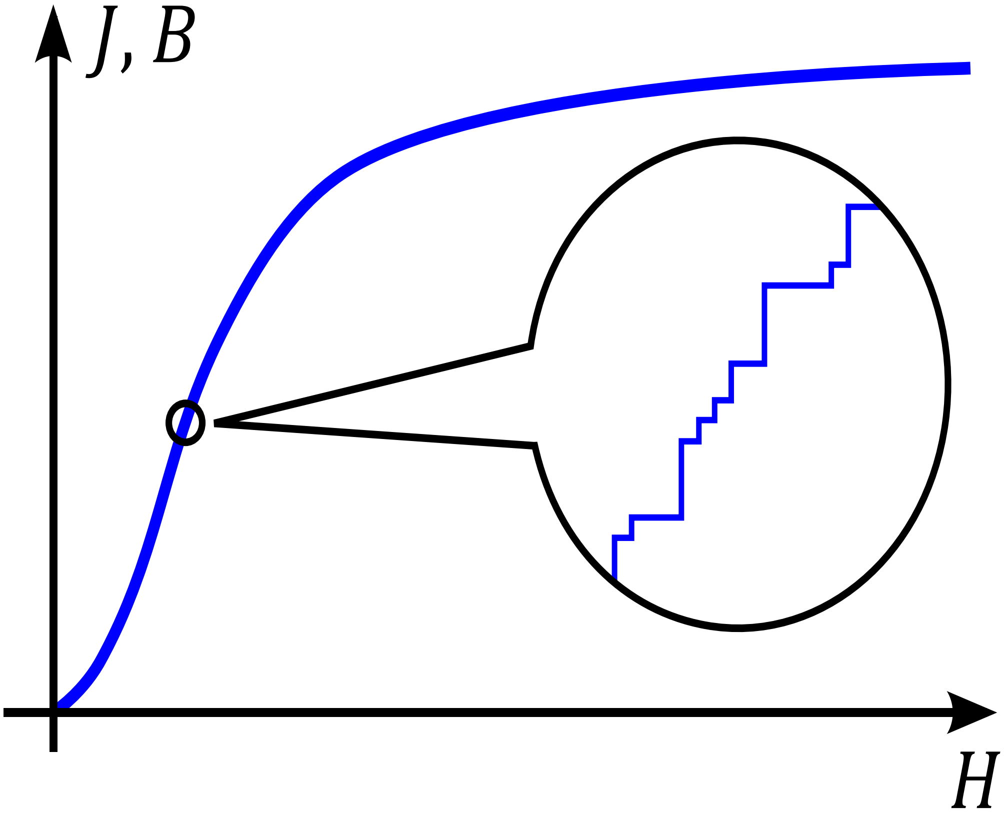

Figure 4.2: Magnetization (J) or flux density (B) curve as a function of magnetic field intensity (H) in ferromagnetic material. The inset shows Barkhausen jumps. Image from [106].

universal signatures similar to granular jamming [47]. However, no evidence of crackling dynamics in force or temperature induced protein unfolding has been observed. The unzipping of dsDNA to ssDNA at room temperature by an external force is another system where crackling can be potentially observed. Intuitively, at extremely large forces, where only the largest barriers survive, we would expect the unzipping to occur almost instantly, like the snapping of chalk, but the dynamics of unzipping at small or intermediate forces is not obvious. At room temperature, double-stranded DNA (dsDNA) is thermally more stable than single-stranded DNA (ssDNA) and single molecule experiments of DNA unzipping under constant force *in-vitro* show that the unzipping fork on DNA moves only when the unzipping force is greater than a critical force $F_c \sim 15$pN for AT base pairs and $F_c \sim 20$ pN for GC base pairs [23]. It was also shown in the experiment of Danilowicz et al. (and further by Weeks et al. [104]) that at near the critical unzipping force, the position of the unzipping fork along the sequence was highly irregular, consisting of long pauses separated by rapid bursts where multiple basepairs unzip in a relatively short timescale. The pause points locations were shown to be mostly sequence dependent at the length ($\sim$ 50,000 bp DNA), timescale (minutes-hours) and spatial resolutions ($\sim$400 bp) of their experiments. However, the authors were unable to explain the large differences in pause times at the same location between various unzipping trajectories of the same sequence: even though the pause locations were reproducible between their trials, the pause times was not. Also, the forces considered were close to the critical forces, where the energy barrier between AT and GC basepairs was significant to cause sequence dependence. Various theoretical models have been proposed which probe the statistical mechanics of DNA unzipping near the critical force [74]. Cocco et al. in an attempt to explain results from single molecule experiments, reported a semimicroscopic theory describing the force and kinetic barriers to DNA unzipping under a constant external force [20, 21], but their theory addresses only the initial unzipping barrier-crossing event and does not study the kinetics of how the unzipping fork moves down the dsDNA. Although most unzipping theories to-date do reveal the fundamental statistical mechanical picture of unzipping, they are not detailed enough to account for subtle energy barriers which could be present on account of the complicated 3-D structure of nucleic acids. MD simulations, which can resolve biomolecular dynamics to the atomic level, can then come as a handy tool to explore such dynamics.

MD simulations have been extensively used as an effective tool to provide basepair-level dynamical details of nucleic acids [16, 95]. However, very few MD simulation studies have probed the micromechanics of force-induced DNA unzipping. Mentes et al. probed the force and kinetic barriers to unzipping the first few DNA basepairs and revelaed that forces much higher than the experimentally calculated critical forces are necessary to unzip the first few terminal basepairs in a DNA oligomer [70]. Santosh amd Maiti reported the first ever MD simulation of a force induced unzipping of a DNA strand where the movement of the unzipping fork along the strand was studies by slowly varying the external unzipping forces in the range of 100-200 pN, but beyond reporting the force-temperature dependence of unzipping initiation, stopped short of a detailed explanation of the pauses and jumps as the unzipping fork moved along the strand [87]. Volkov et al. studied the micromechanics of base pair unzipping of a DNA dodecamer using constant velocity unzipping via steered MD simulations[101]. While constant velocity DNA unzipping leads to some useful understanding of the process, but as the external force adjusts to compensate for any force barriers during unzipping, the distribution of unzipping pauses at a constant external force cannot be studied using steered MD.

In light of the above discussion, we aimed to study, using multiple MD simulations of short DNA strands, the dynamics of DNA unzipping at intermediate external forces where crackling dynamics could possibly be observed. Using an approximation for the unzipping of a large DNA homopolymer using multiple unzipping trajectories of a short DNA homopolymer sequence, we observed that at forces close to double that of the experimentally reported critical force of unzipping, avalanches in the DNA unzipping velocities showed signatures of crackling dynamics. Using analyses similar to those used earlier to observe Barkhausen noises in ferromagnets [12] and crack propagation in materials [13], we calculated avalanche critical coefficients which were found to be in reasonable agreement with those calculated using functional renormalization group theories of Barkhausen noise [25]. In the concluding section, we discuss implications and possibilities of future studies building on the scheme developed in this work.

