UC Irvine
UC Irvine Electronic Theses and Dissertations

Title

Computational Studies of Biomolecular Conformational Changes Through Potential of Mean Force Calculations and Accelerated Molecular Dynamics Simulations

Permalink

https://escholarship.org/uc/item/0xp3z953

Author

Ficici, Emel

Publication Date

2015

Copyright Information

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Computational Studies of Biomolecular Conformational Changes Through Potential of
Mean Force Calculations and Accelerated Molecular Dynamics Simulations

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry with a concentration in Chemical and Materials Physics

by

Emel (Guneyler) Ficici

Dissertation Committee:
Professor Ioan Andricioaei, Chair
Professor Douglas J. Tobias
Professor Craig Martens

2015

DEDICATION

To my family...

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

CURRICULUM VITAE

Emel (Guneyler) Ficici

EDUCATION

Doctor of Philosophy in Chemistry
with a concentration in Chemical and Material Physics **2015**
University of California Irvine *Irvine, CA*

Master of Science in Chemical and Material Physics **2015**
University of California Irvine *Irvine, CA*

Master of Science in Physics **1999**
Middle East Technical University *Ankara, Turkey*

Bachelor of Science in Science Education
Program in Physics **1997**
Bogazici University *Istanbul, Turkey*

REFEREED JOURNAL PUBLICATIONS

"On the Possibility of Facilitated Diffusion of Dendrimers Along DNA" E. Ficici, I. Andricioaei, J. Phys. Chem. B **2015**

"Dendrimers in Nanoscale Confinement: The Interplay of Conformational Change and Pore Blockade", E. Ficici, S. Howorka, I. Andricioaei, submitted for publication **2015**

"The Role of Proton Motive Force in Conformational Transition of SecDF from F-form to I-form", E. Ficici, D. Jeung, I. Andricioaei to be submitted

Molecular Dynamics Simulations of Water Clusters,, S. Erkoc and E. Guneyler Physica E 8, 40-49 (2000)

CONFERENCE PRESENTATIONS

Size Dependent Translocations of PAMAM Dendrimers   **Feb 2015**
Through Alpha-Hemolysin Pore: Ion Conductance and
Dynamics of Translocation
Biophysical Society Meeting

The Interplay Between SecDF Conformational Change   **Feb 2014**
and Proton Motive Force
Biophysical Society Meeting

Biomimetic Strategy for Nanoparticle Sliding Along   **Feb 2013**
DNA Tracks
Biophysical Society Meeting

ABSTRACT OF THE DISSERTATION

Computational Studies of Biomolecular Conformational Changes Through Potential of
Mean Force Calculations and Accelerated Molecular Dynamics Simulations

By

Emel (Guneyler) Ficici

Doctor of Philosophy in Chemistry with a concentration in Chemical and Materials Physics

University of California, Irvine, 2015

Professor Ioan Andricioaei, Chair

Molecular dynamics (MD) simulations have made significant contributions to our under-
standing of biomolecular structures, functions and processes during the last few decades.
This proven success is partly due to the level of detail that MD simulations can offer through
all-atomistic models with femtosecond level time resolutions. However, their real power in
studying increasingly larger systems and biomolecular events with increasingly longer time
scales comes from availability of advanced sampling and accelerating techniques. In this
study, we utilize some of these advanced techniques to investigate energetics, dynamics and
other related properties of diverse biomolecular systems and events as outlined below:

In the first two chapters, we study the interaction of a polyamidoamine (PAMAM) dendrimer,
a biocompatible artificial molecule with a highly branched structure, with (i) DNA (Chapter
1) and (ii) α-hemolysin (α-HL), a transmembrane protein nanopore (Chapter 2). Since den-
drimers are versatile molecules in their potential uses in various fields of bionanomedicine
and bionanotechnology, their interactions with biomolecules are the subject of active re-
search. In Chapter 1, we study the interaction between a third generation (G3) dendrimer
and DNA from a biomimetic perspective. To evaluate the possibility of mimicking the pro-
tein nonspecific search behavior along DNA, we obtained potential of mean force (PMF) of

dendrimer-DNA interaction along a circular reaction coordinate around DNA by using umbrella sampling technique. PMF reveals a free energy minimum when the dendrimer center of mass is on DNA major groove and a free energy energy barrier of 8.5 kcal/mol between dendrimer- DNA minor-groove and dendrimer-DNA major-groove interactions. As such, our result identifies a helical path along DNA major-groove as a possible reaction coordinate for dendrimer sliding along DNA.

Chapter 2 is the result of a collaborative effort with experimentalist Dr. Stephan Howorka from University College London, UK. In this study, we investigate translocation/capture properties of dendrimers of various sizes, G1-G5, through α-hemolysin pore through the use of steered molecular dynamics technique. Computationally, we were able to obtain current blockade values in very close agreement with experimental data. Finally, in Chapter 3, we find a low energy reaction pathway for the large conformational change of a protein, SecDF, using Targeted Molecular Dynamics method, and obtain potential of mean force for this conformational change. This study helps us relate the SecDF conformational change to its proposed function and explain its mechanism of function.

Chapter 1

Introduction

Many biomolecules undergo significant conformational changes to perform their functions [1–4]. Therefore understanding these functional conformational changes is an important goal in the studies of biomolecules. An important thermodynamic quantity used in the studies of protein conformational changes is the free energy profile along a reaction coordinate that connects the two states. Free energy is the reversible work needed to bring the system from the reactant state to the product state along a predefined/reaction coordinate and it includes contributions from enthalpy and entropy. Conformational equilibrium processes as well as reaction rates can be calculated from free energy barriers [5] by using the transition state theory [6, 7]. Free energy profile of a reaction for complex systems can be obtained from simulations [8, 9] through the use of potential of mean force (PMF) formalism established by Kirkwood [10]. PMF can be calculated by special sampling techniques such as free energy perturbation technique [11], Jarzynski equality [12] and umbrella sampling technique [13]. This study will focus on the applications of umbrella sampling technique to events with time scales that are longer than accessible with ordinary molecular dynamics simulations. Choosing a reaction coordinate is an important step in potential of mean force calculations. Ideally, it is the slowest degree of freedom but it can be chosen to be any 1 or 2 dimensional

coordinate such as a distance. angle, dihedral, an RMSD metric, etc. When studying protein conformational changes, if the structures for the two low energy conformations are known, i.e. reactant state and product state, one can find a low energy pathway between the two states by using Targeted Molecular Dynamics [14, 15] method as described in the next section. This pathway can later be used as a reaction coordinate for the conformational transition in umbrella sampling calculations.

1.1 Targeted Molecular Dynamics

Targeted Molecular Dynamics (TMD) is a method commonly used to find a low energy pathway for conformational transitions at ordinary temperatures [14, 15]. It has been successfully implemented to study large scale conformational changes for proteins [16–23] and for nucleic acids [24, 25]. In this method, an evolving holonomic constraint of the following form is used to drive the system from an initial configuration ($\vec{r}_{init}$) to a target configuration ($\vec{r}_{final}$):

$$\phi(\vec{r}_t) \equiv (RMSD(\vec{r}_t, \vec{r}_{final}))^2 - \rho(t)^2 = 0 \tag{1.1}$$

Here $\rho(t)$ is an evolving root-mean-square-distance (RMSD) criteria which is initially set equal to the mass-weighed RMSD of initial structure from the target structure, i.e. $(\text{RMSD}(\vec{r}_{init}, \vec{r}_{final}))$. To drive the structure towards the target structure, $\rho(t)$ is monotonically decreased until it is close to close zero. Here $\vec{r}_{init}$, $\vec{r}_{final}$, and $\vec{r}_t$ are 3N-dimensional configurational vectors of the N-atoms for initial structure, final structure, and intermediate structure at time t, respectively. RMSD between initial and final structures are calculated

by:

$$RMSD(\vec{r}_{init}, \vec{r}_{final}) = \sqrt{\frac{1}{\sum_{k=1}^{N} m_k} \sum_{k=1}^{N} m_k |\vec{r}_{k,init} - \vec{r}_{k,final}|^2} \qquad (1.2)$$

Satisfying the constraint given in equation (1.1) results in an additional constraint force, F^c, which is given by:

$$F^c \equiv \lambda \frac{d\phi(\vec{r}_t)}{d\vec{r}_t} = 2\lambda(\vec{r}_t - \vec{r}_{final}) \qquad (1.3)$$

The value of the parameter λ that satisfies the constraint is determined at each integration step.

1.2 Potential of Mean Force through Umbrella Sampling

Once a minimum energy pathway is obtained for the conformational transition, potential of mean force (PMF) [10] along this pathway can be calculated from the probability distribution functions obtained through extensive sampling [8, 9]. Potential of mean force $(A(\xi))$ is

obtained from equilibrium probability distribution as follows:

$$A(\xi) = -k_B T \ln < p(\xi) > \tag{1.4}$$

where $< p(\xi) >$ is the average distribution function obtained from Boltzman weighted average as follows:

$$p(\xi) = \frac{\int d\vec{r}\, \delta(\xi'(\vec{r}) - \xi)\, e^{-U(\vec{r})/k_B T}}{\int d\vec{r}\, e^{-U(\vec{r})/k_B T}} \tag{1.5}$$

Here $\vec{r}$ represents the set of 3N-dimensional configurational vectors for N atoms of the system and $U(\vec{r})$ is the total energy of the system as a function of $\vec{r}$.

One difficulty in obtaining sufficient sampling along the reaction pathway is the difficulty of sampling high energy sates. In umbrella sampling method [13], sampling of all regions along the reaction coordinate, including the high energy regions, can be achieved by dividing the reaction coordinate into a series of overlapping windows and by using additional restraining potentials (w_b) centered at each window. The resulting states can then be histogrammed into bins along the reaction coordinate to obtain a probability distribution. While the resulting distribution is obtained in the existence of biasing potentials, it can be unbiased through weighted histogram analysis (WHAM) method [26, 27] to give an unbiased distribution, and hence an unbiased potential of mean force. The results from biased simulations can also be

used to obtain equilibrium ensemble averages of other thermodynamic quantities ($< O >$) of interest after unbiasing as follows [28]:

$$< O >= \frac{< O\, e^{\beta w_b} >_b}{< e^{\beta w_b} >_b} \tag{1.6}$$

Here $<>_b$ and $<>$ represent biased and unbiased ensemble averages, respectively.

1.3   Weighted Histogram Analysis

As mentioned previously, once a series of simulations are run under a biasing potential $(w_{b,i})$ used to restrain the system to a given umbrella window, biased probability distributions are obtained, which can later be used to obtain the unbiased probabilities (P(x)):

$$P(x) = \frac{\sum_{i=1}^{N} n_i(x)}{\sum_{i=1}^{N} N_i(x) e^{-\beta(F_i - w_{b,i})}} \tag{1.7}$$

Here $n_i(x)$ is the number of counts that the reaction coordinate value falls into the bin x during simulation i, and N is the number of simulations. F_i is the free energy shift from simulation i. If we were to know the true value of F_i, in principle, we would find the true probability distributions. However, both F_i and $P(x)$ are unknown in this equation. It has been shown [26] that Equation 1.7 forms a set of self-consistent equations with the following

equation, which can be solved for F_i and $P(x)$ iteratively:

$$F_i = -\beta \ln \left(\sum_{x_{bins}} P(x) \, e^{-\beta w_{b,i}(x)} \right) \tag{1.8}$$

The solution for $P(x)$ is the best estimates to the unbiased probability distribution, which gives the best estimate to the free energy profile along the reaction coordinate.

1.4   Steered Molecular Dynamics

One frequently used method to accelerate the kinetics of events in MD simulations is steered MD, in which an additional biasing potential is applied [29][30–32] to drive the reaction in a desired direction. The type of the steering potential can vary according to the type of event to be simulated. For example, to simulate conditions similar to AFM experiments, usually a harmonic potential centered around a point which is moving at a constant speed is applied in a given direction [33]. To simulate the effects of a voltage bias of magnitude V across a membrane, a constant electric field (E) interacting with the partial charges (q_i) is the biasing potential of choice applied in the direction normal to the membrane plane such that a force of magnitude $q_i E$ is applied on each charge [34],[35]. Usually the value of steering potentials used in simulations need to be higher than the actual values used in the experiments to obtain advantage in the speed up. However, when very large steering potentials are used,

there is the risk of causing artificial configurations due to lack of sufficient relaxation of the system components.

Recently, an elegant work is published [36] to refer to this difficulty of increasing the potential while maintaining the advantage in speed up, yet at the same time minimizing the artificial strain on the system. Accordingly, an electrostatic grid potential is obtained for a system of interest [34], which can later be scaled up to yield higher steering potential values [36]. Since the value of steering force felt by every component in the system varies with position, mimicking the variations in the true potential distributions of the system, this method, known as grid steered molecular dynamics (GSMD), provides more realistic conditions when used in combination with constant electric field. We will implement this method in Chapter 3 of this work to study the translocation of dendrimers through a nanopore.

1.5   CHARMM Potential Energy Function

Molecular dynamics simulations involve determining the coordinates and velocity of each atom at a given time. This is achieved by first determining the force on each atom applied by other atoms in the system at time t and then finding the velocity of each atom at a later time, $t + \Delta t$, by numerically integrating Newton's equation of motion [37]. Force on each atom can be calculated from the gradient a potential energy function. In our simulations, we used CHARMM potential energy function [38–40], which has the following form:

$$U(\vec{R}) = \sum_{bonds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \sum_{dihedrals} K_\varphi(1+cos(n\varphi-\delta))$$

$$+ \sum_{improper} K_\omega(\omega-\omega_0)^2 + \sum_{UreyBradley} k_{UB}(S-S_0)^2 + \sum_{residues} U_{CMAP}(\varphi, \psi)$$

$$+ \sum_{nonbonded} \left\{ \epsilon_{i,j}^{min} \left[\left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{min}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon r_{ij}} \right\} \quad (1.10)$$

CHARMM force field is one of the most commonly used potential energy functions in the studies of biomolecules. We refer to Ref.s [38–40] for more details on this potential.

Bibliography

[1] Karplus, M., and Kuriyan, J. (2005) Molecular Dynamics and Protein Function. *Proc. Natl. Acad. Sci. U.S.A. 102*, 6679–6685.

[2] Shea, J. E., and Brooks, C. L. (2001) From Folding Theories to Folding Proteins: A Review and Assessment of Simulation Studies of Protein Folding and Unfolding. *Annu. Rev. Phys. Chem. 52*, 499–535.

[3] Yu, H., Ma, L., Yang, Y., and Cui, Q. (2007) Mechanochemical Coupling in the Myosin Motor Domain. I. Insights from Equilibrium Active-Site Simulations. *PLoS Computational Biology 3*, e21.

[4] Yu, H., Ma, L., Yang, Y., and Cui, Q. (2007) Mechanochemical Coupling in the Myosin Motor Domain. II. Analysis of Critical Residues. *PLoS Comput. Biol. 3*, e23–e23.

[5] Chandler, D. (1978) Statistical Mechanics of Isomerization Dynamics in Liquids and the Transition State Approximation. *J. Chem. Phys. 68*, 2959.

[6] Eyring, H. (1935) The Activated Complex in Chemical Reactions. *J. Chem. Phys. 3*, 107.

[7] Evans, M. G., and Polanyi, M. (1935) Some Applications of the Transition State Method to the Calculation of Reaction Velocities, Especially in Solution. *Trans Faraday Soc*

[8] Roux, B., and Karplus, M. (1991) Ion Transport in a Model Gramicidin Channel. Structure and Thermodynamics. *Biophys. J. 59*, 961–981.

[9] Roux, B. (1995) the Calculation of the Potential of Mean Force Using Computer Simulations. *Comput. Phys. Commun. 91*, 275–282.

[10] Kirkwood, J. G. (1935) Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys. 3*, 300.

[11] Zwanzig, R. W. (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys. 22*.

[12] Jarzynski, C. (1997) Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.*

[13] Torrie, G. M., and Valleau, J. P. (1977) Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys. 23*, 187–199.

[14] Schlitter, J., Engels, M., Krüger, P., Jacoby, E., and Wollmer, A. (1993) Targeted Molecular Dynamics Simulation of Conformational Change-Application to the T-R Transition in Insulin. *Mol. Simul. 10*, 291–308.

[15] Schlitter, J., Engels, M., and Krüger, P. (1994) Targeted Molecular Dynamics: A New Approach for Searching Pathways of Conformational Transitions. *J. Mol. Graph. 12*, 84–89.

[16] Ma, J., Flynn, T. C., Cui, Q., Leslie, A. G. W., Walker, J. E., and Karplus, M. (2002) a Dynamic Analysis of the Rotation Mechanism for Conformational Change in F(1)-ATPase. *Structure 10*, 921–931.

[17] Ma, J., and Karplus, M. (1997) Molecular Switch in Signal Transduction: Reaction Paths of the Conformational Changes in Ras P21. *Proc. Natl. Acad. Sci. U.S.A. 94*, 11905–11910.

[18] van der Vaart, A., Ma, J., and Karplus, M. (2004) The Unfolding Action of GroEL on a Protein Substrate. *Biophys. J. 87*, 562–573.

[19] Ma, J., Sigler, P. B., Xu, Z., and Karplus, M. (2000) A Dynamic Model for the Allosteric Mechanism of GroEL. *J. Mol. Biol. 302*, 303–313.

[20] Barbacid, M. (1987) Ras Genes. *Annu. Rev. Biochem. 56*, 779–827.

[21] Ford, B., Hornak, V., Kleinman, H., and Nassar, N. (2006) Structure of a Transient Intermediate for GTP Hydrolysis by Ras. *Structure 14*, 427–436.

[22] Hall, B. E., Bar-Sagi, D., and Nassar, N. (2002) The Structural Basis for the Transition from Ras-GTP to Ras-GDP. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12138–12142.

[23] Fernando Diaz, J., Wroblowski, B., Schlitter, J., and Engelborghs, Y. (1997) Calculation of Pathways for the Conformational Transition Between the GTP- and GDP- bound States of the Ha-ras-p21 Protein: Calculations with Explicit Solvent Simulations and Comparison with Calculations in Vacuum. *Proteins 28*, 434–451.

[24] Aci, S., Mazier, S., and Genest, D. (2005) Conformational Pathway for the Kissing Complex-Extended Dimer Transition of the SL1 Stem-Loop from Genomic HIV-1 RNA As Monitored by Targeted Molecular Dynamics Techniques. *J. Mol. Biol. 351*, 520–530.

[25] Bascom, G., and Andricioaei, I. (2014) Single-Walled Carbon Nanotubes Modulate the B- to A-DNA Transition. *J. Phys. Chem. C 118*, 29441–29447.

[26] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992) the Weighted Histogram Analysis Method for Free Energy Calculations on Biomolecules. I. the Method. *J. Comput. Chem. 13*, 1011–1021.

[27] Grossfield, A. WHAM: The Weighted Histogram Analysis Method", Version 2.0.6. http://membrane.urmc.rochester.edu/content/wham.

[28] Becker, O. M., MacKerell, A. D., Jr., Roux, B., and Watanabe, M. *Computational Biochemistry and Biophysics*; CRC Press, 2001.

[29] Hu, J., Ma, A., and Dinner, A. R. (2006) Bias Annealing: A Method for Obtaining Transition Paths de novo. *J. Chem. Phys. 125*, 114101–114101.

[30] Grübmuller, H., Heymann, B., and Tavan, P. (1996) Ligand Binding: Molecular Mechanics Calculation of the Streptavidin-Biotin Rupture Force. *Science 271*, 997–999.

[31] Leech, J., Prins, J. F., and Hermans, J. (1996) SMD: Visual Steering of Molecular Dynamics for Protein Design. *IEEE Comput. Sci. Eng.*

[32] Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y., and Schulten, K. (1997) Molecular Dynamics Study of Unbinding of the Avidin-Biotin Complex. *Biophys. J. 72*, 1568–1581.

[33] Paci, E., and Karplus, M. (1999) Forced Unfolding of Fibronectin Type 3 Modules: An Analysis by Biased Molecular Dynamics Simulations. *J. Mol. Biol. 288*, 441–459.

[34] Aksimentiev, A., and Schulten, K. (2005) Imaging α-Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map. *Biophys. J. 88*, 3745–3761.

[35] Gumbart, J., Khalili-Araghi, F., Sotomayor, M., and Roux, B. (2012) Constant Electric Field Simulations of the Membrane Potential Illustrated with Simple Systems. *Biochim. Biophys. Acta. 1818*, 294–302.

[36] Wells, D. B., Abramkina, V., and Aksimentiev, A. (2007) Exploring Transmembrane Transport Through α-Hemolysin with Grid-Steered Molecular Dynamics. *J. Chem. Phys. 127*, 125101.

[37] Allen, M. P., and Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: New York, NY, USA, 1989.

[38] Brooks, B. R. et al. (2009) CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem. 30*, 1545–1614.

[39] Foloppe, N., and Mackerell, A. D. (2000) All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem. 21*, 86–104.

[40] MacKerell, A. D., and Banavali, N. K. (2000) All-Atom Empirical Force Field for Nucleic

Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution.

J. Comput. Chem. **21**, 105–120.

Chapter 2

On the Possibility of Facilitated Diffusion of Dendrimers Along DNA

We investigate the electrostatics, energetics and dynamics of dendrimer-DNA interactions that mimic protein-DNA complexes as means to design facilitated mechanisms by which dendrimers can slide and search DNA for targets. By using all-atom molecular dynamics simulations, we calculated the free energy profiles of dendrimer-binding around the DNA via umbrella sampling. We also calculated electrostatic interaction maps in comparison to proteins, as well as the dynamical changes induced by DNA-dendrimer interactions via NMR-measurable order parameters. Our results show that for dendrimers to go around DNA there is a free energy barrier of 8.5 kcal/mol from the DNA major groove to DNA minor groove, with a minimum in the major groove. This barrier height makes it unlikely for an all-amine dendrimer to slide along DNA longitudinally, but following a helical path may be possible along the major groove. Comparison of the non bonded interaction energy and the interaction free energy profiles reveal a considerable entropic cost as the dendrimer binds to DNA. This is also supported by the mobility patterns obtained from NMR-measurable order parameter values, which show a decreased mobility of the dendrimer N-H bond vectors in

14

the DNA-binding mode.

This chapter is reproduced in part with permission from [E. Ficici and I. Andricioaei, On the Possibility of Facilitated Diffusion of Dendrimers Along DNA, J. Phys. Chem. B, In Press, 2015. http://dx.doi.org/10.1021/acs.jpcb.5b02090] Copyright [2015] American Chemical Society.

2.1   Introduction

Proteins that bind DNA have critical functions in many genetic transactions, including DNA replication, unwinding, supercoiling, recombination, and damage repair, to name just a few [1]. These vital biological processes rely on the fast search for binding targets along DNA by the proteins. For example, DNA transcription factors recognize a target sequence of just a few DNA base pairs within a couple of seconds; after taking into account the length of DNA and typical transcription factor concentration, this corresponds to an average scanning rate of about 10^5 sites per second [2, 3].

The phenomenology of search and recognition has been puzzling for decades due to the remarkable speed and accuracy involved. Such a fast search creates a speed-stability paradox, given the mutually exclusive energy requirements for fast searching and for high stability of the protein-DNA complex [4]. It is thought that behind the observed speed and accuracy of the facilitated search process lies the ability of the proteins to use a combination of four different modes of searching, i.e., sliding, hopping, jumping, and intersegmental transfer [4–7]. A combination of these search modes will determine the speed with which a protein finds its target. Several studies tackled this issue by taking into account factors such as protein conformational change upon binding [4, 8], complementarity of the DNA and protein charge patterns [9], shape of the protein and its charge distribution [10], and the location of the

target site [11]. Some other studies questioned the existence of speed-stability paradox [12] and argued that this so called paradox is an artifact of continuum models of the protein search process, which are only valid for scanning lengths significantly greater than the size of the binding site [12]. By introducing the scanning length as a critical parameter, the authors developed a discrete-state stochastic model which allows different scalings for search times as a function of DNA length for different search regimes. However, existing studies of protein-DNA interactions have not provided, as of yet, an effective atomistic picture of this nonspecific protein-DNA interaction or of the resulting search process.

Our study herein targets a biomimetic approach to the facilitated diffusion of DNA binding proteins in which we are interested in replacing proteins with dendritic nanoparticles that bind DNA. Related studies [13–15] suggest that the protein-DNA interaction is dominated by electrostatic interactions when the protein is in its nonspecific search mode. Therefore, with the proper adjustment of surface charge density and distribution of a charged nanoparticle, the theoretical possibility exists to mimic the facilitated search process of DNA-binding proteins.

Our chosen model for the proteins are poly-amidoamine (PAMAM) dendrimers [16], which are highly charged functional nano-particles that can be considered as artificial proteins for targeted delivery of drugs and genetic material into cells. Their size, charge, and surface properties can be easily controlled synthetically [17]. Because of their interesting properties, they have potential in areas such as gene transfection, drug delivery, *in vivo* imaging, and biosensing [18–21]. The interaction of dendrimers with DNA has been explored recently [22–32].

A factor that affects the sliding ability of a protein is the shape of the free energy landscape along the DNA. Using detailed theoretical analyses, Slutsky et al. [4] showed that the experimentally reported rapid search is possible when a three-dimensional search process is combined with a one-dimensional search along DNA, and that the optimum search time

is obtained when the time spent on the one-dimensional search is equal to the time spent on the three-dimensional search. While the three-dimensional search time is determined by nonspecific interactions between the protein and the DNA phosphate backbone, the one-dimensional search time may additionally depend on sequence-specific interactions.

The nonspecific binding energy contribution to the total binding energy is estimated to be in the range $10k_BT - 15k_BT$ for proteins (from an Arrhenius-type dependence of the dissociation rate) [4]. Moreover, by assuming that the specific binding energies can be described by a Gaussian distribution with variance σ^2, it can be shown [4] that one-dimensional search can be described by normal diffusion with a diffusion constant that depends exponentially on σ, with σ being a measure of the roughness of the specific binding energy landscape as the protein slides along DNA. Because of this exponential dependence, the diffusion constant decays rapidly with σ. For rapid search to be possible, the specific binding energy landscape must be smooth, such that $\sigma \lesssim 2k_BT$ [4, 33]. For example, experiments done with proteins hOgg1 [34] and p53 [35] reported values of $\sigma < k_BT$ for sliding along DNA, supporting this argument. On the other hand, Slutsky and et al. [4] also estimated that the high stability requirement at the target site requires σ to be at least $5k_BT$ for a genome of about 10^6 base pair long. This requirement for stability is obviously in contradiction with the speed requirement. In order to address this paradox, a two state model was developed [4, 8]. This model suggests a conformational difference between non-specific protein-DNA interactions and specific protein-DNA interactions. It emphasizes the significance of conformational flexibility in nonspecific searches. Several studies showed that the existence of disordered domains in protein structures promotes the speed in the nonspecific search process. For example, the p53 core domain that recognizes a specific DNA sequence has a slower search compared with the disordered C-domain [36]. The disordered tails are also shown to promote intersegmental transfer through a 'monkey bar' mechanism, which can help the protein explore more distant regions and to reduce the redundancy in one dimension [37].

17

Besides the free energy landscape and the existence of flexible regions, another factor that affects the search speed is the electrical charge. For example, the disorder as well as the positive charge on certain domains is shown to increase the non-specific search speed for p53 [36, 38]; p53 is shown to maintain constant contact with DNA as it slides [35]. A considerable difference in the electrostatic potential surface of a DNA binding protein when it is in complex with nonspecific DNA sequence versus when it is in complex with specific DNA sequence has been found [39]. In the former case, there are negatively charged pockets scattered among positively charged regions on the DNA binding face of the protein. In the latter case, however, positive charges are observed to focus more on the DNA binding regions. This shows that charge redistribution plays an important role in distinguishing between specific and nonspecific complexes.

Due to the low affinity of nonspecific protein-DNA complexes, there are available only a few nonspecifically-bound protein-DNA structures, including EcoRV [40], BamHI [39], BstYI [41], and Lac Repressor [15]. A structural analysis of these complexes indicates that the interactions are mainly between negatively charged phosphate groups and positively charged protein domains [14]. In addition, a comparison of salt dependence of the interaction, as well as hydration differences between specific and nonspecific complexes show that nonspecific interactions are dominated by electrostatic interactions [14].

With the lessons learned from DNA-sliding proteins in mind, we seek to find if and how nanoparticles can be used, instead of proteins. Functionalized nanoparticles hold great promise for various biomedical applications, which include their potential use as biosensors, cell-labeling agents, imaging or contrast agents, artificial catalytic sites, and DNA/protein microarrays, to name just a few [42, 43]. They can also be used in targeted gene or drug delivery [17–19, 21, 44–47]. For our purposes, we propose that polyamidoamine (PAMAM) dendrimers [16, 17, 19], which are charged nanoparticles, can serve as first model systems to mimic the electrostatic nature of protein-DNA interactions.
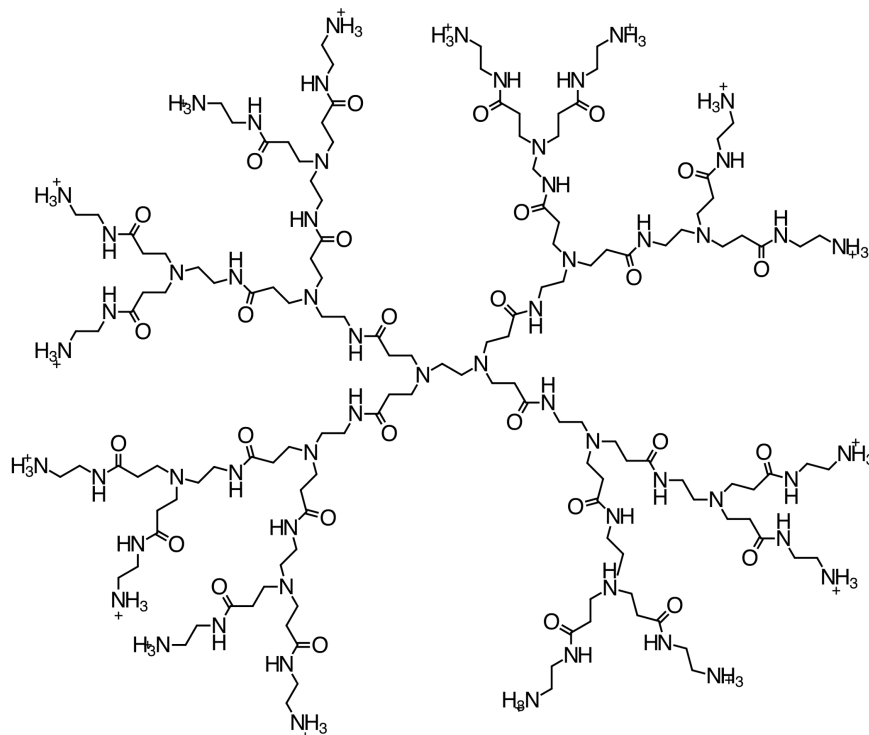
Figure 2.1: Structure of generation 2 dendrimer. A generation number is defined as the number of branching points.

