Long time and large size molecular dynamics simulations made feasible through new TeraGrid hardware and software

Kirby Vandivort* James C. Phillips* Elizabeth Villa* Peter L. Freddolino* James Gumbart*
Leonardo G. Trabuco* Danielle E. Chandler* Jen Hsin* Christopher B. Harrison*
Laxmikant Kalé† Klaus Schulten‡*

1 Introduction

Atomic-level views of biomolecules, like DNA or protein, furnished foremost experimentally, have revolutionized modern science and technology, leading from DNA structure to genomics and from protein structure to modern pharmacology. In order to learn how life "works" and to turn living cells into factories for clean energy and affordable drugs, our century needs to know the entire cell, atom-by-atom. The required resolution can be furnished through advances in crystallography and electron microscopy; the former routinely providing atomic-level views of most biomolecules and the latter moving rapidly towards whole cell views. The data from these experiments are synthesized into atomic-level movies of the cells through the "computational microscope" of molecular dynamics simulations on massively parallel computers.

We present below a brief introduction to molecular dynamics simulation, followed by three current projects demonstrating the capabilities of the method running on TeraGrid hardware. The projects span a wide range of structural systems biology research, focusing on (i) the folding dynamics and mechanisms of small proteins; (ii) the ribosome, a large molecular machine responsible for protein synthesis; and (iii) the chromatophore, an entire organelle in photosynthetic bacteria. All three studies proceed in close collaboration with experimental researchers.

2 Molecular dynamics simulation

Molecular dynamics (MD) simulation provides an essential means of understanding the function of biomolecular machines at the atomic level. To initialize the simulation, full atomic coordinates of proteins, nucleic acids, and/or lipids of interest are modeled based on known crystallographic or other structures and solvated in explicit water and ions. An empirical energy function, which consists of approximations of covalent interactions in addition to nonbonded Lennard-Jones and electrostatic terms, is applied to the system. The resulting Newtonian equations of motion are then integrated by symplectic and reversible methods using a timestep of 1–2 fs. Modifications are made to the equations of motion to control temperature and pressure during the simulation [1, 27]. Long-range electrostatic forces are typically calculated using the particle mesh Ewald (PME) [12] algorithm, which scales as $O(N \log N)$ rather than $O(N^2)$ in the number of atoms. Due to the dominance of the $O(N)$ short-range interaction calculations for typical simulation sizes, measured runtime scales linearly with atom count.

The basic protocol for MD simulations consists of minimization to eliminate initial contacts that would destabilize the integrator, equilibration to a temperature of ~ 300 K and a pressure of 1 atm, and simulation in an isobaric (NPT) ensemble for 10-100 ns or more. Simple MD is sufficient to test the stability of a biomolecular aggregate or to observe the relaxation of the system into a more favorable conformation. However, if the goal is to study events that would not spontaneously occur during the timespan of the simulation, the system may be compelled to undergo a transition via the application of forces in a steering or adaptive biasing protocol. Important observations may be made during such a simulation of non-equilibrium events, even though the simulated timescale is much shorter than the natural one.

All of the simulations described below were performed with NAMD (NAnoscale Molecular Dynamics), a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems [41, 65]. Distributed free of charge since 1995 [56, 57], NAMD is recognized as the leading software for running such simulations on large parallel machines, having demonstrated scaling to thousands of processors in a 2002 paper [64] that received a Gordon Bell Award. NAMD may be used with CHARMM [50], Amber [9], or OPLS [40] parameters and is closely linked with the popular visualization and analysis program VMD [37]. NAMD is available on most TeraGrid machines and its performance continues to improve [45, 76, 4]. For example, a 1.07 M atom virus simulation [25] with PME electrostatics runs on 1,024 cores of the NCSA Abe cluster with 76% parallel efficiency at 884 GFLOP/s and 2,403 GOP/s. Future NAMD development and support is funded by a recently-renewed NIH Resource grant and a NAMD simulation of 100 millions atoms is a design target and acceptance test for the upcoming Blue Waters petascale system.

3 Folding

Predicting the folding mechanism and folded structure of a protein from its amino acid sequence is one of the most important problems in structural biology today. A better theoretical framework describing the protein folding process would both aid in the understanding of diseases caused by point mutations in proteins, and allow the design of protein variants with novel function. While several approaches, including Monte Carlo structure prediction [13] and the use of a large number of short MD trajectories [83], have been applied to the protein folding problem, long timescale molecular dynamics trajectories offer significantly more information on the folding process itself because the entire process can be observed over realistic timescales.

Downhill-folding mutants of several proteins which fold on timescales of 1-3 μs are currently known, and offer an ideal target for MD simulations aiming to observe complete folding trajectories of proteins which can be compared to experiment. Microsecond simulations currently represent the upper limit of timescales accessible to molecular dynamics, and require the application of substantial computing resources and optimization of the simulation software and parameters. However, the wealth of information on the protein folding process that

*Beckman Institute, University of Illinois at Urbana-Champaign.
†Department of Computer Science and Beckman Institute, University of Illinois at Urbana-Champaign.
‡To whom correspondence should be addressed.

would be provided by such a clear comparison between experiment and theory makes the simulation of these proteins an important opportunity.

TeraGrid resources have recently been applied to generate multiple microsecond folding trajectories of the WW domain, a β-sheet protein, and the villin headpiece, an α-helical protein. Both are well-studied experimentally and computationally, and have served as model systems for previous investigation of protein folding through molecular dynamics [46, 66, 81, 99, 7, 47, 83, 16]. In order to make simulation of the complete folding processes of the target proteins (4-5 μs) computationally tractable, NAMD 2.6 [65] was modified to improve scaling through the use of modified load balancers, pencil decomposition for PME electrostatic calculations, and the addition of hand-coded loops with SSE intrinsics to treat constraints; all changes are available in the CVS version of NAMD [26]. These modifications allowed efficient scaling of the MD simulations to roughly one processor per 100 atoms, compared to a limit of one processor per 1,000 atoms for most systems using NAMD 2.6.

Several MD trajectories longer than 1 μs were generated for both the WW domain and villin headpiece, as described below. The Fip35 mutant of human Pin1 WW domain [49] and wild-type chicken villin headpiece subdomain were used [6]. All simulations were performed using a 2.0 fs timestep, with bonded interactions and short range nonbonded interactions calculated every step, and long range electrostatics updated every 3 steps. A cutoff of 7.0 Å was used for short-range interactions, with shifting started at 8.0 Å, and PME was used to treat long-range electrostatics. The CHARMM22 force field [52] with CMAP corrections [53] was used for both cases.

