

Modeling supramolecular assemblages

Adrian H Elcock

There has been some progress (but not much) in simulating supramolecular assemblages in the past year. The two main technical advances have been, firstly, the establishment of a protocol for extracting equilibrium thermodynamic data from forced (i.e. nonequilibrium) simulations and experiments, and, secondly, the development of a method for accurately calculating the electrostatics of enormous systems. Some recent applications have demonstrated the increasing feasibility of performing meaningful simulations of very large systems.

Addresses

Department of Biochemistry, University of Iowa, Iowa City, IA 52242-1109, USA; e-mail: adrian-elcock@uiowa.edu

Current Opinion in Structural Biology 2002, 12:154–160

0959-440X/02/\$ – see front matter

© 2002 Elsevier Science Ltd. All rights reserved.

Abbreviations

BD	Brownian dynamics
EM	electron microscopy
MD	molecular dynamics
PB	Poisson–Boltzmann
PME	particle mesh Ewald
SH	Src homology

Introduction

The word ‘supramolecular’ means different things to different people. Because of this, it is appropriate to begin this review with an apology to those from the discipline of supramolecular chemistry, who may have unwittingly added it to their pile of ‘must-read’ (or ‘must-photocopy’) material. Sitting down to examine this review, they will be disappointed to find that it is of almost no relevance to their own research efforts. Instead, as they will now discover, the supramolecular assemblages that are under discussion here are those formed by the association of multiple biological macromolecular components. As some of the impressive advances that have been made in determining the structures of these assemblages are discussed elsewhere in this issue, an attempt has been made here to remain within the remit and concentrate on reviewing work that is primarily theoretical in nature. A completely clean separation between theoretical and experimental studies is not possible, however, for the following reasons. First, a review that was restricted to truly theoretical studies would be embarrassingly short, perhaps only two paragraphs, for the simple reason that applications of computer simulations to macromolecular assemblages are, as yet, rather scarce. Second, the determination of all supramolecular structures is achieved by employing computational techniques (at some level) to construct molecular models consistent with the available experimental data. Finally, as is detailed below (in a way that may at times appear to be more of a rant than a reasoned scientific thesis), this reviewer is strongly of the opinion

that the agenda for simulation research is often set more by the kinds of structures that have recently become available than by the development of truly testable hypotheses. Each of these points, together with a highly selective list of interesting developments in both methodology and applications of supramolecular modeling, is discussed in more detail below.

Structures

In the past several years, there has been an explosion in the availability of structures of very large macromolecular assemblages [1]. The most prominent recent example is, of course, that of the ribosome [2,3], high-resolution structures of which have finally been obtained by good old-fashioned X-ray crystallographic methods, albeit applied at a previously unheard-of scale. The road to a high-resolution structure has been tortuous — the first crystals were, after all, obtained some 20 years ago — and a final solution has been obtained not simply through dogged perseverance, but also through the incorporation of a number of novel technical developments (all covered in an excellent review in an earlier issue of this journal [4••]). Lest we get carried away and imagine that this ushers in a new age in which X-ray crystal structures of large assemblages will be obtained routinely, it is worth pointing out that the rate-limiting step for such studies remains that of obtaining high-quality, diffracting crystals, a process that, even today, remains more art than science.

From a theoretical perspective, the technical developments in the field of structure determination that are perhaps of most interest are those aimed at facilitating the combination of X-ray crystallography methods with cryoelectron microscopy (cryo-EM) techniques, the former providing a high-resolution picture of isolated components and the latter giving a lower resolution image of an entire supramolecular assemblage [5]. For crystallographers, of course, a combination of the techniques is especially attractive, given that it provides a route to obtaining two major publications for the price of one-and-a-bit. This cynical comment aside, few would argue that the more prominent cases in which high- and low-resolution models have been combined (e.g. the landmark actin filament model developed from fiber diffraction data [6]) have not been of tremendous value for shaping ideas in their respective fields. Recent developments have focused on taking the task of fitting high-resolution structures into low-resolution density out of the hands of the possibly prejudiced human and into those of the more objective (but usually less intelligent) computer [7•]. As a perfect overlap of high- and low-resolution data is rare, an especially interesting development has been the addition of a third element to the fitting process; molecular dynamics (MD) simulations can be used to guide the high-resolution model into a conformation more consistent

with the low-resolution data [8]. Clearly, work along these lines significantly blurs the boundary between experiment and model. It should be remembered, however, that this boundary is also blurred, though the fact is more easily overlooked, in the use of both X-ray crystallography [9] and (especially) NMR spectroscopy [10] for the solution of macromolecular structures.

