Teaching macromolecular modeling

Stephen C. Harvey and Robert K.-Z. Tan
Department of Biochemistry, University of Alabama at Birmingham, Birmingham, Alabama 35294 USA

ABSTRACT Training newcomers to the field of macromolecular modeling is as difficult as is training beginners in x-ray crystallography, nuclear magnetic resonance, or other methods in structural biology. In one or two lectures, the most that can be conveyed is a general sense of the relationship between modeling and other structural methods. If a full semester is available, then students can be taught how molecular structures are built, manipulated, refined, and analyzed on a computer. Here we describe a one-semester modeling course that combines lectures, discussions, and a laboratory using a commercial modeling package. In the laboratory, students carry out prescribed exercises that are coordinated to the lectures, and they complete a term project on a modeling problem of their choice. The goal is to give students an understanding of what kinds of problems can be attacked by molecular modeling methods and which problems are beyond the current capabilities of those methods.

INTRODUCTION

The emergence of macromolecular modeling as a scientific discipline has created the need for resources to introduce modeling concepts and methodology to newcomers to the field. Over the past few years, we have wrestled with the task of teaching macromolecular modeling to graduate students, postdoctoral research associates, and faculty members. Those efforts have included introductory lectures (1–2 h) given as part of other courses, as well the teaching of a one-semester advanced graduate level course in macromolecular modeling.

In our view, the problem of introducing modeling to beginners is a formidable task. Although the material does not seem that complex to those of us in the field, the basic concepts of modeling are as difficult for the novice to grasp as are the ideas behind x-ray crystallography, nuclear magnetic resonance (NMR), and other structural methods. Students and other scientists may have some kind of understanding of the overall goals of modeling (the protein folding problem is an example known to almost everyone) but they usually have no idea of how molecules are represented on a computer, how the dependence of energy on conformation is determined, or how structures are manipulated and optimized. They may be vaguely aware of the difference between structure refinement based on large quantities of experimental data versus attempts at de novo structural prediction, but they do not know how one uses information from a potential energy function in combination with experimental information, particularly when there is only a handful of real data. In short, they generally do not know anything about the philosophy and assumptions that underlie molecular modeling nor about the capabilities and limitations of the methods.

In their invitation to write this article, the editors asked us to consider what can be taught in a brief introduction (one or two lectures) and what can be covered in a longer course. The philosophies must be very different, since several weeks are required to teach the actual methods. We will consider each of these in turn.

WHAT CAN BE TAUGHT IN A BRIEF INTRODUCTION TO MODELING

Probably the most that one can hope for in a brief introduction is that students would recognize that: (*a*) all known "structures," whether described in words, pictures, or sets of atomic coordinates, are, in fact, models; (*b*) the more real data that goes into it, the better the model; (*c*) predictive models are generally better than descriptive models; (*d*) among the best models are those that are eventually destroyed because they make predictions that are tested and found to be wrong (we often learn the most from unexpected results); and (*e*) molecular modeling is a legitimate scientific discipline, with all the accompanying strengths and weaknesses. This last point is particularly important. There are still too many scientists who basically believe that molecular models are garbage and that nothing useful has ever been learned from them. A discussion of the DNA double helix is usually insufficient to disabuse these people of their prejudice. The modeler must be prepared to present a handful of recent examples. On the other hand, there are also too many investigators who do not realize that we cannot reliably predict the effects of point mutations on structure and enzyme activity, much less predict the three-dimensional structure of a protein or nucleic acid from its sequence. If the lecturer has time to present only one set of data on this point, the modeling studies of Novotny et al. (1, 2) on misfolded proteins are particularly instructive. They lead quite naturally to a brief discussion of current limitations and future research directions for the modeling field.

In this regard, the American Chemical Society (ACS) has offered short courses (televised presentations available via satellite and lasting about half a day) on macromolecular modeling. We have presented two of the ACS short courses on modeling at the University of Alabama at Birmingham (UAB), one of them near the end of the

Address correspondence to Stephen C. Harvey.

term when we were teaching our modeling course. On both occasions, the participants were nearly unanimous in reporting that they got very little from the ACS short course. This conclusion was independent of the background of the participant, ranging from those for whom this was a first exposure to modeling up to professional modelers. In our opinion, the failure of these courses was due to their focus on methods, rather than on a global view of modeling as a discipline.