PAMAM dendrimers (Figure 2.1), often referred to as "artificial proteins" [20], resemble proteins in their chemical structure. Their size and surface chemistry can be controlled and their toxicity can be reduced by adjusting the surface charge size or core chemistry [44, 48]. They can be conjugated with different surface modifiers such as amino acids [49] and in principle they can be used to mimic the non-specific search along DNA, the firm binding to a target sequence and possibly subsequent catalytic events. For many such possible applications of dendrimers, it is crucial to explore the dendrimer-DNA interactions in general, and the diffusion properties of dendrimers along DNA in particular, in atomistic detail.

Cationic dendrimers of generation 5 (G5) or larger are shown to bind to DNA irreversibly [24, 26] and can condense DNA. On the other hand, there is evidence from computational studies that cationic dendrimers of generation 3 slightly bend DNA locally, but do not condense it [23]. Moreover, sliding of third generation (G3) and fourth generation (G4) dendrimers along DNA, until they find an ideal binding location, has been reported in a

molecular dynamics study [25]. Considering the fact that G3 is highly flexible and positively charged (at physiological pH), it can be an ideal candidate to mimic the nonspecific search process along DNA. Moreover, importantly, the energy landscape it sees along and around DNA can be modified by adjusting its terminal charges. Therefore, the flexibility of a lower generation dendrimer combined with optimum terminal charges have the potential to overcome the strong attraction due to DNA that would lead to balancing association and dissociation, which is necessary for nonspecific search mechanism.

In a previous study, we computed the binding free energy profile along the radial direction for an all-amine terminated G3 dendrimer interacting with DNA; the profile exhibited a well-depth of -13.5 kcal/mol [23]. When this value is compared with the nonspecific binding energy required for facilitated search of DNA binding proteins (estimated to be in the $10k_BT - 15k_BT$ range, or $5.93 - 8.895$ kcal/mol at room temperature [4]) G3 has about two times higher a nonspecific binding energy than typical proteins. A similar study done with G3 (surface charge $+32$) and higher generation dendrimers G4 (surface charge $+64$) and G5 (surface charge $+128$) reported that the binding free energy increases almost linearly with the generation [27]. While a dendrimer of such charge size is unlikely to dissociate from DNA in the radial direction, we should also note that the free energy depends on the tension on DNA. It has been shown that as the tension on DNA increases, the free energy barrier decreases for a given dendrimer size and charge [23]. This information can be used as a starting point and the linear dependence can be used to estimate an optimum charge size that would facilitate diffusion along DNA. For this purpose, we setup to explore the free energy profiles along longitudinal sliding pathways. The organization of the rest of this paper is as follows: In the next section (Theory and Methods), we describe the system setup and simulation details after which we provide a theoretical framework for potential of mean force and NMR order parameter calculations. We then continue with a Results section, in which we compare electrostatic potential maps for some DNA binding proteins and dendrimer interacting nonspecifically with DNA. We also present potential of mean force and

NMR order parameter calculation results for dendrimer interacting with DNA in the Results section. We end with a Concluding Discussion.

2.2 Theory and Methods

2.2.1 Simulation Methods

In all simulations, CHARMM 27 all-atom force field for nucleic acids [50, 51] was used with the CHARMM software (Chemistry at HARvard Macromolecular Mechanics), version c34b2 [52–54]. Langevin dynamics is used with a friction coefficient of 10 ps^{-1} applied on non-hydrogen atoms and an implicit solvent model (GBMV2) [55] is used. The accuracy of generalized Born methods rely highly on the accuracy of the Born radii, i.e. the distance of a charge location from the solvent boundary in a given molecule. GBMV2 method is shown to give highly accurate Born radii such that the relative error in total electrostatics solvation energy is shown to be $\leq 1\%$ between GBMV and Poisson theory [56, 57]. The nonbonded cutoff distance was 21 Å, with a switching function from 16 Å to 18 Å. The SHAKE algorithm [58] was used to constrain the distance of covalent bonds to hydrogen atoms, enabling a 2 femtosecond time step. A third generation PAMAM dendrimer, G3, with all-amine terminations was generated using parameters obtained as described in C. V. Kelly et al. [59]. Double stranded B-DNA, 48 base pairs in length and a repeat of the CGAT sequence was generated using the program NAB [60]. This sequence was used as a model for a "generic" DNA sequence [61] that mixes all possible base pairings. DNA and G3 were initially prepared separately as follows. Each one was first minimized in vacuum using 1000 steps of steepest descent and 2000 steps of adopted basis Newton-Raphson (ABNR). This was followed by 2000 steps of ABNR in implicit solvent. It was then heated for 50 ps while a harmonic constraint with a force constant of 1 kcal/mol/Å^2 was applied on all atoms.

This was followed by an equilibration for about 2 ns with all the harmonic constraints were gradually removed. For DNA only, after the removal of constraints, the relative distances between complimentary base pairs at each end were restrained with a harmonic potential by using NOE-type restraints in order to prevent fraying.

The DNA axis was defined by the line connecting the center of geometry of the (DNA1)29th-(DNA2)68th base pairs and the center of geometry of the (DNA1)19th-(DNA2)78th base pairs. Three separate systems were prepared by combining the dendrimer and DNA at three distances of 25 Å, 35 Å, and 45 Å, which we label as $r25$, $r35$, and $r45$, respectively. Intermolecular distances were calculated from the dendrimer center of mass (COM) to the middle of the DNA axis. The longest axis of the dendrimer was placed to be parallel to the DNA axis. For each system, the dendrimer was then rotated rigidly around the defined DNA axis by 10 degrees to obtain multiple windows. Rotation was done in such a way that the same side of the dendrimer was facing the DNA. Figure 2.2 shows a snapshot of system $r45$ where the dendrimer is positioned at a dihedral angle value of 180 degrees defined by the dummy atoms and the dendrimer center of mass (P1:P2:P3:G3 COM). The entire range of the dihedral values (0–360°) was sampled. For each system, after a brief equilibration period of 40 ps, DNA atoms were harmonically restrained to their most recent positions with a harmonic force constant of 100 kcal/mol/Å^2 in order to maintain DNA shape and orientation with respect to the dendrimer. Additional windows are added as needed for PMF calculations which resulted in 61 windows for system $r25$, 129 windows for system $r35$, and 155 windows for system $r45$.

2.2.2 Potential of Mean Force

The umbrella sampling method [62] was used to obtain a potential of mean force (PMF) [63] around the DNA. A reaction coordinate was defined between two states, then the reaction
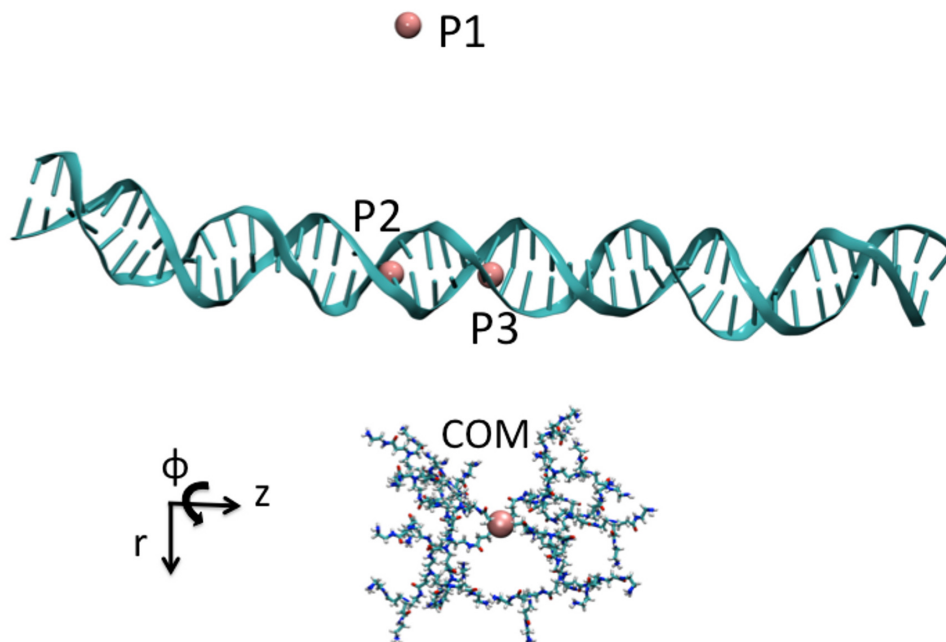
Figure 2.2: The dendrimer is shown at a pseudo dihedral value of $\phi = 180°$ and a distance of $r = 45$ Å, defined by three points and the dendrimer center of mass: P1:P2:P3:G3 COM.

coordinate was divided into multiple overlapping windows and the system was restrained around each window with a harmonic biasing force. This helps to sample all regions, including the higher energy regions which are difficult to sample within the accessible time scale of molecular dynamics simulations.

The biasing potential

$$U_b(r, \phi, z) = \frac{1}{2}k_r(r - r_0)^2 + \frac{1}{2}k_\phi(\phi - \phi_0)^2 + \frac{1}{2}k_z(z - z_0)^2 \tag{2.1}$$

was used, where r is the radial distance between DNA axis and the dendrimer center of mass, z is the displacement of the dendrimer center of mass along the DNA axis, and ϕ is the pseudo dihedral angle as defined by the three fixed points and the dendrimer center of mass (See Figure 2.2). For all PMF simulations, $z_0 = 0$ and r_0=25 Å, 35 Å or 45 Å. The force constants k_r and k_z are 500 kcal/mol/Å^2. We used various values for the force

constant k_ϕ which varied between $2.5, 10, 25, 50, 100, 200$, and 250 kcal/mol/rad^2, depending on the overlap of the probability distribution of the reaction coordinate between neighboring windows. The length of the simulations for the different windows varied between 4 ns and 25 ns because of the difference in the force constant values. The weighted histogram analysis method (WHAM) [64, 65] was used to unbias the simulation results, with multiple restraints handled through a minor modification in the WHAM code.

In order to investigate the sliding of a PAMAM dendrimer along DNA, the free energy profile along one DNA turn was obtained. In our previous study [23], we showed that the interaction of the dendrimer with DNA bases is not significant and is nonspecific. Here, we are mostly interested in the change in interaction due to minor and major groove variations and not due to larger DNA conformational differences. Therefore we want to avoid changes in DNA conformations, such as untwisting, bending, groove width differences etc. along sliding pathways. This can be achieved by taking advantage of the helical symmetry of DNA. Accordingly, for an ideal DNA, the sliding surface along a line parallel to DNA axis is periodic with a period of one DNA turn. The landscape that would be seen for longitudinal sliding should be identical to the surface that would be seen in the case of rotation around the DNA on a plane orthogonal to DNA axis. A similar approach has been previously used in calculating the variation of electrostatic binding energy with respect to the minor and major groove interactions of a BamHI-DNA complex [66]. Therefore, we define our reaction coordinate along a circular path of constant radius around DNA axis on a plane perpendicular to DNA axis. This reaction coordinate is defined to be a pseudo dihedral angle with respect to the fixed points represented in Figure 2.2.

2.2.3 Calculating NMR Order Parameters

The MD simulations were used to compute NMR-based order parameters, which have been previously used to characterize the dynamics of biomolecules [67–69]. According to the Lipari-Szabo model [70], the degree of spatial restriction of motion and the rate of that motion for subnanosecond motion was described by the model-free NMR order parameters S^2 and τ_e respectively, where S^2 is a generalized order parameter and τ_e is an effective correlation time. Order parameter values can further be used to estimate the configurational entropy change due to bond vector motions [71].

NMR order parameter S^2 for the dendrimer N1-H5 bond vectors were calculated from the molecular dynamics trajectories. The equilibrium expression [72, 73],

$$S^2 = \frac{1}{2}[3\sum_{i=1}^{3}\sum_{j=1}^{3}\langle\mu_i\mu_j\rangle^2 - 1] \tag{2.2}$$

was used, where μ_i represent the Cartesian coordinates of the normalized internuclear vector after alignment of the MD snapshots with respect to the dendrimer core carbon and nitrogen atoms. Alternatively, S^2 was also calculated from the tail values of internal correlation functions using the following relation assuming internal and overall motions are not correlated and internal fast and slow motions are not correlated [67, 68]:

$$C_i(t) = S^2 + (1 - S_f^2)e^{-t/\tau_f} + (S_f^2 - S^2)e^{-t/\tau_s} \tag{2.3}$$

Here, $S^2 = S_f^2 S_s^2$ corresponds to the tail value of the correlation function and f and s indicate "fast" and "slow" motions, respectively.

2.2.4 Poisson-Boltzmann Calculations

The PBEQ module in CHARMM was used to solve the linearized Poisson-Boltzmann equation numerically [74]. The dielectric boundary was determined from the molecular surface with a probe radius of 1.4 Å and atomic Born radii values calculated from free energy calculations for standard amino acids and DNA [75, 76] were used. A grid spacing of 1 Å before focusing and 0.5 Å after focusing was used. A dielectric constant of 1 was used for the molecules, and a value of 80 was used for the solvent. Salt concentration was 0.15 M. Images were obtained by using PyMOL [77].

2.3 Results

2.3.1 Electrostatic Potential Surfaces

As Honig and Nicholls [78] point out, the surface electrostatic potential is influenced not only by the charge size or distribution near the molecular surface, but also by surface geometry. For example, narrow regions or clefts (as found in many DNA binding proteins or DNA minor groove) are shown to enhance the electrostatic potential [79], an effect known as "electrostatic focusing" [79]. Therefore, in order to understand the role of electrostatics in protein sliding, it is useful to compare the surface electrostatic potential map of the dendrimer-DNA complex to proteins complexed with DNA nonspecifically. We obtained electrostatic potential surface maps for the following nonspecific protein-DNA structures: BamHI (PDB ID: 1ESG) [39], BstYI (PDB ID: 2POJ) [41] EcoRV (PDB ID: 2RVE) [40], and Lac Repressor (PDB ID: 1OSL) [15].

As shown in Figure 2.4, a large positive charge concentration is evident on the DNA binding surface of proteins. While positively charged regions are dominant on these surfaces, there

are some smaller pockets of negatively charged regions scattered among positive regions as well. It is possible that the existence of these negatively charged regions may be needed to help the protein dissociate easily from a nonspecific DNA site.

For comparison with these proteins, we also generated the electrostatic potential surface map for the G3-DNA system, using the same method as described above for the proteins (See Figure 2.4(e)). Based on the comparison, it can be argued that although a high attraction is expected between the positively charged dendrimer terminals and the negatively charged DNA sugar phosphate backbone, the repulsion due to the observed negative regions located on the branches may reduce this attraction to some extent. Another interesting property of the dendrimer is that since its branches are very flexible, the branches are observed to repel each other when they get closer. This effect also prevents the dendrimer to stick to the DNA. Therefore the theoretical possibility exists for the dendrimer to mimic the sliding motion of proteins along DNA, likely following a helical path.

2.3.2   Interaction Energy

To map the energy profile for the interaction between the two molecules in three dimensional space, the interaction energy was calculated along various radial as well as angular coordinates as follows. First we kept the distance constant at 45 Å, and calculated the interaction energy map around the DNA axis for two different planes orthogonal to the DNA axis positioned at half DNA turn apart. The interaction energy was calculated as the sum of electrostatic and van der Waals energies:

$$E_{int} = E_{elect} + E_{vdw} \tag{2.4}$$

(a) BamHI

(b) BstYI

(c) EcoRV

(d) LacR

(e) G3
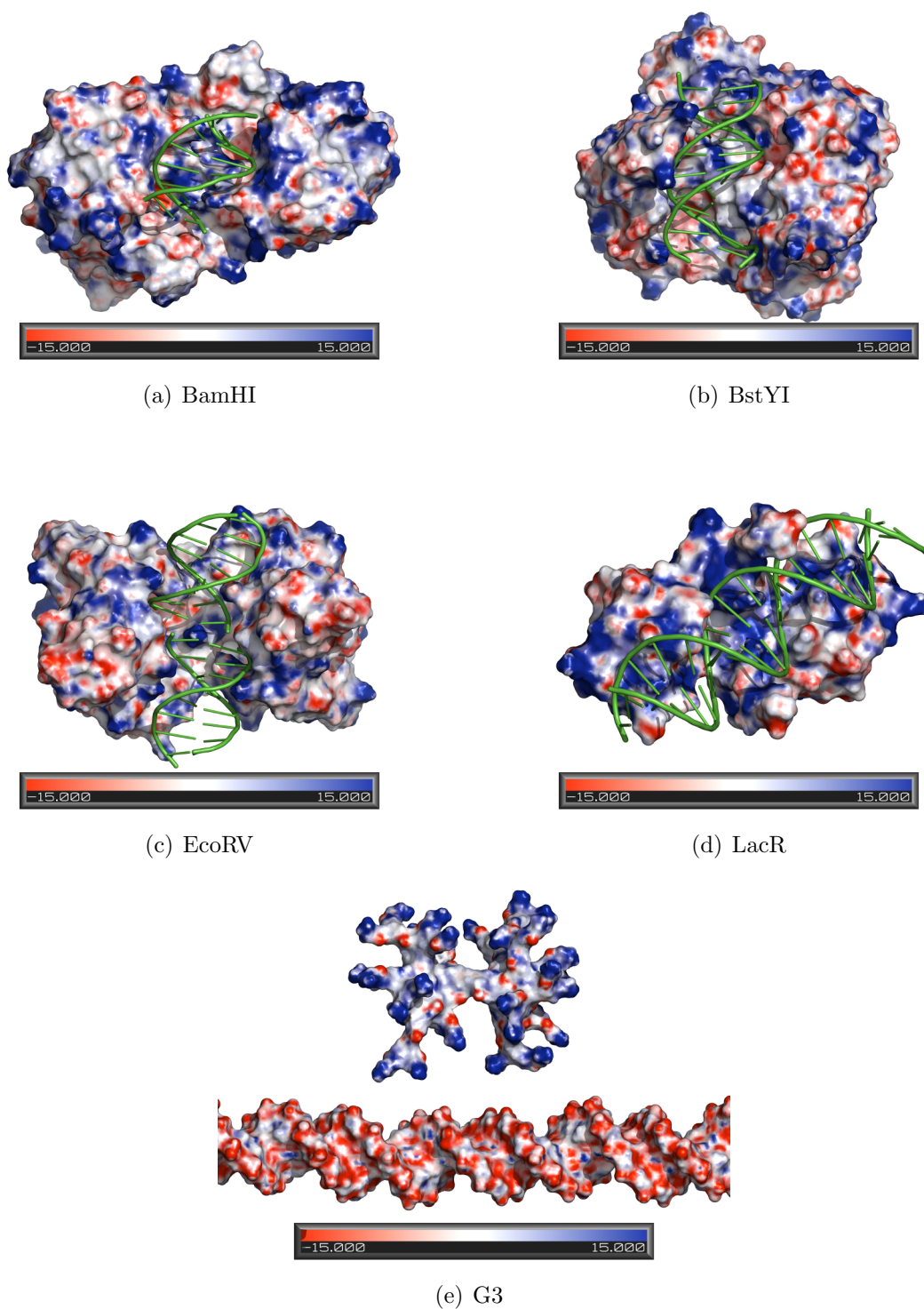
Figure 2.4: Electrostatic potential surfaces as calculated with PBEQ module of CHARMM shown for proteins: 2.4(a) BamHI (PDB ID 1ESG); 2.4(b) BstYI (PDB ID 2POJ); 2.4(c) EcoRV (PDB ID 2RVE); and 2.4(d) Lac Repressor (PDB ID 1OSL) complexed with DNA nonspecifically; and 2.4(e) G3. Images were made by PyMOL. Negative regions are shown with red, and positive regions are shown with blue. Heat map color scale is from -15.0 kT/e to 15.0 kT/e.

Figure 2.3 shows the variation of the interaction energy around DNA at an intermolecular distance of 45 Å at two different planes orthogonal to DNA axis: the $z = 0$ plane (red), which is the plane that intersects the DNA axis orthogonally at its center, and the $z = 17$ plane (blue), which is separated from the former by 17 Å, i.e., half DNA turn. The sugar-phosphate group is marked with arrows to distinguish minor and major groove positions on the $z = 0$ plane. Note that the angle measure on the minor groove side is about 145°, consistent with typical B-form values. The nonbonded interaction, which is predominantly electrostatic in nature, is significantly more favorable (by about 100 kcal/mol) when the dendrimer center of mass is on the minor groove. This is consistent with previous electrostatic potential calculations of B-form DNA, which show that the electronegative potential is enhanced in the narrow minor groove due to electrostatic focusing [79, 80]. Also noticeable is a shift of the pattern when the dendrimer is placed on the $z = 17$ plane. This is expected due to the helical symmetry of the DNA structure, verifying that the observed interaction pattern is due to the positions of negatively charged phosphates on DNA backbone.

(a)   (b)
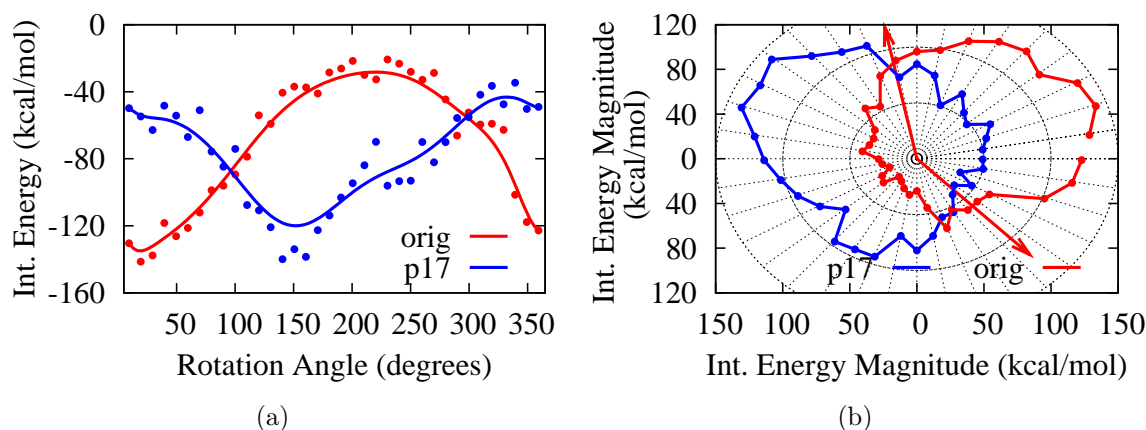
Figure 2.3: Interaction energy plots for G3-DNA separation of 45 Å at two different planes orthogonal to DNA axis: the $z = 0$ plane (red) and the $z = 17$ Å plane (blue) in cartesian 2.3(a) and polar 2.3(b) representations. In 2.3(a) smoothing line is used to guide the eye. In 2.3(b) the positions of sugar-phosphate group is marked with arrows to distinguish minor and major grooves on the $z = 0$ plane. Note that the radius is proportional to the magnitude of the interaction energy.

Interaction energies for intermolecular distances of 25 Å and 35 Å were also calculated on the $z = 0$ plane to gauge how the energy profile changes as the dendrimer gets closer to DNA. Figure 2.5(a) shows that while the interaction energy values are much lower at distances closer to DNA (as expected), the shape of the energy profile in angular directions is maintained. Table 2.1 lists the values of the interaction energy and its components at minor and major grooves for three intermolecular distances. Finally, starting from a distance of 45 Å, the dendrimer is pulled towards the DNA axis along the radial direction with a constant force of magnitude 10 pN for various dihedral angle values (See Figure 2.5(b)). These interaction energy profiles around the DNA and in the radial direction show that the difference between dendrimer-minor groove interaction and dendrimer-major groove interaction is amplified as the dendrimer gets closer to DNA. The difference in the DNA-G3 interaction energy at the minor groove and major groove is about 100 kcal/mol for $r45$, 300 kcal/mol for $r35$, and 600 kcal/mol for $r25$.
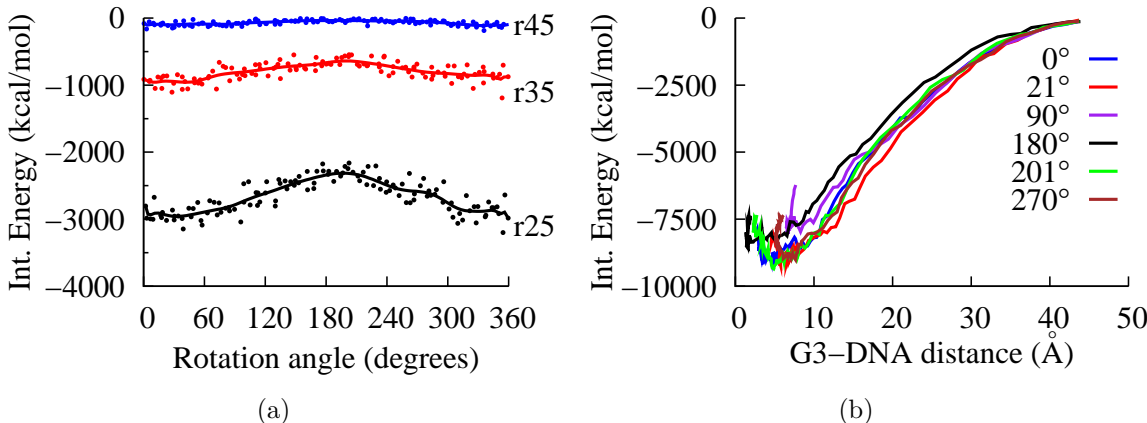
Figure 2.5: 2.5(a) A comparison of interaction energy calculated at three G3-DNA separation around DNA axis. 2.5(b) Dependence of interaction energy on radial distance for various dihedral angle values.

The closest distance at which the dendrimer approaches the DNA when the dendrimer is restrained to the $z = 0$ plane and at various dihedral values when no distance restraint applied, varied between 17 Å and 28 Å over a simulation of length 4 ns (data not shown). Therefore a distance of 25 Å can be safely considered to represent a binding mode.
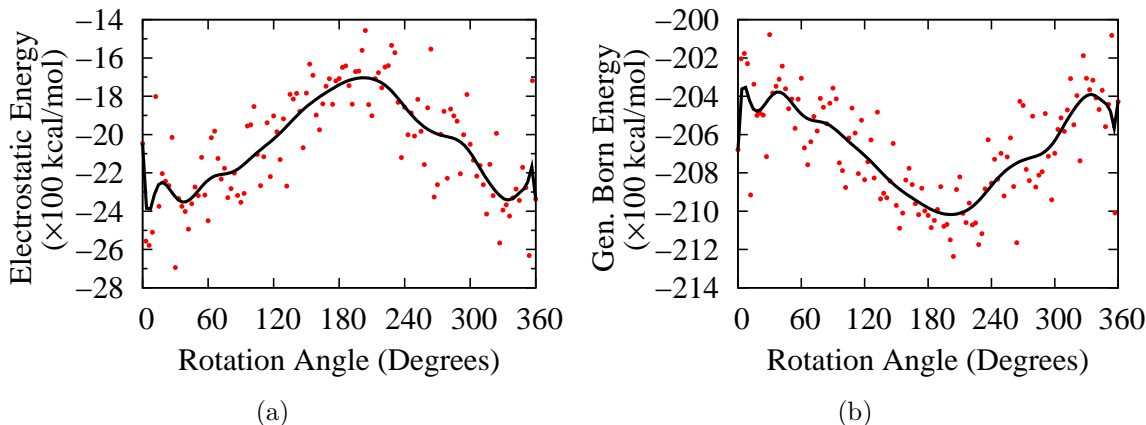
Figure 2.6: Electrostatic 2.6(a) and Generalized Born 2.6(b) components of total energy for G3-DNA separation of 25 Å

Figure 2.6 shows a comparison of the electrostatic component of the total energy value (Figure 2.6(a)) for the dendrimer-DNA system with the electrostatic contribution to the generalized Born solvation energy as obtained from the GBMV2 module [55] in CHARMM (Figure 2.6(b)); anticorrelated behavior is observed. Total electrostatic, van der Waals and generalized Born solvation energy contributions are also calculated for r35 and r45, which are reported in Table 2.2. The generalized Born solvation energy is a continuum approximation to the electrostatic contribution to solvation energy [55]. According to Figure 2.6(b), solvation energy at the minor groove is higher than the solvation energy at the major grove. In other words, in the region where interaction energy, which is predominantly electrostatic in nature, is more favorable (minor groove), solvation energy is less favorable. This opposite behavior of solvation energy and electrostatic energy is in agreement with Honig et al. in that the charged or polar groups are a source of destability for molecules or complexes in aqueous solutions [78]. This is because desolvation costs for binding of opposing charges are larger than Coulombic attractions.

In another computational study of dendrimer-DNA interactions performed in explicit solvent, the authors reported a total electrostatic energy of −22400 kcal/mol for a complex consisting of G3 and a 38 basepair DNA including the interactions due to ions and water [25]. In our

Table 2.1: Contributions of electrostatic and van der Waals energies to interaction energy at minor and major grooves.

	Electrostatic (kcal/mol)	vdW (kcal/mol)	Interaction En. (kcal/mol)
r25			
Minor Groove	−2787.17	−17.5468	−2804.72
Major Groove	−2344.03	−16.5974	−2360.63
r35			
Minor Groove	−937.139	−0.642	−937.139
Major Groove	−624.75	−0.449	−625.199
r45			
Minor Groove	−140.382	−0.028	−140.410
Major Groove	−28.196	−0.0016	−28.194

implicit solvent simulations, this corresponds to the sum of the total system electrostatic energy and the electrostatic contribution to generalized Born energy, which are listed in Table 2.2. Since the DNA we use in our study is longer, i.e. 48 basepairs, we expect the total electrostatic energy in our system to be slightly lower. As expected, for an interaction distance of 25 Å, we find −22720 kcal/mol and −22733 kcal/mol for major groove and minor groove binding modes, respectively. While a one-to-one comparison cannot be made due to the different number of DNA atoms in the two studies, this close agreement in total electrostatic energy values suggest that the use of implicit solvent model used in our study yields results consistent with explicit solvent simulations.

Another similar molecular dynamics study which compared the non-bonded interactions of a charged nano-particle with DNA minor groove versus major groove reported a preference for the major groove [81]. In that study, gold nanoparticles functionalized with thiolated alkane ligands carrying a small charge of +6 (AuNP-NH3) were used. The authors reported an electrostatic energy difference of 658.36 kJ/mol between minor and major groove electrostatic interactions, which corresponds to 157.2 kcal/mol overall or 26.2 kcal/mol per amine charge. Our results for a distance of 25 Å show a difference of 641.36 kcal/mol overall or 20.04

Table 2.2: Contributions of electrostatic and van der Waals energies to total system energy and electrostatic contribution to Generalized Born solvation energy.