4.2 Methods

4.2.1 Initial structure data and simulation protocol

The initial structure of the DNA 18-mer homopolymer with 18 repeats of the AT basepair was obtained from the make-na server [30]. A homopolymer was selected for this analysis to avoid any possibility of sequence dependence of the noise in the force unzipping results, as sequence dependence was not a focus of this study. The strand was minimized, slowly heated with constraints on the backbone, followed by a short equilibration run with constraints slowly released, followed by a 1 ns unconstrained equilibration run. The simulations were done using the CHARMM software and the CHARMM36 force field for nucleic acids [14]. As it has been shown previously that the unzipping of the terminal DNA strand has large force barriers to commence unzipping [70, 20], a large initial force of $\vec{F} = 300$ pN was first applied on backbone O3$'$ atoms of top base pairs in a fixed opposite direction (y direction) and a short (otherwise unconstrained) MD run was performed which was sufficient to completely unzip the first two basepairs. This partially unzipped structure was used as the initial structure for further unzipping analysis (Fig. 4.3).

4.2.2 Gathering unzipping velocity from trajectories

Our aim in this study was to analyze the statistics of jumps during the unzipping of a long homogeneous DNA sequence at a constant external force to verify our hypothesis of the crackling of DNA unzipping. We chose to run multiple unzipping simulations of a short DNA 18-mer, and combine the unzipping data of each individual simulation to approximately mimic the unzipping profile of a single long DNA strand. To illustrate, if D is the overall strand separation of a (hypothetical) long single DNA sequence at any point of time, and

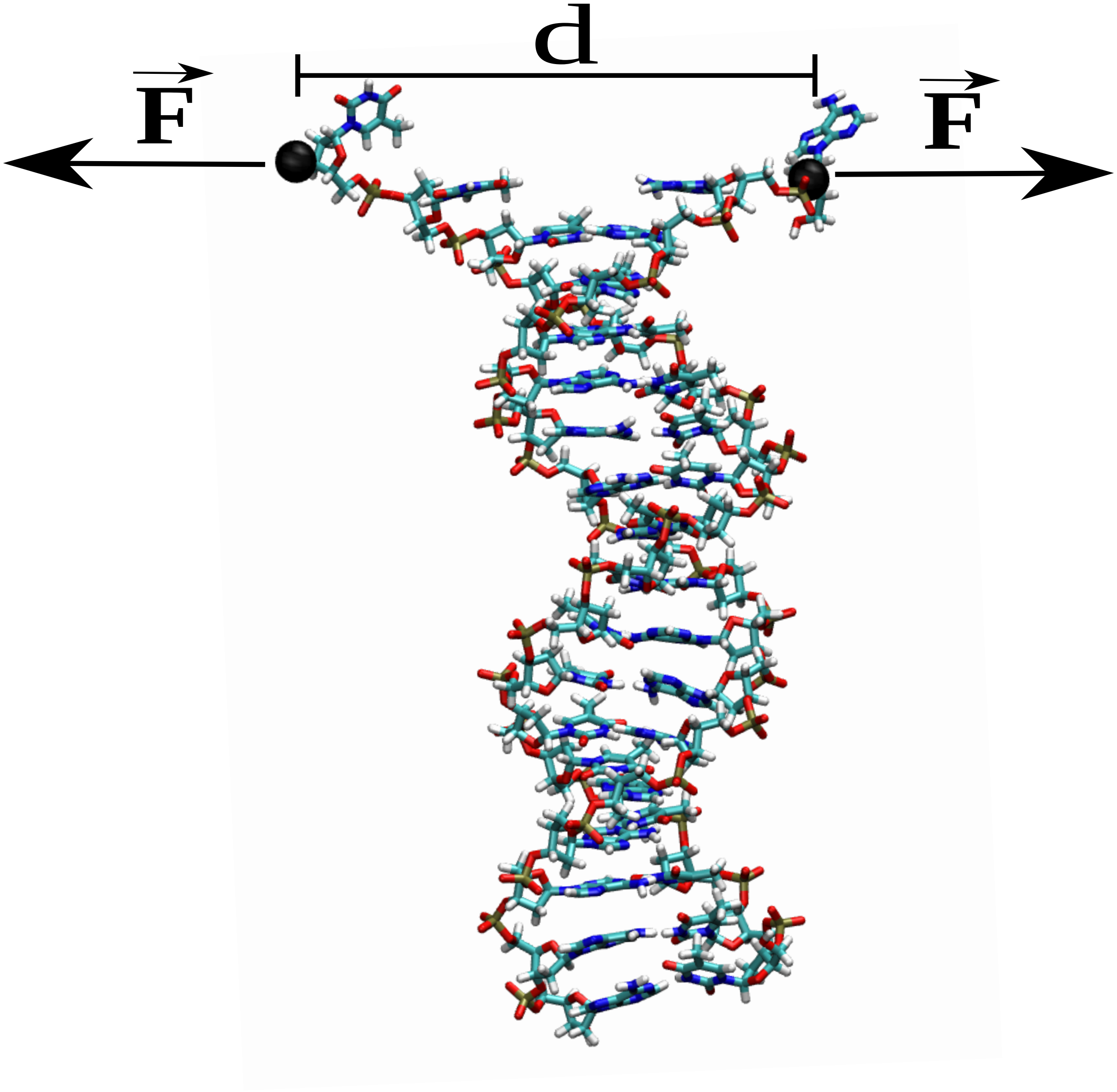

Figure 4.3: The DNA 18-mer homopolymer illustrating the constant applied force for unzipping $\vec{F}$ and the distance d used to identify the strand separation.

d_i is the strand separation of the ith run (defined as the distance between the O3' atoms of the top basepair as in Fig. 4.3), then D is defined as,

$$D = d_i + \sum_{k=1}^{i-1} d_k \tag{4.1}$$

where d_k is the final strand separation at the end of the kth simulation. A total of 800 3 ns long simulations were run for each 18-mer, resulting in a total of $\sim$2.4μs data gathered for a force value of 50 pN. The force value of $\vec{F} = 50$ pN was chosen (after analyzing multiple trajectories with unzipping forces varying from 30pN to 70pN) such that they are large enough to unzip sufficient base-pairs within the simulation time, and small enough to avoid complete unzipping of the 18-mer into two single strands. The simulation data was analyzed and the very few trajectories which were completely or almost completely unzipped were not considered for further analysis. A Weiner filter[1] with a box size of 5 data points was applied to the distance data to remove the thermal noise contribution to unzipping (Fig. 4.4). Noise removal was essential to ensure that the derivative data collected later would include only the contribution due to basepair unzipping and not thermal fluctuations. First derivative of the de-noised distance data D gave the velocity of unzipping time series for the long unzipping trajectories (Fig. 4.5).