A ONE-SEMESTER INTRODUCTORY COURSE ON MODELING

In an introductory course at the advanced graduate level, one can hope to make the points described above while teaching the actual methods that are used. Here we discuss some of the approaches we have taken during the one-semester course that we teach, the issues that have arisen, and what we have learned about teaching modeling to the uninitiated.

The backgrounds of our students are quite varied. Most of them come from departments that participate in UAB's Graduate Program in Cellular and Molecular Biology, so they tend to have stronger backgrounds in biology than in chemistry, physics, or mathematics. We assume that students already know the elements of macromolecular structure and function, presumably from an introductory biochemistry course. No computer programming experience is required as a prerequisite to our course.

Goals for the course

The most important overall goal is to give the students an understanding of the kinds of problems that can be attacked by molecular modeling methods and what kinds of problems are beyond the current capabilities of those methods. We usually begin our first class with a statement something like, "If you are an experimentalist, you know what kinds of questions you can ask using a given technique, be it gel electrophoresis, fluorescence spectroscopy, or whatever. You also know what kinds of questions cannot be addressed by that technique. This is because you have some understanding of how the method works. Our goal is to give you a similar feeling for molecular modeling."

This goal is accomplished if they understand how a molecule is represented on a computer, what a potential function is, how energy minimization and molecular dynamics work, how modeling methodology enters into the refinement of a molecular structure by crystallography or NMR, and why we cannot predict the three-dimensional structure of a protein from its amino acid sequence. The final lecture of the course is devoted to a personal view of the important unsolved problems in macromolecular modeling and the future directions of research in the field.

On the practical side, the students should become familiar with the use of a commercial modeling package. This should involve experience in the building and manipulation of molecules on a computer, refinement of a crude model by energy minimization, and the opportunity to look at different computer graphics representations of a given structure.

Course organization

To achieve the above goals, our modeling course combines formal lectures, discussion sessions, and laboratory sessions. The heart of the course is a term project chosen by each student early in the course and requiring the use of a commercial modeling package. Lectures focus on basic principles. The discussion sessions provide opportunities to consider issues arising in the lectures and to help the students solve problems they have encountered in the lab sessions. To be certain that the students really work in the lab, the course grade is based entirely on a term paper and oral presentation describing the results of the student's modeling project.

We should point out here that our design of the laboratory sessions continues to evolve. The first time that we offered this course, we gave no prescribed lab exercises except the tutorials introducing the use of the workstation and the commercial modeling programs. Students spent the entire term working on their own projects. There were two or three very nice projects, but we were not satisfied with the students' grasp of basic principles at the end of the course. Consequently, when we next offered the course, we introduced some mandatory exercises involving diatomic and triatomic molecules (described below). We found these very helpful in getting across important ideas, and we plan to expand those exercises in the future. The following section of the article does not differentiate between exercises we have actually used in the past and those that we hope to introduce in the future. One extreme possibility would be to eliminate the student projects entirely and devote lab sessions to prescribed exercises coordinated to lecture and discussion sessions. These are "cookbook" exercises like those often used in traditional introductory physics and chemistry labs. This course could then be used as a prerequisite for a follow-on course based almost entirely on projects chosen by the student. The latter could be either a conventional course run for a group of students or it could be an advanced nonthesis research course offered to individuals. The balance between prescribed lab exercises and time for work on student projects has to be adjusted to the background of the students, of course, but it has been our experience that it is very easy to overestimate the rate at which students will grasp basic ideas and the amount that they can actually accomplish in the laboratory.

Modeling laboratory

The laboratory part of the course begins with prescribed exercises designed to introduce the students to the computer and to the use of a commercial modeling package.

Next come a series of mandatory exercises that are used to illustrate basic concepts that are being covered concurrently in the lecture and discussion sessions. Later in the term, laboratory sessions are reserved for work on individual projects. Except for a very brief session at the start of the course when we introduce groups of four to eight students to the computer room, there are no group sessions. Students sign up for 1–2 h on either of our two IRIS 4D workstations (Silicon Graphics Corp., Mountain View, CA) and work at their own pace.

The first two weeks are used to introduce the students to the computer and to the commercial modeling package. We use the QUANTA/CHARMm package from Polygen Corp. (Waltham, MA), but other packages could be used. The only requirements are a good graphics interface, a good user reference manual, and decent tutorials for both the computer system and for the molecular modeling package. (Several years ago we ran a workshop based on AMBER, which had none of those, and the results were very poor.) After the students have run through a tutorial introducing them to the use of the workstation, we devote about an hour of class time to a lecture on the elements of the unix operating system. This is necessary so that students will know how to use a directory structure, to keep track of their disk usage, and to remove files that are no longer needed, particularly core dumps that occasionally occur. We also maintain an electronic bulletin board using a locally developed program based on the Unix *news* function. Students are encouraged to post observations and problems here; answers to problems are also posted, either by the instructors or by other students.