Simulation methods

The availability of medium- to high-resolution structures is not the only prerequisite to meaningful simulation work; the availability of suitable simulation techniques is also obviously important. As several other reviews discuss methods that also have direct implications for the simulation of supramolecular assemblages (see the review by Gabdoulhine and Wade, this issue, pp 204–213; [11]), the list of methodologies discussed here is deliberately selective: a more general discussion of some of the techniques available has also been published [12*]. As there is currently no reason to believe that there will be any conceptual differences between the principles that dictate the stability and dynamics of large assemblages and those that apply to isolated macromolecules (all that appears different is the scale), the methodological problems to be overcome appear primarily technical in nature. Because of this, although some of the recent advances are undoubtedly important, they are often boring in the extreme to those who are not actively involved in their development; an attempt has therefore been made here to restrict the discussion to results that may be of more general interest. In passing, it is worth stating that it would, of course, be utterly refreshing if new underlying physical principles were to be uncovered from the study of complex assemblages [13].

One of the great hopes for simulation methods is to help fill in gaps in our knowledge about the internal motions and conformational fluctuations of macromolecules, motions that appear repeatedly to be of major functional importance (e.g. [14]). The most popular experimental technique for investigating the conformational dynamics of individual proteins is NMR spectroscopy: in addition to providing information on fluctuations occurring on a nanosecond to picosecond timescale [15], more recently developed NMR techniques have allowed much longer timescale fluctuations (microsecond to millisecond) to be studied [16]. In the present context, however, it should be remembered that most applications of these methods have so far involved relatively small proteins. The experimental techniques for examining the range of structures that can be adopted by macromolecules are, of course, X-ray crystallography (assuming that alternative conformations are crystallizable) and cryo-EM — the latter being especially useful for reconstructing conformational variability in very large assemblages (e.g. [17**]).

For simulating conformational fluctuations in macromolecular systems on the timescale of approximately 10 ns, the technique of choice remains MD methods. The basis of

the technique as applied to proteins and nucleic acids has remained more or less unchanged for many years [18]; the one major technical development of the past several years has been the routine incorporation of particle mesh Ewald (PME) methods (and variants) for the description of long-range electrostatic interactions [19]. The need for an accurate treatment of such interactions became apparent in the mid 1990s after repeated (and abject) failures to perform structurally stable MD simulations of DNA. PME methods enabled extremely stable trajectories to be routinely obtained at a minimal additional computational cost. As these methods have been in place for several years and have been well documented [20], they would not be worthy of discussion here but for the fact that some recent results have just begun to call their all-conquering abilities into a little doubt. Interesting work by Hünenberger and colleagues [21,22*] has indicated that the imposition of crystalline order that is implicit in the use of PME methods — which is fine if one is interested in simulating a crystal, but not necessarily if one's intention is to simulate dilute solution — may actually result in too much stability, at least in simulations of an α helix in water [22*]. If this result is also demonstrated to apply to other systems, it will be amusing, though not perhaps for the PME code developers, to find that the solution to one problem (artificially low stability) has simply resulted in the incorporation of a new simulation artifact (artificially high stability). As excessive stability might be expected to place constraints on the conformational freedom of a macromolecule, developments in this area may be worth watching in the near future.

Given the current limitations on the upper timescale of MD simulations, it is generally not feasible to observe events such as ligand binding or large-scale transitions between energetically accessible conformations of even a single macromolecule, far less those of a macromolecular complex. For those interested in studying such processes, two alternative methods are available. If the beginning and end states of a transition are known, one option is to use 'steered' (or 'targeted' [23]) MD techniques, which seek to enforce structural transformations by the imposition of additional forces in the simulations (see also Update): for ligand binding/unbinding or forced unfolding processes, such simulations are, in effect, the theoretical analogs of the real-life single-molecule experiments performed using techniques such as atomic force microscopy (AFM; [24]). As steered MD has already been the subject of a nice review in this journal [25], there would again be little point in discussing it except for the following rather interesting development. As transitions in steered MD simulations (and single-molecule experiments) are only effected by the application of external forces, it might be thought that obtaining equilibrium thermodynamic information for the process would require the extrapolation of data to the (unattainable) case in which either the external force is zero or the entity applying the external force (the 'cantilever') is moved at zero speed. However, building on

a fundamental relationship derived by Jarzynski [26] (and described in a wonderfully readable manner in [27••,28•]), Hummer and Szabo [29••] have developed a recipe for extracting the underlying equilibrium thermodynamics of a forced process by weighted averaging of the results of a number of AFM experiments. Although primarily focused on experimental protocols, what makes their work particularly interesting is the fact that its central results are equally applicable to data obtained from steered MD simulations. It will be interesting to see if the same kind of analysis might be useful for determining the equilibrium thermodynamics of more structurally complicated processes, such as those involved in large-scale conformational transitions of macromolecular assemblages (e.g. [30]).