	Electrostatic	vdW	Total Nonbonded	Gen. Born Electr. Solvation Energy
	(kcal/mol)	(kcal/mol)	(kcal/mol)	(kcal/mol)
r25				
Minor Groove	-2325.35	-464.2	-2789.55	-20407.8
Major Groove	-1683.99	-461.66	-2145.65	-21036
r35				
Minor Groove	-415.27	-442.772	-858.04	-22304.7
Major Groove	-126.82	-444.38	-571.21	-22588.4
r45				
Minor Groove	526.5	-446.88	79.62	-23298.8
Major Groove	581.952	-452.75	129.2	-23251

kcal/mol per amine charge, which is comparable to the above results in magnitude, but contradict them in terms of the preferred region. While we did not observe much difference between the van der Waals interactions for both modes of binding, the gold nanoparticle study reported a difference of 330.08 kJ/mol in the van der Waals interactions between the two modes of interactions, which favor the minor groove. The difference in nano-particle composition, size and flexibility may be a factor in explaining this difference in the two studies. Another explanation of this difference may be that in the mentioned study, DNA was free to move, resulting in DNA bending and adjustment of the groove width. In our study, on the other hand, DNA is kept straight and the widths of the minor and major grooves are maintained.

Based on the results mentioned above, we infer that for a longitudinal sliding motion along DNA to be possible, a free energy barrier is expected between the minor groove and major groove positions. Its height will determine whether or not the barrier can be scaled by thermal fluctuations, resulting in a random or, respectively, a helical sliding motion.

While the calculations reported above are useful to understand binding vs. sliding, the interaction energy values mentioned above are only enthalpic contributions to the free energy profiles. In order to include the entropic contributions, further calculations of the potentials of mean force around DNA are needed.

2.3.3   Potentials of Mean Force

The potential of mean force around DNA for G3-DNA center of mass separations of 25 Å, 35 Å, and 45 Å along the pseudo-dihedral angle reaction coordinate are shown in Figure 2.7. Contrarily to the interaction energy patterns described above, the free energy minimum was located in the major groove. While the free energy is almost smooth at a distance of 45 Å, the preference for the major groove starts to be felt at a distance of 35 Å with a barrier of about 2 kcal/mol. When the dendrimer is at a distance of 25 Å, the free energy barrier gets close to 8.5 kcal/mol.

For comparison of our results with protein-DNA interactions, we need to consider studies of PMF that explore the free energy profiles along possible sliding pathways along DNA. For example, Marklund et al. [82] explored the PMF for LacI head domain nonspecifically interacting with DNA along a helical reaction coordinate and in the radial direction. In the radial direction, they found the barrier to be $12k_BT$ (or 7.15 kcal/mol at room temperature), which is within the theoretically estimated range for facilitated diffusion of proteins [4]. This is almost half of the value we reported for G3 (13.5 kcal/mol) in a previous study [23]. Marklund et al. also obtained the PMF along the helical direction, and they found it to be periodic, with a periodicity of one base pair, and a barrier of no larger than 3.5 k_BT (or 2.1 kcal/mol). This value is close to the estimated value required for facilitated diffusion, which is estimated to be less than $2k_BT$ [4].

Based on the PMF results for G3, it is unlikely that G3 will dissociate from DNA due to the
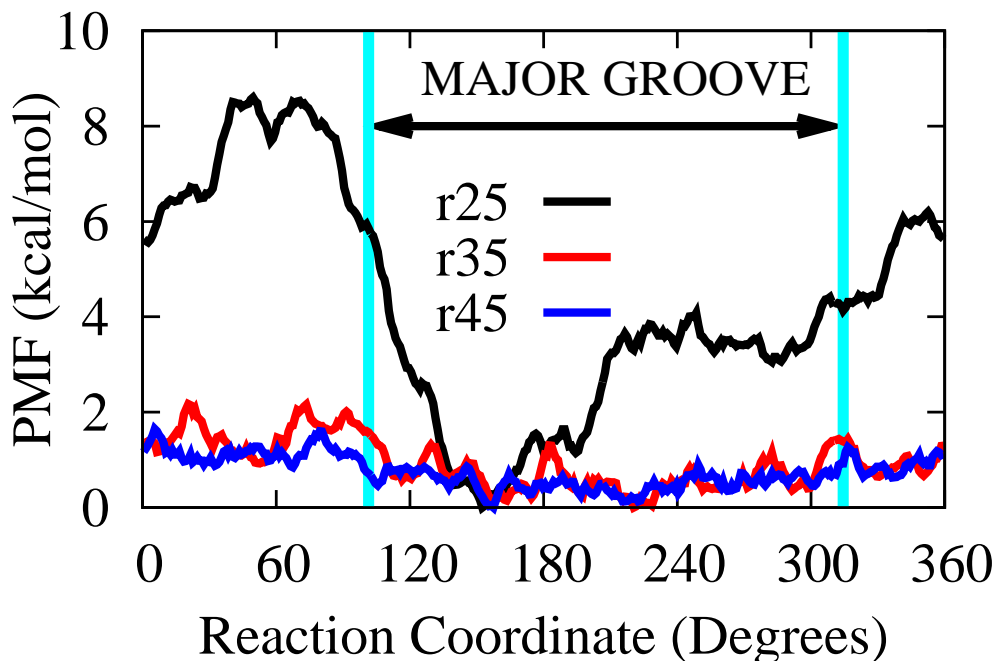
Figure 2.7: Potentials of Mean Force along circular reaction coordinates around DNA axis for dendrimer-DNA distances of 25 Å, 35 Å, and 45 Å.

high barrier in the radial direction. In the longitudinal direction, we identified the major groove as the region of more favorable interaction, with a barrier of 8.5 kcal/mol, which is high enough to make the motion in this direction also unlikely. There exists the possibility, however, that the dendrimer may follow a helical path, similar to some DNA binding proteins such as LacI [82], human hOgg1, BstYI. MutY, E. coli MutM M74A, and BamHI, etc. [83]. Free energy profiles along helical coordinates will need to be determined to verify this. In any case, we believe that based on the assumptions mentioned earlier about the almost linear dependence of PMF barrier along radial direction, if the dendrimer charge size is reduced by half, this may result in cutting the barriers in both radial direction and sliding direction by half, which in turn may result in values close to the ones reported for LacI above.

We note that our study uses implicit solvent and therefore, does not include any structural effects of water or salt. An earlier study showed the important effects of ordered water layers in dendrimer-DNA interactions [22]. Due to the dominant electrostatic nature of the

interactions, the role of salt will also be important. However, these effects can be assumed to have effects of similar magnitude for major and minor grooves such that the relative interactions reported here are still valid.

2.3.4   NMR Order Parameters

In order to understand the dendrimer dynamics as the dendrimer interacts with DNA, we calculated NMR generalized order parameters S^2 for the dendrimer N-H bond vectors at multiple positions around DNA using both the equilibrium expression and the tail value of the autocorrelation functions (See Methods). The tail values were calculated from the average of the last 300 ps of the trajectory. Sample correlation functions are shown in the Supporting Information (Figure S1) for system $r25$. Except for residues 7 and 16, all correlation functions converge on the nanosecond time scale. This is because dendrimer motions are well sampled within this time scale.

According to the model-free approach [70], the lower S^2 value for a given residue corresponds to more mobile motion, whereas a higher S^2 value corresponds to more restricted motion. For a protein, the residues closer to the N-terminal and C-terminal are more mobile while the residues in the interior regions are more restricted. This results in a "frowning" pattern in which the residues closer to the two termini have lower S^2 values and the residues closer to the central regions have higher S^2 values. In contrast, dendrimers have a different pattern of mobility, in which all termini are flexible, not just the C- and N-termini in the protein case. Figure 2.15(b) shows the pattern of order parameter values, which are color coded according to generation as shown in Figure 2.15(a) for one prototypical example. More S^2 plots for various positions of dendrimer around DNA are included in the Figures $2.9 - 2.14$. Accordingly, N-H bonds in the innermost layer (generation 0), have consistently higher S^2 values compared with bonds in the outer layers (generations $1 - 3$). That is, the motion is

more restricted in the inner layers than in the outer layers, as expected from the branching geometry. Moreover, S^2 values approach zero for the terminals showing unrestricted motion.

When the mobility patterns at different dendrimer-DNA interaction distances are compared, it is observed that S^2 values for an interaction distance of 25 Å are usually higher than S^2 values for interaction distances of 35 Å and 45 Å. This shows that mobility is more restricted as the dendrimer gets closer to DNA and as the terminals start to make more contacts with the DNA. This behavior is consistent with that of many (but not all) proteins, which are known to show a rigid behavior upon binding [84–86], resulting in higher S^2 values.

NMR order parameters (S^2) for dendrimer N-H bonds at various dendrimer positions around DNA, as specified by pseudo dihedral angle values (ϕ), are shown in Figures $2.9 - 2.14$ for systems $r25$, $r35$, and $r45$, respectively. Color coding is based on dendrimer generation (See Figure 2.15). S^2 values calculated from equilibrium expression (Figures 2.10, 2.12, and 2.14) and from the tail values of N-H bond autocorrelation functions (Figures 2.9, 2.11, and 2.13) are given below. Sample autocorrelation functions for select residues can be seen in Figure 2.8.
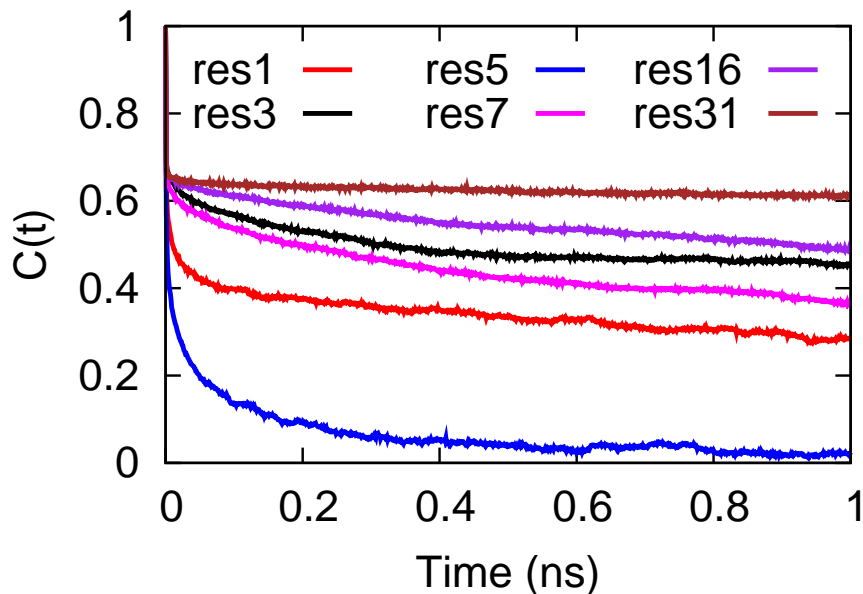
Figure 2.8: Sample correlation functions.

(a) $\phi = 24°$

(b) $\phi = 72°$

(c) $\phi = 102°$

(d) $\phi = 150°$

(e) $\phi = 210°$

(f) $\phi = 225°$

(g) $\phi = 309°$
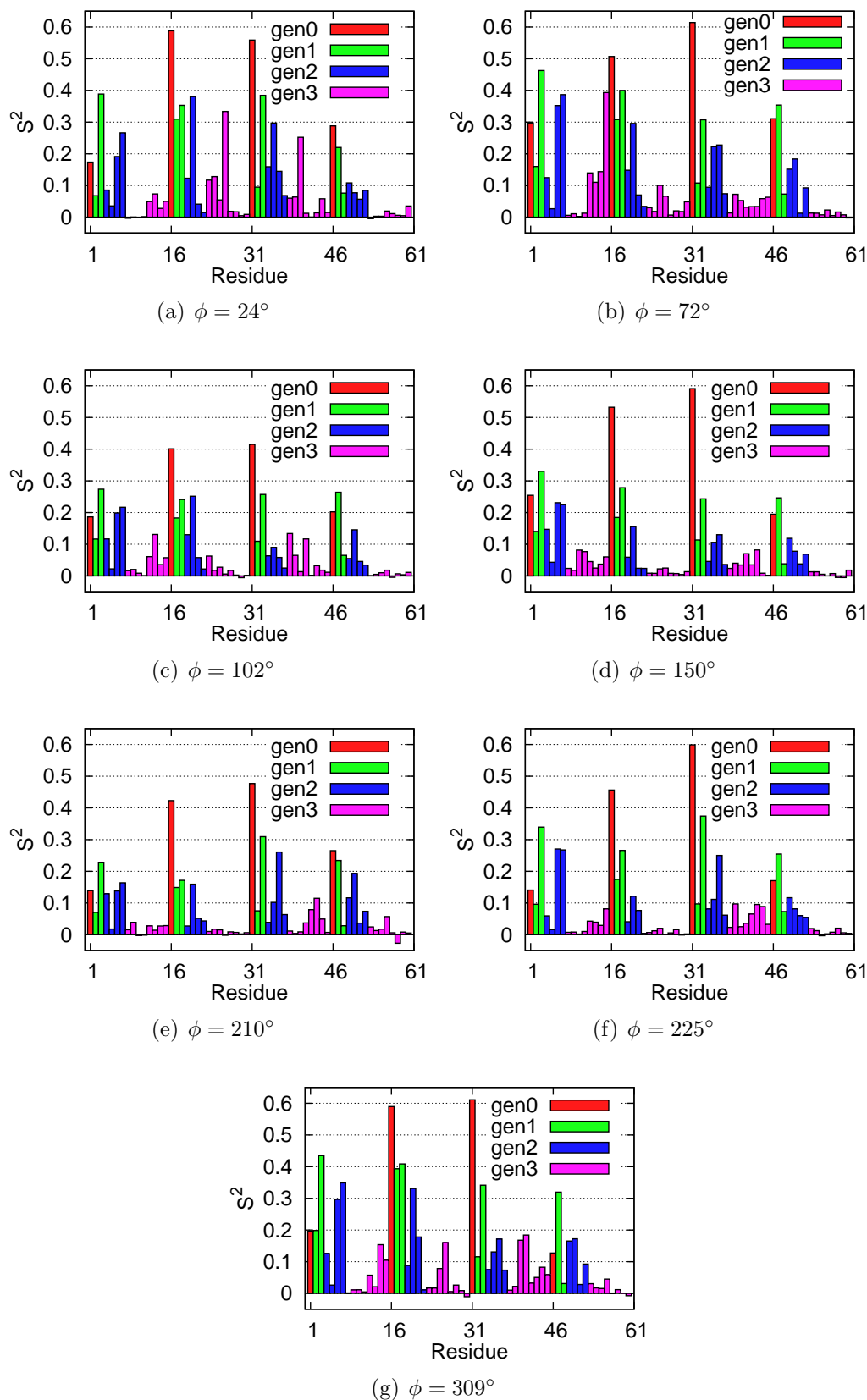
Figure 2.9: S^2 values as calculated from the tail of autocorrelation functions for system $r25$.

(a) $\phi = 24°$

(b) $\phi = 72°$

(c) $\phi = 102°$

(d) $\phi = 150°$

(e) $\phi = 210°$

(f) $\phi = 225°$

(g) $\phi = 309°$
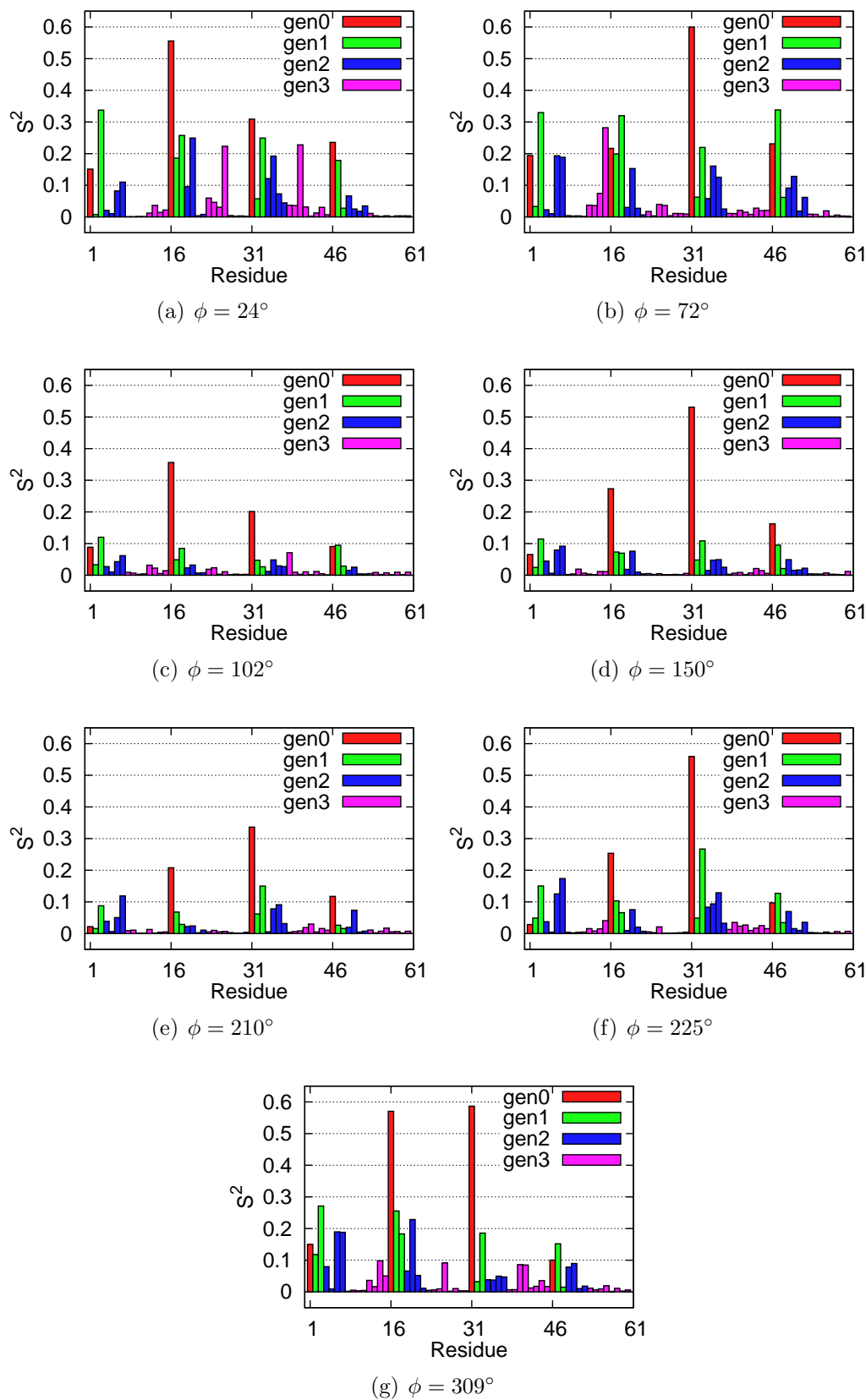
Figure 2.10: S^2 values as calculated from the equilibrium expression for system $r25$.

(a) $\phi = 27°$

(b) $\phi = 72°$

(c) $\phi = 105°$

(d) $\phi = 150°$

(e) $\phi = 210°$

(f) $\phi = 225°$

(g) $\phi = 309°$
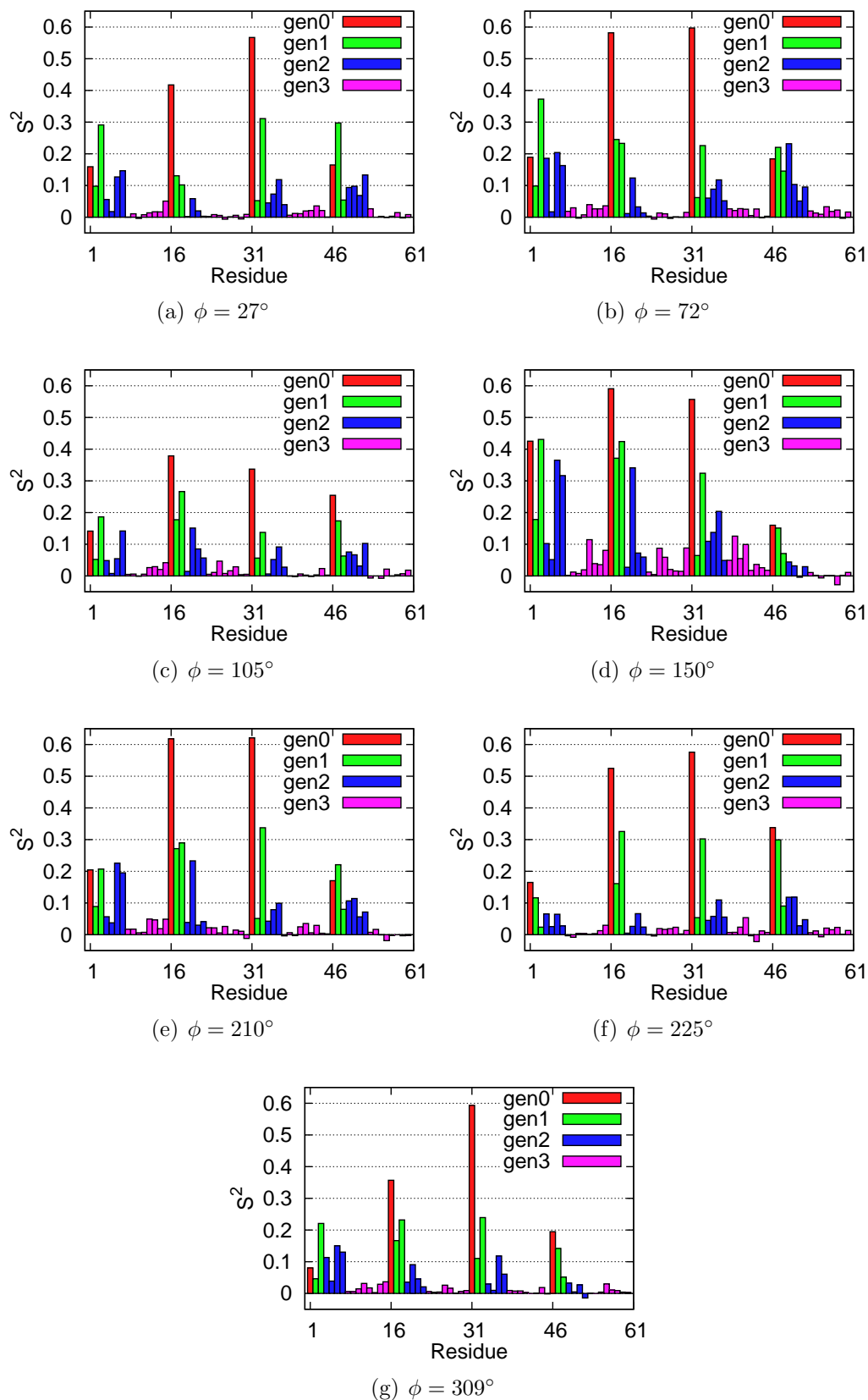
Figure 2.11: S^2 values as calculated from the tail of autocorrelation functions for system $r35$.

(a) $\phi = 27°$

(b) $\phi = 72°$

(c) $\phi = 105°$

(d) $\phi = 150°$

(e) $\phi = 210°$

(f) $\phi = 225°$

(g) $\phi = 309°$
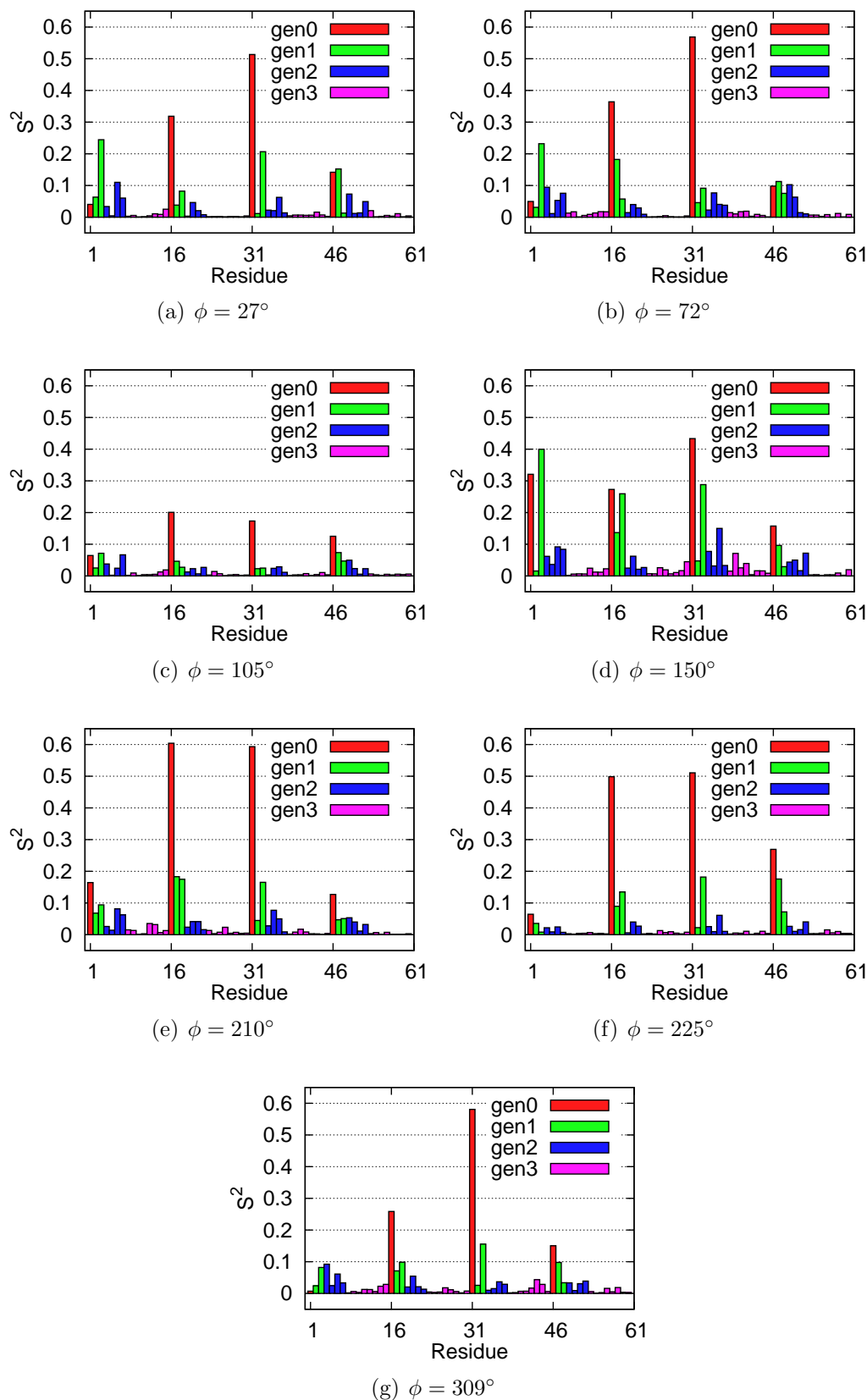
Figure 2.12: S^2 values as calculated from the equilibrium expression for system $r35$.

(a) $\phi = 27°$

(b) $\phi = 72°$

(c) $\phi = 99°$

(d) $\phi = 150°$

(e) $\phi = 215°$

(f) $\phi = 225°$

(g) $\phi = 315°$
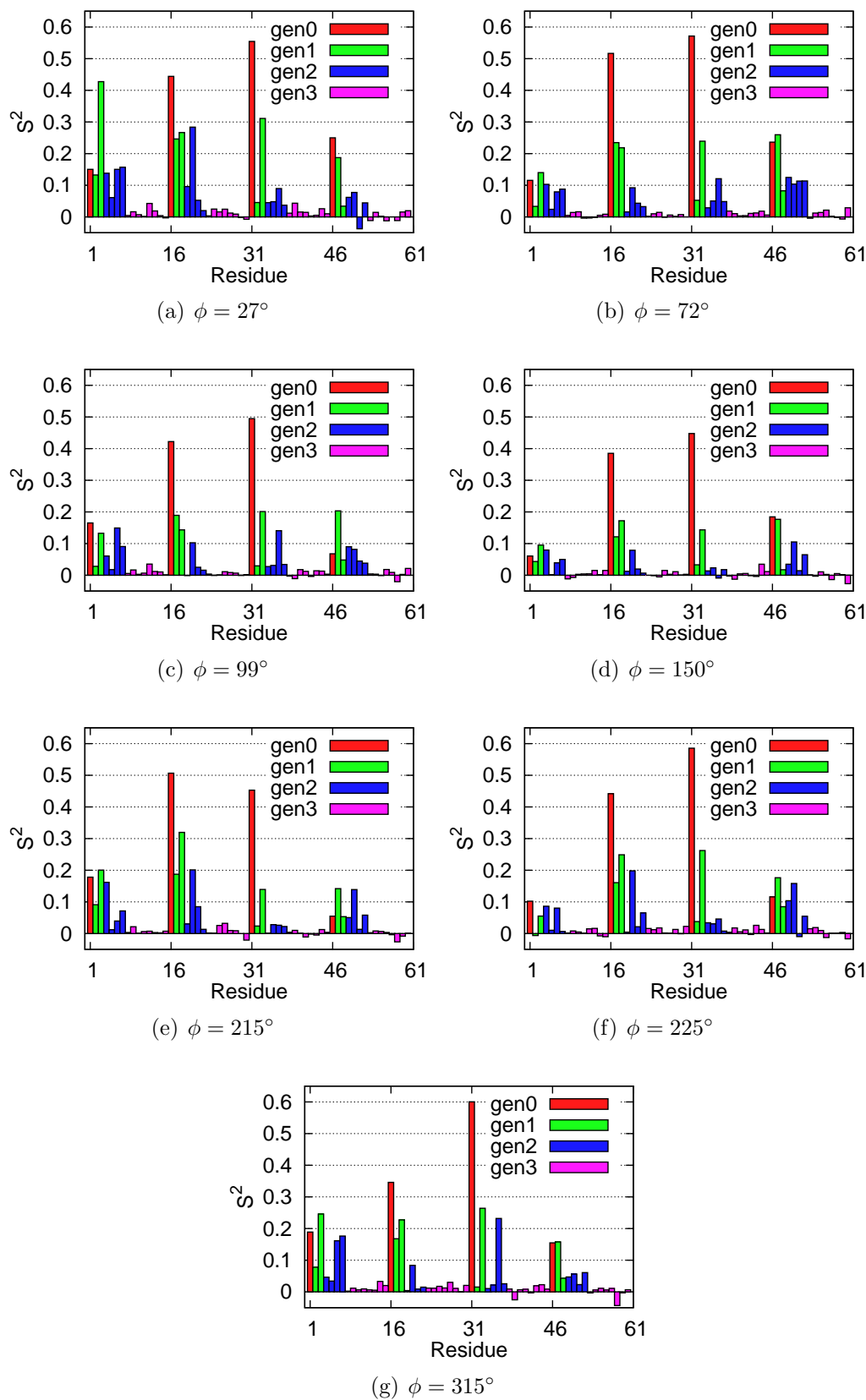
Figure 2.13: S^2 values as calculated from the tail of autocorrelation functions for system $r45$.