By the third week of the course, students are asked to select a term project. They begin the project immediately, and, in the current incarnation of the course, they use lab sessions over the next two to four weeks for both the prescribed exercises and their term projects, and then work only on the projects.

Student projects may repeat a study described in the literature or they may pose new problems. They may deal with small molecules or macromolecules. We give a list of possible projects, even though the majority of students prefer to define their own. Among the projects on the list are the Ramachandran plot for the alanine dipeptide using constrained energy minimization, base stacking in a dinucleotide, melting of an alpha helix using molecular dynamics, intercalation of a drug into a double helix, loop structure in the hypervariable region of an immunoglobulin, stem-loop structure in a nucleic acid hairpin, analysis of a molecular dynamics simulation on a small protein or nucleic acid, solvent structure around a macromolecule, and the effect of a point mutation on local structure in a macromolecule or in a macromolecule/ligand complex.

Since students are inexperienced and cannot evaluate feasibility, they generally have rather naive ideas about what can be accomplished in a few weeks. They often take this course because they hope to apply the results to a specific problem related to their own research, so we advise them about feasibility but allow them to select projects without regard to feasibility; any project is considered appropriate as long as it requires them to use the computer. There are two advantages to this approach: students are generally very interested in their projects, and lessons on the capabilities and limitations of molecular modeling methods are driven home vividly. Students sometimes continue their projects beyond the end of the course. At least two papers have been published about projects begun in the course by students outside our own laboratory.

Basic concepts

The course begins with a description of how a molecule is represented on a computer. This requires a description of the difference between covalent and nonbonded interactions, including a discussion of topology files (bond lists, nonbonded exclusion lists, and so on) and a careful introduction to potential functions, including a discussion of parameter files. Lectures and discussion sections are coordinated with prescribed lab exercises to illustrate these concepts.

If students are to really understand the basic principles of modeling, it is necessary to proceed very slowly. We begin with diatomic molecules (H_2 and the united atom representation of ethane), using bond energy as the paradigm for the relationship between conformation and internal energy. Students are assigned a laboratory exercise in which they build models with different bond lengths and plot the calculated energy versus bond length. When energy minimization and molecular dynamics (MD) are introduced (either at this point or later), the diatomic model is an ideal system for the students to monitor the progressive changes in conformation and energy, since the single degree of freedom makes analysis and graphical description very simple. One useful exercise is a free molecular dynamics simulation, starting with a nonideal bond length. This can be used to illustrate energy conservation and the interconversion of potential and kinetic energy. Students can also measure the observed vibration frequency and, given the relationship between frequency, reduced mass, and force constant, they can calculate the force constant; this supports a discussion of how parameters of the potential energy function can be determined from experimental observables, such as vibrational frequencies. If different students use different initial bond lengths, then it also can be shown that frequency is essentially independent of the amplitude of vibration, at least for small amplitudes. Other concepts also can be illustrated with diatomic molecules. For instance, the issue of how to set the size of the MD timestep can be introduced by asking the students to monitor the behavior of the system as the timestep is varied; they will find that, above some critical value, total energy is no longer conserved, and the system blows up. Another example is the difference be-

tween free MD, MD with velocity reassignment, and MD with velocity rescaling.

We use united atom models for propane and butane to introduce bond angles and torsions, respectively. Pentane or a long linear alkane is then used to introduce van der Waals interactions and the concept of a nonbonded exclusion list. We then go back and introduce nonbonded one to four interactions in the butane model (with and without scaling) to show how this modification to the potential function can improve the quality of the dependence of conformational energy on the torsion angle. Finally, electrostatic interactions are introduced, along with detailed discussions about atom types and parameter files. The alanine dipeptide or any other small peptide can be used for this task; if a united atom representation is used, then improper torsions also can be presented here.

In our experience, students tend to be divided into two types—those for whom mathematical and physical concepts come easily and those for whom the introduction of equations obscures underlying concepts, rather than clarifying them. It is important to make the maximum use of graphical plots of energy versus the relevant conformational variables, rather than simply giving equations. Physical models, especially ball-and-stick models, are also very helpful in lectures and discussions.