An alternative strategy for trying to identify large-scale conformational motions, or for identifying the static and dynamic domains of very large molecules, has come from the intelligent use of simplified representations of protein structure [31,32]. Central to these approaches is the replacement of an all-atom model with one in which each amino acid is represented by a single (or several) pseudo-atom(s). The major advantage of these structural descriptions is that, when combined with equally crude approximations concerning the forces acting between nearby residues, they can yield an analytical description of a macromolecule's internal motions and are therefore free from the statistical uncertainties that can plague analysis of more structurally detailed MD simulations [33]. Despite their approximations, the methods can yield surprisingly useful descriptions of the overall conformational dynamics of macromolecules — as indicated by the good agreement with experimental observables such as crystallographic temperature factors, hydrogen/deuterium exchange rates and NMR order parameters (discussed in [34•]). One of the most effective of these simple models, the Gaussian network model (GNM) pioneered by Bahar and Jernigan [31], has recently been extended [34•] to allow for anisotropic motions; models such as this may represent a more feasible route to simulating conformational fluctuations of very large systems than can currently be obtained using conventional MD approaches.

The sheer size of supramolecular complexes not only introduces obstacles to the treatment of their dynamical properties, but also places limits on the application of the Poisson–Boltzmann (PB) methods that are otherwise commonly used to describe the electrostatic properties of macromolecular systems [35]. This limitation is primarily attributable to the fact that the standard finite difference methods used to solve the equation require enormous three-dimensional grids to encompass entire complexes at a reasonable resolution. In the past year, however, there has been a major breakthrough in this area with direct relevance to the subject of supramolecular assemblages. The groups of McCammon and Holst [36••] have jointly developed the APBS method (for adaptive PB solver)

for solving the PB equation for systems consisting of more than a million atoms. This impressive advance now allows the complicated electrostatics of microtubules, ribosomes and any other outrageously large molecular system one might care to imagine to be calculated comparatively straightforwardly. It is worth pointing out that the real advantage of this method is not to be found in the calculation of electrostatic energies; these could after all, in most cases, be calculated by judicious use of the original technique of ‘focusing’ [37]. Instead, the method's greatest utility is likely to be found in simulations of macromolecular diffusion (see the review by Gabdoulline and Wade, this issue, pp 204–213), in which not only electrostatic interactions play important roles, but also a complete, accurate description of the electrostatic potential around the entire macromolecular complex is a necessity.

Simulation results

There have been few simulations reported in the past year that clearly fit the classification of ‘modeling supramolecular assemblages’, at least as such assemblages are defined in this review. Those that come closest are two Brownian dynamics (BD) studies employing low-resolution structural models of DNA. In the first, Beard and Schlick [38•], building on previous work by their own and other groups, have used BD simulations to simulate the behavior of chromatin models comprising up to 48 nucleosomal particles. The simulations show that a 30 nm diameter model of the chromatin fiber is stable in 50 mM salt, but unravels rapidly in 10 mM salt. Encouragingly, these and other results reported in the same study are shown to be in quite good agreement with the available experimental data. In another interesting BD study, a similar but cruder kind of DNA modeling was used to estimate the forces required to package DNA into a bacteriophage capsid [39•]. In addition to providing force estimates, the simulations also revealed an interesting structural transition in the manner in which the highly concentrated DNA is packaged within the capsid.

In addition to these BD studies, there have been several MD studies that have produced interesting results for large macromolecular systems; even though such systems are not technically assemblages, two of these studies (coincidentally performed by the same group) are discussed in detail below. In the first [40•], a series of MD simulations was performed of the tyrosine kinases c-Src and Hck to investigate changes in the internal dynamics resulting from dephosphorylation of Tyr527, a step necessary for activation of the kinase. As the authors recognize, the timescale of these simulations (5 ns or less) is too short to obtain a completely sampled (and therefore completely reliable) view of the dynamics and, certainly, much too short to observe the known transition in the kinase's activation loop, which is remote from the site of dephosphorylation. Nevertheless, by performing a variety of simulations, the authors obtained interesting and unanticipated evidence

for the existence of strong correlations between the motions of the SH2 and SH3 domains of the kinase — correlations mediated by a linker region previously thought to be rather unimportant to the function of the enzyme. Especially dramatic was the effect of mutating three key residues in c-Src's linker region to glycine. In the resulting mutant, all structure in this region was lost within 4 ns and coupling between the SH2 and SH3 domains was dramatically reduced. Cross-correlations between domains have been studied before, but what sets this study apart from most others is the inclusion of experimental testing of the ideas; support for the linker's importance was obtained from the fact that the three-glycine c-Src mutant was found to be constitutively active (i.e. unregulated) *in vivo*. Of course, this ability to combine simulation and experimental approaches is available to a limited number of research groups; the results obtained, however, provide ample demonstration of the power of such a combination.