(a) $\phi = 27°$

(b) $\phi = 72°$

(c) $\phi = 99°$

(d) $\phi = 150°$

(e) $\phi = 215°$

(f) $\phi = 225°$

(g) $\phi = 315°$
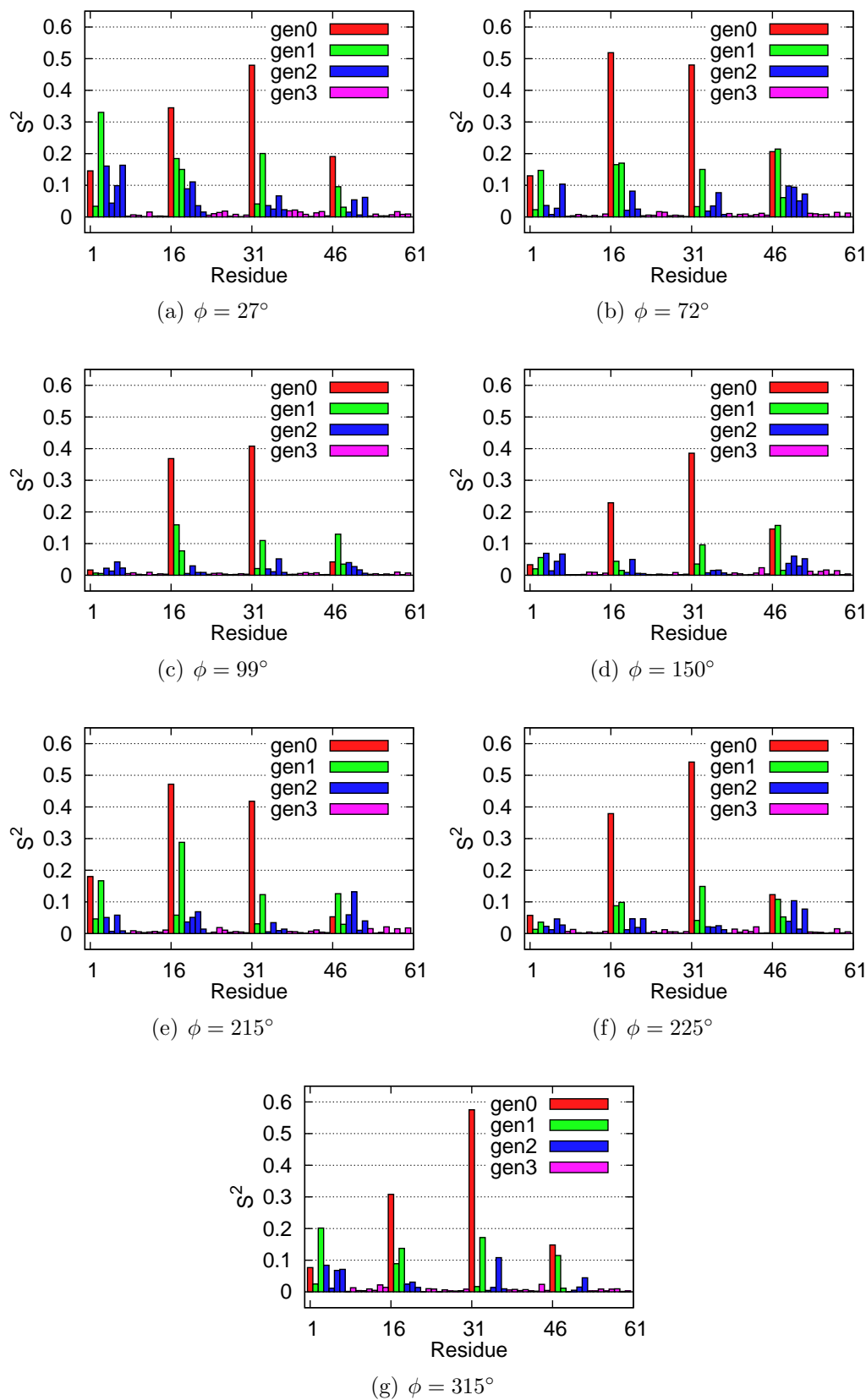
Figure 2.14: S^2 values as calculated from the equilibrium expression for system $r45$.
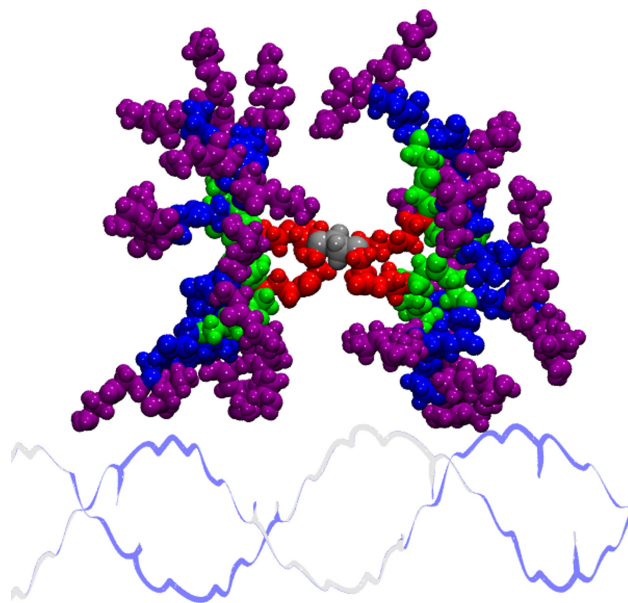
(a)
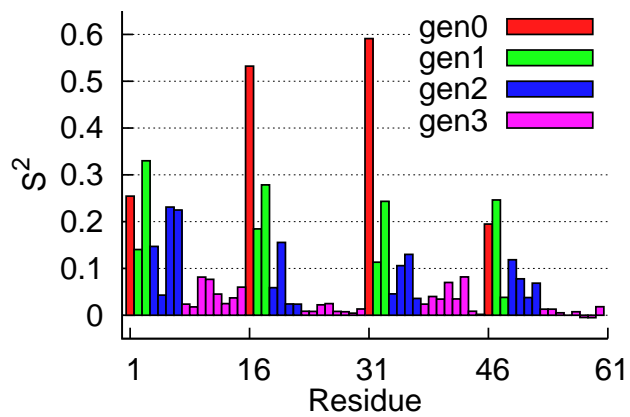
(b)

Figure 2.15: The dendrimer placed at a distance of $r = 25$ Å and a dihedral angle of $\phi = 150°$ is shown in 2.15(a) color coded by generation as follows: generation 0 (red), generation 1 (green), generation 2 (blue), and generation 3 (purple). The core is shown in silver color. Corresponding S^2 values are shown in 2.15(b).

Our findings in the area of bond motion via S^2 could be readily verifiable by NMR liquid state experiments [67, 87], in which one can label the nitrogens with ^{15}N isotopes, and can reveal binding dynamics through the measurement of order parameters.

2.4   Concluding Discussion

In this study, we analyzed the potential use of cationic PAMAM dendrimers to mimic the facilitated search mechanism of proteins along DNA. With this goal in mind, we investigated the energetics and dynamics of dendrimer-DNA binding interactions. We calculated the interaction energy by using all-atom molecular dynamics simulations and mapped the free energy along a circular reaction coordinate around DNA via umbrella sampling. We also calculated NMR order parameters, first time to our knowledge for a PAMAM dendrimer, to understand the dynamics in different binding modes, comparing minor and major groove binding.

Based on a comparison of energetics, we concluded that, despite considerably more favorable electrostatic interactions in the minor groove, the free energy is lower in the major groove than minor groove. This can be explained by a high entropic cost as the dendrimer is attracted to DNA in the minor groove. The higher values of NMR order parameters when the dendrimer is in minor-groove binding mode supported this explanation. The free energy barrier between the minor and major groove interactions was found to be about 8.5 kcal/mol, a value high enough to likely inhibit a possible sliding motion along DNA axis. Our result not only quantifies the free energy barrier introduced due to major groove-minor groove variations in DNA landscape, but also identifies the helical path along DNA major groove as a suitable reaction coordinate to study possible diffusion of dendrimers. This can be tested by experiments in which dendrimers have free access to DNA from all sides such as optical tweezers. Experimental setups in which dendrimer free access to DNA is restricted

may inhibit a possible helical motion. The possibility of a sliding motion along a helical axis (similarly to some DNA-sliding proteins) will be dictated by the PMF along a helical axis spiraling through the major groove. Alternatively, the overall charge size may be reduced to half of its current value to obtain free energy values comparable to proteins. This is the first study to our knowledge, in which a potential of mean force of interaction around DNA is calculated for a cationic, flexible molecule. This information will benefit various fields of DNA nanotechnology, in which not only dynamic, but also static DNA structures are used for various potential applications, some of which involve the interaction of DNA with other molecules similar to dendrimers. Moreover, our PMF profile is a significant contribution in itself in any area that involves dendrimer-DNA interactions. Our results can be used in the fields of bionanomedicine and bionanotechnology that involve dendrimer-DNA interactions and in designing artificial walking molecules.

Bibliography

[1] Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, D. B., and Darnell, J. *Molecular Cell Biology*; Freeman: New York, 2004.

[2] Hu, L., Grosberg, A. Y., and Bruinsma, R. (2008) Are DNA Transcription Factor Proteins Maxwellian Demons? *Biophys. J. 95*, 1151–1156.

[3] Riggs, A. D., Bourgeois, S., and Kohn, M. (1970) The Lac Represser-Operator Interaction: III. Kinetic Studies. *J. Mol. Biol. 53*, 1–17401–417.

[4] Slutsky, M., and Mirny, L. A. (2004) Kinetics of Protein-DNA Interaction: Facilitated Target Location in Sequence-Dependent Potential. *Biophys. J. 87*, 4021–4035.

[5] Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981) Diffusion-Driven Mechanisms of Protein Translocation on Nucleic Acids. 1. Models and Theory. *Biochemistry 20*, 6929–6948.

[6] Winter, R. B., and von Hippel, P. H. (1981) Diffusion-Driven Mechanisms of Protein Translocation on Nucleic Acids. 2. the Escherichia Coli Repressor–Operator Interaction: Equilibrium Measurements. *Biochemistry 20*, 6948–6960.

[7] Halford, S. E. (2004) How Do Site-Specific DNA-Binding Proteins Find Their Targets? *Nucleic Acids Res. 32*, 3040–3052.

[8] Mirny, L., Slutsky, M., Wunderlich, Z., Tafvizi, A., Leith, J., and Kosmrlj, A. (2009) How a Protein Searches for Its Site on DNA: The Mechanism of Facilitated Diffusion. *J. Phys. A: Math. Theor. 42*, 434013.

[9] Cherstvy, A. G., Kolomeisky, A. B., and Kornyshev, A. A. (2008) Protein-DNA Interactions: Reaching and Recognizing the Targets. *J. Phys. Chem. B 112*, 4741–4750.

[10] Florescu, A. M., and Joyeux, M. (2009) Dynamical Model of DNA-Protein Interaction: Effect of Protein Charge Distribution and Mechanical Properties. *J. Chem. Phys. 131*, 105102.

[11] Kolomeisky, A. B., and Veksler, A. (2012) How to Accelerate Protein Search on DNA: Location and Dissociation. *J. Chem. Phys. 136*, 125101.

[12] Veksler, A., and Kolomeisky, A. B. (2013) Speed-Selectivity Paradox in the Protein Search for Targets on DNA: Is It Real or Not? *J. Phys. Chem. B 117*, 12695–12701.

[13] Misra, V. K., Hecht, J. L., Yang, A. S., and Honig, B. (1998) Electrostatic Contributions to the Binding Free Energy of the LambdacI Repressor to DNA. *Biophys. J. 75*, 2262–2273.

[14] Givaty, O., and Levy, Y. (2009) Protein Sliding Along DNA: Dynamics and Structural Characterization. *J. Mol. Biol. 385*, 1087–1097.

[15] Kalodimos, C. G. (2004) Structure and Flexibility Adaptation in Nonspecific and Specific Protein-DNA Complexes. *Science 305*, 386–389.

[16] Tomalia, D. A., Baker, H., Dewald, J., Hall, M., Kallos, G., Martin, S., Roeck, J., Ryder, J., and Smith, P. (1985) A New Class of Polymers: Starburst-Dendritic Macromolecules. *Polym. J. 17*, 117–132.

[17] Tomalia, D. A., Huang, B., Swanson, D. R., Brothers, H. M., and Klimash, J. W. (2003) Structure Control Within Poly(amidoamine) Dendrimers: Size, Shape and Regio-Chemical Mimicry of Globular Proteins. *Tetrahedron 59*, 3799–3813.

[18] Esfand, R., and Tomalia, D. A. (2001) Poly(amidoamine) (PAMAM) Dendrimers: From Biomimicry to Drug Delivery and Biomedical Applications. *Drug Discov. Today 6*, 427–436.

[19] Kukowska-Latallo, J. F., Bielinska, A. U., Johnson, J., Spindler, R., Tomalia, D., and Baker, J. R. (1996) Efficient Transfer of Genetic Material into Mammalian Cells Using Starburst Polyamidoamine Dendrimers. *Proc. Natl. Acad. Sci. U.S.A. 93*, 4897–4902.

[20] Tomalia, D., Reyna, L. A., and Svenson, S. (2007) Dendrimers As Multi-Purpose Nanodevices for Oncology Drug Delivery and Diagnostic Imaging. *Biochem. Soc. Trans. 35*, 61–67.

[21] Bosman, A. W., Janssen, H. M., and Meijer, E. W. (1999) About Dendrimers: Structure, Physical Properties, and Applications. *Chem. Rev. 99*, 1665–1688.

[22] Mills, M., Orr, B. G., Banaszak Holl, M. M., and Andricioaei, I. (2013) Attractive Hydration Forces in DNA–Dendrimer Interactions on the Nanometer Scale. *J. Phys. Chem. B 117*, 973–981.

[23] Mills, M., Orr, B., Banaszak Holl, M. M., and Andricioaei, I. (2010) Microscopic Basis for the Mesoscopic Extensibility of Dendrimer-Compacted DNA. *Biophys. J. 98*, 834–842.

[24] Yu, S., Li, M., Choi, S., Baker, J., and Larson, R. (2013) DNA Condensation by Partially Acetylated Poly(amido Amine) Dendrimers: Effects of Dendrimer Charge Density on Complex Formation. *Molecules 18*, 10707–10720.

[25] Nandy, B., and Maiti, P. K. (2011) DNA Compaction by a Dendrimer. *J. Phys. Chem. B 115*, 217–230.

[26] Ritort, F., Mihardja, S., Smith, S., and Bustamante, C. (2006) Condensation Transition in DNA-Polyaminoamide Dendrimer Fibers Studied Using Optical Tweezers. *Phys. Rev. Lett. 96*, 118301.

[27] Nandy, B., Maiti, P. K., and Bunker, A. (2012) Force Biased Molecular Dynamics

Simulation Study of Effect of Dendrimer Generation on Interaction with DNA . *J. Chem. Theory Comput.* 722–729.

[28] Örberg, M.-L., Schillén, K., and Nylander, T. (2007) Dynamic Light Scattering and Fluorescence Study of the Interaction Between Double-Stranded DNa and Poly(amido-Amine) Dendrimers. *Biomacromolecules 8*, 1557–1563.

[29] Qamhieh, K., Nylander, T., and Ainalem, M.-L. (2009) Analytical Model Study of Dendrimer/DNA Complexes. *Biomacromolecules 10*, 1720–1726.

[30] Chen, W., Turro, N. J., and Tomalia, D. A. (2000) Using Ethidium Bromide to Probe the Interactions Between DNA and Dendrimers. *Langmuir 16*, 15–19.

[31] Ottaviani, M. F., Furini, F., Casini, A., Turro, N. J., Jockusch, S., Tomalia, D. A., and Messori, L. (2000) Formation of Supramolecular Structures Between DNA and Starburst Dendrimers Studied by EPR, CD, UV, and Melting Profiles. *Macromolecules 33*, 7842–7851.

[32] Yu, S., and Larson, R. G. (2014) Monte-Carlo Simulations of PAMAM Dendrimer–DNa Interactions. *Soft Matter 10*, 5325.

[33] Zwanzig, R. (1988) Diffusion in a Rough Potential. *Proc. Natl. Acad. Sci. U.S.A. 85*, 2029–2030.

[34] Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L., and Xie, X. S. (2006) A Base-Excision DNA-Repair Protein Finds Intrahelical Lesion Bases by Fast Sliding in Contact with DNA. *Proc. Natl. Acad. Sci. U.S.A. 103*, 5752–5757.

[35] Tafvizi, A., Huang, F., Leith, J. S., Fersht, A. R., Mirny, L. A., and van Oijen, A. M. (2008) Tumor Suppressor P53 Slides on DNA with Low Friction and High Stability. *Biophys. J. 95*, L01–L03.

[36] Tafvizi, A., Huang, F., Fersht, A. R., Mirny, L. A., and van Oijen, A. M. (2011) A Single-Molecule Characterization of P53 Search on DNA. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 563–568.

[37] Vuzman, D., Azia, A., and Levy, Y. (2010) Searching DNA Via a "Monkey Bar" Mechanism: The Significance of Disordered Tails. *J. Mol. Biol. 396*, 674–684.

[38] Khazanov, N., and Levy, Y. (2011) Sliding of P53 Along DNa Can Be Modulated by Its Oligomeric State and by Cross-Talks Between Its Constituent Domains. *J. Mol. Biol. 408*, 335–355.

[39] Viadiu, H., and Aggarwal, A. K. (2000) Structure of BamHI Bound to Nonspecific DNA: A Model for DNA Sliding. *Mol. Cell 5*, 889–895.

[40] Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., and Wilson, K. S. (1993) The Crystal Structure of EcoRV Endonuclease and of Its Complexes with Cognate and Non-Cognate DNA Fragments. *EMBO J. 12*, 1781–1795.

[41] Townson, S. A., Samuelson, J. C., Bao, Y., Xu, S., and Aggarwal, A. K. (2007) BstYI Bound to Noncognate DNA Reveals a "Hemispecific" Complex: Implications for DNA Scanning. *Structure 15*, 449–459.

[42] Zeng, F., and Zimmerman, S. C. (1997) Dendrimers in Supramolecular Chemistry: From Molecular Recognition to Self-Assembly. *Chem. Rev. 97*, 1681–1712.

[43] Duncan, R. (2003) The Dawning Era of Polymer Therapeutics. *Nat. Rev. Drug Discovery 2*, 347–360.

[44] Malik, N., Wiwattanapatapee, R., Klopsch, R., Lorenz, K., Frey, H., Weener, J. W., Meijer, E. W., Paulus, W., and Duncan, R. (2000) Dendrimers: Relationship Between

Structure and Biocompatibility in Vitro, and Preliminary Studies on the Biodistribution of 125I-Labelled Polyamidoamine Dendrimers in Vivo. *J. Control. Release 65*, 133–148.

[45] Ferruti, P., Marchisio, M. A., and Duncan, R. (2002) Poly (amido-Amine)s: Biomedical Applications. *Macromol. Rapid Commun. 23*, 332–355.

[46] Behr, J. (2001) Synthetic Gene-Transfer Vectors. *Acc. Chem. Res. 26*, 274–278.

[47] Jang, W., Kamruzzaman Selim, K. M., Lee, C., and Kang, I. (2009) Bioinspired Application of Dendrimers: From Bio-Mimicry to Biomedical Applications. *Prog. in Polym. Sci. 34*, 1–23.

[48] Duncan, R., and Izzo, L. (2005) Dendrimer Biocompatibility and Toxicity. *Adv. Drug Delivery Rev. 57*, 23–23.

[49] Navath, R. S., Menjoge, A. R., Wang, B., Romero, R., Kannan, S., and Kannan, R. M. (2010) Amino Acid-Functionalized Dendrimers with Heterobifunctional Chemoselective Peripheral Groups for Drug Delivery Applications. *Biomacromolecules 11*, 1544–1563.

[50] Foloppe, N., and Mackerell, A. D. (2000) All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem. 21*, 86–104.

[51] MacKerell, A. D., and Banavali, N. K. (2000) All-Atom Empirical Force Field for Nucleic Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution. *J. Comput. Chem. 21*, 105–120.

[52] Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (2004) Program for Macromolecular Energy, Minimization,and Dynamics Calculations. *J. Comput. Chem. 4*, 187–217.

[53] Brooks, B. R., Bruccoleri, R. E., Olafson, D. J., States, D. J., Swaminathan, S., and

Karplus, M. (1983) CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem. 4*, 187–217.

[54] MacKerel Jr., A. D., Brooks III, C. L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. In *CHARMM: The Energy Function and Its Parameterization with an Overview of the Program*; v. R. Schleyer et al., P., Ed.; The Encyclopedia of Computational Chemistry; John Wiley & Sons: Chichester, 1998; Vol. 1; pp 271–277.

[55] Lee, M. S., Feig, M., Salsbury, F. R., and Brooks, C. L. (2003) New Analytic Approximation to the Standard Molecular Volume Definition and Its Application to Generalized Born Calculations. *J. Comput. Chem 24*, 1348–1356.

[56] Feig, M., Chocholoušová, J., and Tanizaki, S. (2005) Extending the Horizon: Towards the Efficient Modeling of Large Biomolecular Complexes in Atomic Detail. *Theor. Chem. Acc. 116*, 194–205.

[57] Bashford, D., and Case, D. A. (2000) Generalized Born Models of Macromolecular Solvation Effects. *Annu. Rev. Phys. Chem. 51*, 129–152.

[58] Ryckaert, J., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys. 23*, 327–341.

[59] Kelly, C. V., Leroueil, P. R., Nett, E. K., Wereszczynski, J. M., Baker, J. R., Orr, B. G., Banaszak Holl, M. M., and Andricioaei, I. (2008) Poly(amidoamine) Dendrimers on Lipid Bilayers I: Free Energy and Conformation of Binding. *J. Phys. Chem. B 112*, 9337–9345.

[60] Macke, T. J., and Case, D. A. *Molecular Modeling of Nucleic Acids*; American Chemical Society: Washington, DC, 1997; pp 379–393.

[61] Nummela, J., and Andricioaei, I. (2009) Energy Landscape for DNA Rotation and Sliding Through a Phage Portal. *Biophys. J. 96*, L29–L31.

[62] Torrie, G. M., and Valleau, J. P. (1977) Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys. 23*, 187–199.

[63] Kirkwood, J. G. (1935) Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys. 3*, 300.

[64] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992) The Weighted Histogram Analysis Method for Free Energy Calculations on Biomolecules. I. the Method. *J. Comput. Chem. 13*, 1011–1021.

[65] Grossfield, A. WHAM: The Weighted Histogram Analysis Method", Version 2.0.6. `http://membrane.urmc.rochester.edu/content/wham`.

[66] Sun, J., Viadiu, H., Aggarwal, A. K., and Weinstein, H. (2012) Energetic and Structural Considerations for the Mechanism of Protein Sliding Along DNA in the Nonspecific BamHI-DNa Complex. *Biophys. J. 84*, 3317–3325.

[67] Musselman, C., Zhang, Q., Al-Hashimi, H., and Andricioaei, I. (2010) Referencing Strategy for the Direct Comparison of Nuclear Magnetic Resonance and Molecular Dynamics Motional Parameters in RNA. *J. Phys. Chem. B 114*, 929–939.

[68] Salmon, L., Bascom, G., Andricioaei, I., and Al-Hashimi, H. M. (2013) A General Method for Constructing Atomic-Resolution RNA Ensembles Using NMR Residual Dipolar Couplings: The Basis for Interhelical Motions Revealed. *J. Am. Chem. Soc. 135*, 5457–5466.

[69] Levy, R. M., and Karplus, M. *Trajectory Studies of NMR Relaxation in Flexible Molecules*; ACS: Washington, D.C., 2009; pp 445–468.

[70] Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear

Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc. 104*, 4546–4559.

[71] Wrabl, J. O., Shortle, D., and Woolf, T. B. (2000) Correlation Between Changes in Nuclear Magnetic Resonance Order Parameters and Conformational Entropy: Molecular Dynamics Simulations of Native and Denatured Staphylococcal Nuclease. *Proteins 38*, 123–133.

[72] Markwick, P. R. L., Bouvignies, G., and Blackledge, M. (2007) Exploring Multiple Timescale Motions in Protein GB3 Using Accelerated Molecular Dynamics and NMR Spectroscopy. *J. Am. Chem. Soc. 129*, 4724–4730.

[73] Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M., and Brooks, B. R. (1992) A 500 Ps Molecular Dynamics Simulation Study of Interleukin-1β in Water: Correlation with Nuclear Magnetic Resonance Spectroscopy and Crystallography. *J. Mol. Biol. 226*, 239–250.

[74] Im, W., Beglov, D., and Roux, B. (1998) Continuum Solvation Model: Computation of Electrostatic Forces from Numerical Solutions to the Poisson-Boltzmann Equation. *Comput. Phys. Commun. 111*, 59–75.

[75] Nina, M., Beglov, D., and Roux, B. (1997) Atomic Radii for Continuum Electrostatics Calculations Based on Molecular Dynamics Free Energy Simulations. *J. Phys. Chem. B 101*, 5239–5248.

[76] Banavali, N. K., and Roux, B. (2002) Atomic Radii for Continuum Electrostatics Calculations on Nucleic Acids. *J. Phys. Chem. B 106*, 11026–11035.

[77] The PyMOL Molecular Graphics System, Version 1.5.0.2, Schrödinger, LLC.

[78] Honig, B., and Nicholls, A. (1995) Classical Electrostatics in Biology and Chemistry. *Science 268*, 1144–1149.

[79] Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S., and Honig, B. (2009) The Role of DNA Shape in Protein-DNA Recognition. *Nature 461*, 1248–1253.

[80] Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B., and Mann, R. S. (2010) Origins of Specificity in Protein-DNA Recognition. *Annu. Rev. Biochem. 79*, 233–269.

[81] Railsback, J. G., Singh, A., Pearce, R. C., McKnight, T. E., Collazo, R., Sitar, Z., Yingling, Y. G., and Melechko, A. V. (2012) Weakly Charged Cationic Nanoparticles Induce DNA Bending and Strand Separation. *Adv. Mater. 24*, 4261–4265.

[82] Marklund, E. G., Mahmutovic, A., Berg, O. G., Hammar, P., van der Spoel, D., Fange, D., and Elf, J. (2013) Transcription-Factor Binding and Sliding on DNA Studied Using Micro- and Macroscopic Models. *Proc. Natl. Acad. Sci. U.S.A. 110*, 19796–19801.

[83] Blainey, P. C., Luo, G., Kou, S. C., Mangel, W. F., Verdine, G. L., Bagchi, B., and Xie, X. S. (2009) Nonspecifically Bound Proteins Spin While Diffusing Along DNA. *Nat. Struct. Mol. Biol. 16*, 1224–1229.

[84] Iwahara, J., Zweckstetter, M., and Clore, G. M. (2006) NMR Structural and Kinetic Characterization of a Homeodomain Diffusing and Hopping on Nonspecific DNA. *Proc. Natl. Acad. Sci. U.S.A. 103*, 15062–15067.

[85] van Tilborg, P. J. A., Czisch, M., Mulder, F. A. A., Folkers, G. E., Bonvin, A. M. J. J., Nair, M., Boelens, R., and Kaptein, R. (2000) Changes in Dynamical Behavior of the Retinoid X Receptor DNA-Binding Domain upon Binding to a 14 Base-Pair DNA Half Site. *Biochemistry 39*, 8747–8757.

[86] Lee, A. L., Kinnear, S. A., and Wand, A. J. (2000) Redistribution and Loss of Side Chain Entropy upon Formation of a Calmodulin–Peptide Complex. *Nat. Struct. Mol. Biol. 7*, 72–77.

[87] Nikolova, E. N., Bascom, G. D., Andricioaei, I., and Al-Hashimi, H. M. (2012) Probing Sequence-Specific DNA Flexibility in A-Tracts and Pyrimidine-Purine Steps by Nuclear Magnetic Resonance 13C Relaxation and Molecular Dynamics Simulations. *Biochemistry 51*, 8654–8664.

Chapter 3

Dendrimers in Nanoscale Confinement: The Interplay of Conformational Change and Pore Blockade

This chapter is reproduced in part with permission from Nano Letters, submitted for publication. Unpublished work copyright 2015 American Chemical Society.

3.1   Introduction

Hyperbranched dendrimers are nanocarriers for drugs, imaging agents, and catalysts, and their passage through biological nano-channels or entrapment in inorganic nanoporous support scaffolds are of fundamental interest. However, the molecular process of nanoscale confinement and its effect on dendrimer conformations are poorly understood. Here, we use

single-molecule experiments and molecular dynamics simulations to establish an atomically detailed model consistent with ionic current measurements. We discover and explain that migration into nanopores is not dictated by the diameter of dendrimers but by their size-dependent compressibility. Differences in flexibility also rationalize the apparent anomaly that the experimental ion current read-out depends in non-linear fashion on dendrimer size, and nanoscale confinement is inferred to reduce the protonation of the polycationic structures. Our model can likely be expanded to other dendrimers and be applied to improve the analysis of biophysical experiments, rationally design functional materials such as nanoporous filtration devices or nanoscale drug carriers that effectively pass biological pores.

Dendrimers are nanoscale, star-like, hyperbranched molecules with high functional impact [1, 2]. Their structural hallmarks are (i) a void interior to encapsulate imaging agents or drugs, (ii) a dense outer shell with numerous terminal groups that can be chemically modified with fluorophores, catalytic groups, or molecular receptors, and (iii) a spherical shape that influences the interaction with other objects including nanopores. Biological nanopores help control molecular traffic across cellular membranes, while natural and man-made pores are exploited in nanotechnology for filtration [3] as well as the label-free sensing of analytes [4–6] including electrical sequencing of individual DNA strands [7, 8]. Confining dendrimers temporary or permanently into nanopores is of relevance in basic and applied science. In research, permeating dendrimers of different size are used to probe the lumen of biological and inorganic nanopores [9], while in biomedicine the interaction influences the transport of encapsulated therapeutics across biological membranes [10, 11] or achieves the blocking of pore-forming toxins [12, 13]. In materials science, the permanent logding results in nanoscale filtration devices [14], and reusable catalysts that enclose the dendrimers within nanotubes [15].

