

A World Wide Web Based Textbook on Molecular Simulation

Peter T. Cummings^{1,2}, Hank D. Cochran^{2,1}, Juan J. dePablo³,
Denis J. Evans⁴, David A. Kofke⁵, Athanassios Z. Panagiotopoulos⁶,
Richard L. Rowley⁷

¹University of Tennessee / ²Oak Ridge National Laboratory /
³University of Wisconsin / ⁴Australian National University /
⁵State University of New York at Buffalo / ⁶University of Maryland /
⁷Brigham Young University

Introduction

In this paper, we describe an innovative approach to combining research and curriculum development for the field of chemical engineering. The methodology has the potential to define a new paradigm for instruction in rapidly-evolving fields such as molecular simulation, computational chemistry, biochemical engineering, and materials science. Our immediate aim has been to initiate a World Wide Web (WWW)-based "textbook" on molecular simulation, and to introduce it into the graduate and undergraduate chemical engineering curricula at our respective institutions. The textbook will have the additional role of a refereed electronic journal that elaborates on important new developments and applications as they appear in the research literature, presenting such work in a manner suitable for class instruction and for self-paced learning. Our broader goal is to see that molecular simulation and molecular concepts in general are finely woven into the undergraduate and graduate chemical engineering curricula nationwide. The development of the text is funded by a grant from the Comined Research and Curriculum Development program of the National Science Foundation. The effort to develop the web text is partially the outgrowth of the establishment a Molecular Modeling Task Force within CACHE (Computer Aids for Chemical Engineers). CACHE, Inc. is a not-for-profit organization whose purpose is to promote cooperation among universities, industry, and government in the development and distribution of computer-related educational materials for the chemical engineering profession. More about CACHE activities can be learned from their WWW site <http://www.cache.org/>.

Historically, the primary role of chemical engineering educators at the undergraduate level has been to teach students the fundamental bases (material and energy balances, transport processes, thermodynamics and reaction engineering) for design calculations that ultimately, as practicing chemical engineers, they will perform using design packages such as Aspen Tech's ASPEN and Simulation Sciences' PRO-II. Understanding the fundamental bases for these calculations leads to all-important insight into the limitations of the calculational and correlation techniques employed in these design packages. It can be argued that the widespread use and calculational power of design packages makes it possible for any reasonably intelligent person to design a chemical process beginning with the process chemistry. The hallmark of a well-educated

chemical engineer is discrimination in the use of such packages, understanding of their limitations of applicability and the ability to critically assess the rigor of the various methods and correlations used for a particular application. As preparation for industry, chemical engineering seniors almost universally are exposed to ASPEN or PRO-II and use these systems for their senior capstone design course. Indeed, CACHE Corporation is largely responsible for bringing automated process design packages into the undergraduate chemical engineering curriculum.

The exceptionally rapid growth of the industrial usage of molecular simulation and quantum mechanical methods implies that a similar educational process is required in the area of molecular simulation (which can be thought of as applied statistical mechanics) and computational chemistry (applied quantum mechanics). In order for chemical engineers to be astute users of commercial molecular simulation and computational chemistry software packages, at the undergraduate level they need to be educated in the fundamental principles underlying them, and the limitations and reliability of the of the techniques used. At the graduate level, more detailed instruction is appropriate: students should be able to modify existing codes and develop their own codes based on code elements provided to them, for application to their own area of research.

Molecular modeling methods can and will appear in undergraduate curricula in several ways: as segments of existing core undergraduate courses (for example, Thermodynamics; Kinetics; Transport; Design, in the latter case as molecular product design); within elective courses (*e.g.*, Biochemical Engineering; Materials Science); and as dedicated courses. To ensure that molecular modeling methods enter the undergraduate curriculum in a smooth and useful way, a textbook needs to be developed that can be used by non-experts in the field. However, in a field that develops as rapidly as molecular simulation, any textbook is out of date at publication. For example, Allen and Tildesley's text¹, often cited as the most comprehensive and pedagogical introduction to molecular simulation, was published in 1987. One notable omission from Allen and Tildesley's text is the GEMC simulation method, which first appeared in the published literature in the same year^{2,3}. It is only now beginning to appear in textbooks on molecular simulation⁴.

These considerations provide the motivation to develop the WWW-based textbook on molecular simulation. It is designed to meet the following constraints: It should cover the fundamentals as well as incorporate significant new developments much faster than printed textbooks can; it should be useful to both undergraduate students (by teaching and illustrating the principles of applied statistical mechanics, and providing the knowledge base necessary for discriminating use of commercial packages) and graduate students (by introducing the techniques at a level that makes it possible for them to apply molecular methods in their own research endeavors); it should be usable by instructors with limited backgrounds in molecular simulation (just as a good thermodynamics or reaction engineering text makes it possible for an instructor who is not an expert in those fields to be an effective teacher of those subjects); it should take advantage of all the possibilities engendered by distribution over the WWW; it should not be susceptible to obsolescence by advances in molecular simulation or in WWW technologies.