Another very recent study by the same group [41**] has combined X-ray crystallographic studies with MD simulations to gain insight into the process of loading the ring-shaped clamp protein (the β subunit) necessary for promoting processive replication by *Escherichia coli* DNA polymerase III. The clamp consists of a highly stable dimer of crescent-shaped β monomers and placing it around DNA requires that this arrangement is broken open through disruption of one of the monomer–monomer interfaces. Such a process, known as clamp loading, is mediated by the so-called γ complex. The crystal structure of the essential δ subunit of γ complexed with a single β monomer indicated that the curvature of the crescent-shaped β monomer was significantly decreased from that observed in the β dimer. In fact, this loss of curvature was such that a model of the β dimer constructed from two such monomer structures would be sufficiently open at one end that the passage of at least a single strand of DNA into the inside of the ring would now be possible. This suggested to the authors that the conformation of the β monomers observed in the dimeric structure might be intrinsically strained. To test this hypothesis, they conducted two separate MD simulations, one of the β dimer and one of a single β monomer (initially in the same conformation as found in the dimer). Amazingly, the MD simulations confirmed their ideas; although stable in the dimer, the conformation of the isolated monomer spontaneously relaxed to a conformation almost identical to that observed in the β – δ complex. As the authors rightly point out, because such large-scale transitions are rarely observed in MD simulations conducted on a nanosecond timescale, this provides strong evidence for the presence of considerable conformational strain in the monomer. As such, this study clearly demonstrates that fundamental insight into the operation of a large macromolecular ‘machine’ (if that is what we must call it) can be obtained from simulations — especially when guided by hypotheses based on experimental results.

Where are we going? (And do we really want to go there?)

Simulations of supramolecular assemblages are at an embryonic stage of development. Although this makes it more or less impossible to write a satisfactory review, it does lend the reviewer an opportunity to outline their own biased view of how they hope the field will develop. The starting point for what follows is this reviewer's desire for simulation techniques to reach a stage of development at which they are seen universally in the research community as true complements to conventional techniques, rather than as tools for simply recapitulating results already obtained from experiments, as is so often the case currently. An essential step in this direction is to show, for as many different observables as possible, that the results of simulations are consistent with experimental data; only once this is demonstrated can one have true confidence in applying simulation methods to problems not directly addressable by experimental techniques.

Most current applications of MD simulations provide a vivid illustration that this requirement is being largely ignored. In a recent review of the field, Daggett [42*] noted that the upper time limit of published MD simulations is not keeping pace with increases in available computer power. A sympathetic interpretation of this observation is that researchers are now conducting multiple simulations of a given system when previously they were performing only a single simulation. A less sympathetic interpretation (my own) is that the increased computational resources are being invested in conducting simulations of progressively larger macromolecules for the same length of time, rather than being invested in simulations of small macromolecules for progressively longer times. This would not be a problem if one could be confident that the accessible timescales of current simulations were sufficiently long that they could be considered to be fully sampled, but this is not the case. In the same review [42*], Daggett noted that simulations of 50–100 ns might have been conducted, but may not have been stable: this clearly raises the question of whether not reporting such results is the correct course to take. In fact, finding out that stability is not maintained on these timescales (or that appropriate fluctuations do not occur on an even longer timescale) would actually be a rather important negative result, given the hope to eventually use MD simulations to study more functionally relevant long-timescale motions. However, with one notable exception [43], instead of making determined efforts to routinely achieve microsecond timescales, tacit agreement appears to have been reached within the MD community that 5–10 ns is somehow an acceptable upper limit for a simulation: perhaps this is because so many of the current MD simulations are not aimed at simulating stable behavior at all, but are instead intended to provide insight into unfolding (and therefore folding) transitions [42*]. For simulations in supposedly stable conditions, 10 ns is perhaps long enough to begin making comparisons to certain NMR observables, such as spin relaxation behavior, which is

usually interpreted in terms of order parameters for the appropriate bond vectors (via the very useful but poorly named ‘model-free’ approach of Lipari and Szabo [44]). At long last, and many years after the first attempts to do so (e.g. [45]), an increasing number of studies are making detailed comparisons between simulated and experimental relaxation behaviors (reviewed in [46]). That said, 10 ns is clearly far below the microsecond timescale at which the truly interesting conformational behavior of proteins is expected to emerge and for which exciting experimental results have recently started to appear [47,48]. A crucial question is, therefore, whether computational resources should continue to be expended accruing incompletely sampled simulation data for larger and larger systems, or whether concerted efforts should be taken to investigate the very long timescale behavior of smaller, model proteins.