Despite these applications, a fundamental understanding of dendrimer behavior under nanoscale confinement is missing. For example, there is currently no definitive answer to the seem-

ingly simple question about a cut-off size for the permeation of dendrimers into pores of given diameter, particularly under the ubiquitous transmembrane voltage present in biological cells as well as in research settings. Hence, it is not clear to which degree the structural flexibility of differently sized dendrimers influences their entrance into pores; their rigidity usually increases with bigger diameter. Furthermore, there is no detailed information on the mechanism for pore insertion, i.e., whether the dendrimers insert completely or partly into a pore. Finally, it is not known whether the nanoscale confinement of dendrimers has an effect on the charge state of the ionizable functional groups of the molecules. A detailed molecular study that settles many of these questions will undoubtedly achieve a step-change in our fundamental understanding of dendrimer dynamics under confinement. Unfortunately, the large body of studies on the nanopore-mediated capturing or transport of linear organic polymers, DNA and polypeptides [16–31] cannot be directly applied due to the different molecular shapes and properties compared to dendrimers.

Here we develop a coherent and detailed molecular model of dendrimers under nanoscale confinement. We apply the widely used protein pore α-hemolysin (αHL) (Figure 3.2) of known atomistic structure [32] to examine the prominent polyamido-amine (PAMAM) dendrimers [33] of increasing size, generation 1 to generation 5 (G1- G5). We measure the electrophoretically induced permeation of positively charged PAMAM at the level of individual pores, and thereby avoid problems of static or temporal heterogeneity which can hobble conventional ensemble measurements [34]. The specific experimental output is the ionic current flowing through single pores as a function of PAMAM size. Our detailed experimental data are synergistically combined with molecular dynamics simulations to probe how far individual dendrimers permeate into the pore lumen. The simulations are validated by a comparison of the simulated to the experimental pore currents. Additionally, the simulations provide unprecedented molecular insight into the mechanism of PAMAM insertion, as well as the structural conformations and the charge of confined dendrimers as a function of their size and nanoscale confinement, something which could not be achieved solely with experiments.

Our coherent understanding will help tailor the molecular properties of dendrimers and pores to suit specific applications.

3.2 Description of the molecular components

In our work on nanoscale confinement, we examined whether PAMAM dendrimers G1 to G5 permeate into the αHL pore. We selected these molecular components for three reasons. Firstly, the size of the dendrimers is easily tunable. With each number of synthetic cycles (that is generation), twice the number of tertiary amine branching points are added, and the number of terminal primary amines doubles (Figure 3.2; Figure 3.1) [33]. The hydrodynamic diameter of PAMAM thereby increases from 2.2 nm for G1 to 2.9 nm for G2, 3.6 nm for G3, 4.5 nm for G4, and 5.4 nm for G5. Secondly, the PAMAM dendrimer and the αHL pore have matching nanoscale dimensions (Figure 3.2) [32]. Finally, the pore lumen at the trans side is of simple cylindrical shape which was anticipated to ease studying the interaction with PAMAM. In particular, at the trans side, the αHL pore features a cylinder-like transmembrane β-barrel of approximately 2 nm inner width and 5.0 nm height which narrows to a 1.0 nm wide inner constriction (Figure 3.2). The pores cis side is geometrically less regular and has a sphere-like chamber and a 2.3 nm-wide cis entrance, which may cause further undesired complexity in our study of pore confinement. Hence, dendrimers were placed at the trans side.

3.3 Experimental Procedures and Data

Electrical recordings of single pores establish dendrimer permeation, with much larger diameters than expected from their hydrodynamic diameter. The generation-dependent movement of PAMAM into the nanoscale confinement was probed by recording the ionic current flowing
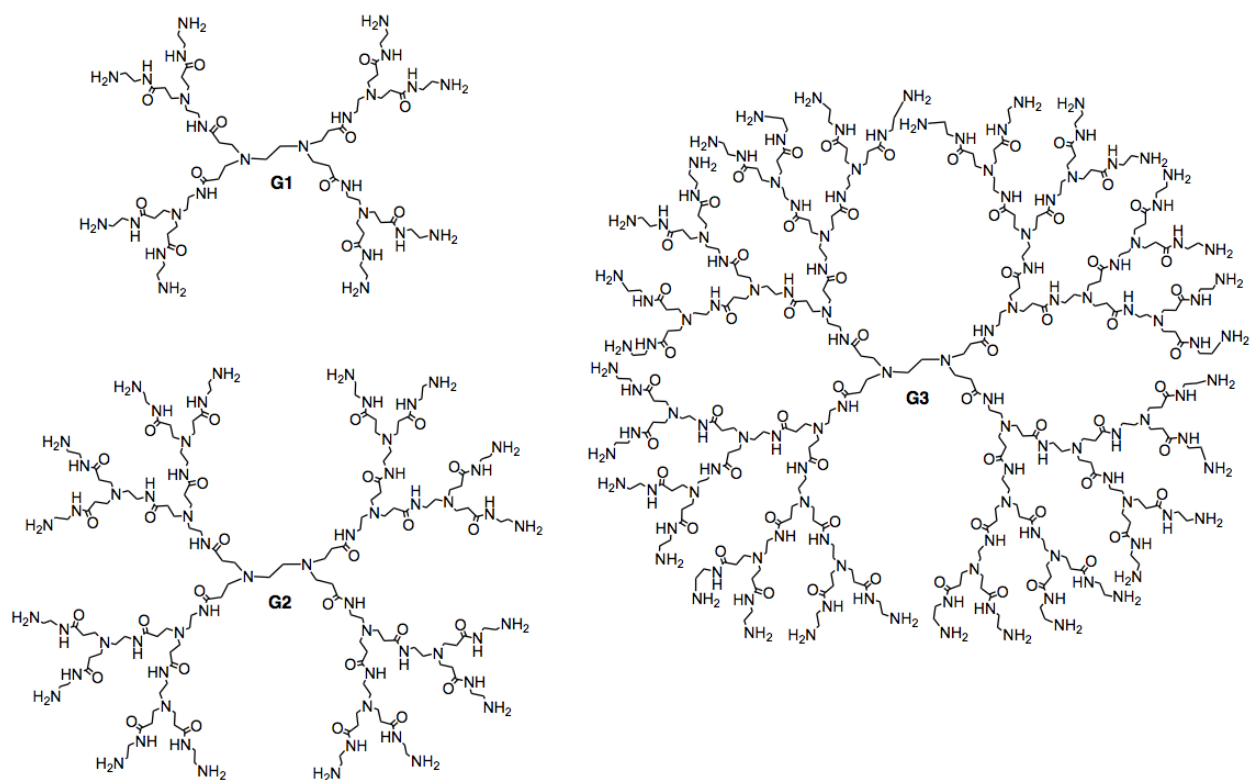
Figure 3.1: Chemical diagrams of PAMAM dendrimers of generation G1, G2 and G3 in the non-protonated form.
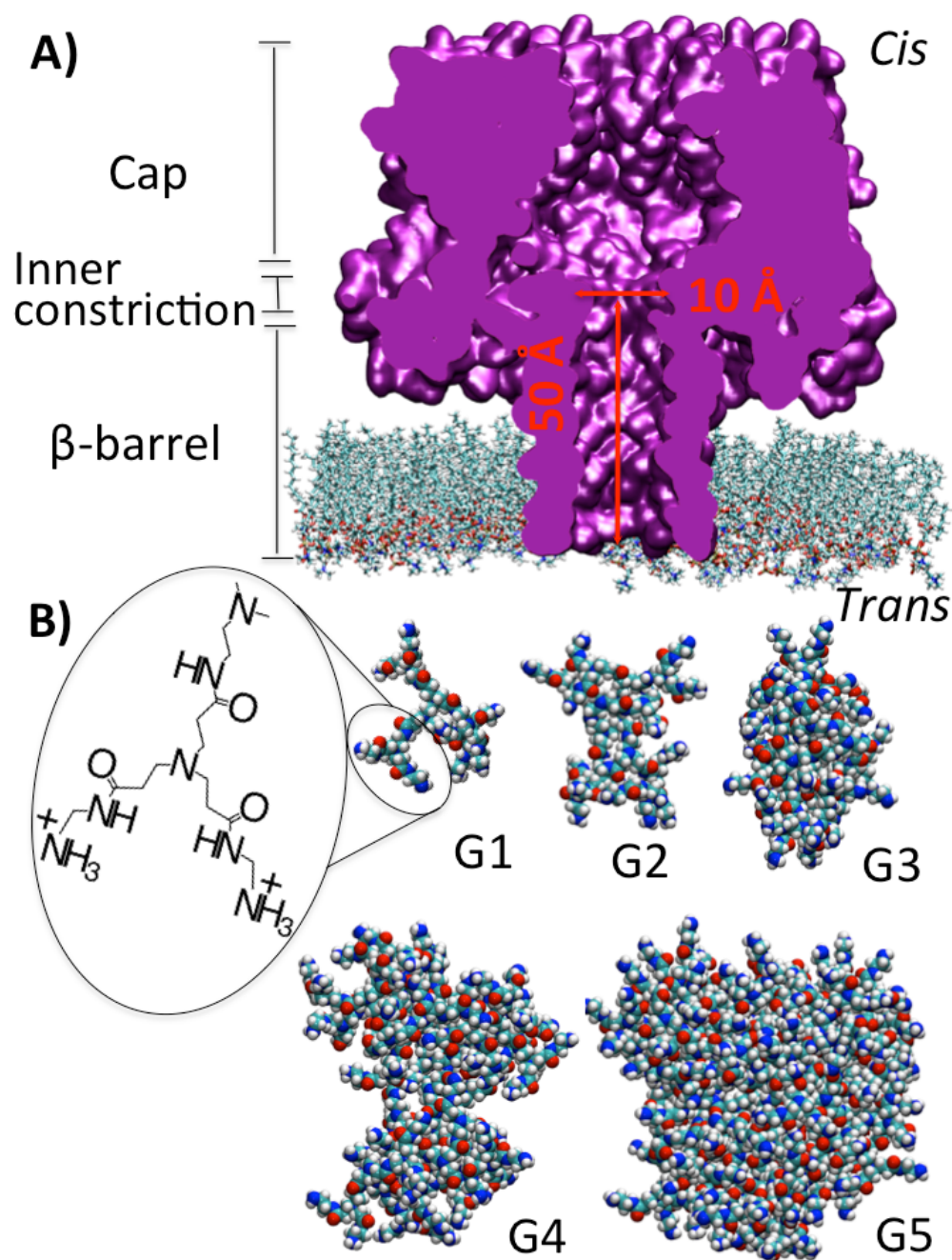
Figure 3.2: Structures of the molecular components used in our study. (A) The lipid bilayer-embedded αHL pore features a transmembrane β-barrel of approximately 2 nm width and an 1.0 nm wide inner constriction. (B) PAMAM dendrimers of generations 1 to 5. The inset shows the chemical structure of a PAMAM branch with protonated terminal primary amines. To study the interaction, dendrimers are placed on the trans side and an effective negative transmembrane potential is applied at the cis side to induce the electrophoretic movement of the dendrimers into the β-barrel.

through single protein channels. Single-channel current recording is a widely used technique to understand the structural dynamics of individual molecules that are captured or passing the pore and can be examined in a label-free fashion under high temporal and conductance resolution [4–6]. Our measurements with αHL were carried out by applying a potential across a membrane with an embedded pore, and by recording the ionic current of the solvated electrolytes. Under standard electrolyte conditions of 1 M KCl [16, 20, 35], a potential of -100 mV relative to the cis side, the blank αHL pore exhibited a conductance of 1020 ± 90 pS ($n = 7$, number of independent recordings) which is in line with the literature [35–37].

3.3.1 Nanopore recordings

Single-channel current recordings were performed by using a planar lipid bilayer apparatus as described [38]. Briefly, a bilayer of *1,2- diphytanoyl-sn-glycero-3-phosphocholine* (Avanti Polar Lipids) was formed on an aperture (100 μm in diameter) in a Teflon septum (Goodfellow Corporation) separating the cis and trans chambers of the apparatus. Each compartment contained 1 M KCl, 50 mM Tris-HCl, pH 8.2. Gel-purified heptameric αHL protein (final concentration $0.01 - 0.1$ ng ml^{-1}) was added to the cis compartment to achieve insertion of a single channel into the bilayer. Subsequently, PAMAM samples (10 μL, G1-G3, 20 wt. %; G4 and G5, 5 wt. %) were added to the trans side containing 1 mL of electrolyte. Transmembrane currents were recorded at a holding potential of $+100$ mV (with the cis side grounded), unless stated otherwise, by using a patchclamp amplifier (Axopatch 200B, Axon Instruments, Union City, CA). For analysis, currents were low-pass filtered at 20 kHz and sampled at 50 kHz using a Digidata 1200 A/D converter (Axon Instruments), as described in Ref. [39].

3.3.2 Single-channel current data

To probe dendrimers under nanoscale confinement, PAMAM were added at a final concentration of 50 μM to the trans side of the αHL. Short current blockades were observed as exemplarily shown for PAMAM G2 (Figure 3.3A; Figure 3.4). The blockades are likely caused when the positively charged PAMAM is electrophoretically driven towards the negatively polarized *cis* side and temporarily resides within the β-barrel (Figure 3.2). The blockade ends when the PAMAM reversibly exits the pore likely to avoid the energetically unfavorable nanoscale confinement and to regain conformational and translational freedom. Remarkably, dendrimers of generation G1 to G4 entered the pore even though their hydrodynamic radius (2.2 nm for G1, 4.5 nm for G4) is larger than the trans-entrance of the β-barrel at 2.1 nm. The diameter of the dendrimers is hence an unsuitable guide to predict their electrophoretic movement into a narrow pore.

3.3.3 Non-monotonous dependence of current blockade on dendrimer generation

To further examine the permeation of large dendrimers, pore blockades were characterized in terms of their blockade amplitude, A, and duration, t_{off}, (Figure 3.3B). The amplitude is the difference between the ionic current at the bottom of the event and the current of the open channel, I_o, normalized to I_o (Figure 3.3B). To obtain statistically relevant data, the distribution of A was plotted in a histogram as exemplarily shown for the G2 dendrimer (Figure 3.3C). The peak of the distribution has a maximum of 46%. The average blockade amplitudes for each PAMAM generation were acquired using at least three independent recordings with a total of around 10000 individual events.

The results are summarized in a plot of amplitude vs. dendrimer generation (Figure 3.3D).
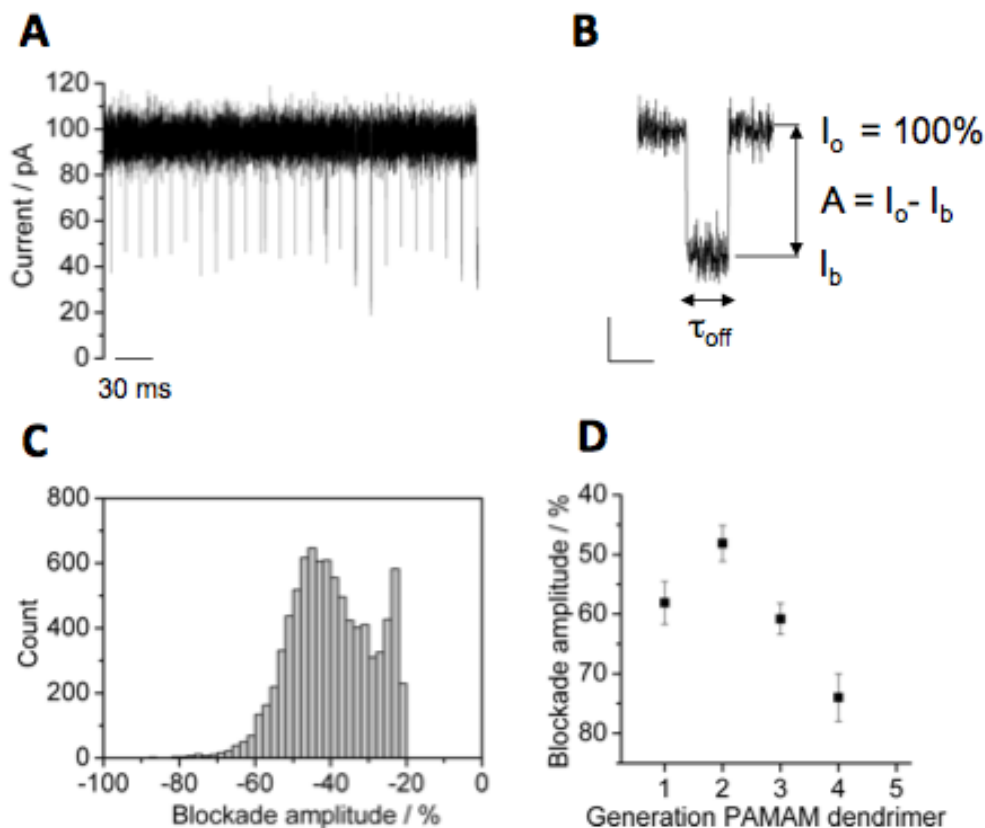
Figure 3.3: PAMAM dendrimers temporarily block the αHL pore with large dendrimer diameter and in an unusual non-linear dependence on generation. (A) Current blockades caused by PAMAM G2 added to the *trans* side of αHL recorded at 1 M KCl, at an effective potential of -100 mV at the *cis* pore side. (B) A single blockade event with duration t_{off} and amplitude A, defined by the difference between the open-channel current I_o and the current level for the blocked channel, I_b. (C) Histogram of amplitude A with a peak at 46 %. The peak at 22% stems from short noise fluctuations of the open channel and was included to avoid the artificial trimming of the main blockade peak. (D) The blockade amplitude depends in a non-monotonous fashion on the dendrimer generation.
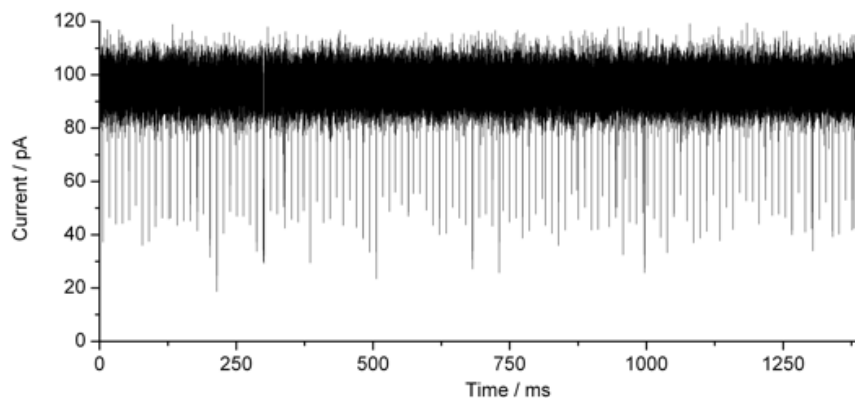
Figure 3.4: Single-channel current trace of αHL in the presence of 50 μM PAMAM G2 at the trans side at -100 mV at the *cis* side. The trace was acquired in the variable-length mode of nanopore acquisition software pClamp. In this mode, segments of the trace without blockade events are removed to yield a concatenated series of blockades. The trace displays events recorded within 10 s which is different to the duration of the concatenated event segments.

Several remarkable features are noted. Firstly, the plot shows an unusual non-linear dependence of A vs. dendrimer generation. The amplitude of the current blockade first lowers from 58.1 ± 3.6 % (G1) to 48.1 ± 3.0 % (G2) before increasing to over 60.8 ± 2.8 % (G3) and 74 ± 4.4 % (G4). This behavior is independent of the transmembrane voltage (Figure 3.5, Figure 3.6) and likely implies that the dendrimers lodged in the pore are not in a meta-stable state but at the energetic minimum. Secondly, G4 PAMAM migrated into the pore only at very high potentials of 160 mV. This can suggest that the lower voltages were not enough to force the dendrimer into the pore. Finally, G5 PAMAM was not electrophoresed into the pore, as judged by the lack of discernible clear current blockades. When considering dendrimers G1 to G3, the blockade duration, t_{off}, depended in non-linear fashion on the dendrimer size, similar to the amplitude (Figure 3.5) while it did not depend on voltage (Figure 3.7). We did not pursue the analysis of t_{off} data any further as we were able to computationally simulate the blockade amplitude.

67
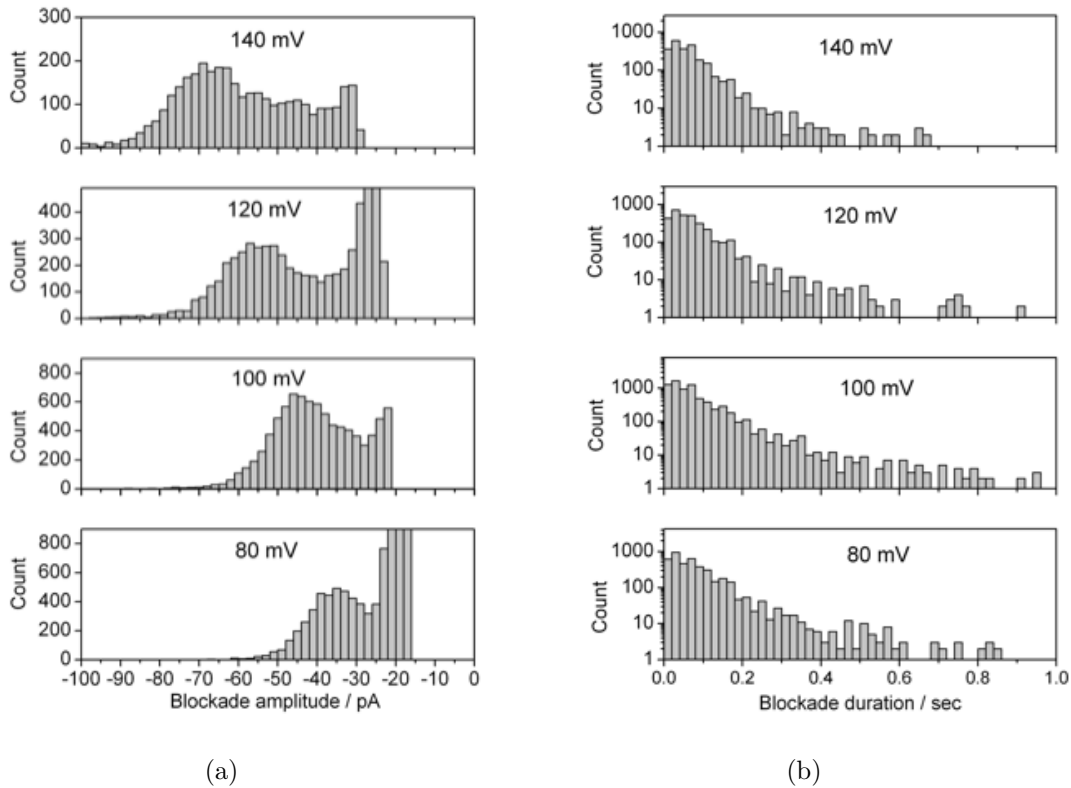
Figure 3.5: Analysis of PAMAM G2 events in terms of blockade amplitude (A) and blockade duration (B) as a function of transmembrane voltage ranging -80 to -140 mV at the *cis* side.

3.4 Computational Simulations of PAMAM Permeation into αHL

In order to explain the unusual blockade of high-generation dendrimers and the non-monotonous dependence of blockade amplitude, we atomistically simulated the electrophoretically driven permeation of PAMAM into αHL using molecular dynamics as described in the following sections.
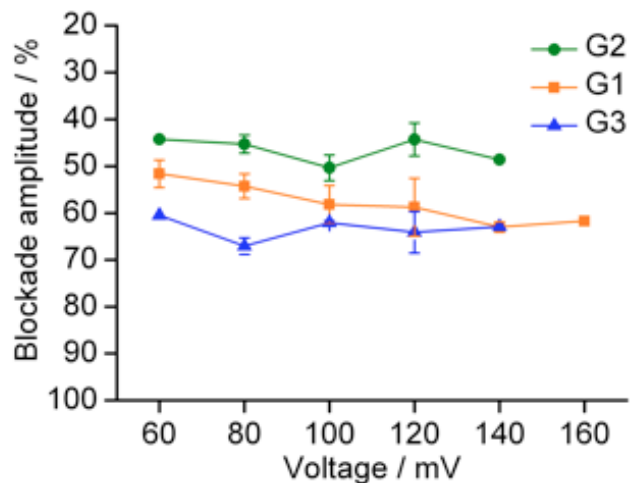
Figure 3.6: The amplitude of the PAMAM-induced blockade events does not show a clear dependence on voltage.

3.4.1 System Preperation

All-amine terminated dendrimers of generations G1-G5 were modeled as described in a previous study [40]. It has been shown that at pH values between 7 and 8, only dendrimer primary amines are protonated [41, 42]. Therefore we used this convention in modeling the dendrimers at neutral pH used in the experiment. Dendrimers were initially minimized for 5000 steps followed by gradually heating over the course of 100 ps, and equilibrating for 30 ns. All simulations were run with a time step of 1 fs at a temperature of 300 K. Non-bonded interactions had a cut-off of 16 with a switching function turned on at 14 . The equilibration of the dendrimers was monitored by the convergence of the radii of gyration, illustrated in Figure 3.8. Representative structures are shown in Figure 3.2, and the convergence of radii of gyration is demonstrated in Figure 3.8.

The equilibrated dendrimer radii were found to be 9.7 Å, 12.1 Å, 13.3 Å, 18.7 Å, and 21.0 Å for G1 through G5, respectively, within the range of previously published values [43]. The dendrimers were then added to the trans side of αHL pore at a distance equivalent to twice the radius of gyration of dendrimer. The protein pore was embedded in a pre-

69
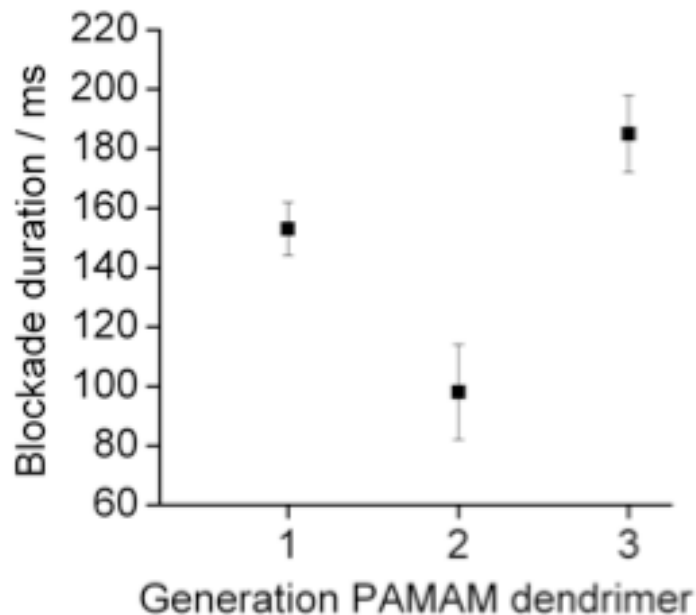
Figure 3.7: The duration of PAMAM-induced blockades depends in a non-linear fashion on the PAMAM generation.

equilibrated, circular patch of a lipid DPPC leaflet with a radius of 60 Å. In order to reduce the computation time while maintaining the key interactions, all atoms were fixed except for the dendrimer and the charged αHL residues at the pore entrance, i.e. ASP127 and LYS131 on each of the seven monomers near $z = 0$ Å, which co-indices with the opening of the β-barrel. Solvent was represented with a continuum model by using generalized Born implicit solvent as implemented in NAMD [44, 45]. All simulations were run by using NAMD software [45] with the CHARMM27 force field [46].

3.4.2 Grid Steered Molecular Dynamics

The computational electrophoresis of the positively charged PAMAM dendrimers was achieved by applying a transmembrane potential following an established protocol [47–49] and grid-steered molecular dynamics (GSMD) [50]. In this approach, a constant electric field is used to simulate transmembrane potential in combination with an electrostatic grid potential ob-
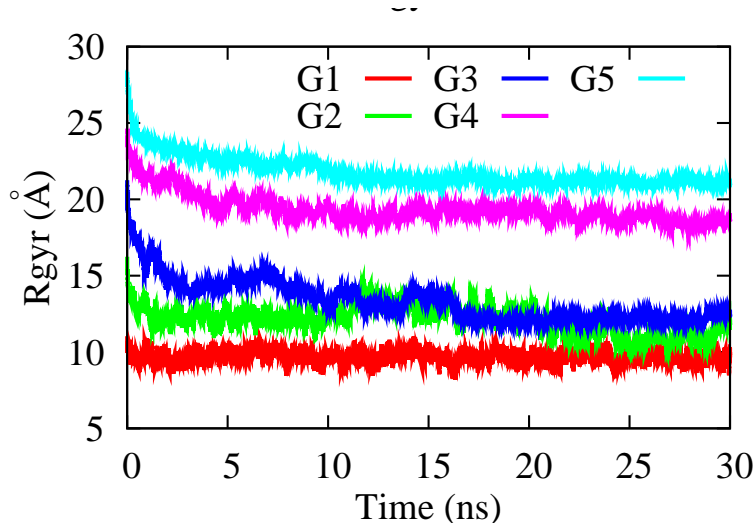
Figure 3.8: Time series for radii of gyration for dendrimers G1-G5 during equilibration.

tained at low potentials, which can then be scaled up to accelerate the molecular movement to times scales accessible to simulations while reducing artificial distortions in molecular configurations [50].

GSMD has been successfully applied to simulate the translocation of DNA strands, DNA hairpins and alpha-helical peptides through αHL [50]. In our simulations, we initially used an effective voltage of 30 V for 500 ps, only $12V$ of which is due to the contribution from constant electric field. We then equilibrated the structure at a lower effective voltage of 4.8 V for 14 ns and finally 500 ps at a potential of 120 mV for 500 ps to obtain meaningful conformations. We used an electrostatic grid potential which was previously calculated for αHL in the presence of an external electric field of 1.2 V from time and space average of a 5.3 ns molecular dynamics simulation in explicit solvent [48]. The grid potential was used as an external potential in addition to a uniform external electric field applied on the dendrimer in $z-$direction as described in Ref. [50]. Corresponding pdb files for the equilibrated αHL and membrane complex were provided by Dr. Aksimentiev and were modified to contain the αHL atoms and a circular patch of single layer of DPPC membrane with a radius of 60 Å.

Each dendrimer was pulled into the pore with a total electric field equivalent to a voltage

of 30 V, (12 V constant electric field, 28 V contribution from electrostatic grid potential after scaling by 15). Molecular dynamics simulation snapshots resulting from the pulling trajectory were saved and used to then calculate (with the model explained below) the degree to which the PAMAM reduced the open pore volume in dependence of the $z-$position along the channel axis, and hence to compute the current blockade as described in Eqs. 3.1 and 3.4.