Design Goals for the WWW-Based Textbook on Molecular Simulation

We have established several design goals for authoring the WWW-based text:

1. Through extensive review by educators both within the project and external to it, the WWW-based textbook conforms to the highest standards of scholarship.
2. Using professional web designers at various stages of the project, the WWW-based textbook is designed to be visually appealing, making full use of the capabilities of the WWW (through integration of text, audio and video) to explain complex concepts, and is designed so that it will not be made obsolete by advances in Web standards and browser design.

3. Each of the universities involved is establishing a new graduate chemical engineering course which will use the WWW-based textbook as its primary text. The course will also be open to senior undergraduates as an elective.
4. For undergraduates using the text (as a supplementary text in undergraduate thermodynamics, applied physical chemistry and transport phenomena courses, or as the major text of an elective course), the primary goal is to educate the students in the principles, capabilities and limitations of molecular-based methods for the prediction of physical properties and phase equilibrium in systems of industrial interest. Lessons and assignments for undergraduates do not require knowledge of a programming language such as FORTRAN or C.
5. For graduate students using the text (as the text for the elective course described in goal 3, or as a supplementary text in various graduate courses), an additional goal (beyond educating the students in the principles, capabilities and limitations of molecular-based methods) will be to provide the necessary background for the students to apply molecular simulation in their own area of research. This will require that the students be able to understand working code elements downloaded from the textbook.
6. Instructors using the WWW-based textbook will be invited to submit outstanding projects completed by students as part of their courses for inclusion in the text as case studies.

Beyond the period of the NSF grant, we are committed to continual expansion and refinement of the WWW-based textbook. As significant new developments in molecular simulation appear in the traditional published literature, authors of such papers will be invited to contribute a more detailed version of their work (which includes code elements and details of analysis methods) to the WWW-based textbook (we stress that we do not intend to develop an electronic journal in the sense of one that publishes original research). Such contributions will be refereed by the principal investigators of this proposal, who will form the nucleus of an editorial board for the ongoing WWW-based textbook. Thus the WWW-based textbook will evolve into both an instructional text accessible to the neophyte and an electronic journal capturing the most important research developments in the field, and providing for all users of the text a bridge between the most fundamental aspects of molecular simulation and its most advanced applications. In order to encourage participation in the WWW-based textbook, each contribution will carry the by-line of the author(s), much like an encyclopedia.

Features of the WWW-Based Textbook on Molecular Simulation

A “WWW-based” book exists exclusively in electronic form, and is read via an electronic browser, such as Netscape or Microsoft Explorer. The text has been conceived, designed and executed to exploit the unique features offered by the WWW format. The electronic medium invites the use of novel tools that manipulate text, present graphics and audio, perform computations, and provide interactivity in ways that significantly enhance the learning experience. Such advantage can be realized only through careful thought to questions that have never before been encountered by the authors of a traditional textbook. Thus the “textbook authoring” endeavor in fact presents three intellectual challenges:

- consideration of how to exploit the WWW-format to meet our ends
- development or collection of supplemental “textbook tools” that do the actual work of exploiting the Web-based features
- organizing and writing the text

A properly-designed and executed Web-based text will have value that not only endures, but grows. Indeed, our aim is to use the NSF support to *initiate* rather than *complete* this text, because we expect that its content will expand and evolve well beyond the term of this proposed work. Growth is expected on two fronts. First, as the field of molecular simulation continues its high-paced development, advances will be rapidly introduced into the text; significantly, these advances will be incorporated in a form that does not narrowly cater to the expert simulationist. Second, as concepts and techniques associated with molecular simulation become widely adopted in the chemical engineering curriculum, student and faculty users will contribute “case studies” (subject to quality review) that represent a novel, research-oriented application of molecular simulation to an engineering or scientific problem. Both types of contribution can serve the interests of the practicing engineer, who would consult the text to understand and solve routine physical-property problems while on the job (having been educated in a department that instituted the proposed curricular advance). The “user/contributor” universe for the text can be summarized with a figure