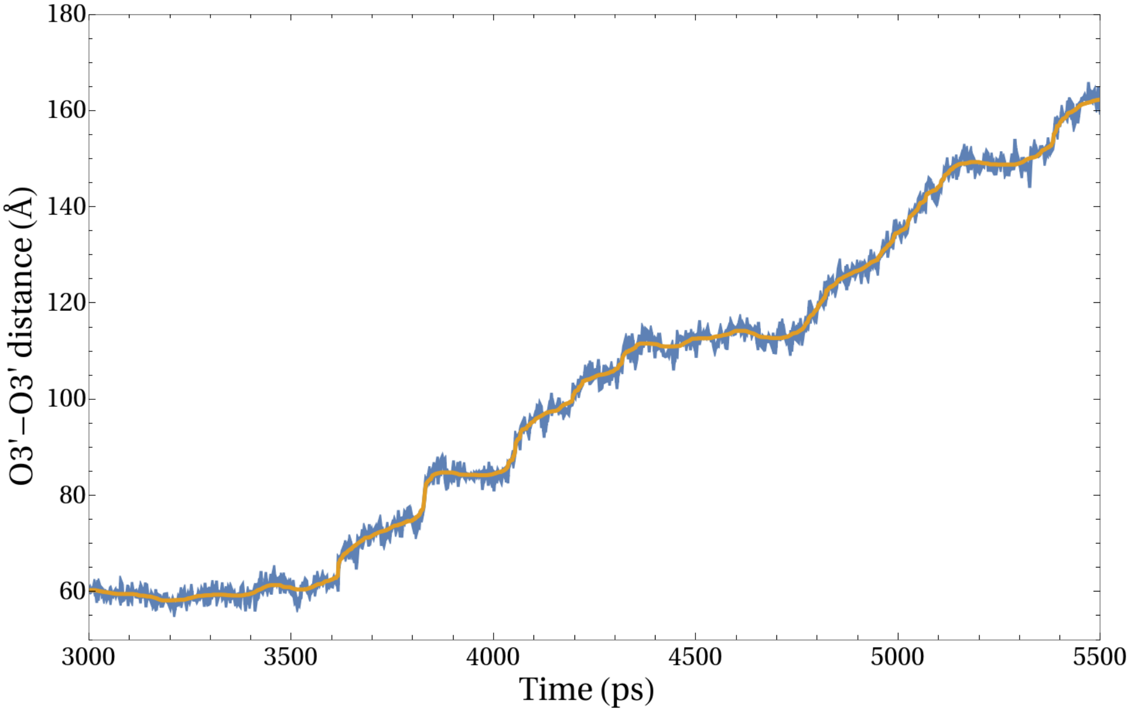

Figure 4.4: Distance between pulled O3' atoms during unzipping with a smooth trend-line obtained after filtering the thermal noise contribution

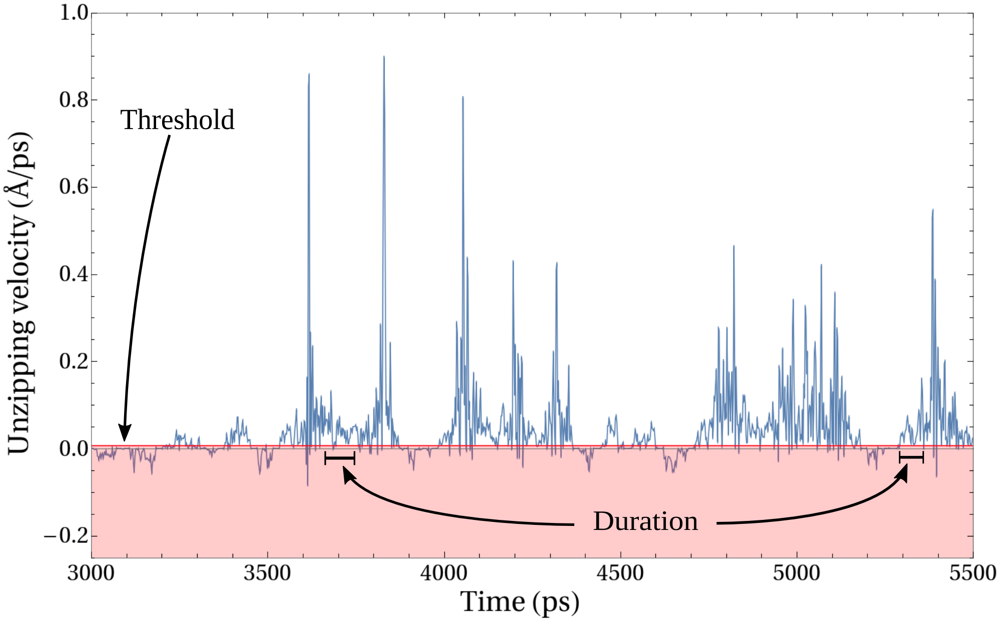

Figure 4.5: Unzipping velocity with the red threshold line for defining avalanche clusters.

4.3 Results and Discussion

Avalanches in DNA unzipping velocity were defined as all clusters of velocities above a velocity threshold (see Fig. 4.5). A threshold is required to properly define the beginning and end of an avalanche: an avalanche begins when the unzipping fork leaves and ends when it arrives at the value of the threshold. Different thresholds were considered ranging from $0.1\langle v\rangle$ to $0.25\langle v\rangle$, where $\langle v\rangle$ refers to the average velocity over time. In Figs. 4.6 and 4.7 we analyze the distribution of both avalanche time and avalanche size, defined as the area under the velocity-time curve for each avalanche event. We considered 3 different threshold values $v_{th} = 0.1\langle v\rangle$, $0.15\langle v\rangle$ and $0.2\langle v\rangle$. The slope of a linear fit line to the log-log plots of the distributions is also displayed on the graphs. For the fitting procedure, very small avalanche sizes much smaller than the inter-basepair distance of ~ 3.4 Å were neglected, and very large avalanches were neglected due to lack of sufficient sampling. The avalanches are seen to be distributed according to a power law with a cutoff at large values, for both avalanche time and avalanche size. The avalanche durations are also found to go as a power law with mean avalanche sizes for all but the smallest of avalanches (Fig. 4.8). For the linear sections of the above three plots, the scaling relations are as follows,

$$P(s) \propto s^{-\tau}, \tag{4.2}$$

$$P(T) \propto T^{-\alpha}, \tag{4.3}$$

$$\langle s \rangle \propto T^{a} \tag{4.4}$$

where τ is the scaling exponent for avalanche size, α is the exponent for avalanche time and $1/a$ is the exponent for the variation of avalanche time with mean avalanche size. We verified that the exponents were independent of the choice of the threshold for the entire range of considered threshold values.