3.5 Calculation of Current Blockade from Simulations

The simulations allow to calculate the ionic current blockade and thereby explain the unusual experimental non-monotonous dependence of blockade amplitude on dendrimer generation. The open cross-sectional area S(z) of the lumen was used to calculate the ionic current blockade. S(z) integrated over z represents the volume of pore lumen which is filled by electrolytes, which is in turn directly proportional to the ionic current. While more sophisticated means to compute current by representing ions explicitly exists [48, 51, 52] they are computationally more challenging to converge because of the direct calculation of the time-dependent flow of ions, particularly for large ionic flows. Moreover, because the current are so large for the system at hand, the approximation we use is expected to be appropriate. Following our previous work[53], the ratio of blocked channel, I_b and the open-channel current I_o was calculated using

$$\frac{I_b}{I_o} = \frac{\int_0^L \frac{dz}{S(z)} + \left(\frac{1}{2d_1} + \frac{1}{2d_2}\right)}{\int_0^L \frac{dz}{S'(z)} + \delta_{3,4}\left(\rho\left(\frac{1}{2d_1'} - \frac{1}{\pi D}\right) + \frac{1}{\pi D}\right) + \delta_{1,2}\left(\frac{1}{2d_1'}\right) + \left(\frac{1}{2d_2}\right)} \tag{3.1}$$

Here, S(z) and S(z) in the numerator and denominator are the open cross-sectional areas when the pore is open and when the pore is dendrimer-blocked, respectively. The terms $1/2d_1$
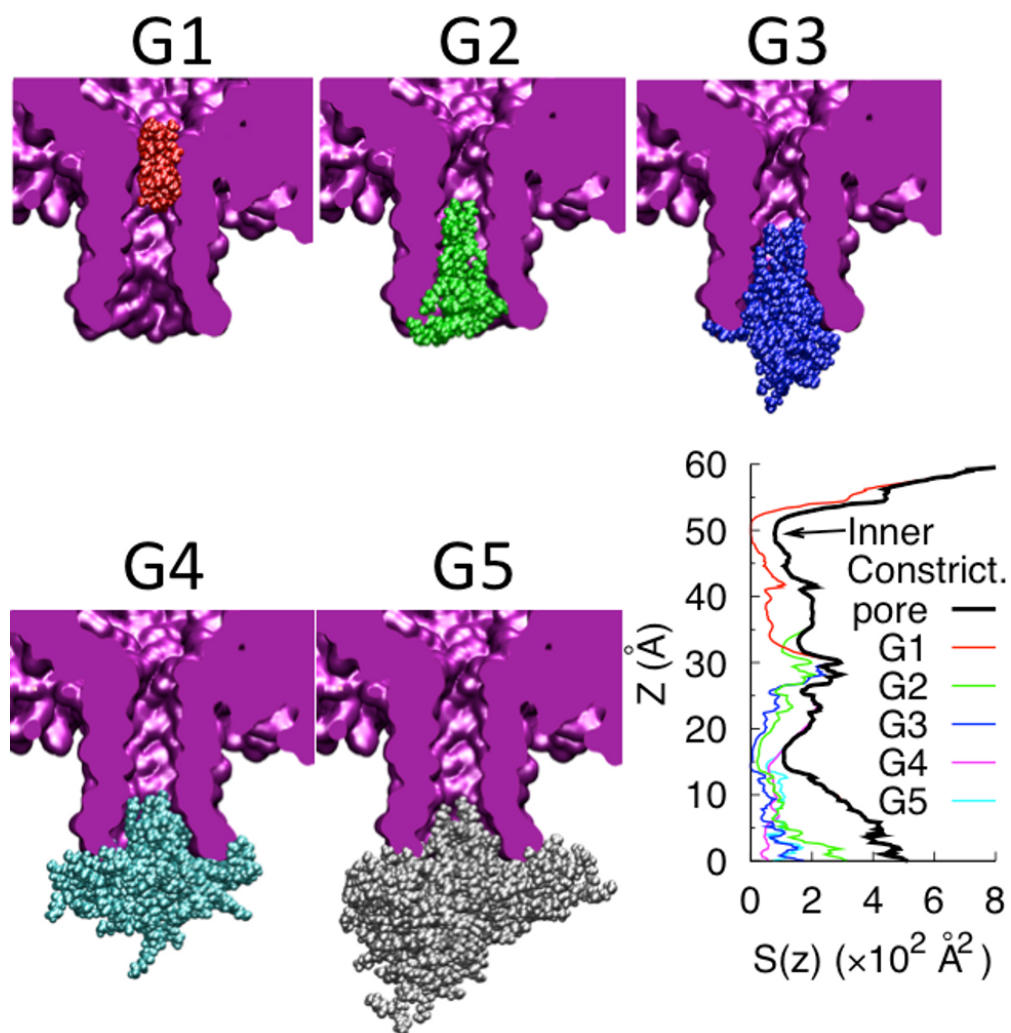
Figure 3.9: Snapshots of PAMAM inserted into the β-barrel of αHL after electrophoresis and equilibration. The dendrimer generations are provided. In the bottom right is a plot of the open cross-sectional area $(S'(z))$ whereby the z-axis position is 0 Å at the trans entrance of the β-barrel and 49.5 Å at the inner constriction.

and $1/2d_2$ in the numerator are the access resistances for the trans and cis pore opening of the unblocked channel with diameters d_1 and d_2, respectively, following Halls model [54].

The trans access resistance of pores blocked by G1 and G2 was calculated similarly using the effective pore diameter d_1' at the trans side. A different approach was used to describe the access resistance of dendrimers G3, G4 and to G5, which, due to their larger size, reside to an increasing degree outside the pore, within a roughly hemispherical region. In our calculations, we treated the region outside the pore as a combination of (i) a hemispherical region with a diameter D large enough to enclose the dendrimer partition outside the pore and (ii) the rest of the region from this hemisphere up to the electrode at infinity. The contribution to access resistance from the latter term is given by access resistance to a hemispherical boundary, which is $1/\pi D$ as described in Ref. [55]. In the region from the surface of this hypothetical hemisphere up to the pore entrance, which has an effective diameter of d_1 when the pore is clogged by the dendrimer, the access resistance can be calculated by the term $\rho(\frac{1}{2d_1'} - \frac{1}{\pi D})$, where ρ is the effective resistivity introduced by the dendrimer, a relative resistivity normalized with respect to the resistivity of the pure solution. As a result, the total relative access resistance in the trans side half-space is $\rho(\frac{1}{2d_1'} - \frac{1}{\pi D}) + \frac{1}{\pi D}$. This treatment of access resistance for different dendrimer generations can be summarized as follows:

$$\delta_{1,2} = \begin{cases} 1 & \text{for G1 and G2} \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

$$\delta_{3,4} = \begin{cases} 1 & \text{for G3, G4, and G5} \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

The additional parameter ρ accountingt for the relative increase in resistivity caused by PAMAM outside the pore region,. The parameter can be further factor out into two components, $\rho = \rho_1(V, V')\rho_2(q)$, i.e., into a steric and an electrostatic component to mirror the reduction in free solubilized electrolyte ions within the region. The steric component is deduced from the reduction of the solvent accessible (free) volume by the dendrimer, and a linear relationship between electrolyte concentration and resistivity was assumed from from a first-order expansion of the by Kohlrausch law [56]. This leads to an expression for the steric relative resitivity $\rho = V/V'$, where V is hemispherical volume in which the dendrimer in enveloped, and V is portion of this volume unoccupied by the dendrimer (solvent accessible). The second component ρ_2 reflects the lowering of electrolyte cation concentration caused by the electrostatic repulsion against the positively charged, protonated PAMAM terminal amine groups at the end of dendrimer branches positioned outside the pore and is assumed to increase linearly with the number of charges outside. The protonation in this structurally non-compressed PAMAM parts differs to the dendrimer parts which are confined within the pore. There, the protonation of their terminal primary amines is, in agreement with other polyelectrolytes [57, 58], assumed to be suppressed to avoid mutual energetically unfavorable electrostatic repulsion, as reflected by the solely steric term $S'(z)$ for calculating the ionic current. A more detailed description of how parameter ρ is derived is included in the next section.

The blockade amplitude was deduced from the ratio of $\frac{I_b}{I_o}$ using equation 4.

$$\%\text{Blockade amplitude} = 100 \times (1 - \frac{I_b}{I_o}) \tag{3.4}$$

3.5.1 Estimating Effective Resistivity (ρ)

The resistivity increase in the hemispherical region outside the pore occupied by the large dendrimers is expected to be due not only to steric but also electrostatic effects. Therefore our relative resistivity estimate has two components. We estimated the steric component from a ratio (V/V') of the ion accessible free volumes in the absence (V) and in the presence (V') of the dendrimer. This estimate is based on a first order approximation to Kohlrausch's law for dilute solution, which shows a direct relationship between conductivity and the concentration of the ions in the solution [56]. For G3, the computed value was 1.66. This estimate is in accord with a previous study done with neutral water-soluble polymers, for which the effect of polymers on bulk solution conductivity was found to result in a decrease of the conductivity by a factor of about 2 [59]. Our result, which is based only on steric effects, has a similar order of magnitude in the ratio of resistivites.

In order to estimate the electrostatic contribution factor to the resistivity, we used the best overall fit to the experiment and found that a value of 2.5 yielded a good fit for G3. By using the two components (steric and electorstatic), the overall effective resistivity (ρ) for G3 was 4.15. For G4, the steric component of the resistivity was calculated to be 2.81 with the same method. To estimate the scaling needed to calculate electrostatic component for G4, we compared the number of terminal charges outside the pore. Accordingly, G4 has 3 times the number of charges outside the pore (48) than G3 (16). This led us to estimate electrostatic component of resistivity for G4 by multiplying with 3 the corresponding term for G3. After we determined the electrostatic component to be 7.5, we estimated the overall resistivity to be 21 for G4.

It has previously been shown that hydration forces due to ordering of water molecules caused by charged molecules in general, and dendrimers in particular, can extend beyond the range of electrostatic forces.6 Therefore the charged dendrimer terminals staying outside the pore

are expected not only to modulate the local ion concentration but also to extend the range of electrostatic forces through ordered water molecules.

3.5.2 Understanding Dendrimer Compressibility

Relative Angular Mobility

To compare the varying degrees of structural flexibility of G1 to G5, we quantified the mobility of the PAMAM branches with respect to each other with the Relative Angular Mobility (RAM) estimator:

$$\mathrm{RAM} = \frac{\theta(t) - \theta(0)}{\theta(0)} \tag{3.5}$$

where $\theta(t)$ is the angle between one terminal nitrogen, one central carbon, and another terminal nitrogen at time t as shown in Figure 3.10, with $\theta(0)$ is the corresponding angle at the beginning of the simulation. Positive RAM values indicate increasing separation of the involved terminals compared with the initial separation and negative values indicate decreasing distance between the terminals. We calculated RAM for select terminals during equilibrations of dendrimers for each generation. As seen in Figure 3.11, the fluctuations in the relative mobilities consistently decrease with increasing generation. This is consistent with the stronger intramolecular interactions at higher dendrimer generation, which in turn limit mobility of the branches and reduce overall dendrimer compressibility.

(a)   (b)
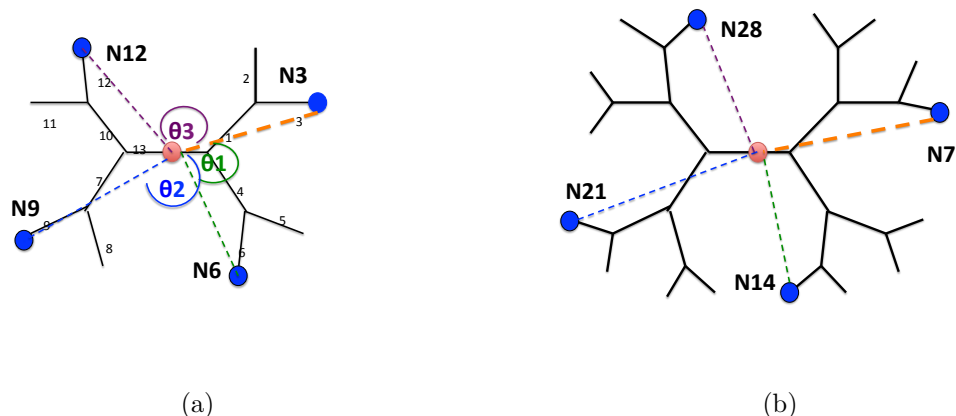
Figure 3.10: Sketch illustrating the angle descriptions for Relative Angular Mobility (RAM) calculations for 3.10(a) G1 and 3.10(b) G2.

Dendrimer Self Energy

In order to quantify the overall energetics of PAMAM permeation into the protein pore, we calculated the z-dependent self energy. The self energy is the total intramolecular potential energy calculated from the sum of all bonded and non-bonded interactions between the dendrimer atoms. For this purpose, the simulations were conducted using a voltage of 4.8 V for 1 ns and 120 mV for 0.5 ns. Figure 3.12 shows a plot of self energy normalized to the number of the dendrimer atoms, as a function of dendrimer center of mass position. Figure 3.13 is the analogous plot against simulation time. The plots illustrate that dendrimers tend to stay upon application of the lower potential- within $5 - 10$ Å of their recent positions reached under the higher potential. The general trend is that the normalized self-energy increases with dendrimer generation in line with higher steric hindrance. Against this trend, G1 has a higher self-energy than G2, likely because G1 is being quickly drawn into the pore constriction where the electrostatic potential gradient has its maximum [48]. In other words, the strong electrostatically-induced attraction compensates the considerable steric compression.

Molecular dynamics establish that pore permeation is not directly determined by dendrimer diameter but by the generation-dependent compressibility of PAMAM. The results of our
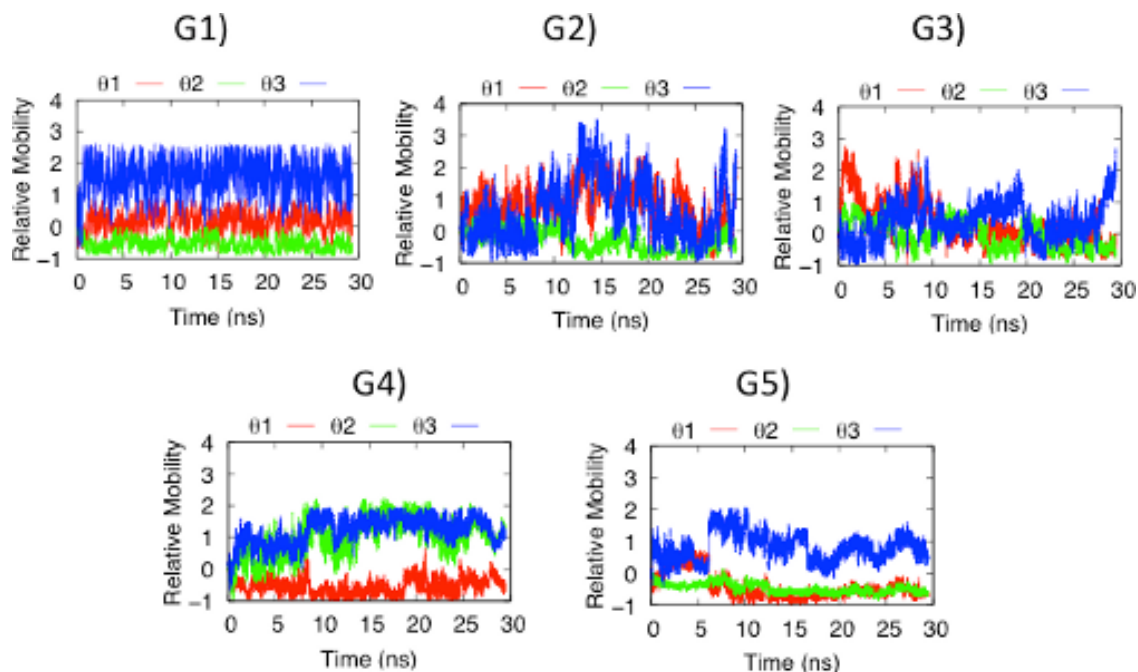
Figure 3.11: Relative Angular Mobility time series for dendrimers G1-G5 during equilibration.

simulations are summarized in Figure 3.9. They reveal that the permeation into the pore strongly depends on the dendrimer generation. G1 completely entered the β-barrel to block the inner constriction while G2 did not reach as deeply. By contrast, G3 did not fully enter the barrel, and only the smaller portion of G4 and G5 was inside the pore.

The simulations are highly valuable as they firstly replicate the experimentally found permeation of large dendrimers. Similar to nanopore recordings, all PAMAM dendrimers from G1 to G4 at least partly if not completely entered the pore in the simulations (Figure 3.9) even though the hydrodynamic diameter for all is larger then the trans-entrance. Secondly, the computational data distinguish two molecular modes for pore entrance for pore entrance. G1 and G2 enter the pore completely (Figure 3.9). By contrast, only several branches of the dendrimer G4 are able to reach into the pore lumen (Figure 3.9). G3 occupies a middle position between the two extremes. The simulations, thirdly, provide a reason for the generation-dependent entrance mode in terms of structural flexibility and compressibility of the outermost dendrimer shell [60]. The different degrees of flexibilities are apparent in
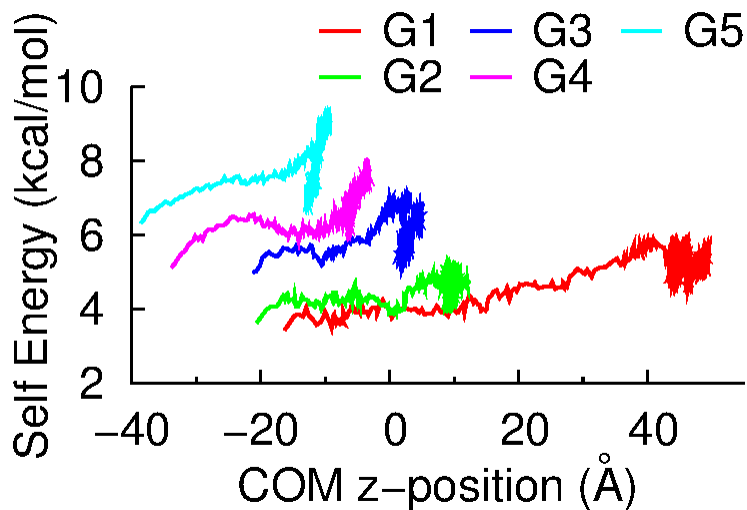
Figure 3.12: Dendrimer self energy values normalized by the number of dendrimer atoms.

the computationally derived dendrimers terminal relative mobilities which represent the size dependent change in the angle from the dendrimer center to two terminal branches (Figures 3.10 and 3.11). As a further illustration that rigidity increases with dendrimer size, we show in Figures 3.12 and 3.13 that the intramolecular potential energy required to push the dendrimer into the pore scales with generation.

Using this insight, we can explain that G1 and G2 can enter so deeply into the pore because the required associated rearrangement of the PAMAM branches is energetically feasible due to the flexible and non-crowded outer dendrimer shell. By contrast, only a few branches of the G4 dendrimer (Figure 3) enter the pore because the crowded outer shell of the dendrimer does not allow for a larger structural reorganization as found for G1 or G2. Indeed, G4 cannot reach deeper into the pore as this would require the energetically costly pushing aside the other branches of the dendrimer. One discrepancy is that PAMAM G5 entered the pore in the simulations but not in the experiments. A possible reason for this mismatch could be that a PAMAM G5 branch resides too short in the barrel to be detected amid the current noise of the open channel.

To achieve a quantitative measure of the varying degrees of pore blockade, we computed from
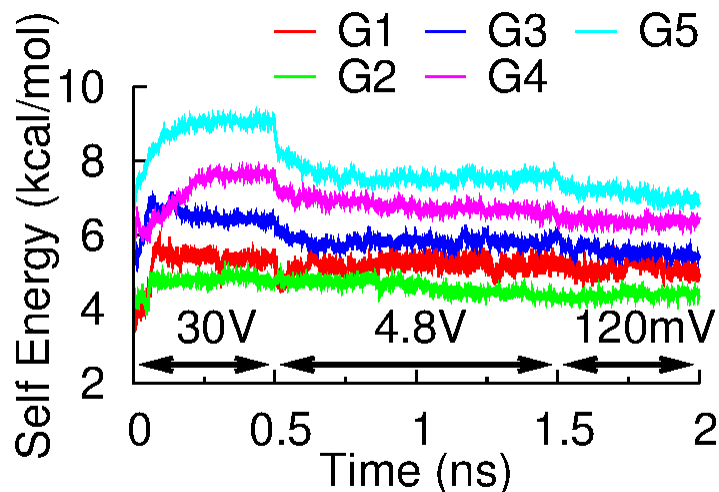
Figure 3.13: Dendrimer self energy time series normalized by dendrimer atom numbers.

the molecular dynamics simulations the open cross-sectional area of the lumen, S(z), against the z coordinate running along the central pore axis (Figure 3.9). S(z) is the area of the pore which is not blocked by a PAMAM dendrimer. For reference, the open (non-PAMAM blocked) pore area is in black (Figure 3). The plot highlights that G1 blocks the pore at the inner constriction while G2 and G3 constrict mostly the middle part of the β-barrel.

3.6   Results and Discussions

As shown in Figure 3.15, the theory was able to capture well the experimentally found non-monotonous dependence of pore blockade on dendrimer generation. This is of note, as our model does not use any phenomenological fit to experimental data but rather calculates the current from dendrimer-induced and position-dependent blocked pore and access volumes, and electrostatic factors. The small offset of approximately 10% between the computational and experimental blockade levels could have been overcome by including a mathematical fitting factor in Eq. 3.1, but doing so would have diluted the simplicity and biophysical coherence of our model. The computational model correctly predicts the experimentally

observed trend, yet it slightly and consistently overestimates the blockade amplitude. A possible explanation for this consistent shift could be the different temporal scales underlying the simulated dynamics vs. the experimental measurement of the pore-blocking events. As the temporal resolution of the measurement is about 50 microseconds, time-averaged partially blocked conformations could be present in the measured data, but not in the simulation, which accounts for nanosecond dynamics in the thermodynamically dominant state of dendrimer-pore interactions.

Nevertheless, the agreement between theory and experiment is remarkable considering that the model is relatively simple compared to several more complex computational strategies [48, 51, 52]. A key reason contributing to the agreement, particularly for G3 and G4, has to do with the relative resistivities of PAMAM inside and outside the channel. The difference in the resistivities is biophysically justified by the steric component, and by the fact that squeezing PAMAM inside the channel pushes charged terminal amino groups together and hence lowers their protonation state to avoid energetically unfavorable electrostatic repulsion under confinement. Consequently, PAMAM branches inside the pore are considered to solely act as steric but not as an electrostatic barrier for the passage of electrolyte cations, which in turn results in a higher ionic current compared to flux through PAMAM outside the pore. When expressed via the parameter ρ, the resistivity of non-confined PAMAM is about 4 times bigger for G3 and 21 times bigger for G4 than of dendrimer branches confined inside the pore. This parameter is key for obtaining a good match to the data, as setting it to equal in and outside the pore resulted in a mismatch between experiment and data as shown by a comparison of Figures 3.14(b) and 3.15.
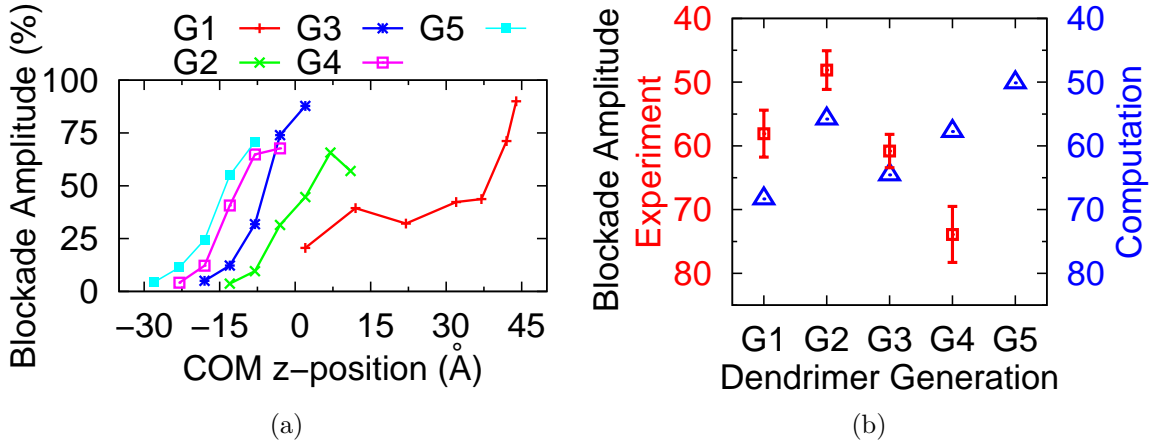
Figure 3.14: 3.14(a) Position dependence of current blockade % for each dendrimer generation during the first 500 ps simulation under a bias of 30 V voltage ($\rho = 1$). 3.14(b) Comparison of experimental % current blockade values to the simulation results as obtained from the average of the last three data points.

3.7 Conclusions

We have measured and simulated how PAMAM dendrimers permeate into the widely used model pore αHL to uncover fundamental insight into dendrimer-pore interactions at the nanoscale. The key findings of our study are threefold. Firstly, we show that permeation is not governed by dendrimer size but by the generation-dependent conformational flexibility, i.e., by the details of if and how dendrimers can modulate their shape and internal structure to enter the pore. The result is important for basic and applied science as it is now easier to predict the optimal dendrimer generation for a given nanopore diameter. Secondly, our data suggest that pore-confinement is reducing the ionization state of PAMAM which, in turn, can help anticipate the permeation properties through charged nanopores. Thirdly, we demonstrate that the permeation-induced current blockade exhibits a non-trivial dependence on dendrimer generation. This is caused by an intriguing interplay pore geometry, different dendrimer compressibilities, and confinement-induced changes in PAMAM properties including ionization. Small, easily compressible dendrimers like G1 and G2 completely move into the cylinder-shaped pore to cause a large blockage caused by steric factors. By contrast, at the
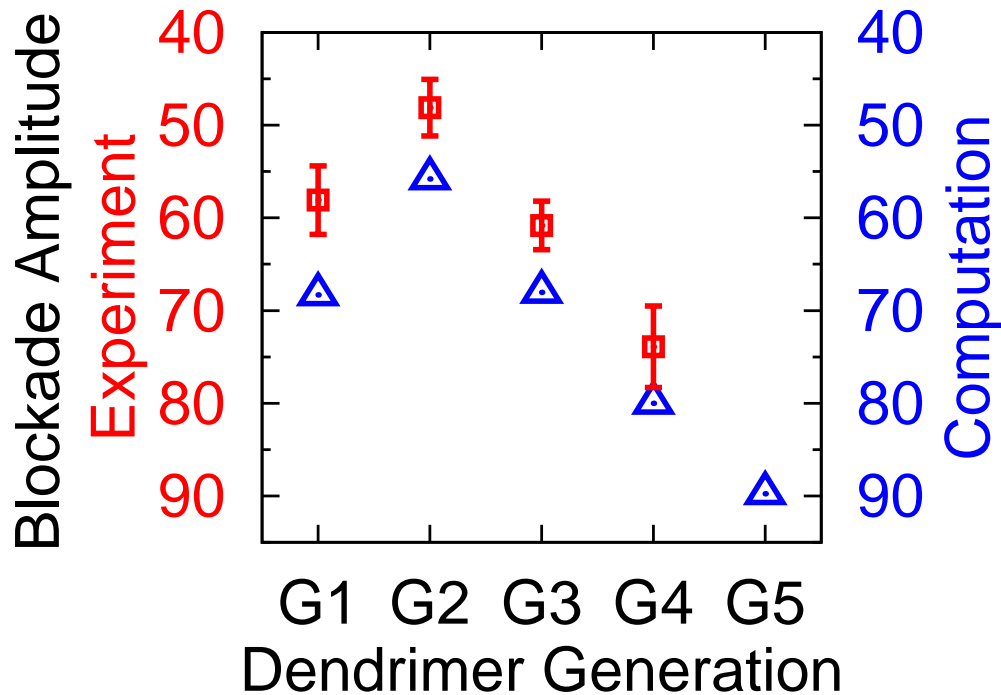
Figure 3.15: A comparison of experimental and computational results for % current blockade values calculated with estimated effective resistance (ρ) values.

other extreme, larger poorly compressible dendrimer G4 does not completely enter the pore but causes a larger current blockade as the PAMAM section outside the pore is fully protonated to electrostatically repel electrolyte cations. In conclusion, our scientifically rich study highlights the importance of molecular flexibility and non-trivial nanoparticle nanopore interactions in modulating pore transport, and can likely be expanded to other dendrimers or molecules designed to permeate or lodge inside nanoporous structure to improve applications in biotechnology, biomedicine, material science, and biophysical research.

Bibliography

[1] Astruc, D., Ornelas, C., Diallo, A. K., and Ruiz, J. (2010) Extremely Efficient Catalysis of Carbon-Carbon Bond Formation Using "Click" Dendrimer-Stabilized Palladium Nanoparticles. *Molecules 15*, 4947–4960.

[2] Lee, C. C., MacKay, J. A., Fréchet, J. M. J., and Szoka, F. C. (2005) Designing Dendrimers for Biological Applications. *Nature Biotechnol. 23*, 1517–1526.

[3] Baker, L. A., and Bird, S. P. (2008) Nanopores - A makeover for membranes. *Nature Nanotechnol. 3*, 73–74.

[4] Howorka, S., and Siwy, Z. (2009) Nanopore Analytics: Sensing of Single Molecules. *Chem. Soc. Rev. 38*, 2360.

[5] Bayley, H., and Cremer, P. S. (2001) Stochastic Sensors Inspired by Biology. *Nature 413*, 226–230.

[6] Mayer, M., and Yang, J. (2013) Engineered Ion Channels as Emerging Tools for Chemical Biology. *Acc. Chem. Res. 46*, 2998–3008, PMID: 23932142.

[7] Cherf, G. M., Lieberman, K. R., Rashid, H., Lam, C. E., Karplus, K., and Akeson, M. (2012) Automated Forward and Reverse Ratcheting of DNA in a Nanopore at 5-Å Precision. *Nature Biotechnol. 30*, 344–348.

[8] Manrao, E. A., Derrington, I. M., Laszlo, A. H., Langford, K. W., Hopper, M. K., Gillgren, N., Pavlenok, M., Niederweis, M., and Gundlach, J. H. (2012) Reading DNA at Single-Nucleotide Resolution with a Mutant MspA Nanopore and Phi29 DNA Polymerase. *Nature Biotechnol. 30*, 349–353.

[9] Ignacio-de Leon, P. A. A., and Zharov, I. (2011) Size-Selective Molecular Transport Through Silica Colloidal Nanopores. *Chem. Commun. 47*, 553–555.

[10] Sarin, H., Kanevsky, A. S., Wu, H., Brimacombe, K. R., Fung, S. H., Sousa, A. A., Auh, S., Wilson, C. M., Sharma, K., Aronova, M. A., Leapman, R. D., Griffiths, G. L., and Hall, M. D. (2008) Effective Transvascular Delivery of Nanoparticles Across the Blood-Brain Tumor Barrier into Malignant Glioma Cells. *J. Transl. Med. 6*, 80.