A similar but less drastic situation arises in the area of macromolecular electrostatics. As outlined earlier, a major advance in the past year was the development of a method capable of solving the PB equation for very large systems. The availability of such a method will now undoubtedly focus attention on the electrostatic properties of very large macromolecular assemblages, but to what end? Certainly, there are interesting diffusional issues that can now be investigated, such as the effects of electrostatic interactions on the processivity of kinesin motors as they move along microtubules [49]. Beyond this, however, it should be remembered that the new method does not represent a revolution in terms of fundamental insight into electrostatic interactions; it simply represents a revolution in terms of the scale at which such interactions can be computed. Much of the analysis of electrostatic effects will therefore continue to consist of staring wide-eyed and open-mouthed at blue and red blobs distributed over the surface of macromolecules, when the same information could, in some cases (but by no means all), be obtained simply by coloring surfaces according to the identity of the exposed residues. Instead, more fundamental insights into electrostatic effects are likely to be found at smaller scales, either by continuing to argue over the meaning of dielectric constants [50•] or by making attempts to improve agreement with experiments that, as near as possible, provide direct measurements of electrostatic interactions [51].

If it has not by now become apparent, this reviewer is strongly of the opinion that, in our desire to apply simulations to progressively larger, more elaborate systems, we may forget that many of the methods in use have remained fundamentally untested at one level or another. In fact, the pressure to study more complex (and visually impressive) systems is actually a serious impediment to conducting proper validation studies. Of course, it is difficult to imagine a grant-reviewer leaping out of their chair in excitement at reading of plans to conduct microsecond simulations of bovine pancreatic trypsin inhibitor (BPTI) or lysozyme, but in terms of settling

outstanding questions regarding the accuracy of MD simulations, such studies are likely to be of more use than a 10 ns simulation of a very large macromolecular system. Of course, it is difficult to make this point without appearing to be an incorrigible curmudgeon — especially as the purpose of this review was specifically to focus on supramolecular assemblages — and to assuage such impressions let me state my own belief that there is genuine excitement to be experienced in the application of simulation techniques to supramolecular assemblages. But this excitement should not blind us to the fact that the real challenge to the simulator is to uncover or develop specific hypotheses relating to the operation of such systems that can be meaningfully tested with the extant methods. For supramolecular assemblages, this is likely to be achieved by the use of more simplified structural models than those used in MD; deficient as they are in terms of describing atomic-level interactions, residue-level models can much more easily yield the kinds of converged results that are needed for one to properly assess the usefulness of a theoretical model. Rant over, I will now unwisely wager half of my meager start-up funds that, somewhere in the academic community, plans to conduct MD simulations of the ribosome have already been drawn up and that, even as this review is being completed, preliminary efforts are being made to redimension arrays to prevent repeated system crashes. I will wager the remainder of my funds that such simulations will be motivated more by a desire to conduct a simulation of a ‘really big’ system than by a wish to prove or disprove a clear hypothesis relating to the assemblage’s function.

Update

Several very recent simulations have demonstrated the potential for performing meaningful simulations of very large macromolecular systems. Bernèche and Roux [53••] have reported interesting simulations aimed at understanding the efficiency and selectivity of ion transfer through the potassium channel. de Groot and Grubmüller [54••] have provided an atomically detailed view of water diffusion through two transmembrane channels: aquaporin-1 and the glycerol facilitator GlpF. A nice additional report from the same group [55•] provides an illustration of the possibilities for using steered MD to perform provocative simulations — this time simulating the primary mechanical energy transfer steps in F_1 -ATP synthase.

Finally, my largely incoherent screams calling for the proper testing of MD simulations have received a timely boost thanks to the recent report of the extremely fast-folding prototypical β -sheet protein domains [56•]. The observed folding times of less than 20 μ s raise the possibility that, in addition to providing views of internal conformational dynamics of proteins, MD simulations may soon yield a detailed picture of *bona fide* protein folding events — assuming of course that existing parameter sets and simulation protocols are of sufficient accuracy.