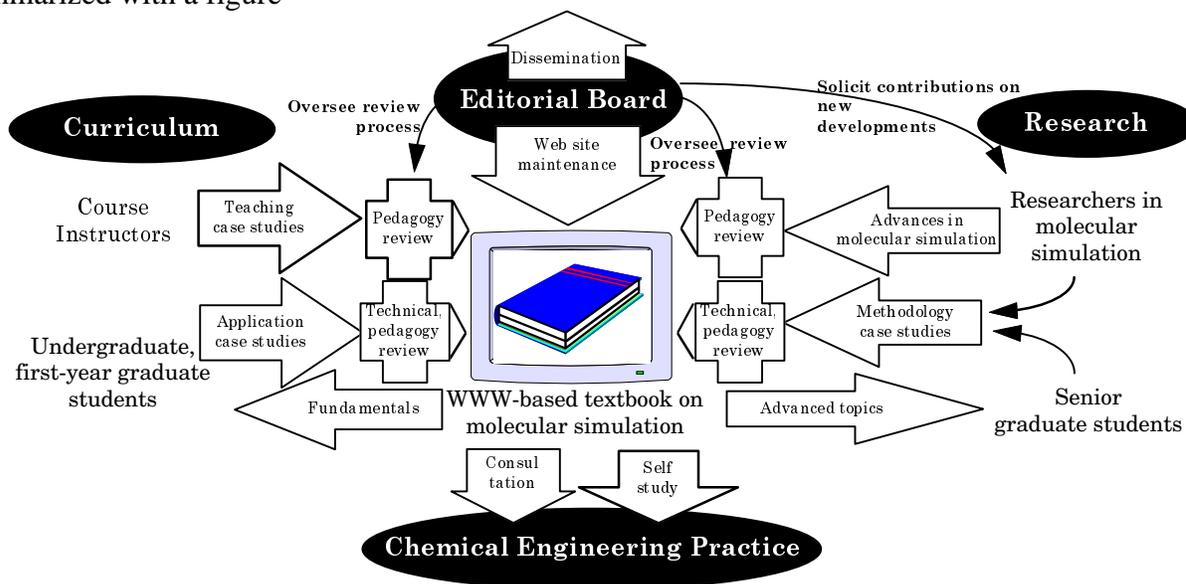


Figure 1. The user/contributor universe of the WWW-based textbook on molecular simulation

It should be evident from this description that the WWW-based text would represent a dynamic, perpetual interface between the world of research in molecular modeling and simulation on the one hand, and the undergraduate and graduate curricula in chemical engineering on the other.

Novel features of a Web-based book

The Web-based nature of the book invites a vast array of possibilities that enhance its design and use. The single most important feature has already been mentioned, namely the ability to update the text rapidly (which is not to be confused with casually) to incorporate research advances and curricular developments.

The second most important feature arises from our ability to process the text in real time just before presenting it to the reader. This capability is very important, because it greatly enhances

the versatility of the text, and because it precludes the possibility that the book would be rendered obsolete by evolving WWW standards or improvements in Web-browser design. A figure is helpful in explaining the concept.

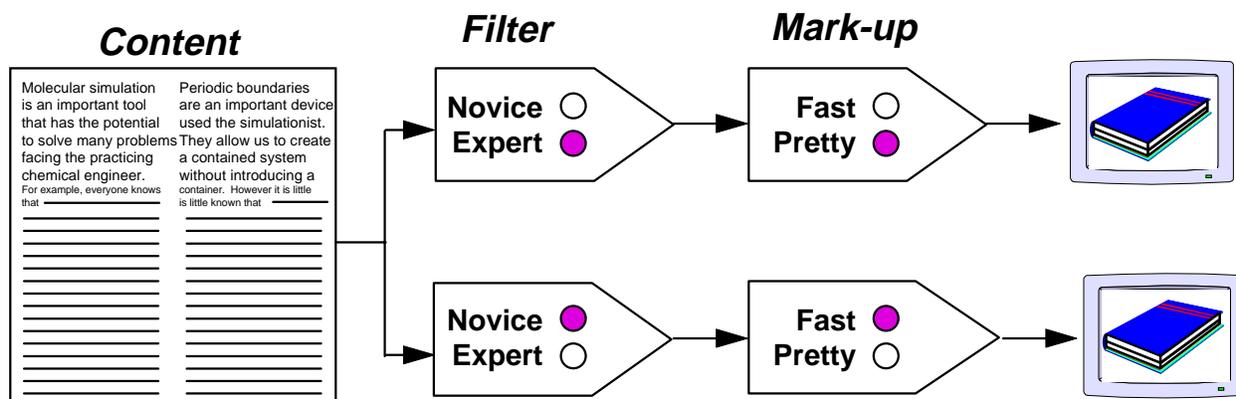


Figure 2. The user/text interface

The content of the book resides in a structure that is well indexed and content-coded. Such codes indicate primarily the level of presentation of the material (*e.g.*, elementary or advanced), but also other features, such as whether it relies on programming knowledge, and in which language. The key element of the interface is a configurable filter that processes the text before presenting it to the reader. The filter is responsible for screening the content and structuring the text and graphics for presentation via the Web browser. The reader (or his classroom instructor) have the option to indicate his or her level of expertise in molecular simulation, along with a few other details regarding level of interest in programming details, the speed of the connection being used (which would cause large graphics to be minimized for slow connections), *etc.* The material that would be presented to the reader would then be processed to tailor it to his or her ability and interests.