[11] Sarin, H., Kanevsky, A. S., Wu, H., Sousa, A. A., Wilson, C. M., Aronova, M. A., Griffiths, G. L., Leapman, R. D., and Vo, H. Q. (2009) Physiologic Upper Limit of Pore Size in the Blood-Tumor Barrier of Malignant Solid Tumors. *J. Transl. Med. 7*, 51.

[12] Kong, L., Harrington, L., Li, Q., Cheley, S., Davis, B. G., and Bayley, H. (2013) Single-Molecule Interrogation of a Bacterial Sugar Transporter Allows the Discovery of an Extracellular Inhibitor. *Nat. Chem. 5*, 651–659.

[13] Förstner, P., Bayer, F., Kalu, N., Felsen, S., Förtsch, C., Aloufi, A., Ng, D. Y. W., Weil, T., Nestorovich, E. M., and Barth, H. (2014) Cationic PAMAM Dendrimers As Pore-Blocking Binary Toxin Inhibitors. *Biomacromolecules 15*, 2461–2474.

[14] Martin, H., Kinns, H., Mitchell, N., Astier, Y., Madathil, R., and Howorka, S. (2007) Nanoscale Protein Pores Modified with PAMAM Dendrimers. *J. Am. Chem. Soc. 129*, 9640–9649.

[15] Lindner, J.-P., R ben, C., Studer, A., Stasiak, M., Ronge, R., Greiner, A., and Wendorff, H.-J. (2009) Reusable Catalysts Based on Dendrimers Trapped in Poly(p-xylylene) Nanotubes. *Angew. Chem. Int. Ed. 48*, 8874–8877.

[16] Kasianowicz, J., Brandin, E., Branton, D., and Deamer, D. (1996) Characterization of Individual Polynucleotide Molecules Using a Membrane?Channel. *Proc. Nat. Acad. Sci. U. S. A. 93*, 13770–13773.

[17] Vercoutere, W., and Akeson, M. (2002) Biosensors for DNA Sequence Detection. *Curr. Opin. Chem. Biol. 6*, 816–822.

[18] Deamer, D. W., and Branton, D. (2002) Characterization of Nucleic Acids by Nanopore Analysis. *Acc. Chem. Res. 35*, 817–825.

[19] Maglia, G., Restrepo, M. R., Mikhailova, E., and Bayley, H. (2008) Enhanced Translocation of Single DNA Molecules through α-Hemolysin Nanopores by Manipulation of Internal Charge. *Proc. Nat. Acad. Sci. U. S. A. 105*, 19720–19725.

[20] Butler, T. Z., Gundlach, J. H., and Troll, M. A. (2006) Determination of RNA Orientation during Translocation through a Biological Nanopore. *Biophys. J. 90*, 190–199.

[21] Wang, H., Dunning, J. E., Huang, A. P.-H., Nyamwanda, J. A., and Branton, D. (2004) DNA heterogeneity and phosphorylation unveiled by single-molecule electrophoresis. *Proc. Nat. Acad. Sci. U. S. A. 101*, 13472–13477.

[22] Wanunu, M., Chakrabarti, B., Mathé, J., Nelson, D. R., and Meller, A. (2008) Orientation-Dependent Interactions of DNA with an α-Hemolysin Channel. *Phys. Rev. E 77*, 031904.

[23] Wiggin, M., Tropini, C., Tabard-Cossa, V., Jetha, N. N., and Marziali, A. (2008) Nonexponential Kinetics of DNA Escape from α-Hemolysin Nanopores. *Biophys. J. 95*, 5317–5323.

[24] Meller, A., Nivon, L., and Branton, D. (2001) Voltage-Driven DNA Translocations through a Nanopore. *Phys. Rev. Lett. 86*, 3435–3438.

[25] Howorka, S., and Bayley, H. (2002) Probing Distance and Electrical Potential within a Protein Pore with Tethered DNA. *Biophys. J. 83*, 3202–3210.

[26] Sutherland, T., Long, Y., Stefureac, R., Bediako-Amoa, I., Kraatz, H., and Lee, J. (2004) Structure of Peptides Investigated by Nanopore Analysis. *Nano Lett. 4*, 1273–1277.

[27] Movileanu, L., Schmittschmitt, J. P., Martin Scholtz, J., and Bayley, H. (2005) Interactions of Peptides with a Protein Pore. *Biophys. J. 89*, 1030–1045.

[28] Huang, L., Kirmizialtin, S., and Makarov, D. E. (2005) Computer Simulations of the Translocation and Unfolding of a Protein Pulled Mechanically through a Pore. *J. Chem. Phys. 123*, 124903.

[29] Oukhaled, G., Mathé, J., Biance, A. L., Bacri, L., Betton, J. M., Lairez, D., Pelta, J., and Auvray, L. (2007) Unfolding of Proteins and Long Transient Conformations Detected by Single Nanopore Recording. *Phys. Rev. Lett. 98*, 158101.

[30] Talaga, D. S., and Li, J. (2009) Single-Molecule Protein Unfolding in Solid State Nanopores. *J. Am. Chem. Soc. 131*, 9287–9297.

[31] Stefureac, R., Long, Y.-t., Kraatz, H.-B., Howard, P., and Lee, J. S. (2006) Transport of α-Helical Peptides through α-Hemolysin and Aerolysin Pores. *Biochemistry 45*, 9172–9179.

[32] Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J. E. (1996) Structure of Staphylococcal α-Hemolysin, a Heptameric Transmembrane Pore. *Science 274*, 1859–1866.

[33] Bauer, B. J., and Amis, E. J. *Dendrimers and Other Dendritic Polymers*; John Wiley Sons, Ltd, 2002; pp 255–284.

[34] Zlatanova, J., and van Holde, K. (2006) Single-Molecule Biology: What is it and How Does It Work? *Mol. Cell 24*, 317–329.

[35] Mitchell, N., and Howorka, S. (2008) Chemical Tags Facilitate the Sensing of Individual DNA Strands with Nanopores. *Angew. Chem. Int. Ed. 47*, 5565–5568.

[36] Howorka, S., Movileanu, L., Braha, O., and Bayley, H. (2001) Kinetics of Duplex Formation for Individual DNA Strands Within a Single Protein Nanopore. *Proc. Nat. Acad. Sci. U.S.A. 98*, 12996–13001.

[37] Howorka, S., Cheley, S., and Bayley, H. (2001) Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnology 19*, 636–639.

[38] Braha, O., Walker, B., Cheley, S., Kasianowicz, J., Song, L., Gouaux, J., and Bayley, H. (1997) Designed protein pores as components for biosensors. *Chem. Biol. 4*, 497–505.

[39] Movileanu, L., Howorka, S., Braha, O., and Bayley, H. (2000) Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nature Biotechnology*

[40] Kelly, C. V., Leroueil, P. R., Nett, E. K., Wereszczynski, J. M., Baker, J. R., Orr, B. G., Banaszak Holl, M. M., and Andricioaei, I. (2008) Poly(amidoamine) Dendrimers on Lipid Bilayers I: Free Energy and Conformation of Binding. *J. Phys. Chem. B 112*, 9337–9345.

[41] Cakara, D., Kleimann, J., and Borkovec, M. (2003) Microscopic Protonation Equilibria of Poly(amidoamine) Dendrimers from Macroscopic Titrations. *Macromolecules 36*, 4201–4207.

[42] Lee, I., Athey, B. D., Wetzel, A. W., Meixner, W., and Baker, J. R. (2002) Structural Molecular Dynamics Studies on Polyamidoamine Dendrimers for a Therapeutic Application: Effects of pH and Generation. *Macromolecules 35*, 4510–4520.

[43] Maiti, P. K., Çağın, T., Wang, G., and Goddard, W. A. (2004) Structure of PAMAM Dendrimers: Generations 1 Through 11. *Macromolecules 37*, 6236–6254.

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

[45] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C.,

Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable Molecular Dynamics with NAMD. *J. Comput. Chem. 26*, 1781–1802.

[46] Mackerell, A. D. et al. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins . *J. Phys. Chem. B 102*, 3586–3616.

[47] Roux, B. (2008) The Membrane Potential and Its Representation by a Constant Electric Field in Computer Simulations. *Biophys. J. 95*, 4205–4216.

[48] Aksimentiev, A., and Schulten, K. (2005) Imaging α-Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map. *Biophys. J. 88*, 3745–3761.

[49] Gumbart, J., Khalili-Araghi, F., Sotomayor, M., and Roux, B. (2012) Constant Electric Field Simulations of the Membrane Potential Illustrated with Simple Systems. *Biochim. Biophys. Acta. 1818*, 294–302.

[50] Wells, D. B., Abramkina, V., and Aksimentiev, A. (2007) Exploring Transmembrane Transport Through α-Hemolysin with Grid-Steered Molecular Dynamics. *J. Chem. Phys. 127*, 125101.

[51] Noskov, S. Y., Im, W., and Roux, B. (2004) Ion Permeation through the α-Hemolysin Channel: Theoretical Studies Based on Brownian Dynamics and Poisson-Nernst-Plank Electrodiffusion Theory. *Biophys. J. 87*, 2299–2309.

[52] Freites, J. A., Schow, E. V., White, S. H., and Tobias, D. J. (2012) Microscopic Origin of Gating Current Fluctuations in a Potassium Channel Voltage Sensor. *Biophys. J. 102*, L44–L46.

[53] Mereuta, L., Roy, M., Asandei, A., Lee, J. K., Park, Y., Andricioaei, I., and Luchian, T. (2014) Slowing down Single-Molecule Trafficking Through a Protein Nanopore Reveals Intermediates for Peptide Translocation. *Sci. Rep. 4*.

[54] Hall, J. E. (1975) Access Resistance of a Small Circular Pore. *J. Gen. Physiol. 66*, 531–532.

[55] Hille, B. (1970) Ionic Channels in Nerve Membranes. *Prog. Biophys. Mol. Biol. 21*, 1–32.

[56] Atkins, P., and De Paula, J. *Atkins' Physical Chemistry*; Oxford University Press: Oxford, 2006.

[57] Nap, R. J., Tagliazucchi, M., and Szleifer, I. (2014) Born Energy, Acid-Base Equilibrium, Structure and Interactions of End-Grafted Weak Polyelectrolyte Layers. *J. Chem. Phys. 140*, 024910.

[58] Longo, G. S., Olvera de la Cruz, M., and Szleifer, I. (2014) Equilibrium Adsorption of Hexahistidine on pH-Responsive Hydrogel Nanofilms. *Langmuir 30*, 15335–15344.

[59] Bezrukov, S. M., and Vodyanoy, I. (1993) Probing Alamethicin Channels with Water-Soluble Polymers. Effect on Conductance of Channel States. *Biophys. J. 64*, 16–25.

[60] Hierlemann, A., Campbell, J. K., and Baker, L. A. (1998) Structural Distortion of Dendrimers on Gold Surfaces: a Tapping-Mode AFM Investigation. *J. Am. Chem. Soc.*

Chapter 4

The Role of Proton Motive Force in Conformational Transition of TtSecDF from F-form to I-form

4.1 Introduction

SecDF is an important component of the Sec protein translocation machinery embedded in the bacterial membrane. In combination with its other components, namely SecYEG and SecA, SecDF is associated with many functions such as stabilization of the other components within the membrane, maintaining the transmembrane potential which is crucial for the function of this machinery, and facilitating the ATP independent stage of the protein translocation mechanism. Related studies suggest that SecDF undergoes functionally important conformational changes which involve mainly the P1-head domain and that these changes are coupled with the proton motive force. However, there is still not a clear understanding of how SecDF functions, its exact role in the translocation machinery, and how

its function is related to the proton motive force. In order to understand the nature of this suggested conformational change and how it may be coupled with the proton motive force in atomistic detail, we used targeted molecular dynamics (TMD) to find low energy pathways between the two conformations, i.e. F-form and I-form. Additionally we also obtained potentials of mean force (PMF) along a root-mean-square-distance (RMSD) based reaction coordinate. A comparison of the free energy profiles obtained in the existence and absence of a transmembrane potential indicate that the interaction of the P1 domain dipole moment with a transmembrane potential considerably lowers the free energy barrier in the direction of F-form to I-form transition. Our results also show that I-form to F-form transition is likely only in the absence of a transmembrane potential.

4.2   Background on SecDF

An important mechanism in the infection of host cells by bacterial virulence factors is the translocation of unfolded secretory proteins across bacterial membrane [1]. One such translocation mechanism is the Sec pathway [2–5]. Sec pathway is the common pathway between prokaryotes and eukaryotes in the translocation of secretory or plasma membrane proteins across the membrane. In prokaryotes, translocation is through the plasma membrane, while in eukaryotes it is through the endoplasmic reticulum membrane [4]. Sec translocon machinery is a protein-conducting channel which functions not only to translocate unfolded proteins across the plasma membrane but also to insert proteins into the membrane [2].

In many bacteria, the Sec machinery is composed of SecYEG [6, 7], the core of the complex forming a protein conducting channel, Sec A ATPase [1, 2], which initiates the translocation through binding of ATP, and SecDF, which is thought to facilitate the translocation in the later stages. In some organisms, SecDF is found as closely interacting two proteins, i.e. Sec D and SecF, and in some others it is formed by a single chain [8].

Studies show that SecDF is functional in various stages of protein translocation. For example, it is thought to stabilize other components of the Sec machinery, i.e. SecY [9] and membrane-inserted state of SecA [10–12]. It also regulates the protein translocation by stabilizing the translocation intermediates and by preventing their backward movement [12]. SecDF is also important in the release of preprotein to the periplasmic side [13]. Besides its direct role in protein translocation as mentioned, SecDF is also required for the maintenance of proton motive force levels, which is needed for efficient translocation [12, 14]. It is due to these and related functions that the lack of SecDF has been shown to result in severe growth retardation and defects in protein-export *in vivo* [15, 16].

While there are mixed results in the role of SecDF in translocation in *in vitro* studies, *in vivo* studies show an indispensable role in effective translocation of proteins in a fully active state [15–17]. Some studies show that the rate of protein translocation depends on the levels of SecD and SecF and that lack of SecDF results in secretion defect and growth inhibition [15]. In one study with *Listeria monocytogenes*, a gram-positive bacteria, SecDF is shown to be dispensable for growth *in vitro*, but it is required *in vivo* for bacterial infection in facilitating the translocation of virulence factors in a fully active state [17]. SecDF mutants resulted in reduced activity of translocated proteins, presumably due to misfolding. Its important role in the translocation of several virulence factors, and hence in the bacterial infection process, makes SecDF an important possible drug target in preventing bacterial infections. Also because of its possible role in protein folding, understanding how SecDF functions is important.

Despite its significant role in facilitating protein translocation, protein folding, and membrane protein insertion, the mechanism of SecDF function remains largely unexplored. It is shown that SecDF functions in combination with proton motive force, Δp, which is the electrochemical potential difference of proton across the membrane [18]. It is shown that both chemical and electrical components of proton motive force, i.e. pH gradient, ΔpH and

94

transmembrane electric potential , $\Delta\psi$, respectively, have a direct role in the effective functioning of SecDF [19]. ATP and proton motive force act at different parts of the preprotein translocation catalytic cycle [19, 20]. ATP binding to SecA initiates the translocation. ATP hydrolysis later provides the release of the unfolded protein. Proton motive force acts at later stages in preventing the backward movement of the preprotein and it drives the rapid and efficient forward translocation reaction. The coupling of SecDF function to proton motive force suggests that SecDF undergoes a conformational change driven by proton motive force to function.

In a recent study, structure of SecDF from *Thermus thermophilus* HB8, or TtSecDF, show, indeed, two different conformations, F-form and I-form [21, 22] (See Figure 4.1). The two conformations vary by the relative positioning of P1-head, i.e. the large periplasmic domain of TtSecDF, which is suggested to undergo a hinge-motion. An analysis of SecDF structure using electron tomography and single particle reconstruction [23] also supports the existence of two conformations consistent with the F-form and I-form structures reported [22] earlier. The latter study also confirms the conformational flexibility of P1 head relative to other domains, as predicted by the former study.

While it is known that proper functioning of SecDF depends on proton motive force and a conformational change driven by proton motive force, the exact mechanism and how the two are coupled are not known. In this study, we find the possible low energy pathways for the SecDF conformational transitions in both directions, i.e. F-form to I-form and I-form to F-form, by using all atom molecular dynamics simulations and Targeted Molecular Dynamics (TMD) method. In order to understand the role of proton motive force and how it influences the transition, we also study the conformational transitions in the existence of the transmembrane electrostatic potential component of the proton motive force, which we modeled through a constant electric field. In addition, through potential of mean force calculations along the conformational transition reaction pathways, we show that the interaction of the

dipole moment of SecDF P1-head with the transmembrane potential significantly lowers the barrier for F-to-I transition, and makes the I-form energetically more favorable. This can explain the source of the driving force needed for SecDF to carry the translocation intermediates from inside the channel after capturing them while in the F-form, into the periplasm, as it transitions into the I-form conformation. With the depletion of the transmembrane potential, i.e. in the absence of potential, the barriers for transition in both directions seem to have similar barriers.

In the following sections, we first give information about the targeted molecular dynamics (TMD) and umbrella sampling methods we used in this study. Next we explain the details of the structure setup and simulations. Finally, we present and discuss our results followed by concluding remarks.

4.3 Theory and Methods

4.3.1 Targeted Molecular Dynamics

In this study, targeted molecular dynamics (TMD) method is used to find a low energy pathway for conformational transitions between F-form and I-form. In this method, an evolving holonomic constraint is used to drive the system from an initial configuration to a target configuration. A more detailed explanation of this method is included in Section 1.1.

4.3.2 Umbrella Sampling

Once a minimum energy reaction pathway is obtained for the conformational transition, potential of mean force (PMF) [36] along this pathway can be calculated from the probability distribution functions obtained through extensive sampling [37, 38]. One advanced sampling

technique used to overcome the problem of insufficient sampling in the regions of high energy is the umbrella sampling technique [39]. In this method, reaction coordinate is treated as a series of adjacent and overlapping windows. Separate simulations for each window is performed while the system in each window is restrained to the corresponding reaction coordinate value by using an additional harmonic biasing potential centered at that window. After the biased probability distributions are obtained by histogramming, they are unbiased by using weighted histogram method (WHAM) [40, 41].

In this study, potential of mean force along the pathway that connects the two SecDF conformations, i.e. F-form and I-form, is calculated with the use of the following biasing potential, $(w_{b,i})$, for each i-th umbrella sampling window:

$$w_{b,i} = K_i \big(\Delta RMSD(\vec{r_t}) - \Delta RMSD(\vec{r_0}) \big)^2 \tag{4.1}$$

where ΔRMSD is defined as:

$$\Delta RMSD(\vec{r}) = RMSD(\vec{r_t}, \vec{r_F}) - RMSD(\vec{r_t}, \vec{r_I}) \tag{4.2}$$

Here $RMSD(\vec{r_t}, \vec{r_F})$ and $RMSD(\vec{r_t}, \vec{r_I})$ represent the RMSD of the instantaneous structure from initial equilibrated F-form and I-form structures, respectively. $\Delta RMSD(\vec{r_0})$ is the minimum of the harmonic biasing potential. The choice of this one dimensional order parameter, $\Delta RMSD(\vec{r})$, maintains the difference between RMSD values with respect to two reference structures around the given minimum while allowing relaxation in RMSD value with respect to each reference structure. This form of biasing potential is previously used to obtain free

energy profile between B-form and A-form DNA structures [35, 42].

4.3.3 Dipole Moment Calculation

To understand how the SecDF conformational change involving P1-head subdomain may be coupled with proton motive force, we monitored the dipole moment of this subdomain along the TMD trajectories. P1-head includes four alpha helices and protein alpha-helices are known to have net dipole moments along the helix axis arising from the additivity of the individual peptide bonds [43–45]. Dipole moment $(\vec{\mu})$ of a protein or a domain can be calculated from the position vector $(\vec{r})$ and the partial charge of each atom (q_j) by using [46]:

$$\vec{\mu} = \sum_j q_j \vec{r}_j \qquad (4.3)$$

When the sum of the charges is zero, dipole moment is origin independent [46].

While the dipole moment values obtained from TMD trajectories are helpful in gaining insight about how different transition pathways compare, they are obtained under nonequilibrium conditions. In order to retrieve equilibrium values of ensemble averaged dipole moment along the reaction coordinate, trajectories from umbrella sampling simulations can be used by unbiasing [47]. The unbiased $< \mu_z >_i$, i.e. the time averaged $z-$component of dipole moment of P1-head subdomain, for each simulation window i after reweighting is calculated

as follows:

$$< \mu_z >_i = \frac{\sum\limits_{j=1}^{N_i} \mu'_{z,j} \; e^{\beta w_{b,j}}}{\sum\limits_{j=1}^{N_i} e^{\beta w_{b,j}}} \tag{4.4}$$

Here $w_{b,j}$ is the biasing potential for j^{th} snapshot in window i, $\mu'_{z,j}$ is the z component of biased dipole moment value for the same snapshot, and N_i is the total number of snapshots within window i.

4.3.4 Simulation Methods

Structure Setup

Coordinates for full-length F-form is taken from Protein Data Bank with PDB code 3AQP [22]. The coordinates of the missing residues in P4 domain are repaired after overlaying P4 domain with the NMR solution structure of P4 with PDB code 3AQO. The rest of the missing coordinates are built by using INTCOR and positions of all hydrogens were determined from HBUILD facilities in CHARMM. Full-length F-form is then oriented along z-direction such that the two transmembrane helices TM4 (residues 323-353) and TM10 (residues 619-651) are centered at $z = 0$ Å (See Figure 4.1(a)). P1 domain of I-form structure is available with PDB code 3AQO as shown in Figure 4.1(b). Full-length I-form is modeled by first superimposing heavy backbone atoms of P1 base (from PDB ID 3AQO) with corresponding region in the full-length F-from followed by appending the missing regions from coordinates of the repaired F-form structure. As seen in Figure 4.1(c), I-form differs from F-form in the way P1-head subdomain is positioned onto the P1-base subdomain.
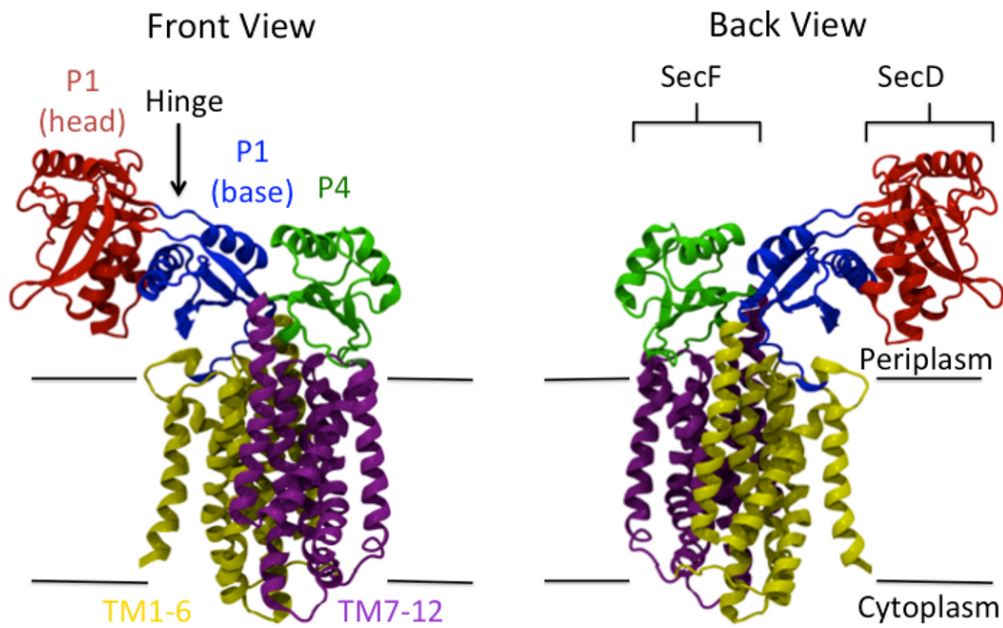
(a) Complete F-form structure (PDB ID: 3AQP)
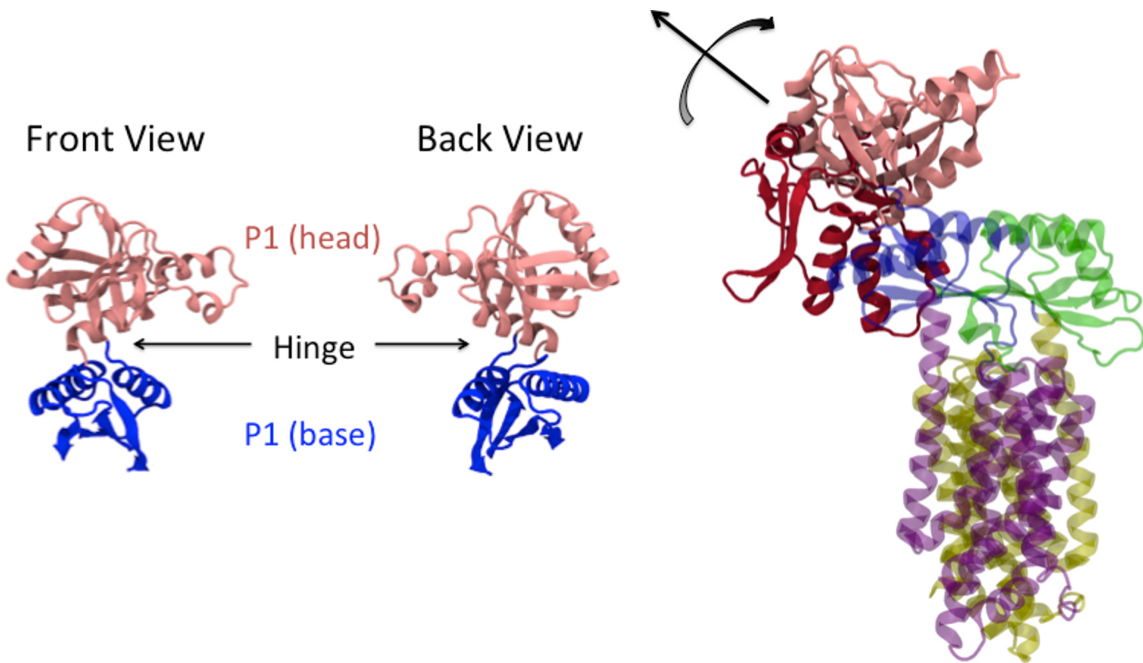
(b) I-form P1 domain (PDB ID: 3AQO)

(c) I-form P1 base aligned with F-form P1 base to build complete I-form structure

Figure 4.1: The two conformations of SecDF, 4.1(a) F-form and 4.1(b) I-form, differ by the positioning of P1-head domain on P1-base domain. F-form and I-form P1-head subdomains are shown in red and pink colors, respectively. 4.1(c) Complete I-form structure is built by superimposing I-form P1 base subdomain onto the corresponding subdomain of F-form structure and appending the missing domains from F-form structure.

Equilibration

Both F-form and I-form are equilibrated under the same conditions. SHAKE [48] is used to fix the hydrogen bonds to allow a time step of 2 fs. Velocity Verlet integrator is used with Nose thermostat to keep the temperature constant at 298.15 K. Generalized Born with a simple switching (GBSW) model [49, 50] is used to simulate the solvent and membrane effects. Membrane thickness is set to be 31 Å centered at $z = 0$ Å with a membrane switching length of 5 Å on each side. Surface tension coefficient is set to a value of 0.03 kcal/mol/$\mathring{A}^2$. Non-bonded interactions are cutoff at 12 Å with smoothing function turned on at 10 Å. CHARMM22 all-hydrogen parameters for proteins with CMAP corrections optimized specifically for GBSW is used with CHARMM software (Chemistry at HARvard Macromolecular Mechanics), version c34b2 [51–53].

First the system is minimized briefly with 50 steps of steepest descent followed by 50 steps of ABNR in implicit solvent and membrane while the backbone heavy atoms were fixed. Then a harmonic restraining force with a force constant of 60 kcal/mol/$\mathring{A}^2$ was applied to backbone heavy atoms and the system was again minimized until change in energy is 0.01 kcal/mol and then heated briefly. This cycle of minimization and heating is repeated while gradually reducing the force constant at each cycle until all restraint is removed after 300 ps. The system is then equilibrated at constant temperature for about 30 ns. The same procedure is used separately to prepare the structures for both conformations while an electric field equivalent to a transmembrane voltage of 240 mV across the membrane thickness is applied in the $-z$ direction. RMSD of the backbone heavy atoms with respect to the original structure after reorienting each frame with respect to the original structure is calculated to monitor equilibration (See Figure 4.2).

TMD Simulation

The RMSD between equilibrated F-form and I-form structures is calculated to be 12.8 Å for the case when no electric field is applied, and 13.5 Å for the case when electric field is applied. For each of these cases, I-form and F-form structures are initially placed after a best fit alignment of all atoms between the two structures. Then separate simulations are run to pull each structure towards the other, yielding four simulations: (i) F-form to I-form with electric field (f2i E), (ii) I-form to F-form with electric field (i2f E), (iii) F-form to I-form no electric field (f2i), and (iv) I-form to F-form no electric field (i2f). TMD method as implemented in CHARMM [54] is used with a pulling speed of 5 Å/ns and a time step of 1 fs. Artificial rotation is removed every 10 simulation steps. The system is weakly coupled to a heat bath at 298.15 K with a coupling constant of 0.5 ps [55]. Equilibrated structures from above are used. For each case, initial and target structures are aligned with respect to TM4 and TM10, two longest helices at the center of the transmembrane region. All the transitions studied were smooth with no unusually high energy structures encountered during transition. P1-head dipole moment is calculated along each trajectory for a selection of residues $58 - 261$, which has a zero net charge.

Umbrella Sampling

Initial structures for each umbrella window is obtained from TMD trajectories by selecting snapshots with 0.2 Å RMSD between adjacent windows. To equilibrate each structure in each window, initial simulations are run with a higher biasing force constant of 200 kcal/mol/$\mathring{A}^2$ for 20 ps. The force constant is gradually reduced to 10 kcal/mol/$\mathring{A}^2$ over the course of 80 ps after which additional equilibration is performed for 50 ps using this force constant. Finally, each umbrella sampling window is sampled for 1 ns. After obtaining biased probability distributions from the resulting trajectories, WHAM [40, 41] is used to unbias the distributions

and to obtain PMF profile. Equilibrium averages of dipole moment $z-$component, μ_z, is calculated from umbrella sampling simulations after unbiasing according to equation (4.4).

4.4 Results and Discussions

RMSD of equilibrated F-form structures from the crystal structures reached a plateau at 4 Å when electric field is applied and 5 Å when electric field is not applied (See Figure 4.2). On the other hand, RMSD for I-form structures reached a plateau at 6 Å when no electric field is applied and 7 Å when electric field is applied. The relatively higher RMSD for I-form structures can be understood by the lack of a complete I-form structure. Since the only available structure for I-form is the P1 domain, which has only P1-base subdomain in common with the complete F-form structure, the I-form structure we built assumes that P1-base subdomain in I-form has the same positioning with respect to the other domains as in F-form. However, as pointed out by another study [23], it is possible that P1 base subdomain in the two structures may have different positions relative to the other domains which may explain the relatively high RMSD of equilibrated I-form with respect to the originally built structure.