Acknowledgements

The author is very grateful for financial support from the Irene Wells Medical Research Fund administered through the University of Iowa. He is also grateful for a critical reading of parts of this manuscript by Kip Murphy and Barry Honig.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Nogales E, Grigorieff N: **Molecular machines: putting the pieces together.** *J Cell Biol* 2001, **152**:F1-F10.
2. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: **The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution.** *Science* 2000, **289**:905-920.
3. Wimberly BT, Brodersen DE, Clemons WMJ, Morgan-Warren R, von Rhein C, Hartsch T, Ramakrishnan V: **Structure of the 30A ribosomal subunit.** *Nature* 2000, **407**:327-339.
4. Ramakrishnan V, Moore PB: **Atomic structures at last: the ribosome in 2000.** *Curr Opin Struct Biol* 2001, **11**:144-154.
- An engaging and authoritative review of the exciting progress in the understanding of ribosome function prompted by recent structural advances.
5. Rossmann MG: **Fitting atomic models into electron-microscopy maps.** *Acta Crystallogr* 2001, **56**:1341-1349.
6. Holmes KC, Popp D, Gebhard W, Kabsch W: **Atomic model of the actin filament.** *Nature* 1990, **347**:44-49.
7. Wriggers W, Chacón P: **Modeling tricks and fitting techniques for multiresolution structures.** *Structure* 2001, **9**:779-788.
- A nice assessment of the strengths and weaknesses of existing techniques for constructing models of large macromolecular complexes by combining high- and low-resolution structural data.
8. Wriggers W, Birmanns S: **Using Situs for flexible and rigid-body fitting of multi-resolution single molecule data.** *J Struct Biol* 2001, **133**:193-202.
9. Brandén C-I, Jones TA: **Between objectivity and subjectivity.** *Nature* 1990, **343**:687-689.
10. Doreleijers JF, Rullmann JAC, Kaptein R: **Quality assessment of NMR structures: a statistical survey.** *J Mol Biol* 1998, **281**:149-164.
11. Smith GR, Sternberg MJ: **Prediction of protein-protein interactions by docking methods.** *Curr Opin Struct Biol* 2001, **12**:28-35.
12. Elcock AH, Sept D, McCammon JA: **Computer simulation of protein-protein interactions.** *J Phys Chem* 2000, **105**:1504-1518.
- At the risk of being accused of shameless self-promotion, this article may be useful to those looking for a reasonably comprehensive review of the use of simulation methods for studying protein-protein interactions.
13. Laughlin RB, Pines D, Schmalian J, Stojkovic BP, Wolynes P: **The middle way.** *Proc Natl Acad Sci USA* 2000, **97**:32-37.
14. Wang L, Pang Y, Holder T, Brender JR, Kurochkin AV, Zwietering ERP: **Functional dynamics in the active site of the ribonuclease binase.** *Proc Natl Acad Sci USA* 2001, **98**:7684-7689.
15. Kay LE: **Protein dynamics from NMR.** *Nat Struct Biol* 1998, **5**:513-517.
16. Palmer AG, Kroenke CD, Loria JP: **Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules.** *Methods Enzymol* 2001, **339**:204-238.
17. Zhou ZH, Liao W, Cheng RH, Lawson JE, McCarthy DB, Reed LJ, Stoops JK: **Direct evidence for the size and conformational variability of the pyruvate dehydrogenase complex revealed by three-dimensional electron microscopy.** *J Biol Chem* 2001, **276**:21704-21713.
- A fascinating study showing that the size of this classic multienzyme complex fluctuates by as much as $\pm 10\%$ from its mean value. In addition to taking care to demonstrate that the observed fluctuations are not experimental artifacts, the authors also present intriguing evidence that increased size of the cage-like structure is accompanied by increased numbers of the enzyme's flexible N-terminal domains inside the central cavity.
18. McCammon JA, Harvey SC: *Dynamics of Proteins and Nucleic Acids.* Cambridge, UK: Cambridge University Press; 1997.
19. Darden T, Perera L, Li LP, Pedersen L: **New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations.** *Structure* 1999, **7**:R55-R60.