3.1 WW DOMAIN FOLDING

Three WW domain trajectories were generated, of durations 10.0, 2.0, and 1.5 μs. In all three cases a similar pattern is observed: the protein undergoes rapid hydrophobic collapse over the first 500 ns, after which a series of interconverting α-helical intermediates are observed for the remainder of the simulation. The folded state, a sample molten globule immediately following collapse, and representative structures from clustering analysis of the three trajectories are shown in Fig. 1. While the presence of α-helical intermediates might be expected even in the folding process of a β-sheet protein [68], the persistence of this structure well past the expected folding time (4-5 μs) is likely indicative of a failure of the force field to properly weight α-helical and β-sheet conformations, a problem that has been previously observed in unfolding simulations of a WW domain using AMBER force fields [93] and in general assays of the α-helical propensity of the CHARMM force field prior to the addition of CMAP cross-terms [97].

3.2 VILLIN HEADPIECE

Three trajectories simulating folding of the villin headpiece are in progress, with current durations of 4.8 μs, 1.7 μs, and 0.7 μs; the expected folding time is on the order of 4 μs [44]. All three simulations show near-native secondary structure (three α-helices) and general topology, although none have, as of yet, reached a native state. The folded structure, snapshots of structures at the end of currently obtained trajectories, and plots of C_α root mean square deviation (RMSD) to the native state throughout each trajectory are shown in Fig. 2. As can be seen, trajectory 2 is currently closest to the native state. Trajectory 1 is in a near-native state but helix III is in a reversed orientation; it is unclear how long the structure will remain kinetically trapped in this conformation. Both trajectories 2 and 3 show correct relative orientations of the secondary structure elements, and it appears likely that they will reach native states within the experimental folding time.

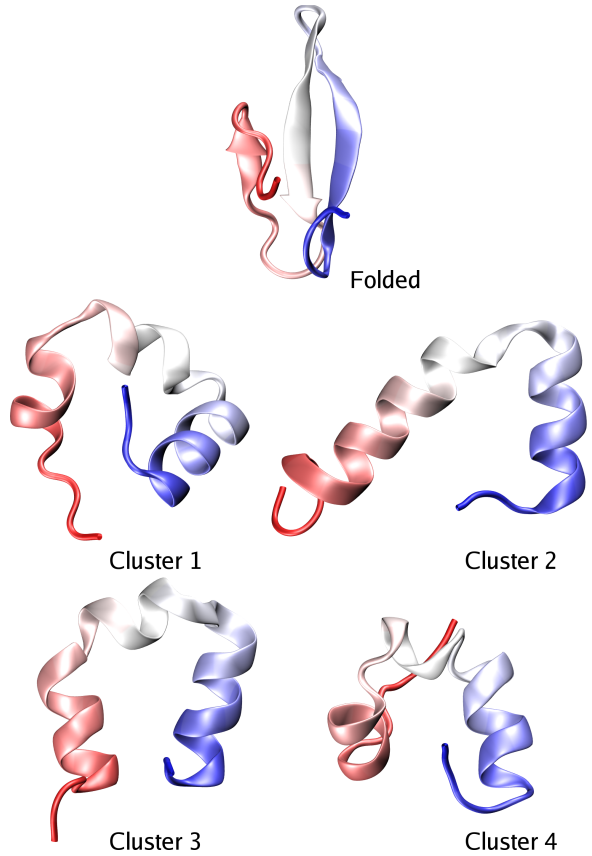

Figure 1: Folded conformation and four most occupied clusters from three Fip35 folding trajectories. Clustering was performed as described in [14] with a cutoff of 5.0 Å.

It appears likely that the differences in the success of folding simulations of the WW domain and villin headpiece are due largely to differences in the force field's ability to properly recognize the folded state of the proteins as a global free energy minimum; this would be consistent with past findings on the relative stabilities of β-sheet and α-helical proteins in the CHARMM force field. The effects of the force field on WW domain folding will be further tested in the future through the use of replica exchange simulations [84] to better probe the free energy surface, and attempts to fold the protein from the same starting conformation using the AMBER ff03 force field [17].

4 RIBOSOME

The translation of genetic information into protein is an essential process in all cells. At the core of this process lies the ribosome, a 2.5–4.5 MDa ribonucleoprotein complex, where decoding of genetic information and protein synthesis take place. Bacterial ribosomes consist of three ribosomal RNAs and 55 proteins. During the process of translation, the ribosome undergoes several conformational changes and binds to different factors that catalyze specific reactions, including tRNA molecules charged with amino acids, elongation and release factors, among others [48]. Over the past few years, landmark progress in the area of structural biology of translation has been made thanks to both cryo-electron microscopy (cryo-EM) [23] and X-ray crystallography [77]. Cryo-EM offers snapshots of the ribosome in functional states [88, 89, 24, 90, 30, 55], currently with resolutions of 7–12 Å. This work is complemented by X-ray crystallography, which can now resolve the whole ribosome at atomic resolution [77, 43, 78]. However, it

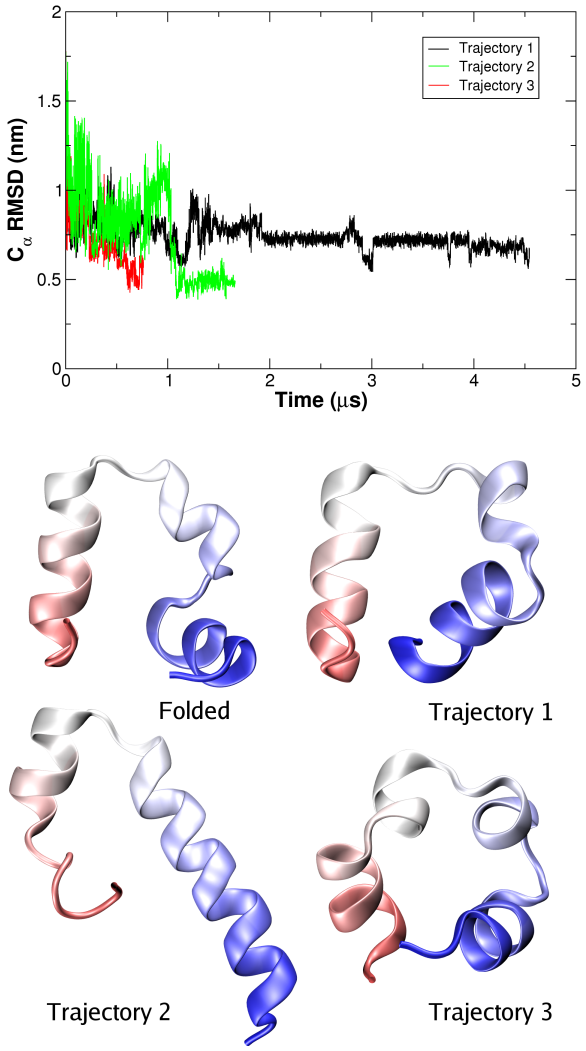

Figure 2: (top) Graph of C_α-RMSD to the native state for three villin headpiece folding trajectories. (bottom) Folded structure of villin headpiece, and structures at the current endpoint of all three folding trajectories.

is unclear what functional states the X-ray structures represent, if any, due to the non-natural conditions needed to form crystals; additionally, such crystallization conditions cannot be achieved currently for important functional states, e.g., when elongation factors bind to the ribosome. Atomic details of the interactions between the factors and the ribosome, along with the conformational changes of the ribosome itself, are crucial for understanding its function.