Figure 4.3 shows superimposed structures of before and after equilibration for all cases. It can be seen that P1-head subdomain in the equilibrated F-form structure lies slightly closer to the TM region than the original F-form structure. For the I-form case, equilibrated P1 domain appears closer to P4 domain than it is in the original structure.

Finding and Characterizing Conformational Transition Pathway

TMD simulations reveal that SecDF conformational change involve the proposed hinge-motion of P1-head as expected through a rotational motion between F-form and I-form
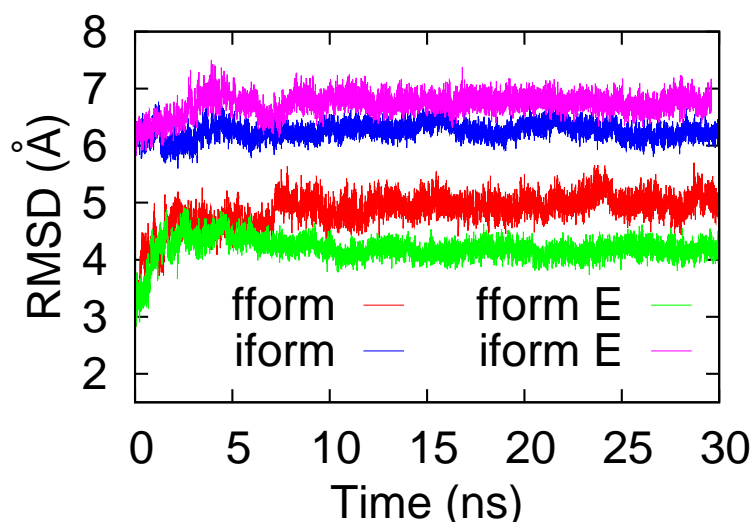
Figure 4.2: Rmsd time series for the backbone heavy atoms during equilibration. Each frame is reoriented to give the best fit to the original structure.

No Electric Field

F-form

I-form

Periplasm

Cytoplasm

(a)

Electric Field

F-form

I-form

Periplasm

+

V
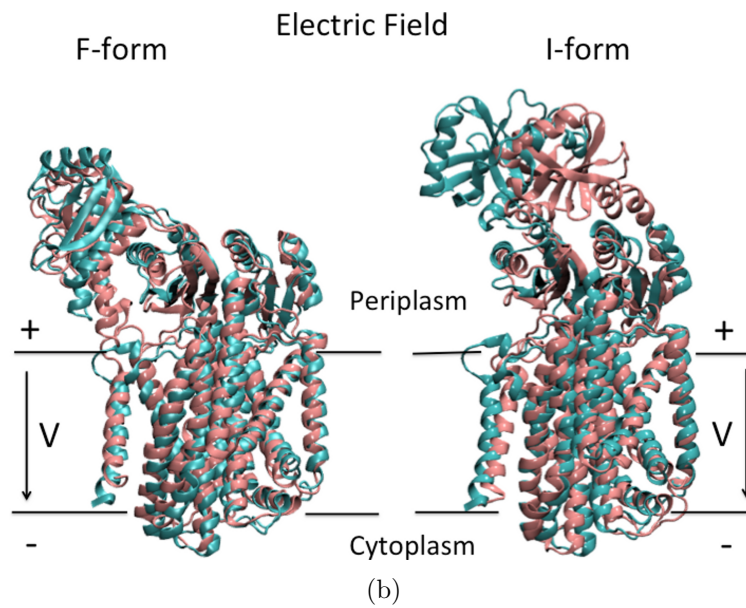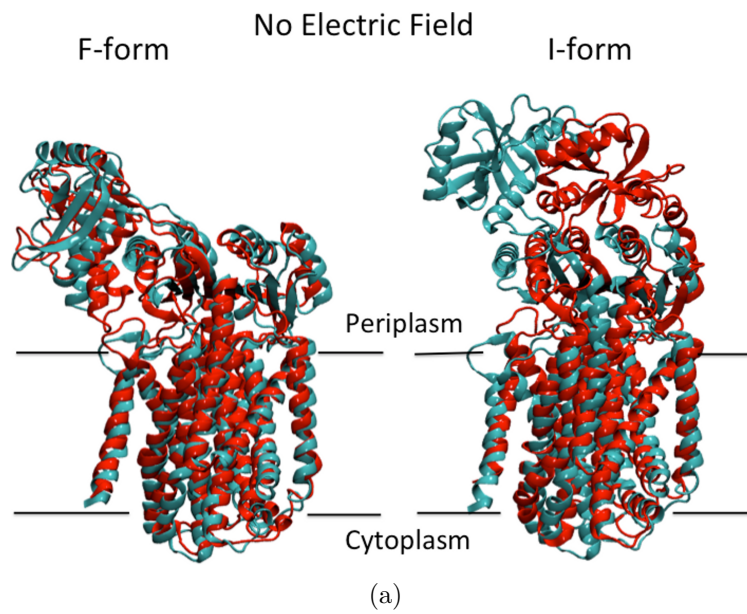
−

Cytoplasm

+

V

−

(b)

Figure 4.3: SecDF 4.3(a) F-form and 4.3(b) I-form equilibrated structures (red) superimposed on the original structures (cyan).

structures. P1-head includes four alpha helices, which are known to have considerable dipole moments. Since P1-head domain is basically neutral, it will interact with the electric field through its dipole moment. Therefore in order to investigate how P1-head dipole moment ($\vec{\mu}$) would interact with the electric field we calculated P1-head dipole moment for residues 58-261. This selection has a zero net charge and therefore dipole moment values are independent of the choice of coordinate system. Since μ_x and μ_y (Figure 4.4) are orthogonal to the electric field applied in the $-z$ direction, only z component, μ_z will interact with the electric field. As seen in Figure 4.5, all transition paths except for the case of F-to-I transition in the existence of electric field yield similar trends of μ_z. The observed hysteresis indicate a favorable interaction of P1-head dipole moment with the electric field during F-to-I transition when electric field is applied.
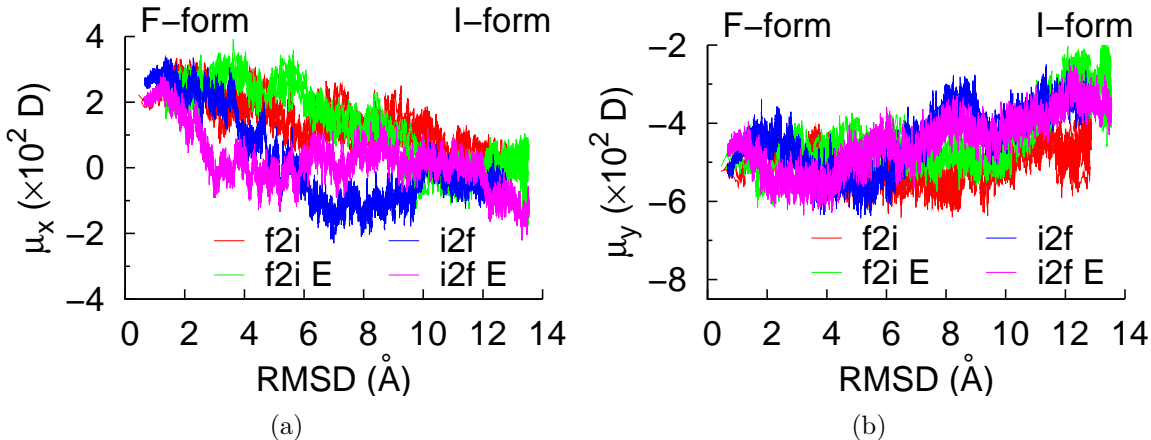
Figure 4.4: 4.4(a) x component (μ_x) and 4.4(b) y component (μ_y) of the dipole moment for P1-head subdomain.

Potential of Mean Force

In order to investigate the extent to which the interaction between μ_z and electric field may help the transition in F-to-I direction, we obtained the potentials of mean force (PMF) along the conformational transition pathways defined by a ΔRMSD metric as explained earlier in the Methods section.
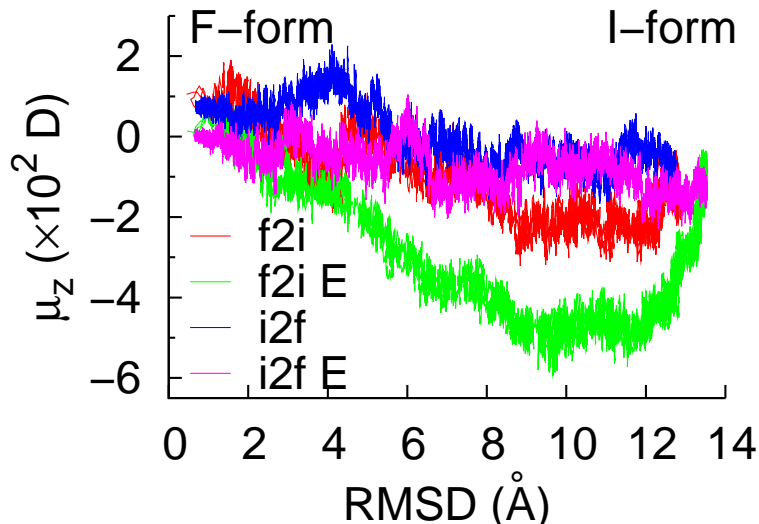
Figure 4.5: Dipole moment z-component (μ_z) for P1-head from targeted molecular dynamics simulation trajectories.

Accordingly, reaction coordinate takes values between $\Delta RMSD = -12$ Å and 12 Å, which correspond to F-form and I-form conformations, respectively.

As seen in Figure 4.6, PMF obtained when no electric field is applied shows two minima near -9.2 Å (F-form-like structure) and 6.4 Å (I-form-like structure) with nearly identical energies separated by a wide free energy barrier of 7.5 kcal/mol. The minimum near F-form structure is narrower indicating that F-form conformation is more restricted in its motion than I-form-like structures. This is possibly due to the strong interaction of the two helices, shown in cyan in Figure 4.7, with the transmembrane helices in the F-form conformation. PMF profiles also reveal that F-to-I transition is almost barrierless in the existence of electric field. In this case, I-form conformation has a much lower energy (-18.3 kcal/mol) with a wide low energy basin. This shows that P1-head is highly mobile in this region, more so than the previous case where no electric field is applied, with multiple low energy structures. Sample structures from each free energy minimum are shown in Figure 4.7. The very high energy barrier in the reverse direction makes F-form conformation highly unlikely in the existence of electric field.
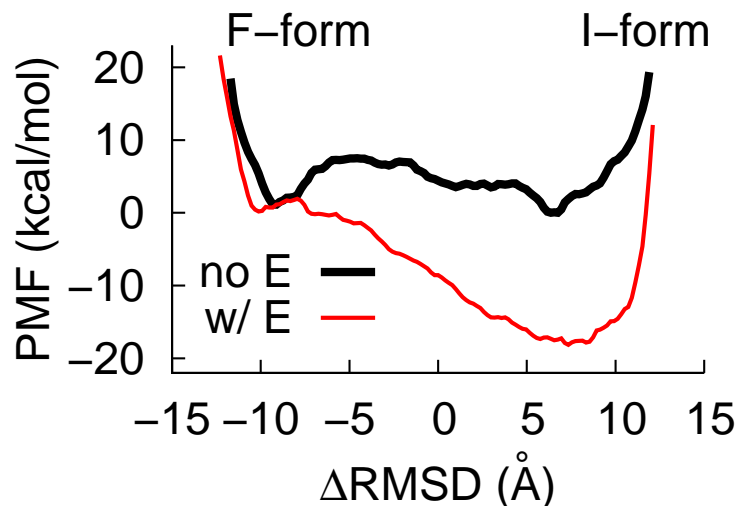
Figure 4.6: Potential of mean force profiles for the conformational transition between SecDF F-form and I-form when electric field is applied (red) and when no electric field is applied (black).

We also calculated $< \mu_z >$, averaged dipole moment z component for each umbrella sampling window after unbiasing. As shown in Figure 4.8, the $z-$component of the P1-head dipole moment is about $100D$ lower when there is electric field than when there is no electric field. Also in the I-form conformation, P1-head dipole moment is oriented such that its interaction with the electric field through its dipole moment will be more than in the F-form conformation.

4.5 Concluding Discussions

In this study we obtained low energy transition pathways between SecDF F-form and I-form conformations by using all-atomistic targeted molecular dynamics simulations. An analysis of P1-head dipole moment along the transition pathways revealed a hysteresis between F-to-I and I-to-F transitions in the existence of a transmembrane potential. This indicates a coupling between transmembrane potential and P1-head dipole moment. Free energy profiles along a ΔRMSD based reaction coordinate used to describe the transitions showed
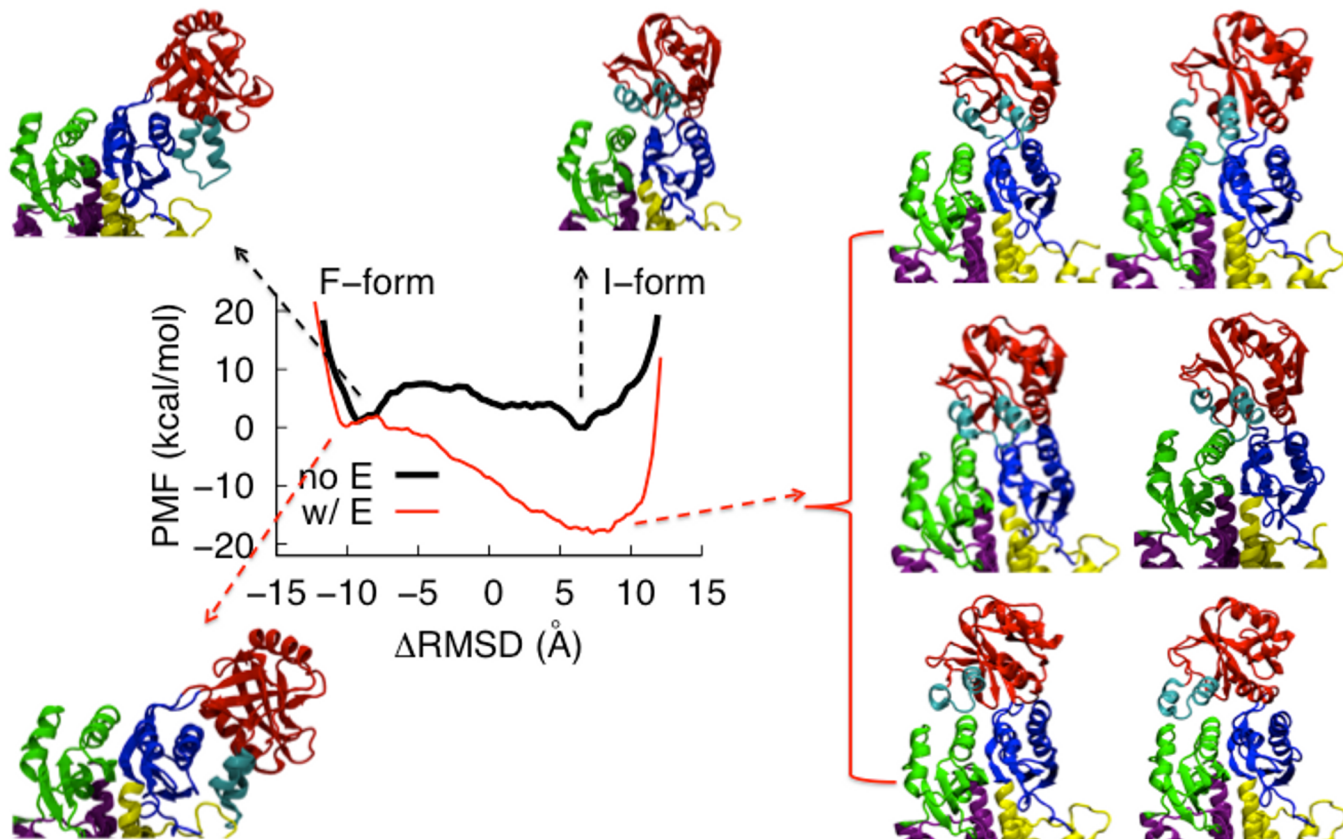
Figure 4.7: Sample minimum energy structures near each minima of the potential of mean force profiles.

that transmembrane potential indeed lowers the barrier significantly in the F-to-I transition pathway. A free energy difference of about -18.3 kcal/mol between F-form and I-form conformations may provide the driving force needed for P1-head to carry unfolded preproteins into the periplasm after capturing them in the SecYEG protein conducting channel.

The potential of mean force profiles help us explain the coupling between proton motive force and SecDF conformational change. They further provide strong evidence for the following mechanism of function: SecDF captures preproteins from the channel in its F-form when the transmembrane potential is in a depleted state. As the transmembrane potential is formed, SecDF undergoes a conformational change from F-form to I-form and stays in the I-form conformation as long as transmembrane potential is maintained. The high free energy barrier in the reverse direction makes I-to-F transition highly unlikely when there is
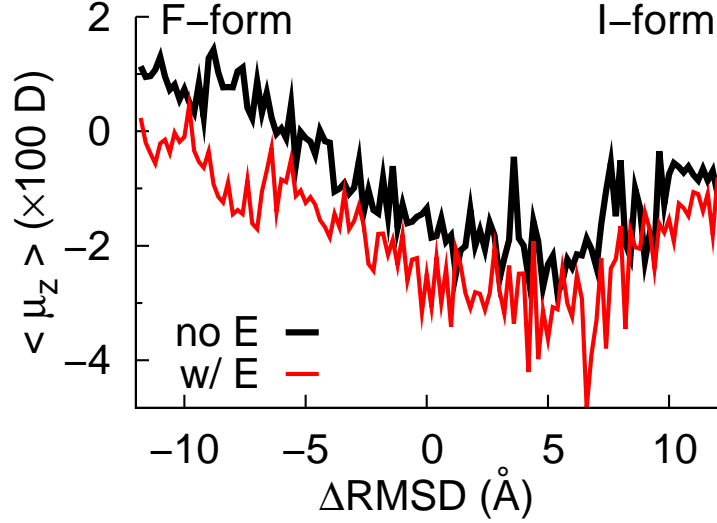
Figure 4.8: Unbiased averages for μ_z, i.e. $z-$component of dipole moment from umbrella sampling simulations.

a transmembrane potential. As the transmembrane potential is depleted, F-form conformation becomes accessible as much as I-form conformation, since the two conformations have similar free energy minima with a barrier of about 7.5 kcal/mol. While in the absence of transmembrane potential F-form and I-form have similar free energy values, the barrier in the F-to-I direction is steeper than in the reverse direction, which may promote the F-form conformation compared to the I-form conformation.

As a result we can conclude that the favorable interaction of P1-head dipole moment with the electric field formed due to the transmembrane potential may explain the SecDF functional conformational change and how it is coupled to the proton motive force. Here we need to note that this study includes only the electric transmembrane potential component ($\Delta\psi$) of proton motive force, Δp, and does not include the effects of pH gradient component (ΔpH). The effects of ΔpH can be studied by explicit solvent simulations and by taking into account different protonation states of ionizable residues in the SecDF structure.

Bibliography

[1] Zimmer, J., Nam, Y., and Rapoport, T. A. (2008) Structure of a Complex of the ATPase SecA and the Protein-Translocation Channel. *Nature 455*, 936–943.

[2] du Plessis, D. J. F., Nouwen, N., and Driessen, A. J. M. (2011) The Sec Translocase. *BBA Biomembranes 1808*, 851–865.

[3] Hand, N. J., Klein, R., Laskewitz, A., and Pohlschroder, M. (2006) Archaeal and Bacterial SecD and SecF Homologs Exhibit Striking Structural and Functional Conservation. *J. Bacteriol. 188*, 1251–1259.

[4] Rapoport, T. A. (2007) Protein Translocation Across the Eukaryotic Endoplasmic Reticulum and Bacterial Plasma Membranes. *Nature 450*, 663–669.

[5] Rapoport, T. A. (2008) Protein Transport Across the Endoplasmic Reticulum Membrane. *FEBS J. 275*, 4471–4478.

[6] Van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) X-Ray Structure of a Protein-Conducting Channel. *Nature 427*, 36–44.

[7] Tsukazaki, T., Mori, H., Fukai, S., Ishitani, R., Mori, T., Dohmae, N., Perederina, A., Sugita, Y., Vassylyev, D. G., Ito, K., and Nureki, O. (2008) Conformational Transition of Sec Machinery Inferred from Bacterial SecYE Structures. *Nature 455*, 988–991.

[8] Bolhuis, A., Broekhuizen, C. P., Sorokin, A., van Roosmalen, M. L., Venema, G., Bron, S., Quax, W. J., and van Dijl, J. M. (1998) SecDF of Bacillus Subtilis, a Molecular Siamese Twin Required for the Efficient Secretion of Proteins. *J. Biol. Chem. 273*, 21217–21224.

[9] Sagara, K., Matsuyama, S., and Mizushima, S. (1994) SecF Stabilizes SecD and SecY, Components of the Protein Translocation Machinery of the Escherichia Coli Cytoplasmic Membrane. *J. Bacteriol. 176*, 4111–4116.

[10] Economou, A., and Wickner, W. (1994) SecA Promotes Preprotein Translocation by Undergoing ATP-Driven Cycles of Membrane Insertion and Deinsertion. *Cell 78*, 835–843.

[11] Economou, A., Pogliano, J. A., Beckwith, J., Oliver, D. B., and Wickner, W. (1995) SecA Membrane Cycling at SecYEG Is Driven by Distinct ATP Binding and Hydrolysis Events and Is Regulated by SecD and SecF. *Cell 83*, 1171–1181.

[12] Duong, F., and Wickner, W. (1997) The SecDFyajC Domain of Preprotein Translocase Controls Preprotein Movement by Regulating SecA Membrane Cycling. *EMBO J. 16*, 4871–4879.

[13] Matsuyama, S., Fujita, Y., and Mizushima, S. (1993) SecD Is Involved in the Release of Translocated Secretory Proteins from the Cytoplasmic Membrane of Escherichia Coli. *EMBO J. 12*, 265–270.

[14] Arkowitz, R. A., and Wickner, W. (1994) SecD and SecF Are Required for the Proton Electrochemical Gradient Stimulation of Preprotein Translocation. *EMBO J. 13*, 954.

[15] Pogliano, J. A., and Beckwith, J. (1994) SecD and SecF Facilitate Protein Export in Escherichia-Coli. *EMBO J. 13*, 554–561.

[16] Echizen, Y., Tsukazaki, T., Dohmae, N., Ishitani, R., and Nureki, O. (2011) Crystallization and Preliminary X-Ray Diffraction of the First Periplasmic Domain of SecDF, a Translocon-Associated Membrane Protein, from Thermus Thermophilus. *Acta Crystallogr. F 67*, 1367–1370.

[17] Burg-Golani, T., Pozniak, Y., Rabinovich, L., Sigal, N., Nir Paz, R., and Herskovits, A. A. (2013) Membrane Chaperone SecDF Plays a Role in the Secretion of Listeria Monocytogenes Major Virulence Factors. *J. Bacteriol. 195*, 5262–5272.

[18] Shiozuka, K., Tani, K., Mizushima, S., and Tokuda, H. (1990) The Proton Motive Force Lowers the Level of ATP Required for the in Vitro Translocation of a Secretory Protein in Escherichia Coli. *J. Biol. Chem. 265*, 18843–18847.

[19] Driessen, A. J., and Wickner, W. (1991) Proton Transfer Is Rate-Limiting for Translocation of Precursor Proteins by the Escherichia Coli Translocase. *Proc. Natl. Acad. Sci. U.S.A. 88*, 2471–2475.

[20] Schiebel, E., Driessen, A. J., Hartl, F. U., and Wickner, W. (1991) Delta Mu H+ and ATP Function at Different Steps of the Catalytic Cycle of Preprotein Translocase. *Cell 64*, 927–939.

[21] Tsukazaki, T., Mori, H., Fukai, S., Numata, T., Perederina, A., Adachi, H., Matsumura, H., Takano, K., Murakami, S., Inoue, T., Mori, Y., Sasaki, T., Vassylyev, D. G., Nureki, O., and Ito, K. (2006) Purification, Crystallization and Preliminary X-Ray Diffraction of SecDF, a Translocon-Associated Membrane Protein, from Thermus Thermophilus. *Acta Crystallogr. F 62*, 376–380.

[22] Tsukazaki, T., Mori, H., Echizen, Y., Ishitani, R., Fukai, S., Tanaka, T., Perederina, A., Vassylyev, D. G., Kohno, T., Maturana, A. D., Ito, K., and Nureki, O. (2011) Structure and Function of a Membrane Component SecDF That Enhances Protein Export. *Nature 474*, 235–238.

[23] Mio, K., Tsukazaki, T., Mori, H., Kawata, M., Moriya, T., Sasaki, Y., Ishitani, R., Ito, K., Nureki, O., and Sato, C. (2013) Conformational Variation of the Translocon Enhancing Chaperone SecDF. *J. Struct. Funct. Genomics 15*, 107–115.

[24] Schlitter, J., Engels, M., Krüger, P., Jacoby, E., and Wollmer, A. (1993) Targeted Molecular Dynamics Simulation of Conformational Change-Application to the T-R Transition in Insulin. *Mol. Simul. 10*, 291–308.

[25] Schlitter, J., Engels, M., and Krüger, P. (1994) Targeted Molecular Dynamics: A New Approach for Searching Pathways of Conformational Transitions. *J. Mol. Graph. 12*, 84–89.

[26] Ma, J., Flynn, T. C., Cui, Q., Leslie, A. G. W., Walker, J. E., and Karplus, M. (2002) a Dynamic Analysis of the Rotation Mechanism for Conformational Change in F(1)-ATPase. *Structure 10*, 921–931.

[27] Ma, J., and Karplus, M. (1997) Molecular Switch in Signal Transduction: Reaction Paths of the Conformational Changes in Ras P21. *Proc. Natl. Acad. Sci. U.S.A. 94*, 11905–11910.

[28] van der Vaart, A., Ma, J., and Karplus, M. (2004) The Unfolding Action of GroEL on a Protein Substrate. *Biophys. J. 87*, 562–573.

[29] Ma, J., Sigler, P. B., Xu, Z., and Karplus, M. (2000) A Dynamic Model for the Allosteric Mechanism of GroEL. *J. Mol. Biol. 302*, 303–313.

[30] Barbacid, M. (1987) Ras Genes. *Annu. Rev. Biochem. 56*, 779–827.

[31] Ford, B., Hornak, V., Kleinman, H., and Nassar, N. (2006) Structure of a Transient Intermediate for GTP Hydrolysis by Ras. *Structure 14*, 427–436.

[32] Hall, B. E., Bar-Sagi, D., and Nassar, N. (2002) The Structural Basis for the Transition from Ras-GTP to Ras-GDP. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12138–12142.

[33] Fernando Diaz, J., Wroblowski, B., Schlitter, J., and Engelborghs, Y. (1997) Calculation of Pathways for the Conformational Transition Between the GTP- and GDP- bound

States of the Ha-ras-p21 Protein: Calculations with Explicit Solvent Simulations and Comparison with Calculations in Vacuum. *Proteins 28*, 434–451.

[34] Aci, S., Mazier, S., and Genest, D. (2005) Conformational Pathway for the Kissing Complex-Extended Dimer Transition of the SL1 Stem-Loop from Genomic HIV-1 RNA As Monitored by Targeted Molecular Dynamics Techniques. *J. Mol. Biol. 351*, 520–530.

[35] Bascom, G., and Andricioaei, I. (2014) Single-Walled Carbon Nanotubes Modulate the B- to A-DNA Transition. *J. Phys. Chem. C 118*, 29441–29447.

[36] Kirkwood, J. G. (1935) Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys. 3*, 300.

[37] Roux, B., and Karplus, M. (1991) Ion Transport in a Model Gramicidin Channel. Structure and Thermodynamics. *Biophys. J. 59*, 961–981.

[38] Roux, B. (1995) The Calculation of the Potential of Mean Force Using Computer Simulations. *Comput. Phys. Commun. 91*, 275–282.

[39] Torrie, G. M., and Valleau, J. P. (1977) Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys. 23*, 187–199.

[40] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992) The Weighted Histogram Analysis Method for Free Energy Calculations on Biomolecules. I. the Method. *J. Comput. Chem. 13*, 1011–1021.

[41] Grossfield, A. WHAM: The Weighted Histogram Analysis Method", Version 2.0.6. `http://membrane.urmc.rochester.edu/content/wham`.

[42] Banavali, N. K., and Roux, B. (2005) Free Energy Landscape of A-DNA to B-DNA Conversion in Aqueous Solution. *J. Am. Chem. Soc. 127*, 6866–6876.

[43] Hol, W. G. (1985) The Role of the Alpha-Helix Dipole in Protein Function and Structure. *Prog. Biophys. Mol. Bio. 45*, 149–195.

[44] Hol, W. G., van Duijnen, P. T., and Berendsen, H. J. (1978) The Alpha-Helix Dipole and the Properties of Proteins. *Nature 273*, 443–446.

[45] Wada, A. (1976) The Alpha-Helix As an Electric Macro-Dipole. *Adv. Biophys.* 1–63.

[46] Antosiewicz, J. (1995) Computation of the Dipole Moments of Proteins. *Biophys. J. 69*, 1344.

[47] Becker, O. M., MacKerell, A. D., Jr., Roux, B., and Watanabe, M. *Computational Biochemistry and Biophysics*; CRC Press, 2001.

[48] Ryckaert, J., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys. 23*, 327–341.

[49] Im, W., Lee, M. S., and Brooks, C. L. (2003) Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem. 24*, 1691–1702.

[50] Im, W., Feig, M., and Brooks, C. L., III (2003) An Implicit Membrane Generalized Born Theory for the Study of Structure, Stability, and Interactions of Membrane Proteins. *Biophys. J. 85*, 2900–2918.

[51] Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (2004) Program for Macromolecular Energy, Minimization,and Dynamics Calculations. *J. Comput. Chem. 4*, 187–217.

[52] Brooks, B. R., Bruccoleri, R. E., Olafson, D. J., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem. 4*, 187–217.

[53] MacKerel Jr., A. D., Brooks III, C. L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. In *CHARMM: The Energy Function and Its Parameterization with an Overview of the*

Program; v. R. Schleyer et al., P., Ed.; The Encyclopedia of Computational Chemistry; John Wiley & Sons: Chichester, 1998; Vol. 1; pp 271–277.

[54] Brooks, B. R. et al. (2009) CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem. 30*, 1545–1614.

[55] Berendsen, H. J., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys. 81*, 3684–3690.