20. Beveridge DL, McConnell KJ: **Nucleic acids: theory and computer simulation, Y2K.** *Curr Opin Struct Biol* 1998, **10**:182-196.
21. Hünenberger PH, McCammon JA: **Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatics study.** *Biophys Chem* 1999, **78**:69-88.
22. Weber W, Hünenberger PH, McCammon JA: **Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: influence of artificial periodicity on peptide conformation.** *J Phys Chem* 2000, **104**:3668-3675.
- Following up on the ideas expressed in [21], the authors conducted MD studies of a model α helix in periodic boxes of varying size. In a result that may have some people squirming, it is found that, whereas the helix rapidly unfolds in the largest simulation box (40 Å length), it remains stable in boxes of smaller size (30 and 20 Å lengths). The authors present compelling evidence that these differences in behavior (of which the former is probably the more realistic) are artifacts of the periodicity implicit in Ewald methods.
23. Schlitter J, Engels M, Kruger P, Jacoby EU, Wollmer A: **Targeted molecular dynamics simulation of conformational change: application to the T-R transition in insulin.** *Mol Sim* 1993, **10**:291-308.
24. Fisher TE, Marszalek PE, Fernandez JM: **Stretching single molecules into novel conformations using the atomic force microscope.** *Nat Struct Biol* 2000, **7**:719-724.
25. Isralewitz B, Gao M, Schulten K: **Steered molecular dynamics and mechanical functions of proteins.** *Curr Opin Struct Biol* 2001, **11**:224-230.
26. Jarzynski C: **Nonequilibrium equality for free energy differences.** *Phys Rev Lett* 1997, **78**:2690-2693.
27. Jarzynski C: **How does a system respond when driven away from thermal equilibrium?** *Proc Natl Acad Sci USA* 2000, **98**:3636-3638.
- An excellent commentary describing (in admirably straightforward language) the theoretical result that inspired the work in [29**].
28. Hendrix DA, Jarzynski C: **A 'fast growth' method of computing free energy differences.** *J Chem Phys* 2001, **114**:5974-5981.
- Building on the same idea behind the work described in [29**], Hendrix and Jarzynski describe the potential application of the idea in the area of calculating free energy differences between molecular states using computer simulations. A highly readable work in a usually unreadable journal.
29. Hummer G, Szabo A: **Free energy reconstruction from nonequilibrium single-molecule pulling experiments.** *Proc Natl Acad Sci USA* 2000, **98**:3658-3661.
- An important paper for researchers of both experimental and simulation persuasions. Although much of the mathematics is difficult for a nonphysicist to follow, the essential results and their application to experimental data are well described.
30. Ma JP, Sigler PB, Xu ZH, Karplus M: **A dynamic model for the allosteric mechanism of GroEL.** *J Mol Biol* 2000, **302**:303-313.
31. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG: **Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential.** *Fold Des* 1997, **2**:173-181.
32. Hinsen K: **Analysis of domain motions by approximate normal mode calculations.** *Proteins* 1998, **33**:417-429.
33. Berendsen HJC, Hayward S: **Collective protein dynamics in relation to function.** *Curr Opin Struct Biol* 2000, **10**:165-169.
34. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I: **Anisotropy of fluctuation dynamics of proteins with an elastic network model.** *Biophys J* 2001, **80**:505-515.
- The authors describe the extension of a simple but successful model for protein conformational dynamics to account for anisotropic positional fluctuations. In addition to providing a good source of references for others' work, the paper also provides perhaps the most accessible description of the authors' original Gaussian network model (GNM).
35. Honig B, Nicholls AJ: **Classical electrostatics in biology and chemistry.** *Science* 1995, **268**:1144-1149.

36. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA:
 •• **Electrostatics of nanosystems: application to microtubules and the ribosome.** *Proc Natl Acad Sci USA* 2001, **98**:10037-10041.
 A breakthrough in the computation of PB electrostatics of very large macromolecular systems (or 'nanosystems' as they are somewhat mystifyingly named by the authors). Imaginatively combining a variety of numerical techniques, the authors developed a highly efficient method of solving the PB equation and demonstrate the power of the method by application to the two ribosomal subunits and (amazingly) to a microtubule model consisting of more than a million atoms.
37. Gilson MK, Honig B: **Calculation of the total electrostatic energy of a macromolecular system; solvation energies, binding energies and conformational analysis.** *Proteins* 1988, **4**:7-18.
38. Beard DA, Schlick T: **Computational modeling predicts the structure and dynamics of chromatin fiber.** *Structure* 2001, **9**:105-114.
 One might argue that the title somewhat overstates their case, but this work represents an important step forward in the modeling of nucleosomal DNA. Especially interesting is the use of a comparatively detailed representation of the electrostatics of the nucleosomal core particles: the method used is one independently developed by the authors after a more or less identical technique was reported by Gabdoulline and Wade [52].
39. Kindt J, Tzili S, Ben-Shaul A, Gelbart WM: **DNA packaging and ejection forces in bacteriophage.** *Proc Natl Acad Sci USA* 2001, **98**:13671-13674.
 The system studied and the results obtained are both interesting, but the level of modeling here is more crude than that described in [38*] – and it is surprising that many of the simulation details have only been added as footnotes to the main text.
40. Young MA, Gonfloni S, Superti-Furga G, Roux B, Kuriyan J: **Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation.** *Cell* 2001, **105**:115-126.
 Nothing new from the point of view of simulation methodology, but a nice demonstration that MD techniques, when combined with intelligent analysis and experimental testing of resulting hypotheses, can yield real insight into the functioning of a complex protein system.
41. Jeruzalmi D, Yurieva O, Zhao Y, Young M, Stewart J, Hingorani M,
 •• O'Donnell M, Kuriyan J: **Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III.** *Cell* 2001, **106**:417-428.
 In common with [40*], this study highlights how even MD simulations limited to several nanoseconds in length can be used to investigate important structural hypotheses that are not easily amenable to conventional experimental study. It was, of course, entirely possible that, over the course of these simulations, no interesting conformational behavior might have been observed – that such spectacular effects were observed is perhaps indicative of a combination of good fortune and great insight on the part of the authors.
42. Daggett V: **Long timescale simulations.** *Curr Opin Struct Biol* 2000, **10**:160-164.
 An important review discussing the state of MD simulations in the year 2000 and rightly questioning the meaning of the word 'long' when applied to the duration of a MD simulation.
43. Duan Y, Kollman PA: **Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution.** *Science* 1998, **282**:740-744.
44. Lipari G, Szabo A: **Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity.** *J Am Chem Soc* 1982, **104**:4546-4559.
45. Brooks CL, Karplus M, Pettitt BM: *Proteins. A Theoretical Perspective of Dynamics, Structure and Thermodynamics.* New York: Wiley-Interscience; 1988.
46. Spyropoulos L, Sykes BD: **Thermodynamic insights into proteins from NMR spin relaxation studies.** *Curr Opin Struct Biol* 2001, **11**:555-559.
47. Mulder FAA, Mittenmaier A, Hon B, Dahlquist FW, Kay LE: **Studying excited states of proteins by NMR spectroscopy.** *Nat Struct Biol* 2001, **8**:932-935.
48. Cavanagh J, Vinters RA: **Protein dynamic studies move to a new time slot.** *Nat Struct Biol* 2001, **8**:912-914.
49. Thorn KS, Ubersax JA, Vale RD: **Engineering the processive run-length of the kinesin motor.** *J Cell Biol* 2000, **151**:1093-1100.
50. Schutz CN, Warshel A: **What are the dielectric 'constants' of proteins and how to validate electrostatic models?** *Proteins* 2001, **44**:400-417.
 A comprehensive review of the effects of dielectric constants on electrostatic calculations, emphasizing the fact that different values may be more appropriate for different processes. The exhaustive referencing is characteristically fair to other researchers in the field: fewer than 40% of the citations are to the senior author's own work.
51. Sanchez-Ruiz JM, Makhatazde GI: **To charge or not to charge?** *Trends Biotechnol* 2001, **19**:132-135.
52. Gabdoulline RR, Wade RC: **Effective charges for macromolecules in solvent.** *J Phys Chem* 1996, **100**:3868-3876.
53. Bernèche S, Roux B: **Energetics of ion conduction through the K⁺ channel.** *Nature* 2001, **414**:73-76.
 The authors use umbrella sampling techniques to compute a free energy profile for the passage of several potassium ions through a highly detailed model of the potassium channel. In addition to yielding (correctly) very low free energy barriers for ion conduction, the simulations also reproduce the selectivity of the channel for potassium in preference to sodium.
54. de Groot BL, Grubmüller H: **Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF.** *Science* 2001, **294**:2353-2357.
 In this very interesting work, the authors describe the spontaneous (i.e. unforced) diffusion of water molecules through an atomically detailed model of the aquaporin-1 channel. During the course of a 10 ns simulation, a total of 16 such permeation events are observed, a result that is in good agreement with the very high rate of permeation known from experimental work. Detailed analysis of diffusion events provides a wealth of information; especially interesting is the tentative identification of factors responsible for the different selectivities of the two channels studied.
55. Böckmann RA, Grubmüller H: **Nanoseconds molecular dynamics simulation of primary mechanical energy transfer steps in F1-ATP synthase.** *Nat Struct Biol* 2002, **9**:198-202.
 In this paper, steered MD simulations were used to study conformational rearrangements in the F₁-ATP synthase as its central γ -stalk region is forced to undergo rotation. Despite the fact that rotation is made to occur at a rate several orders of magnitude faster than it happens in reality, a number of features of the simulations appear realistic. More importantly, they suggest a plausible (and testable) explanation for the release of newly synthesized ATP molecules.
56. Ferguson N, Johnson CM, Macias M, Oschkinat H, Fersht A: **Ultrafast folding of WW domains without structured aromatic clusters in the denatured state.** *Proc Natl Acad Sci USA* 2001, **98**:13002-13007.
 As suggested by the title, extremely fast folding rates were measured for truncated WW domains of between 38 and 44 residues. The authors rightly stress the fact that the timescale of folding of these domains is achievable with MD simulations.