Computer simulation is currently the only means to bridge the resolution gap between these techniques, by flexibly fitting crystal structures into cryo-EM maps, which provides atomistic structures of the ribosome in different functional states. We have developed a method, called molecular dynamics flexible fitting (MDFF), to fit atomic structures into cryo-EM maps using molecular dynamics simulations that incorporate the EM data into the simulation as an external potential added to the MD force field, resulting in forces proportional to the density gradient in the EM data being applied to the atomic structure, effectively biasing it toward the region of conformational space consistent with the density distribution of the EM map. In this way, all the internal features present in the EM map are used in the fitting process, while the model remains fully flexible and stereochemically correct through the MD force field [85]. MDFF is implemented in NAMD, and thus the computational cost of the MDFF method scales linearly with system size, permitting the application of MDFF to large macromolecular complexes. Atomic models furnished by MDFF can be used in subsequent MD simulations to study the dynamics of the assemblies in more detail through the use of TeraGrid resources.

In order to obtain structures of the ribosome in functional states at atomic resolution using MDFF, a necessary step for further computational studies, an all-atom model of the *E. coli* ribosome was needed to furnish the initial coordinates. Such model was built [85] from available crystal structures [3, 78], with missing parts obtained through homology modeling [98, 86, 59, 73, 58]. TeraGrid resources allowed us to use MDFF to obtain atomistic structures from twelve different cryo-EM 3-D reconstructions that represent several functional states of the elongation cycle [89, 71, 70] (manuscript in preparation). Due to its sheer size, its lack of symmetry, and its mixed composition of protein and RNA, MDFF simulations of the ribosome require multi-step protocols, as described in [85]. All simulations were run using NAMD 2.6, with the CHARMM27 force field [50, 51, 21, 53] in vacuo, using a dielectric constant of 80. A multiple time-stepping integration scheme was used, calculating bonded interactions every 1 fs and nonbonded interactions every 2 fs; a cutoff distance of 10 Å was used for the nonbonded interactions. Harmonic restrained were imposed to preserve secondary structure during the fitting [85]. The resulting structures can be placed in their native environment (i.e., with both solvent and ions) yielding systems of millions of atoms (see below), and simulated to study the conformational dynamics of the ribosome in each of the functional states, shedding light on the interaction of the ribosome with the several factors bound, as well as between different domains of the complex. As an example, we present all-atom simulations of the SecY-ribosome complex.

In order to insert nascent proteins into membranes, or translocate them across them, the ribosome binds and feeds the protein into SecY, a membrane channel that helps nascent, unfolded proteins cross the membrane or insert into it [91, 62]. Binding of the ribosome to SecY induces SecY to open, at least partially, although the precise interaction remains unknown [82, 69, 34]. After opening of SecY, the translocating polypeptide has to be somehow read in order to determine if each part should cross the channel to the other side or be inserted into the membrane through a "lateral gate". Hypotheses for how this determination takes place include interaction of the polypeptide with lipids or the ribosome modulating SecY [35, 55, 80, 15].

Previous simulations pulling polypeptide segments across the SecY channel and then allowing it to relax revealed the structural response of SecY during translocation, although without the presence of the ribosome [33]. The availability of a new cryo-EM map of the ribosome-channel complex at 9-Å resolution [54], depicted in Fig. 3, reveals the interactions between the ribosome and the channel. MDFF simulations of an atomic model of the translocon [91] combined with our ribosome model and the cryo-EM map have produced an atomic model of the complex. Placement in the right membrane and solvent environment (Fig. 3), yielded a system of 2.7 M atoms, that was simulated using NAMD 2.6, with the CHARMM27 force field [50, 51, 21, 53], with 701,732 TIP3 water molecules and 2,393 lipid molecules. A multiple time-stepping integration scheme was used, calculating bonded interactions every 1 fs and short-range nonbonded interactions every 2 fs with a cutoff distance of 12 Å for the nonbonded interactions. Long-range electrostatic interactions were calculated using PME with a grid density of $\sim 1/1\text{Å}^3$. A constant temperature of 310 K was maintained using Langevin dynamics; a constant pressure of 1 atm was enforced in the z dimension only. Simulations were run on 1,024 cores of the Abe cluster.

The simulations were performed with the goal of illustrating at atomic detail how the ribosome binding affects the channel. After simulating the complex for over 15 ns, we found certain regions of SecY