If time permits, a linear triatomic molecule such as CO_2 provides an opportunity to show the behavior of a coupled harmonic oscillator. Students can plot the length of the two bonds as a function of time, and normal mode analysis can be introduced.

The next important concept is that of conformational space. This can be introduced by discussing examples with one significant degree of freedom (the bond stretch of ethane or the torsional rotation in butane), two degrees of freedom (the two bonds in CO_2 or the Ramachandran plot for the alanine dipeptide), and then the general n-dimensional case. The students should already be familiar with the conformational energy surface of a Ramachandran plot, and a similar plot can easily be constructed for CO_2 bond lengths. A variety of trajectories can be shown on such plots, including pathways of energy minimization, free MD, and MD with periodic reassignment or rescaling of velocities. Students with strong backgrounds in physical chemistry may also benefit from a trajectory in phase space for the bond stretch of a diatomic molecule.

A detailed analysis of the two-dimensional conformational energy surfaces facilitates the discussion of the difference between local and global minima. The multiple minimum problem is nicely introduced with pentane, since one expects nine local minima involving the permutations of *trans, gauche+*, and *gauche−* for the two torsions associated with the C2—C3 and C3—C4 bonds. The students should be able to appreciate how, for a long chain molecule with N torsions, there are something on the order of 3^N local minima (this is actually a wild underestimate) and only a single global minimum. This is also an opportunity to explain conformational searches and the difference between the prediction of a structure and the refinement of a structure. The former requires both an accurate potential function and the ability to find the global minimum without any recourse to experimental data. In its simplest form, structure refinement means the energy minimization of a trial conformation, which will generally be far from the global minimum. In its most useful form, refinement uses as much experimental data as possible, usually from x-ray diffraction or NMR. The data drive the model into the relevant region of conformational space and, in the best of cases, are sufficient to determine the structure. The most common case lies between these two extremes. Information from a potential function is added to the experimental data, usually to guarantee reasonable values for bond lengths, bond angles, and nonbonded contact distances, whereas refinement attempts to satisfy all the experimental constraints.

We have developed one lab exercise that illustrates some of the above concepts quite nicely. The class is given a sequence of a small polypeptide, typically 12–15 amino acids long, and students are then challenged to find the lowest energy conformation for this molecule. A prize of some value is given to the winner. To facilitate comparisons between alternative structures, we agree on a standard form for the potential function (usually united atoms except for explicit polar hydrogens, no explicit solvent, a very long cutoff distance to include all nonbonded pairs, and a uniform dielectric constant), and candidate structures are presented as protein data bank files whose energies are remeasured by us. Students may use any initial structure and any method to refine it. We suggest they consider an extended conformation, an alpha helix, an antiparallel beta sheet with a turn in the middle, and some sort of amorphous blob structure. They can search the Brookhaven database for a likely conformation, they can use an automated structure builder, or they can build the model manually. Any refinement method is considered appropriate, but we do place limits on the total amount of central processing unit time each student can use. If 10 or more students produce models, this exercise shows quite vividly how difficult the de novo prediction of a structure is, and it offers a very good chance to discuss the reliability of both structures and energies produced by computer modeling.

There are a variety of other topics that can be covered in the lecture and discussion periods. Among the most important of these is the difference between free energy and the potential (internal) energy calculated by the potential function. In presenting this, the similarities and differences between energy minimization (which produces steps in conformational space) and MD (which produces steps in time) should be discussed. With more sophisticated students, one can present the basic princi-

ples of various free energy methods, particularly perturbation and thermodynamic cycles.

We also consider Brownian dynamics, Monte Carlo, simulated annealing (with either MD or Monte Carlo), and distance geometry sufficiently important that we spend some time describing these algorithms, although we do not offer laboratory exercises on them. We do not go into detail about alternative energy minimization procedures, but students with strong mathematical backgrounds could certainly understand the differences between them. Similarly, we do not discuss the details of MD, although a simple example using a projectile in a uniform gravitational field could be developed to illustrate the leapfrog algorithm and to present an opportunity for considering numerical errors. Since a fair part of our own research efforts have focused on the development of low resolution models for DNA supercoiling (3) and the structure of the ribosome (4), we spend some time on reduced representations. There are two kinds of reduced representations that we consider important, lattice models (especially as applied to protein folding, e.g., reference 5) and "succinct models." The latter refer to models that use pseudoatoms to represent pieces of the structure. Some of these are generalizations of the familiar united atom models, but others use pseudoatoms that are not simply related to the chemical structure of the molecules being studied (3). More important than the details of reduced representations is an understanding that, in designing a modeling study, it is important to consider computational resources and to use an appropriate level of detail in the study.