Our analysis shows that at an external unzipping force of 50 pN, we can see almost two decades of power law scaling for both avalanche duration and size. In terms of base-pairs unzipped, we can see that avalanches resulting from the unzipping of a single base pair going on to $\sim$ 16 base pairs, all show power law scaling. This is a unique an incredible result, and implies that a model which can explain the force induced unzipping of DNA close to its critical unzipping force can be coarse grained to a length scale of 16 base pairs and still be able to explain crackling dynamics at the larger length scale. Another important result from our simulations, which would be contrary to what one would naively expect, is that there was no clear location dependence of the pause points during unzipping, implying that unzipping times varied from very short to very long for almost all base-pair locations throughout the short DNA strand. We verified this by analyzing the H-bond breaking times at different basepair locations for all of the 800 simulations of the short 18-mer strand. We only considered basepairs from no. 3 to 14 to avoid including any effects of terminal basepair fraying. Although the results of the analysis are not presented in this thesis, but it was fairly clear that the unzipping times were distributed all over the full range of 0 to ∼3ns for almost all basepairs, which implies that at a basepair which unzips almost instantly in one simulation may not unzip at all within the 3ns simulation time in another simulation.

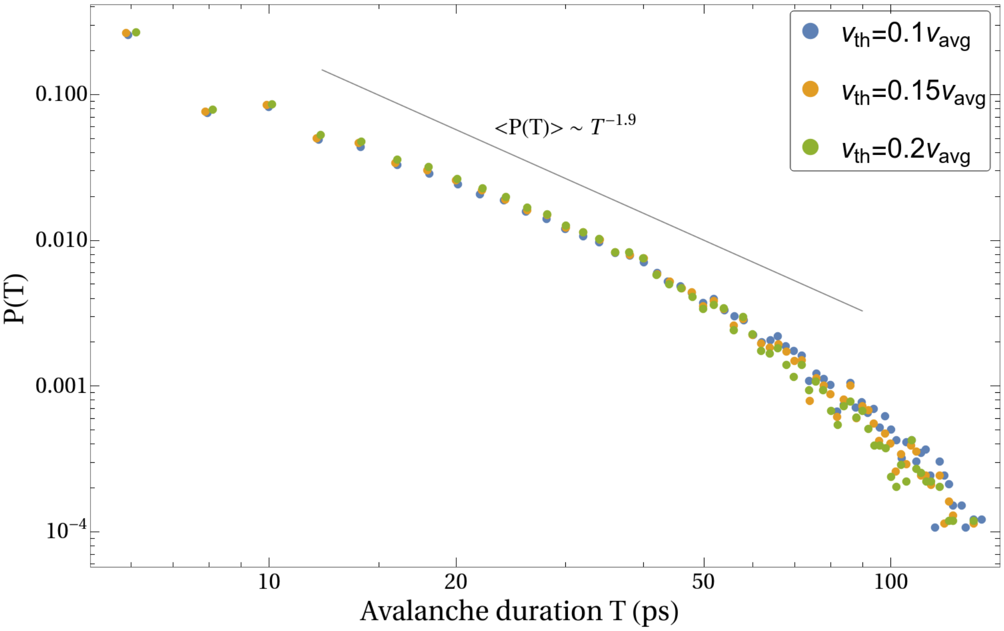

Figure 4.6: Distribution of avalanche duration with best-fit curve and straight line with slope equal to the critical exponent α.

The above results might lead the reader to ask a natural question: why are we able to see crackling dynamics at an applied external force which is about 2-3 times the experimentally observed critical force of DNA unzipping? To answer this, we refer to an earlier seminal work of crackling dynamics by Sethna [92], where the author showed that the theoretical scaling curves predicted from a mean field soluton to the 3-D Ising model work well even far from criticality (the scaling far from the critical point is still power law with a cutoff for larger values; see Fig. 2 of reference [92]). It should therefore not be surprising that even for intermediate unzipping force values we are able to see power law variations with a cutoff. The very fact that these scaling laws predict scaling to hold even far from criticality is what enables us to use MD simulation for this analysis. We hypothesize that as we use lower unzipping forces much closer to the critical unzipping force $F_c \sim 15$ pN, we should be able to see power law scaling for avalanches much larger than those lasting for 10 base pairs. However, very close to the critical force, the critical force, the timescale for unzipping enough basepairs to be able to study their distribution would be in the order of seconds to minutes [23], a timescale currently inaccessible to all-atom MD simulations. Our attempts to use an external unzipping force close to the critical force of 15 pN did not lead to any practically useful results: we could not see a single basepair unzipping even for a simulation lengths upto 10ns if the external unzipping force was equal to 15 pN.

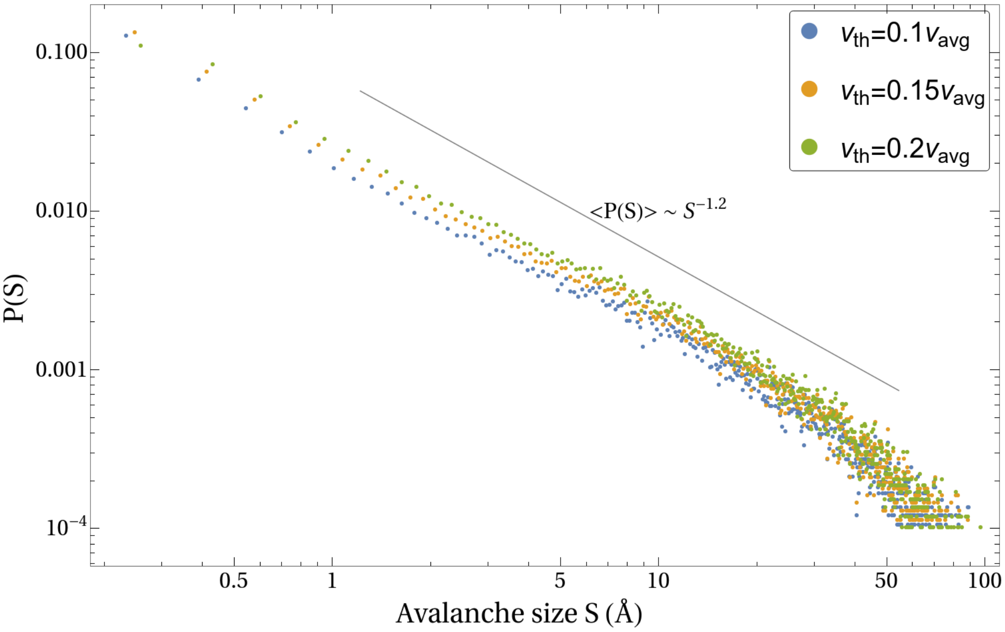

Figure 4.7: Distribution of avalanche size with best-fit curve and straight line with slope equal to the critical exponent τ.