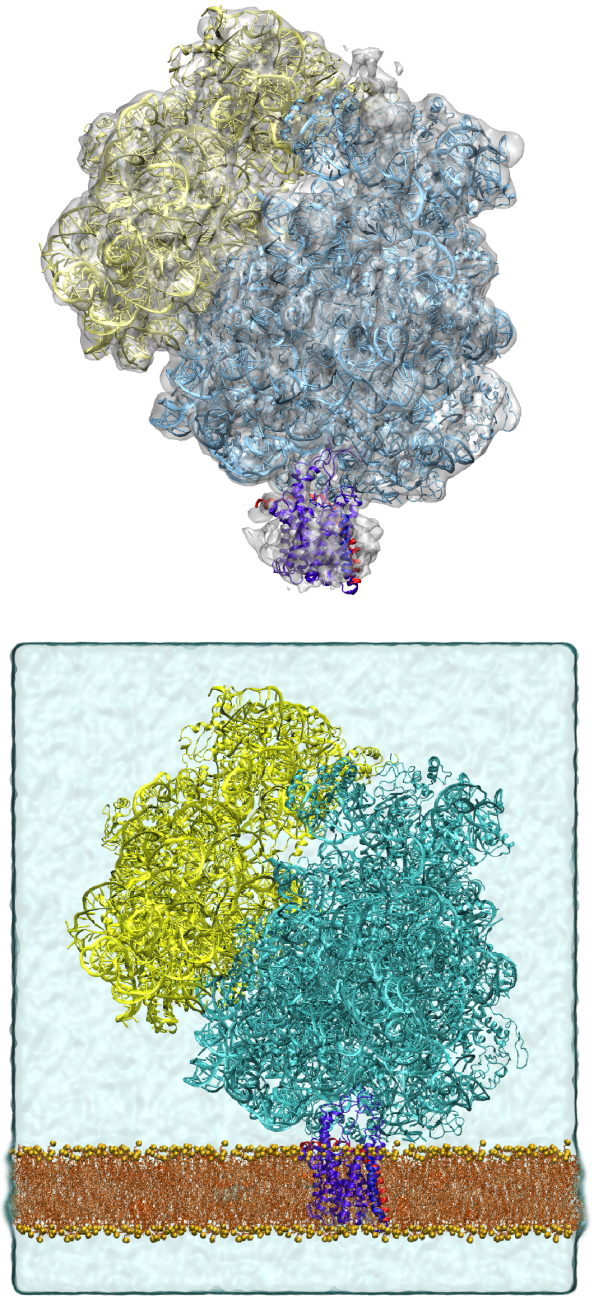

Figure 3: The ribosome bound to SecY. (top) The all-atom model of the *E. coli* ribosome, described in [85], bound to a model of the *E. coli* SecY protein based on the crystallographic structure of the *Methanococcus jannaschii* SecY. The model is fitted to a 9-Å resolution EM density map [54] using MDFF. (bottom) The fitted model of the ribosome-SecY complex is placed in a membrane-water environment and equilibrated using conventional molecular dynamics.

were disturbed more than in the monomer alone, as measured by the root mean-square fluctuations (RMSF) of all the residues. Approximately half the channel (transmembrane helices 5-9) had increased fluctuations. In particular, the plug, a small helix blocking the channel center, became slightly more mobile with the ribosome bound than without, suggesting that, over time, ribosome binding can induce partial channel opening.

5 Chromatophore

Purple bacteria are some of the oldest, and simplest, photosynthetic organisms known [96, 36]. As such they present an opportunity to study photosynthesis at its most basic, reduced to the bare necessities. In purple bacteria, photosynthesis takes place in membrane-bound pseudo-organelles called chromatophores. These chromatophores, as invaginations of the cytoplasmic membrane, come in different shapes such as flat sheets or vesicular bulbs, depending on the species; they also contain all the proteins necessary for converting sunlight into chemical energy, namely ATP. As a typical example, spherical chromatophores are ∼70 nm in diameter and contain approximately 200 membrane proteins in total. However, there are only five unique proteins, listed here in order of decreasing number: Light Harvesting Complex II (LH2, about 150 copies), Light Harvesting Complex I (LH1, 10 dimers), the Reaction Center (RC, 20 copies), cytochrome bc_1 (five), and ATP synthase (only one). LH1 and LH2 act as scaffolds for numerous bacteriochlorophylls, which harvest and transfer the light energy, whereas the RC, bc_1, and ATP synthase carry out the chemical reactions leading to ATP.

While the individual proteins of the chromatophore have been studied extensively [61, 32, 75, 92, 74, 8, 67, 11, 10, 20], including the solution of atomic-level structures of all of them [42, 63, 19, 8, 87, 18, 5], it is still insufficient for understanding their large-scale organization and combined function. To that end, a full-scale model of the chromatophore of *Rb. sphaeroides* was recently developed [79]. Starting from Atomic Force Microscopy (AFM) data on flattened chromatophores, the model was built by first projecting the flat sheet into a sphere and then placing the atomic structures into the data. The AFM data contained only the LH2, LH1, and RC proteins, suggesting that the bc_1 and ATP synthase are not found in the sphere proper, instead possibly residing at the neck or even directly outside the chromatophore (see Fig. 4A). The resulting model precisely places over 4,000 bacteriochlorophylls, which are the primary components for light absorption and transfer. Calculations of the energy transfer efficiency for the network of bacteriochlorophylls demonstrated that the chromatophore transfers the captured energy with over 95% efficiency, thus connecting the organization of the proteins with the overall function [79].

The full model of the chromatophore has helped to address some questions, while also permitting new ones, e.g., how does an initially flat membrane become a spherical chromatophore? Known examples of proteins shaping membranes are increasing, although the mechanisms are not always well understood [22, 94, 72, 28]. Interactions between the proteins, and between the proteins and the membrane, have been suggested previously [38, 28, 29, 2]. Now, in order to explicitly test this suggestion, we have simulated most of the proteins of the chromatophore to determine if they can individually curve membranes. Systems composed of seven LH2s are found to curve the membrane with the resulting curvature dependent on their density. Similarly, the S-shaped LH1-RC dimer develops a bend at the dimerizing interface in simulation. In contrast, dimers of the bc_1 complex do not cause net curvature. All systems are among the largest ever simulated, ranging from a minimum of 400,000 atoms up to over a million atoms, and thus make extensive use of TeraGrid resources. Nearly all simulations were run on Abe, using up to 1,024 processors at a time. In total, the simulations required nearly 700,000 SUs to demonstrate for the first time, to our knowledge, that integral membrane proteins can induce membrane curvature in equilibrium simulations.

LH2 proteins are complexes formed by two rings of eight or nine single transmembrane helices, the inner ring being denoted α and the outer ring β. Between these two rings rests three bacteriochlorophylls per α/β pair. Currently, there are atomic structures for two LH2 proteins, yet both come from organisms having flat chromatophores [60]. For comparison, we built a homology model of LH2