Other important features are made possible in this paradigm. On the curricular side, it becomes possible for a course instructor to tailor the book to his or her needs, even to the point of setting up a specialized table of contents for the course. On the “publishing” side, the proposed arrangement ensures a consistency in the formatting of the text, and allows for it to be easily modified. In the same vein, changes in HTML standards (which govern the “language” read by the Web browsers) can be easily accommodated, and improvements in Web-browsers designs can be rapidly exploited. All of this is possible because the text of the book is not written with all of the HTML “mark-up” included; rather these formatting directives are inserted by the text filter before sending the text to the browser. On the “authoring” side, we then have the enormous benefit of not requiring that our contributors provide text that has been written and extensively formatted with HTML mark-up (or equally, that we do not require a dedicated staff person to do extensive clean-up and formatting of new contributions).

The third important feature of the Web format arises with the advent of Java, the revolutionary programming language for the Web. Java permits authors to write programs that actually run on the user’s machine, and thereby permits (for example) the presentation of an interactive, graphical, working molecular simulation program, without requiring any programming knowledge of the reader. “Conceptual bugs” can be programmed into some simulations to show convincingly why simulations must be conducted as they are. Students can use (correctly

programmed) simulators to examine the behavior of model systems, and compare them to experimental data from the literature; we expect that the outcome of some such studies would be suitable for “publishing” in the book, making for a meaningful undergraduate research experience. Many other pedagogical opportunities accompany the use of Java. Plotted graphs can be made dynamic, with animation, or the capability to expand regions for more detail, or even the ability to select different views of a three-dimensional surface, all at the direction of the reader. Quizzes and self-study questions can be made interactive, with hints provided as the reader needs them. The only limits are the author’s imagination.

We have discussed extensively the three most important advantages that a Web-based text on molecular simulation has over a printed one. There remain many additional capabilities that are unique to the Web-based format. Examples include:

- The “hypertext” capacity of web browsers invites extensive cross-referencing, and frees us from the traditionally linear or hierarchical book structure. Literature citations or topics alluded to elsewhere in the document may be easily accessed by the reader. In addition, cross-references can extend beyond the text itself—hypertext links may be inserted to relevant places on the WWW (indeed, as more journals go “on-line”, citations may bring up the actual article). Search capabilities can be used to supplement (or replace) the traditional indexing system.
- Feedback on the text can be solicited from its users (faculty and students) and promptly acted upon, allowing for continual improvement.
- The text can make significant use of graphics, animation and audio. Thus we can present “movies” of molecular dynamics simulations to convey an idea of how they work, or how a system behaves. A difficult point can be understood by downloading a small movie of an expert in the field describing the concept.
- Large amounts of data (including program listings) can be made available for downloading and analysis by the reader.

Content

The exciting features offered by the Web-based format should not obscure the most important component of the text: its content. The actual technical content of the book—the breadth and depth of the topics it covers, and its effectiveness in conveying them to a diverse readership—is the paramount concern. It is useful to conceive the structure and content of the book as spanning three dimensions. In one direction we measure the breadth of content, the range of categories that one might cover while discussing the field of molecular simulation (*e.g.*, intermolecular potentials; molecular dynamics simulation; evaluation of phase equilibria). In another direction, we measure the depth to which a topic is discussed (*e.g.*, within the topic of intermolecular potentials, we could begin with purely repulsive forms, then attractive, then fixed electrostatic, then polarizable, *etc*). The third direction measures degree of difficulty, in the sense of whether the content is suited for the novice or the expert (or somewhere in between); thus the expert would get an explanation of configurational-bias sampling methods (a “deep” topic) that is couched in the language of partition functions, sampling bias, and Fortran, while the novice might access an explanation of the same topic presented by analogy to the problem of seating a

group of ten people versus two at a restaurant. The point to realize is that depth is not synonymous with degree of difficulty, or complexity.