The critical exponents calculated for DNA unzipping are compared to corresponding theoretical and experimental exponents for Barkhausen noise in polycrystalline and amorphous ferromagnets [12], slow crack growth [13] and neuronal avalanches [84] (Table 4.1). We can see that except for a consistent underestimation of the critical exponent a in our work, the other two exponents agree reasonably well with the exponents from the other systems, suggesting the possibility that DNA constant force unzipping belongs to a universality class common to some of these systems. This is a striking result: It implies that close to the critical unzipping force, mechanical force-induced DNA unzipping can possibly be explained

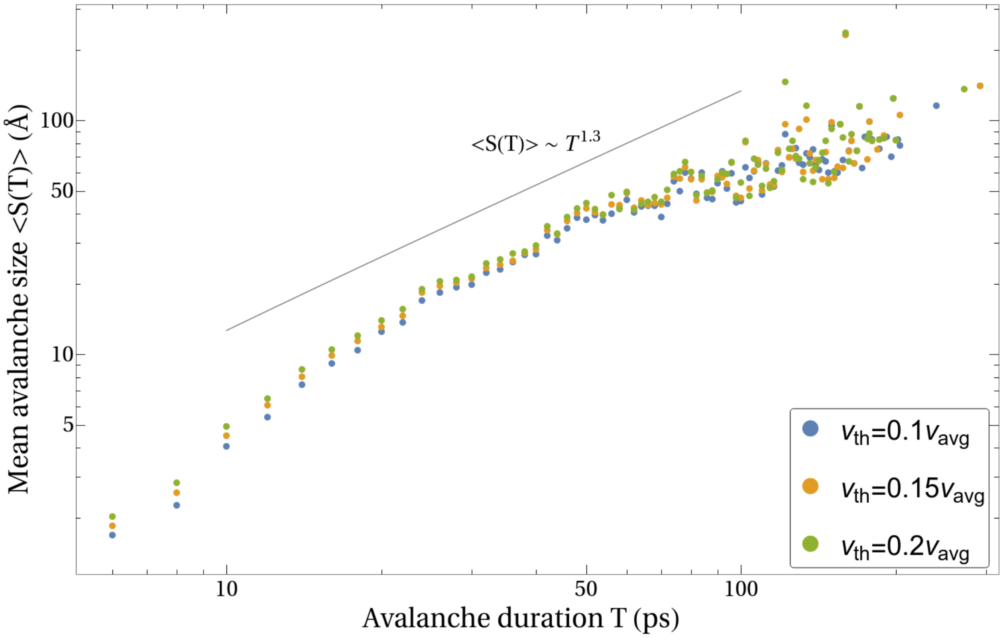

Figure 4.8: Avalanche duration v/s mean size with best-fit curve and straight line with slope equal to the critical exponent $1/a$.

using similar models that can describe crackling in other systems. However, more exhaustive theoretical and experimentcal studies will be required to conclusively predict if DNA unzipping belongs to an entirely different universality class of its own.

Table 4.1: Critical exponents τ, α and a calculated for Barkhausen noise, crack growth and neuronal avalanches. PC refers to polycrystalline and AM refers to amorphous ferromagnetic films.

System	τ	α	a
DNA (this work)	1.2	1.9	1.3
NiFe 100nm film (PC)[12]	1.5	1.9	1.9
B4 100nm film (AM) [12]	1.3	1.5	1.8
Crack Growth [13]	1.3	1.4	1.7
Neuronal avalanches [84]	2.0	3.0	1.9

4.4 Conclusion

The results presented in this chapter indicate that dsDNA homopolymer unzipping under a constant external force of $\vec{F}$=50 pN shows crackling type dynamics, as indicated by an approximately cutoff-limited power-law distribution of unzipping velocity avalanche sizes and duration. The distribution of avalanche sizes indicates that the power law variation holds for DNA strand lengths ranging from a single basepair unzipping to $\sim$ 16 basepairs. The calculated power law exponents are in close agreement with those calculated for crackling noise in other unrelated systems. Going further, there is scope to test for lower force values closer to the critical force of unzipping. Closer to the critical force, we would expect the power law variation to last for longer length scales before being limited by the cutoff. Strand sequence effects can be probed by using non-homogeneous sequences. It will also be interesting to test whether crackling extends to the unzipping of nucleic acids other than DNA, like double-stranded RNA which has a similar double helical structure but different elastic properties. In addition to the possibility of expanding the universality class of crackling, the above results also expand our knowledge base of the mechanism of force-induced DNA unzipping. Essential genetic processes which involve dsDNA unzipping by proteins include replication, transcription, DNA repair and recombination. The results presented in this chapter suggest a new area of exploration relevant to these processes.

Bibliography

[1] *Wiener Filters*, chapter 6, pages 178–204. John Wiley and Sons, Ltd, 2001.

[2] A. Aggarwal, J. Rudnick, R. F. Bruinsma, and W. S. Klug. Elasticity theory of macromolecular aggregates. *Phys. Rev. Lett.*, 109(14):148102, Oct 2012.

[3] A. A. Aksyuk, L. P. Kurochkina, A. Fokine, F. Forouhar, V. V. Mesyanzhinov, L. Tong, and M. G. Rossmann. Structural conservation of the myoviridae phage tail sheath protein fold. *Structure*, 19(12):1885–1894, Dec 2011.

[4] A. A. Aksyuk, P. G. Leiman, L. P. Kurochkina, M. M. Shneider, V. A. Kostyuchenko, V. V. Mesyanzhinov, and M. G. Rossmann. The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. *EMBO J.*, 28(7):821–829, Apr 2009.

[5] M. Allen and D. Tildesley. *Computer Simulation of Liquids*. Oxford: Clarendon Pr, 1987.

[6] M. Antognozzi, A. D. L. Humphris, and M. J. Miles. Observation of molecular layering in a confined water film and study of the layers viscoelastic properties.

[7] F. Arisaka, J. Engel, and H. Klump. Contraction and dissociation of the bacteriophage T4 tail sheath induced by heat and urea. *Prog. Clin. Biol. Res.*, 64:365–379, 1981.

[8] G. S. Ayton, W. G. Noid, and G. A. Voth. Multiscale modeling of biomolecular systems: in serial and in parallel. *Curr. Opin. Struct. Biol.*, 17(2):192–198, Apr 2007.

[9] M. Basler. Type VI secretion system: secretion by a contractile nanomachine. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 370(1679), Oct 2015.

[10] M. Basler, M. Pilhofer, G. P. Henderson, G. J. Jensen, and J. J. Mekalanos. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature*, 483(7388):182–186, Feb 2012.

[11] M. Bathe. A finite element framework for computation of protein normal modes and mechanical response. *Proteins*, 70(4):1595–1609, Mar 2008.

[12] F. Bohn, G. Durin, M. A. Correa, N. R. Machado, R. D. Della Pace, C. Chesman, and R. L. Sommer. Playing with universality classes of Barkhausen avalanches. *Sci Rep*, 8(1):11294, Jul 2018.