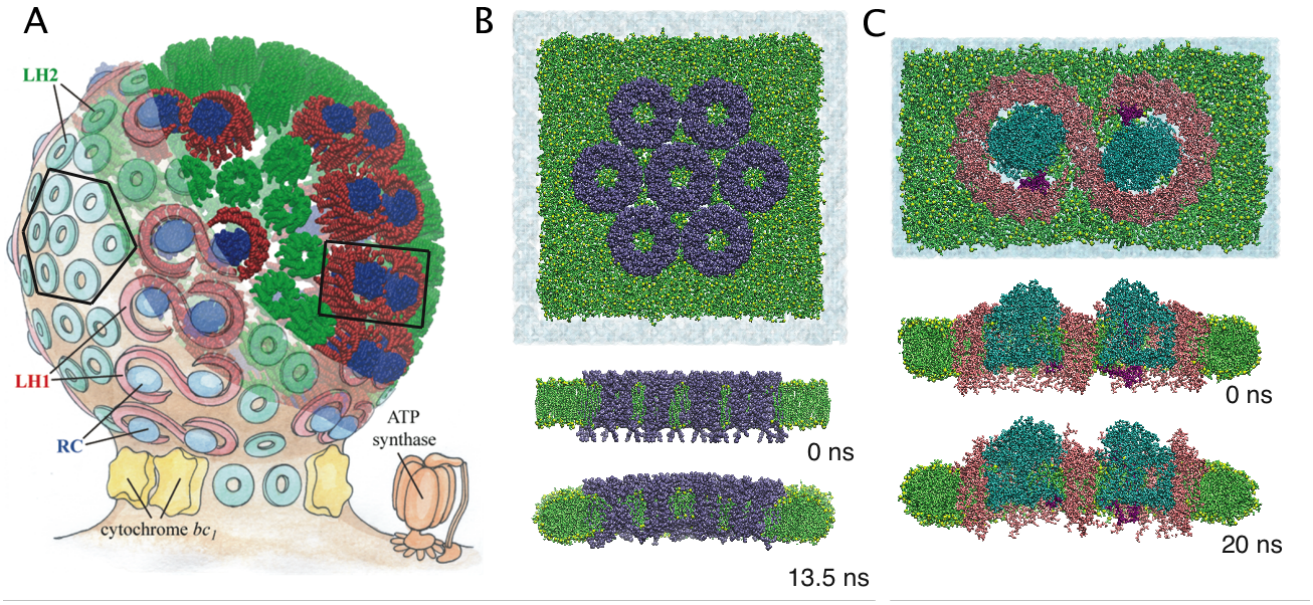

Figure 4: Photosynthetic chromatophore. *A.* Full chromatophore. The chromatophore is shown as both an artist's depiction in the lower half, combined with the modeled version in the upper half. LH2 is in green, LH1-RC in red and blue, bc_1 in yellow, and ATP synthase in orange. *B, C.* Membrane bending by an array of LH2s and an LH1-RC dimer. In the upper panels, both systems are shown from the cytoplasmic side with the membrane colored in green and the water in light blue. LH2s (*B*) are colored purple and LH1-RC (*C*) is colored pink and blue. In the lower panels, both systems are shown in the membrane plane, cut through the middle in order to display the curvature that develops by the end of the simulation.

from *Rb. sphaeroides*, one of the most extensively studied species of purple bacteria which also has spherical chromatophores. Modeling was straightforward given the significant sequence identity between LH2 from *Rb. sphaeroides* and *Rps. acidophila* (44% for the α proteins and 52% for the β proteins), the only exception being a 10-residue cytoplasmic extension of the β proteins. This extension was modeled as a helix based on structure prediction methods using PSIPRED [39], and was placed at the interface between subunits where equilibrium simulations indicated it was stable.

Possessing structures of LH2 from both a flat chromatophore (*Rps. acidophila*) and a spherical one (*Rb. sphaeroides*), we built systems of seven LH2s, arranged in a hexagonal fashion with a single layer of lipids separating them. Surprisingly, LH2 from both species induced membrane curvature, as shown in Fig. 4B. In simulations of 13 ns each, both systems developed an equivalent curvature of $\sim$6 degrees, giving them an effective radius of 79 nm. While this curvature is shallower than the full chromatophore, interactions with other proteins (e.g., LH1-RC) may affect the result. To determine whether packing density affects curvature, another simulation using LH2 from *Rps. acidophila* was carried out in which the LH2s were placed directly adjacent to one another. This system curved much faster, and resulted in a larger curvature of $\sim$8 degrees and a radius of 53 nm. Based on these simulations, we conclude that curvature induced by LH2 is species independent, but is dependent on protein density.

Similar to LH2, LH1 is actually composed of single transmembrane helices, also denoted α and β, which are quite homologous to LH2α and LH2β, respectively. AFM and EM images indicated that LH1 forms an S-shaped dimer with two RCs in *Rb. sphaeroides* for which no full atomic level structure exists (only the components have known structures). Therefore, 28 pairs of LH1α/β subunits were arranged around two RCs in accordance with the low-resolution EM data [67]. PufX, a single-transmembrane-helix accessory protein, was also placed at the RC/LH1 interface based on the location suggested by the EM data [67].

Given that mutants of *Rb. sphaeroides* lacking LH2 still form cylindrical chromatophores, we suspected that LH1-RC dimers can independently induce membrane curvature. Therefore, we simulated the LH1-RC dimer in a full membrane/water system for 20 ns. As shown in Fig. 4C, the dimer develops a bend at the dimer interface toward the periplasmic side. We also calculated the angle of bending to be $\sim$170 degrees, corresponding to a radius of curvature of 72 nm, similar to that found for the LH2 system. This bending is in agreement with single-particle cryo-EM studies, which also show the dimers bend, although to a greater degree (personal communication, C. N. Hunter).

As bc_1 was not resolved in the AFM images of flattened chromatophores, its placement in the chromatophore is still uncertain. Suggestions have included the neck which would imply a saddle-shaped curvature is preferred by the bc_1 [79] or at the pole, indicating a preference for positive curvature [31]. To determine if bc_1 itself induces curvature, we simulated both a single bc_1 dimer and a pair of bc_1 dimers for 10 ns each. In all cases, curvature of the membrane was not observed. While the bc_1 does not cause curvature, it may still respond favorably to a specific curvature, leading to its localization at a particular place in the chromatophore.

Together, the simulations performed indicate how the chromatophore shape develops and what factors govern it. We found no species' dependence for LH2 curvature, although it is known that some species form flat as well as spherical chromatophores. The governing factor for flat vs. spherical chromatophores may instead lie with the organization of the proteins. In the model built for an *Rb. sphaeroides* chromatophore, the LH1-RC-PufX dimers aggregate into stacks; alternatively, in species with flat chromatophores, LH1-RC monomers are more dispersed among the LH2s. The LH1-RC monomers may disrupt long-range curvature that the LH2s would otherwise develop, although such a hypothesis remains to be tested. We found with LH2s a noticeable dependence on density, which may serve a functional role since dense packing of the proteins is necessary for efficient energy transfer [95, 79]. While the simulations performed already are of a large scale, we plan to go to even larger systems in the future, particularly combined arrangements of the LH2, LH1-RC, and bc_1 proteins. This

will allow us to determine the role of the interplay between different proteins in shaping the chromatophore. Finally, we will move to simulations of the full chromatophore which will require on the order of 40 million atoms, but will allow us to describe the entire photosynthetic process at unprecedented detail and scale. Such a simulation will only become possible with the advent of petascale computers, expected to arrive within the next three years.

6 Conclusion

We have in this work given the reader only a small taste of the great variety of biomolecular simulation methods in use today on the TeraGrid. Furthermore, the scope of biological complexity provides a truly inexhaustible supply of worthy research topics that would benefit from computational approaches. Such simulations are rapidly moving from the domain of simulation experts to become yet another tool on the lab bench of the experimentalist. Shared resources such as the TeraGrid will therefore become increasingly important, as they provide the capability for scientists with groundbreaking research but modest local resources to complete simulations in a timely manner, quickening the pace of both theoretical and experimental discovery.

Acknowledgements

This work is supported by grant(s) from the National Institutes of Health P41-RR05969 and the National Science Foundation MCB-0744057. The authors gladly acknowledge supercomputer time provided by Pittsburgh Supercomputing Center and the National Center for Supercomputing Applications via Large Resources Allocation Committee grant MCA93S028 and the Turing Xserve Cluster.

References

[1] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, New York, 1987.