Another topic of considerable current importance in the modeling community is the treatment of solvent effects. We discuss the difference between explicit and implicit approaches, with the aim of sensitizing students to the disadvantages of each of these. Including explicit solvent in a simulation offers the prospect of greater accuracy, but at considerable cost in CPU time; there are also limitations to the accuracy of the most commonly used water models, particularly since they do not include the induced polarization effects that are responsible for the well-known cooperativity of hydrogen bonding. The use of gas phase calculations is very common because of the lower cost, but students should be aware of the problems of such methods. This issue is raised again when we discuss future improvements in modeling methodology (see below).

We think it is important to connect the results of simulations to experiment, so we present a discussion of what kinds of intramolecular motions can be detected by x-ray crystallography (temperature factors), NMR (particularly order parameters and proton exchange), fluorescence depolarization, Raman spectroscopy, electric birefringence decay, and so on. Table 3.1 from McCammon and Harvey (6) shows the time scales of various kinds of intramolecular motions and is a natural point of departure for this discussion.

At the end of the course, we devote a lecture to our personal view of the current strengths and limitations of modeling methodology, important unsolved problems, and likely future directions for research in the field. There are several strengths of the discipline. One of these is the role that modeling plays in the refinement of structures being determined by crystallography and NMR. Another has been the general recognition that molecules are dynamic, not static, and that intramolecular motions play crucial roles in many biological processes. Modelers, along with other structural biologists, have developed a range of graphical representations that have increased the general level of understanding of macromolecular structure. Finally, and most important, we emphasize to our students that the "structures" they see in textbooks and in scientific publications are all models; the best of those models are very accurate, because they are based on substantial quantities of data, but they are models none the less.

With regard to current problems and future research directions, we believe that the most important challenge facing our discipline is the need to develop truly predictive methods. Because of current limitations on potential functions and the finite capabilities of computers, too many modeling studies are simply descriptive. In contrast with theoreticians (who are expected to produce testable hypotheses) and experimentalists (who are expected to test hypotheses), modelers too often carry out studies that neither produce nor test hypotheses. This situation is improving, however. Among the important specific issues are continued improvements in potential functions (we discuss continuum electrostatic methods in some detail), free energy methods, and mixed classical/quantum descriptions for enzyme-substrate simulations. We also think it is important to continue developing the capability of attacking larger systems and doing more extensive surveys of conformational space. This will involve both brute force methods (particularly massive parallelization) and more sophisticated reduced representations.

Reading material

One of the difficulties faced by someone teaching macromolecular modeling is the scarcity of appropriate literature. Neither of the two best-known books (6, 7) is really an introductory textbook. Although our own graduate students have found them to be useful learning aids within the environment of our research group, others have generally reported that neither book has been particularly helpful in the beginning. Ironically, one of the best books (8) is one that deals with small molecules and does not cover macromolecular simulations. The authors state in their introduction that "this is a 'how-to-do-it' book for people who want to use computers to simulate the behavior of atomic and molecular liquids," and they present a logical step-by-step introduction to molecular simulations. This would be a suitable text for a very thorough course, since it presents the computer algorithms in

considerable detail. It is somewhat too detailed for a general introductory course, but selected readings from it can be used.

There are a handful of articles from the research literature that nicely illustrate some important modeling concepts. An excellent brief introduction to potential functions is found in the appendix of a classic article on molecular dynamics (9); the more sophisticated student will benefit from the more extensive treatment given in the papers that describe CHARMm (10) and AMBER (11). Pearlman and Kollman (12) show conformational energy plots for both potential and free energies and discuss the issue of transferability of parameters.

We devote a substantial part of one lecture to the limitations of existing potential functions for the prediction of protein structures, using two classic articles on the misfolding issue (1, 2). In our view, the principle shortcoming of existing potential functions has to do with errors in the treatment of solvent effects, particularly electrostatics, and one of our own reviews (13) was intended to discuss this issue at a level appropriate for students in a course like this; there are other more recent reviews (14, 15). With regard to reduced representations, lattice models and reduced representations for proteins have been reviewed recently (5, 16), and we use other articles from our own research to introduce succinct models for nucleic acids (3, 4, 17–19). A recent review (20) covers the basic principles of free energy methods and recent advances in that area.