[13] D. Bonamy, S. Santucci, and L. Ponson. Crackling dynamics in material failure as the signature of a self-organized dynamic phase transition. *Phys. Rev. Lett.*, 101(4):045501, Jul 2008.

[14] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: the biomolecular simulation program. *J Comput Chem*, 30(10):1545–1614, Jul 2009.

[15] D. L. Caspar. Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys. J.*, 32(1):103–138, Oct 1980.

[16] T. E. Cheatham and P. A. Kollman. Molecular dynamics simulation of nucleic acids. *Annu Rev Phys Chem*, 51:435–471, 2000.

[17] Y. J. Chen, S. Papanikolaou, J. P. Sethna, S. Zapperi, and G. Durin. Avalanche spatial structure and multivariable scaling functions: sizes, heights, widths, and views through windows. *Phys Rev E Stat Nonlin Soft Matter Phys*, 84(6 Pt 1):061103, Dec 2011.

[18] Z. Chen, L. Sun, Z. Zhang, A. Fokine, V. Padilla-Sanchez, D. Hanein, W. Jiang, M. G. Rossmann, and V. B. Rao. Cryo-EM structure of the bacteriophage T4 isometric head at 3.3- resolution and its relevance to the assembly of icosahedral viruses. *Proc. Natl. Acad. Sci. U.S.A.*, 114(39):E8184–E8193, 09 2017.

[19] K. Christensen, L. Danon, T. Scanlon, and P. Bak. Unified scaling law for earthquakes. *Proc. Natl. Acad. Sci. U.S.A.*, 99 Suppl 1:2509–2513, Feb 2002.

[20] S. Cocco, R. Monasson, and J. F. Marko. Force and kinetic barriers to unzipping of the DNA double helix. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8608–8613, Jul 2001.

[21] S. Cocco, R. Monasson, and J. F. Marko. Force and kinetic barriers to initiation of DNA unzipping. *Phys Rev E Stat Nonlin Soft Matter Phys*, 65(4 Pt 1):041907, Apr 2002.

[22] K. Dahmen. Universal clues in noisy skews. *Nature Physics*, 1(1):13–14, 2005.

[23] C. Danilowicz, V. W. Coljee, C. Bouzigues, D. K. Lubensky, D. R. Nelson, and M. Prentiss. DNA unzipped under a constant force exhibits multiple metastable intermediates. *Proc. Natl. Acad. Sci. U.S.A.*, 100(4):1694–1699, Feb 2003.

[24] D. J. De Rosier and A. Klug. Reconstruction of three dimensional structures from electron micrographs. *Nature*, 217(5124):130–134, Jan 1968.

[25] G. Durin and S. Zapperi. Scaling exponents for barkhausen avalanches in polycrystalline and amorphous ferromagnets. *Phys. Rev. Lett.*, 84(20):4705–4708, May 2000.

[26] S. W. Englander and L. Mayne. The nature of protein folding pathways. *Proceedings of the National Academy of Sciences*, 111(45):15873–15880, 2014.

[27] W. Falk and R. D. James. Elasticity theory for self-assembled protein lattices with application to the martensitic phase transition in bacteriophage T4 tail sheath. *Phys Rev E Stat Nonlin Soft Matter Phys*, 73(1 Pt 1):011917, Jan 2006.

[28] S. C. Flores, J. Bernauer, S. Shin, R. Zhou, and X. Huang. Multiscale modeling of macromolecular biosystems. *Brief. Bioinformatics*, 13(4):395–405, Jul 2012.

[29] A. Fokine, Z. Zhang, S. Kanamaru, V. D. Bowman, A. A. Aksyuk, F. Arisaka, V. B. Rao, and M. G. Rossmann. The molecular architecture of the bacteriophage T4 neck. *J. Mol. Biol.*, 425(10):1731–1744, May 2013.

[30] R. Galindo-Murillo, C. Bergonzo, and T. E. Cheatham. Molecular modeling of nucleic Acid structure: setup and analysis. *Curr Protoc Nucleic Acid Chem*, 56:1–21, Mar 2014.

[31] P. Ge, D. Scholl, P. G. Leiman, X. Yu, J. F. Miller, and Z. H. Zhou. Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat. Struct. Mol. Biol.*, 22(5):377–382, May 2015.

[32] M. G. Ghequire and R. De Mot. The Tailocin Tale: Peeling off Phage Tails. *Trends Microbiol.*, 23(10):587–590, Oct 2015.

[33] F. Gittes, B. Mickey, J. Nettleton, and J. Howard. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J. Cell Biol.*, 120(4):923–934, Feb 1993.

[34] M. P. Goertz, J. E. Houston, and X. Y. Zhu. Hydrophilicity and the viscosity of interfacial water. *Langmuir*, 23(10):5491–5497, May 2007.

[35] P. Gonzalez de Prado Salas, I. Horger, F. Martin-Garcia, J. Mendieta, A. Alonso, M. Encinar, P. Gomez-Puertas, M. Velez, and P. Tarazona. Torsion and curvature of FtsZ filaments. *Soft Matter*, 10(12):1977–1986, Mar 2014.

[36] S. Goyal, T. Lillian, S. Blumberg, J.-C. Meiners, E. Meyhöfer, and N. Perkins. Intrinsic curvature of dna influences lacr-mediated looping. *Biophysical Journal*, 93(12):4342–4359, 2007.

[37] S. Goyal, N. Perkins, and C. Lee. Nonlinear dynamics and loop formation in kirchhoff rods with implications to the mechanics of {DNA} and cables. *Journal of Computational Physics*, 209(1):371 – 389, 2005.

[38] S. Goyal, N. C. Perkins, and J.-C. Meiners. Resolving the sequence-dependent stiffness of dna using cyclization experiments and a computational rod model. *Journal of Computational and Nonlinear Dynamics*, 3(1):011003, 2008.

[39] K. F. Graff. *Wave motion in elastic solids.* Dover books on physics. Dover, New York, NY, 1991.

[40] W. I. Hagens, A. G. Oomen, W. H. de Jong, F. R. Cassee, and A. J. Sips. What do we (need to) know about the kinetic properties of nanoparticles in the body? *Regul. Toxicol. Pharmacol.*, 49(3):217–229, Dec 2007.

[41] S. D. Hicks and C. L. Henley. Coarse-grained protein-protein stiffnesses and dynamics from all-atom simulations. *Phys Rev E Stat Nonlin Soft Matter Phys*, 81(3 Pt 1):030903, Mar 2010.

[42] A. D. Hirsh, M. Taranova, T. A. Lionberger, T. D. Lillian, I. Andricioaei, and N. Perkins. Structural ensemble and dynamics of toroidal-like dna shapes in bacteriophage ϕ29 exit cavity. *Biophysical journal*, 104(9):2058–2067, 2013.

[43] P. A. Houle and J. P. Sethna. Acoustic emission from crumpling paper. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 54(1):278–283, Jul 1996.