[2] S. Bahatyrova, R. N. Frese, C. A. Siebert, J. D. Olsen, K. O. van der Werf, R. van Grondelle, R. A. Niederman, P. A. Bullough, C. Otto, and C. N. Hunter. The native architecture of a photosynthetic membrane. *Nature*, 430:1058–1062, 2004.

[3] V. Berk, W. Zhang, R. D. Pai, and J. H. D. Cate. Structural basis for mRNA and tRNA positioning on the ribosome. *Proc. Natl. Acad. Sci. USA*, 103:15830–15834, 2006.

[4] A. Bhatele, S. Kumar, C. Mei, J. C. Phillips, G. Zheng, and L. V. Kale. Overcoming Scaling Challenges in Biomolecular Simulations across Multiple Platforms. In *Proceedings of IEEE International Parallel and Distributed Processing Symposium 2008*, 2008.

[5] A. Camara-Artigas, D. Brune, and J. P. Allen. Interactions between lipids and bacterial reaction centers determined by protein crystallography. *Proc. Natl. Acad. Sci. USA*, 99:11055–11060, 2002.

[6] T. K. Chiu, J. Kubelka, R. Herbst-Irmer, W. A. Eaton, J. Hofrichter, and D. R. Davies. High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci. USA*, 102(21):7517–7522, May 2005.

[7] S. Chowdhury, M. C. Lee, G. Xiong, and Y. Duan. Ab initio folding simulation of the trp-cage mini-protein approaches nmr resolution. *J. Mol. Biol.*, 327(3):711–717, Mar 2003.

[8] M. J. Conroy, W. Westerhuis, P. S. Parkes-Loach, P. A. Loach, C. N. Hunter, and M. P. Williamson. The solution structure of the *Rhodobacter sphaeroides* LH1 β reveals two helical domains separated by a more flexible region: structural consequences for the LH1 complex. *J. Mol. Biol.*, 298:83–94, 2000.

[9] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.

[10] A. Crofts, S. Meinhardt, K. Jones, and M. Snozzi. The role of the quinone pool in the cyclic electron transfer chain on Rhodopseudomonas-sphaeroides - a modified Q-cycle mechanism. *Biochim. Biophys. Acta*, 723:202–218, 1983.

[11] A. R. Crofts, B. Hacker, B. Barquera, C.-H. Yun, and R. B. Gennis. Structure and function of the bc-complex of *Rhodobacter sphaeroides*. *Biochim. Biophys. Acta*, 1101:162–165, 1992.

[12] T. Darden, D. York, and L. Pedersen. Particle mesh Ewald. An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092, 1993.

[13] R. Das, B. Qian, S. Raman, R. Vernon, J. Thompson, P. Bradley, S. Khare, M. D. Tyka, D. Bhat, D. Chivian, D. E. Kim, W. H. Sheffler, L. Malmström, A. M. Wollacott, C. Wang, I. Andre, and D. Baker. Structure prediction for CASP7 targets using extensive all-atom refinement with Rosetta@home. *Proteins*, 69 Suppl 8:118–128, 2007.

[14] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark. Peptide folding: When simulation meets experiment. *Angewandte Chemie International Edition*, 38:236–240, 1999.

[15] S. Dorairaj and T. W. Allen. On the thermodynamic stability of a charged arginine side chain in a transmembrane helix. *Proc. Natl. Acad. Sci. USA*, 104:4943–4948, 2007.

[16] Y. Duan and P. Kollman. Pathways to a protein folding intermediate observed in a 1 microsecond simulation in aqueous solution. *Science*, 282:740–744, 1998.

[17] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comp. Chem.*, 24(16):1999–2012, Dec 2003.

[18] U. Ermler, G. Fritzsch, S. K. Buchanan, and H. Michel. Structure of the photosynthetic reaction center from *Rhodobacter sphaeroides* at 2.65 Å resolution: Cofactors and protein-cofactor interactions. *Structure*, 2:925–936, 1994.

[19] L. Esser, X. Gong, S. Yang, L. Yu, C. A. Yu, and D. Xia. Surface-modulated motion switch: Capture and release of iron-sulfur protein in the cytochrome bc_1 complex. *Proc. Natl. Acad. Sci. USA*, 103:13045–13050, 2006.

[20] B. Feniouk, D. Cherepanov, W. Junge, and A. Mulkidjanian. Coupling of proton flow to atp synthesis in rhodobacter capsulatus: Fof1-atp synthase is absent from about half of chromatophores. *Biochim. Biophys. Acta – Bioener.*, 1506:189–203, 2001.

[21] N. Foloppe and A. D. MacKerrell Jr. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.*, 21:86–104, 2000.

[22] M. Fošnarič, A. Iglič, and S. May. Influence of rigid inclusions on the bending elasticity of a lipid membrane. *Phys. Rev. E*, 74:051503–1–051503–12, 2006.

[23] J. Frank. Cryo-electron microscopy as an investigative tool: The ribosome as an example. *BioEssays*, 23:725–732, 2001.

[24] J. Frank and R. K. Agrawal. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, 406:318–322, 2000.

[25] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, and K. Schulten. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14:437–449, 2006.

[26] P. L. Freddolino, F. Liu, M. Gruebele, and K. Schulten. Ten-microsecond MD simulation of a fast-folding WW domain. *Biophys. J.*, 94:L75–L77, 2008.

[27] D. Frenkel and B. Smit. *Understanding Molecular Simulation From Algorithms to Applications*. Academic Press, California, 2002.

[28] R. N. Frese, J. C. Pàmies, J. D. Olsen, S. Bahatyrova, C. D. van der Weij-de Wit, T. J. Aartsma, C. Otto, C. N. Hunter, D. Frenkel, and R. van Grondelle. Protein shape and crowding drive domain formation and curvature in biological membranes. *Biophys. J.*, 94:640–647, 2008.

[29] R. N. Frese, C. A. Siebert, R. A. Niederman, C. N. Hunter, C. Otto, and R. van Grondelle. The long-range organization of a native photosynthetic membrane. *Proc. Natl. Acad. Sci. USA*, 101:17994–17999, 2004.

[30] I. S. Gabashvili, S. T. Gregory, M. Valle, R. Grassucci, M. Worbs, M. C. Wahl, A. E. Dahlberg, and J. Frank. The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22. *Mol. Cell*, 8:181–188, 2001.

[31] T. Geyer and V. Helms. A spatial model of the chromatophore vesicles of *Rhodobacter sphaeroides* and the position of the cytochrome bc_1 complex. *Biophys. J.*, 91:921–926, 2006.

[32] R. P. Gonçalves, A. Bernadac, J. N. Sturgis, and S. Scheuring. Architecture of the native photosynthetic apparatus of *Phaeopirillum molischianum*. *J. Struct. Biol.*, 152:221–228, 2005.