FINAL REMARKS

Given the growth of macromolecular modeling, it is important to develop pedagogical materials at several levels. Here we have discussed approaches for a very brief introduction (one to two lectures) and for a full introduction at the level of an advanced graduate course. We would be very interested in hearing from others who have taught such courses, both with regard to materials they have used and with regard to their experiences in teaching students with various backgrounds.

We are grateful to the students who have taken our courses and have provided thoughtful evaluations and suggestions.

This work was supported by grants from the National Institutes of Health (GM-34015) and the National Science Foundation (DMB-90-05767).

Received for publication and in final form 13 July 1992.

REFERENCES

1. Novotny, J., R. Bruccoleri, and M. Karplus. 1984. An analysis of incorrectly folded protein models: implications for structure predictions. *J. Mol. Biol.* 177:787–818.
2. Novotny, J., A. A. Rashin, and R. E. Bruccoleri. 1988. Criteria that discriminate between native proteins and incorrectly folded models. *Proteins Struct. Funct. Genet.* 4:19–30.
3. Tan, R. K.-Z., and S. C. Harvey. 1989. Molecular Mechanics Model of Supercoiled DNA. *J. Mol. Biol.* 205:573–591.
4. Malhotra, A., R. K.-Z. Tan, and S. C. Harvey. 1990. Prediction of the three-dimensional structure of *Escherichia coli* 30S ribosomal subunit: a molecular mechanics approach. *Proc. Natl. Acad. Sci. USA.* 87:1950–1954.
5. Levitt, M. 1991. Protein folding. *Curr. Opin. Struct. Biol.* 1:224–229.
6. McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge University Press, London.
7. Brooks, C. L., M. Karplus, and B. M. Pettitt. 1988. Proteins: a theoretical perspective of dynamics, structure, and thermodynamics. *Adv. Chem. Phys.* Vol. 71. I. Prigogine and S. Rice, editors. John Wiley, New York.
8. Alden, M. P., and D. J. Tildsley. 1987. Computer Simulation of Liquids. Clarendon Press, Oxford.
9. McCammon, J. A., P. G. Wolynes, and M. Karplus. 1979. Picosecond dynamics of tyrosine side chains in proteins. *Biochemistry.* 18:927–942.
10. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comp. Chem.* 4:187–217.
11. Weiner, P. K., and P. A. Kollman. 1981. AMBER: Assisted model building with energy refinement: A general program for modeling molecules and their interactions. *J. Comp. Chem.* 2:287–299.
12. Pearlman, D. A., and P. A. Kollman. 1991. Evaluating the assumptions underlying force field development and application using free energy conformational maps for nucleosides. *J. Am. Chem. Soc.* 113:7167–7177.
13. Harvey, S. C. 1989. Treatment of elecrostatic effects in macromolecular modeling. *Proteins Struct. Funct. Genet.* 5:78–92.
14. Bashford, D. 1991. Electrostatic effects in biological molecules. *Curr. Opin. Struct. Biol.* 1:175–184.
15. Moult, J. 1992. Electrostatics. *Curr. Opin. Struct. Biol.* 2:223–229.
16. Jernigan, R. L. 1992. Protein folds. *Curr. Opin. Struct. Biol.* 2:248–256.
17. Harvey, S. C., R. K.-Z. Tan, and A. Malhotra. 1990. Modeling of nucleic acids in the size range 100–10,000 nucleotides: DNA plasmids and ribosomal RNAs. *In* Structure and Methods. Vol. 3. DNA and RNA. R. H. Sarma and M. Sarma, editors. Adenine Press, Guilderland, NY. 283–295.
18. Tan, R. K.-Z., and S. C. Harvey. 1990. Succinct macromolecular models: application to supercoiled DNA. *In* Theoretical Chemistry and Molecular Biophysics. Vol. 1. DNA. D. L. Beveridge and R. Lavery, editors. Adenine Press, Guilderland, NY. 125–137.
19. Malhotra, A., R. K.-Z. Tan, and S. C. Harvey. 1991. Prediction of the three-dimensional structures of ribonucleic acids: from tRNA to 16S ribosomal RNA. *In* Molecular Dynamics: Applications in Molecular Biology. J. M. Goodfellow, editor. Macmillan Press, London. 107–136.
20. McCammon, J. A. 1991. Free energy from simulations. *Curr. Opin. Struct. Biol.* 1:196–200.