[44] W. Humphrey, A. Dalke, and K. Schulten. VMD: visual molecular dynamics. *J Mol Graph*, 14(1):33–38, Feb 1996.

[45] B. Jagannathan, P. J. Elms, C. Bustamante, and S. Marqusee. Direct observation of a force-induced switch in the anisotropic mechanical unfolding pathway of a protein. *Proc. Natl. Acad. Sci. U.S.A.*, 109(44):17820–17825, Oct 2012.

[46] J.-W. Jiang, J.-S. Wang, and B. Li. Youngs modulus of graphene: a molecular dynamics study. *Physical Review B*, 80(11):113405, 2009.

[47] P. P. Jose and I. Andricioaei. Similarities between protein folding and granular jamming. *Nat Commun*, 3:1161, 2012.

[48] S. Kanamaru and F. Arisaka. The structural biology and infection mechanism of bacteriophage T4. *Seikagaku*, 74(2):131–135, Feb 2002.

[49] M. Karplus and J. Kuriyan. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U.S.A.*, 102(19):6679–6685, May 2005.

[50] M. Karplus and J. A. McCammon. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.*, 9(9):646–652, Sep 2002.

[51] A. Kishino and T. Yanagida. Force measurements by micromanipulation of a single actin filament by glass needles. *Nature*, 334(6177):74–76, Jul 1988.

[52] V. A. Kostyuchenko, P. R. Chipman, P. G. Leiman, F. Arisaka, V. V. Mesyanzhinov, and M. G. Rossmann. The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat. Struct. Mol. Biol.*, 12(9):810–813, Sep 2005.

[53] E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, 372(3):774–797, Sep 2007.

[54] E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, 372(3):774–797, Sep 2007.

[55] L. M. Kroon-Batenburg, P. H. Kruiskamp, J. F. Vliegenthart, and J. Kroon. Estimation of the persistence length of polymers by md simulations on small fragments in solution. application to cellulose. *The Journal of Physical Chemistry B*, 101(42):8454–8459, 1997.

[56] C. A. Laughton and S. A. Harris. The atomistic simulation of dna. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(4):590–600, 2011.

[57] P. G. Leiman, P. R. Chipman, V. A. Kostyuchenko, V. V. Mesyanzhinov, and M. G. Rossmann. Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell*, 118(4):419–429, Aug 2004.

[58] P. G. Leiman and M. M. Shneider. Contractile tail machines of bacteriophages. *Adv. Exp. Med. Biol.*, 726:93–114, 2012.

[59] T. D. Lillian and N. Perkins. Electrostatics and self-contact in an elastic rod approximation for dna. *Journal of Computational and Nonlinear Dynamics*, 6(1):011008, 2011.

[60] T. D. Lillian, M. Taranova, J. Wereszczynski, I. Andricioaei, and N. Perkins. A multiscale dynamic model of dna supercoil relaxation by topoisomerase ib. *Biophysical journal*, 100(8):2016–2023, 2011.

[61] N. S. Lossi, E. Manoli, A. Forster, R. Dajani, T. Pape, P. Freemont, and A. Filloux. The HsiB1C1 (TssB-TssC) complex of the Pseudomonas aeruginosa type VI secretion system forms a bacteriophage tail sheathlike structure. *J. Biol. Chem.*, 288(11):7536–7548, Mar 2013.

[62] T. Lux and M. Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(6719):498–500, 1999.

[63] A. Maghsoodi, A. Chatterjee, I. Andricioaei, and N. C. Perkins. A first model of the dynamics of the bacteriophage t4 injection machinery. *Journal of Computational and Nonlinear Dynamics*, 11(4):041026–041026, May 2016.

[64] A. Maghsoodi, A. Chatterjee, I. Andricioaei, and N. C. Perkins. Dynamic Model Exposes the Energetics and Dynamics of the Injection Machinery for Bacteriophage T4. *Biophys. J.*, 113(1):195–205, Jul 2017.

[65] A. Maghsoodi and N. Perkins. Shear Deformation Dissipates Energy in Biofilaments. *Sci Rep*, 8(1):11684, Aug 2018.

[66] R. C. Major, J. E. Houston, M. J. McGrath, J. I. Siepmann, and X. Y. Zhu. Viscous water meniscus under nanoconfinement. *Phys. Rev. Lett.*, 96(17):177803, May 2006.

[67] S. Matsushita, T. Adachi, Y. Inoue, M. Hojo, and M. Sokabe. Evaluation of extensional and torsional stiffness of single actin filaments by molecular dynamics analysis. *J Biomech*, 43(16):3162–3167, Dec 2010.

[68] E. R. May, A. Aggarwal, W. S. Klug, and C. L. Brooks. Viral capsid equilibrium dynamics reveals nonuniform elastic properties. *Biophys. J.*, 100(11):59–61, Jun 2011.

[69] D. McQuarrie. *Statistical Mechanics*. University Science Books, 2000.

[70] A. Mentes, A. M. Florescu, E. Brunk, J. Wereszczynski, M. Joyeux, and I. Andricioaei. Free-energy landscape and characteristic forces for the initiation of DNA unzipping. *Biophys. J.*, 108(7):1727–1738, Apr 2015.

[71] Y. Michel-Briand and C. Baysse. The pyocins of Pseudomonas aeruginosa. *Biochimie*, 84(5-6):499–510, 2002.

[72] M. F. Moody. Sheath of bacteriophage T4. 3. Contraction mechanism deduced from partially contracted sheaths. *J. Mol. Biol.*, 80(4):613–635, Nov 1973.

[73] M. Neek-Amal, F. M. Peeters, I. V. Grigorieva, and A. K. Geim. Commensurability Effects in Viscosity of Nanoconfined Water. *ACS Nano*, 10(3):3685–3692, Mar 2016.

[74] D. R. Nelson. Biophysical dynamics in disorderly environments. *Annu Rev Biophys*, 41:371–402, 2012.

[75] M. N. Nguyen and M. S. Madhusudhan. Biological insights from topology independent comparison of protein 3D structures. *Nucleic Acids Res.*, 39(14):e94, Aug 2011.

[76] J. Novacek, M. Siborova, M. Benesik, R. Pantucek, J. Doskaf, and P. Plevka. Structure and genome release of Twort-like Myoviridae phage with a double-layered baseplate. *Proc. Natl. Acad. Sci. U.S.A.*, 113(33):9351–9356, Aug 2016.

[77] A. Parisien, B. Allain, J. Zhang, R. Mandeville, and C. Q. Lan. Novel alternatives to antibiotics: bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. *J. Appl. Microbiol.*, 104(1):1–13, Jan 2008.

[78] V. A. Petrenko and V. J. Vodyanoy. Phage display for detection of biological threat agents. *J. Microbiol. Methods*, 53(2):253–262, May 2003.

[79] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem*, 25(13):1605–1612, Oct 2004.