[33] J. Gumbart and K. Schulten. Molecular dynamics studies of the archaeal translocon. *Biophys. J.*, 90:2356–2367, 2006.

[34] D. Heritage and W. F. Wonderlin. Translocon pores in the endoplasmic reticulum are permeable to a neutral, polar molecule. *J. Biol. Chem.*, 276:22655–22662, 2001.

[35] T. Hessa, H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S. H. White, and G. von Heijne. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, 433:377–381, 2005.

[36] X. Hu, T. Ritz, A. Damjanović, F. Autenrieth, and K. Schulten. Photosynthetic apparatus of purple bacteria. *Quart. Rev. Biophys.*, 35:1–62, 2002.

[37] W. Humphrey, A. Dalke, and K. Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[38] C. N. Hunter, J. D. Tucker, and R. A. Niederman. The assembly and organisation of photosynthetic membranes in *Rhodobacter sphaeroides*. *Photochem. Photobiol. Sci.*, 4:1023–1027, 2005.

[39] D. T. Jones. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, 292:195–202, 1999.

[40] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118:11225–11236, 1996.

[41] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comp. Phys.*, 151:283–312, 1999.

[42] J. Koepke, X. Hu, C. Muenke, K. Schulten, and H. Michel. The crystal structure of the light harvesting complex II (B800-850) from *Rhodospirillum molischianum*. *Structure*, 4:581–597, 1996.

[43] A. Korostelev, S. Trakhanov, M. Laurberg, and H. F. Noller. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell*, 126:1–13, 2006.

[44] J. Kubelka, W. A. Eaton, and J. Hofrichter. Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.*, 329(4):625–630, Jun 2003.

[45] S. Kumar, C. Huang, G. Zheng, E. Bohm, A. Bhatele, J. C. Phillips, H. Yu, and L. V. Kalé. Scalable Molecular Dynamics with NAMD on Blue Gene/L. *IBM Journal of Research and Development: Applications of Massively Parallel Systems*, 52(1/2):177–187, 2007.

[46] H. Lei and Y. Duan. Ab initio folding of albumin binding domain from all-atom molecular dynamics simulation. *J. Phys. Chem. B*, 111(19):5458–5463, May 2007.

[47] H. Lei and Y. Duan. Two-stage folding of hp-35 from ab initio simulations. *J. Mol. Biol.*, 370(1):196–206, Jun 2007.

[48] A. Liljas. *Structural Aspects of Protein Synthesis*. World Scientific, Singapore, 2004.

[49] F. Liu, D. Du, A. A. Fuller, J. E. Davoren, P. Wipf, J. W. Kelly, and M. Gruebele. An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc. Natl. Acad. Sci. USA*, 105(7):2369–2374, Feb 2008.

[50] A. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586–3616, 1998.

[51] A. D. MacKerell Jr. and N. K. Banavali. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comp. Chem.*, 21(2):105–120, 2000.

[52] A. D. MacKerell Jr., B. Brooks, C. L. Brooks, III, L. Nilsson, B. Roux, Y. Won, and M. Karplus. CHARMM: The energy function and its parameterization with an overview of the program. In P. Schleyer et al., editors, *The Encyclopedia of Computational Chemistry*, pages 271–277. John Wiley & Sons, Chichester, 1998.

[53] A. D. MacKerrell Jr., M. Feig, and C. L. Brooks III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem.*, 25:1400–1415, 2004.

[54] J. F. Ménétret, J. Schaletzky, W. M. Clemons, Jr., A. R. Osborne, S. S. Skånland, C. Denison, S. P. Gygi, D. S. Kirkpatrick, E. Park, S. J. Ludtke, T. A. Rapoport, and C. W. Akey. Ribosome binding of a single copy of the SecY complex: implications for protein translocation. *Mol. Cell*, 28:1083–1092, 2007.

[55] K. Mitra, C. Schaffitzel, T. Shaikh, F. Tama, S. Jenni, C. L. Brooks, N. Ban, and J. Frank. Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature*, 438:318–324, 2005.

[56] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. Skeel, K. Schulten, and R. Kufrin. MDScope – A visual computing environment for structural biology. *Comput. Phys. Commun.*, 91(1-3):111–134, 1995.

[57] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. D. Skeel, and K. Schulten. NAMD – A parallel, object-oriented molecular dynamics program. *Int. J. Supercomp. Appl. High Perform. Comp.*, 10:251–268, 1996.

[58] N. Nevskaya, S. Tishchenko, S. Volchkov, V. Kljashtorny, E. Nikonova, O. Nikonov, A. Nikulin, C. Köhrer, W. Piendl, R. Zimmermann, P. Stockley, M. Garber, and S. Nikonov. New insights into the interaction of ribosomal protein L1 with RNA. *J. Mol. Biol.*, 355:747–759, 2006.

[59] A. Nikulin, I. Eliseikina, S. Tishchenko, N. Nevskaya, N. Davydova, O. Platonova, W. Piendl, M. Selmer, A. Liljas, D. Drygin, R. Zimmermann, M. Garber, and S. Nikonov. Structure of the L1 protuberance in the ribosome. *Nat. Struct. Biol.*, 10:104–108, 2003.

[60] J. Oelze and G. Drews. Membranes of photosynthetic bacteria. *Biochim. Biophys. Acta*, 265:209–239, 1972.

[61] J. D. Olsen, J. N. Sturgis, W. H. J. Westerhuis, G. J. S. Fowler, C. N. Hunter, and B. Robert. Site-directed modification of the ligands to the bacteriochlorophylls of the light-harvesting LH1 and LH2 complexes of *Rhodobacter sphaeroides*. *Biochemistry*, 36:12625–12632, 1997.

[62] A. R. Osborne, T. A. Rapoport, and B. van den Berg. Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell. Dev. Biol.*, 21:529–550, 2005.

[63] M. Z. Papiz, S. M. Prince, T. Howard, R. J. Cogdell, and N. W. Isaacs. The structure and thermal motion of the B800-850 LH2 complex from *Rps. acidophila* at 2.0 Å resolution and 100 K: New structural features and functionally relevant motions. *J. Mol. Biol.*, 326:1523–1538, 2003.

[64] J. Phillips, G. Zheng, S. Kumar, and L. Kale. NAMD: Biomolecular simulation on thousands of processors. In *Proceedings of the IEEE/ACM SC2002 Conference, Technical Paper 277*. IEEE Press, 2002.

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

[66] J. W. Pitera and W. Swope. Understanding folding and design: replica-exchange simulations of "trp-cage" miniproteins. *Proc. Natl. Acad. Sci. USA*, 100(13):7587–7592, Jun 2003.

[67] P. Qian, C. N. Hunter, and P. A. Bullough. The 8.5 Å projection structure of the core RC-LH1-PufX dimer of *Rhodobacter sphaeroides*. *J. Biol. Chem.*, 349:948–960, 2005.

[68] Z. Qin, J. Ervin, E. Larios, M. Gruebele, and H. Kihara. Formation of a compact structured ensemble without fluorescence signature early during ubiquitin folding. *J. Phys. Chem. B*, 106:13040–13046, 2002.

[69] T. A. Rapoport, B. Jungnickel, and U. Kutay. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu. Rev. Biochem.*, 65:271–303, 1996.

[70] U. Rawat, H.Gao, A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank. Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM. *J. Mol. Biol.*, 357:1144–1153, 2006.

[71] U. Rawat, A. Zavialov, J. Sengupta, M. Valle, R. Grassucci, J. Linde, B. Vestergaard, M. Ehrenberg, and J. Frank. A cryo-electron microscopic study of ribosome-bound termination factor RF2. *Nature*, 321:87–90, 2003.

[72] B. J. Reynwar, G. Illya, V. A. Harmandaris, M. M. Müller, K. Kremer, and M. Deserno. Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature*, 447:461–464, 2007.

[73] A. Sali and T. L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, 234:779, 1993.

[74] S. Scheuring, F. Francia, J. Busselez, B. A. Melandris, J.-L. Rigaud, and D. Levy. Structural role of PufX in the dimerization of the photosynthetic core complex of *Rhodobacter sphaeroides*. *J. Biol. Chem.*, 279(5):3620–3626, 2004.

[75] S. Scheuring, J. Seguin, S. Marco, D. Levy, C. Breyton, B. Robert, and J.-L. Rigaud. AFM characterization of tilt and intrinsic flexibility of *Rhodobacter sphaeroides* light harvesting complex 2 (LH2). *J. Mol. Biol.*, 325:569–580, 2003.

[76] K. Schulten, J. C. Phillips, L. V. Kalé, and A. Bhatele. Biomolecular modeling in the era of petascale computing. In D. Bader, editor, *Petascale Computing: Algorithms and Applications*, pages 165–181. Chapman and Hall/CRC Press, Taylor and Francis Group, New York, 2008.

[77] B. S. Schuwirth, M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. M. Holton, and J. H. D. Cate. Structures of the bacterial ribosome at 3.5 Å resolution. *Science*, 310:827–834, 2005.

[78] M. Selmer, C. M. Dunham, F. V. Murphy IV, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, and V. Ramakrishnan. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science*, 313:1935–1942, 2006.

[79] M. K. Sener, J. D. Olsen, C. N. Hunter, and K. Schulten. Atomic level structural and functional model of a bacterial photosynthetic membrane vesicle. *Proc. Natl. Acad. Sci. USA*, 104:15723–15728, 2007.

[80] D. Shental-Bechor, S. J. Fleishman, and N. Ben-Tal. Has the code for protein translocation been broken? *Trends Biochem. Sci.*, 31:192–196, 2006.

[81] C. Simmerling, B. Strockbine, and A. E. Roitberg. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.*, 124(38):11258–11259, Sep 2002.

[82] S. Simon and G. Blobel. A protein-conducting channel in the endoplasmic-reticulum. 65:371–380, 1991.

[83] C. D. Snow, H. Nguyen, V. S. Pande, and M. Gruebele. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature*, 420(6911):102–106, Nov 2002.

[84] Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314:141–151, 1999.

[85] L. G. Trabuco, E. Villa, K. Mitra, J. Frank, and K. Schulten. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure*, 2008. In press.

[86] C. S. Tung and K. Y. Sanbonmatsu. Atomic model of the *Thermus thermophilus* 70S ribosome developed in silico. *Biophys. J.*, 87:2714–2722, 2004.

[87] R. B. Tunnicliffe, E. C. Ratcliffe, C. N. Hunter, and M. P. Williamson. The solution structure of the PufX peptide from *Rhodobacter sphaeroides*. *FEBS Lett.*, 580:6967–6971, 2006.

[88] M. Valle, J. Sengupta, N. K. Swami, R. A. Grassucci, N. Burkhardt, K. H. Nierhaus, R. K. Agrawal, and J. Frank. Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J.*, 21:3557–3567, 2002.

[89] M. Valle, A. Zavialov, W. Li, S. M. Stagg, J. Sengupta, R. C. Nielsen, P. Nissen, S. C. Harvey, M. Ehrenberg, and J. Frank. Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy. *Nat. Struct. Biol.*, 10:899–906, 2003.

[90] M. Valle, A. Zavialov, J. Sengupta, U. Rawat, M. Ehrenberg, and J. Frank. Locking and unlocking of ribosomal motions. *Cell*, 114:123–134, 2003.

[91] B. van den Berg, W. M. Clemons, Jr., I. Collinson, Y. Modis, E. Hartmann, S. C. Harrison, and T. A. Rapoport. X-ray structure of a protein-conducting channel. *Nature*, 427:36–44, 2004.

[92] T. Walz, S. J. Jamieson, C. M. Bowers, P. A. Bullough, and C. N. Hunter. Projection structures of three photosynthetic complexes from *Rhodobacter sphaeroides*: LH2 at 6 Å, LH1 and RC-LH1 at 25 Å. *J. Mol. Biol.*, 282:833–845, 1998.

[93] T. Wang and R. Wade. Force field effects on a B-sheet protein domain structure in thermal unfolding simulations. *J. Chem. Theory Comput.*, 2(1):140–148, 2006.

[94] T. R. Weikl, M. M. Kozlov, and W. Helfrich. Interactions of conical membrane inclusions: Effect of lateral tension. *Phys. Rev. E*, 57(6):6988–6995, 1998.

[95] W. H. J. Westerhuis, M. Vos, R. van Grondelle, J. Amesz, and R. A. Niederman. Altered organization of light-harvesting complexes in phospholipid-enriched *Rhodobacter sphaeroides* chromatophores as determined by fluorescence yield and singlet-singlet annihilation measurements. *Biochim. Biophys. Acta – Bioener.*, 1366:317–329, 1998.

[96] J. Xiong, W. M. Fischer, K. Inoue, M. Nakahara, and C. E. Bauer. Molecular evidence for the early evolution of photosynthesis. *Science*, 289:1724–1730, 2000.

[97] T. Yoda, Y. Sugita, and Y. Okamoto. Secondary-structure preferences of force fields for proteins evaluated by generalized-ensemble simulations. *Chem. Phys.*, 307(2-3):269–283, Dec. 2004.

[98] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. D. Cate, and H. F. Noller. Crystal structure of the ribosome at 5.5 Å resolution. *Science*, 292:883–896, 2001.

[99] R. Zhou, B. J. Berne, and R. Germain. The free energy landscape for beta hairpin folding in explicit water. *Proc Natl Acad Sci U S A*, 98(26):14931–14936, Dec 2001.