[80] J. Pfaendtner, E. Lyman, T. D. Pollard, and G. A. Voth. Structure and dynamics of the actin filament. *J. Mol. Biol.*, 396(2):252–263, Feb 2010.

[81] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with NAMD. *J Comput Chem*, 26(16):1781–1802, Dec 2005.

[82] S. Piana, K. Lindorff-Larsen, and D. E. Shaw. Protein folding kinetics and thermodynamics from atomistic simulation. *Proceedings of the National Academy of Sciences*, 109(44):17845–17850, 2012.

[83] M. G. Poirier and J. F. Marko. Effect of internal friction on biofilament dynamics. *Phys. Rev. Lett.*, 88(22):228103, Jun 2002.

[84] A. Ponce-Alvarez, A. Jouary, M. Privat, G. Deco, and G. Sumbre. Whole-Brain Neuronal Activity Displays Crackling Noise Dynamics. *Neuron*, 100(6):1446–1459, Dec 2018.

[85] U. Raviv, P. Laurat, and J. Klein. Fluidity of water confined to subnanometre films. *Nature*, 413(6851):51–54, Sep 2001.

[86] O. Salih, S. He, S. Planamente, L. Stach, J. T. MacDonald, E. Manoli, S. H. W. Scheres, A. Filloux, and P. S. Freemont. Atomic Structure of Type VI Contractile Sheath from Pseudomonas aeruginosa. *Structure*, 26(2):329–336, Feb 2018.

[87] M. Santosh and P. K. Maiti. Force induced DNA melting. *J Phys Condens Matter*, 21(3):034113, Jan 2009.

[88] A. Schlaich, J. Kappler, and R. R. Netz. Hydration Friction in Nanoconfinement: From Bulk via Interfacial to Dry Friction. *Nano Lett.*, 17(10):5969–5976, Oct 2017.

[89] D. Scholl, M. Cooley, S. R. Williams, D. Gebhart, D. Martin, A. Bates, and R. Mandrell. An engineered R-type pyocin is a highly specific and sensitive bactericidal agent for the food-borne pathogen Escherichia coli O157:H7. *Antimicrob. Agents Chemother.*, 53(7):3074–3080, Jul 2009.

[90] I. I. Serysheva, A. I. Tourkin, I. V. Bartish, and B. F. Poglazov. GTPase activity of bacteriophage T4 sheath protein. *J. Mol. Biol.*, 223(1):23–25, Jan 1992.

[91] J. P. Sethna. Materials science. Crackling wires. *Science*, 318(5848):207–208, Oct 2007.

[92] J. P. Sethna, K. A. Dahmen, and C. R. Myers. Crackling noise. *Nature*, 410(6825):242–250, Mar 2001.

[93] A. Shamloo and B. Mehrafrooz. Nanomechanics of actin filament: A molecular dynamics simulation. *Cytoskeleton (Hoboken)*, Dec 2017.

[94] M. A. Sheikh, R. L. Weaver, and K. A. Dahmen. Avalanche Statistics Identify Intrinsic Stellar Processes near Criticality in KIC 8462852. *Phys. Rev. Lett.*, 117(26):261101, Dec 2016.

[95] J. Sponer, P. Banas, P. Jurecka, M. Zgarbova, P. Kuhrova, M. Havrila, M. Krepl, P. Stadlbauer, and M. Otyepka. Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome. *J Phys Chem Lett*, 5(10):1771–1782, May 2014.

[96] D. E. Tanner, K. Y. Chan, J. C. Phillips, and K. Schulten. Parallel Generalized Born Implicit Solvent Calculations with NAMD. *J Chem Theory Comput*, 7(11):3635–3642, Nov 2011.

[97] N. M. Taylor, N. S. Prokhorov, R. C. Guerrero-Ferreira, M. M. Shneider, C. Browning, K. N. Goldie, H. Stahlberg, and P. G. Leiman. Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature*, 533(7603):346–352, May 2016.

[98] R. Toussaint and S. R. Pride. Interacting damage models mapped onto Ising and percolation models. *Phys Rev E Stat Nonlin Soft Matter Phys*, 71(4 Pt 2):046127, Apr 2005.

[99] Y. Tsuda, H. Yasutake, A. Ishijima, and T. Yanagida. Torsional rigidity of single actin filaments and actin-actin bond breaking force under torsion measured directly by in vitro micromanipulation. *Proc. Natl. Acad. Sci. U.S.A.*, 93(23):12937–12942, Nov 1996.

[100] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen. Gromacs: fast, flexible, and free. *Journal of computational chemistry*, 26(16):1701–1718, 2005.

[101] S. N. Volkov, E. V. Paramonova, A. V. Yakubovich, and A. V. Solov'yov. Micromechanics of base pair unzipping in the DNA duplex. *J Phys Condens Matter*, 24(3):035104, Jan 2012.

[102] N. R. Voss and M. Gerstein. 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.*, 38(Web Server issue):W555–562, Jul 2010.

[103] B. Webb and A. Sali. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci*, 86:1–2, Nov 2016.

[104] J. D. Weeks, J. B. Lucks, Y. Kafri, C. Danilowicz, D. R. Nelson, and M. Prentiss. Pause point spectra in DNA constant-force unzipping. *Biophys. J.*, 88(4):2752–2765, Apr 2005.

[105] D. B. Wells and A. Aksimentiev. Mechanical properties of a complete microtubule revealed through molecular dynamics simulation. *Biophys. J.*, 99(2):629–637, Jul 2010.

[106] Wikipedia. Crackling noise — Wikipedia, the free encyclopedia. `http://en.wikipedia.org/w/index.php?title=Crackling%20noise&oldid=880080902`, 2019. [Online; accessed 04-May-2019].

[107] S. R. Williams, D. Gebhart, D. W. Martin, and D. Scholl. Retargeting R-type pyocins to generate novel bactericidal protein complexes. *Appl. Environ. Microbiol.*, 74(12):3868–3876, Jun 2008.

[108] M. Yaghoubi, T. de Graaf, J. G. Orlandi, F. Girotto, M. A. Colicos, and J. Davidsen. Neuronal avalanche dynamics indicates different universality classes in neuronal cultures. *Sci Rep*, 8(1):3417, Feb 2018.

[109] M. L. Yap, T. Klose, F. Arisaka, J. A. Speir, D. Veesler, A. Fokine, and M. G. Rossmann. Role of bacteriophage T4 baseplate in regulating assembly and infection. *Proc. Natl. Acad. Sci. U.S.A.*, 113(10):2654–2659, Mar 2016.

[110] M. L. Yap and M. G. Rossmann. Structure and function of bacteriophage t4. *Future Microbiol*, 9:1319–1327, Oct 2014. 25517898[pmid].

[111] C. Yui. Structure of pyocin R. I. Isolation of sheath from pyocin R by alkali treatment and its properties. *J. Biochem.*, 69(1):101–110, Jan 1971.