Our plan is ultimately to include all topics of practical relevance in molecular simulation, and to do so to a level of depth and complexity that would interest the most expert simulationist. At the same time, to satisfy the needs of the novice we must include lucid explanations of many of the same topics at a level that assumes little prior experience in simulation. Intuitive explanations of difficult concepts can serve a useful purpose to the student who plans to use rather than to program a molecular simulation. At all times, the student retains the ability to delve into a more complex explanation of a topic if the top-level presentation is not sufficiently thorough for his or her purposes (and the expert may move to a simpler level to see something familiar in a different light). This is a great advantage of a text that strives to meet the needs of readers from diverse backgrounds, and it is made feasible through the Web-based features described above.

Present Status and Release Date

Some elements of the text are in place, although the overall look-and-feel is still evolving. To give an idea of the present status of the interface, the Figures 3 and 4 show snapshots from sessions with the text. Figure 3 shows a picture of the text itself, in a Netscape window. Navigation buttons are present on the left-hand side, while the content is shown in a panel on the right. A large amount of HTML “Table” formatting was used to present the text in the column format shown, but all of this is transparent to the author, who provides only the actual text and pointers to figures and other “sidebar” objects. Handling of equations is not yet in place; we are awaiting further developments in HTML and browser standards rather than attempting ourselves to develop a patch for this problem. Citations are handled by the author rather easily, requiring only a simple mark-up in their text to flag the citation. Formatting and other things needed to create a link to the citation are introduced by the filters. We also plan automatic generation and linking to a glossary.

Clicking on the “Configure” button in the navigation panel brings up the page for configuring the text to the user’s background and interest. Here the reader indicates their level of expertise, desires for the presentation of informative side items (which some users may instead find distracting), the language for presentation of computer code, and so on.



Figure 3. Snapshot from the text.

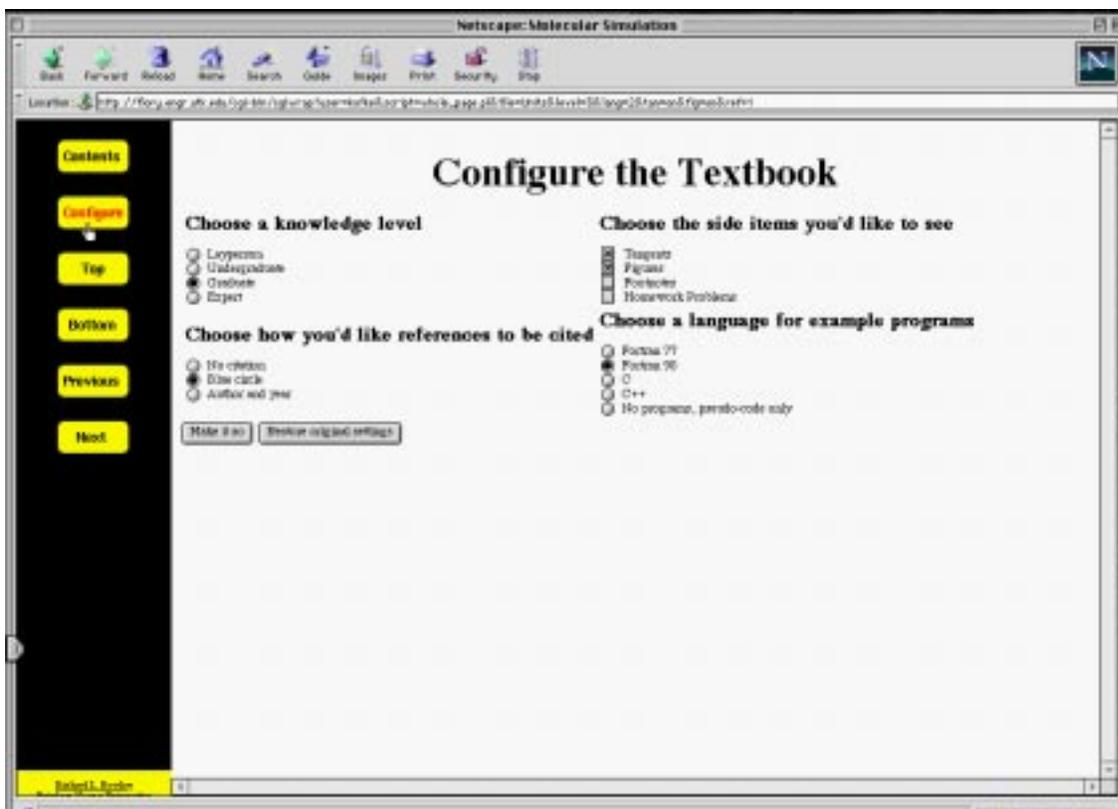


Figure 4. Snapshot of the text configuration screen.

The initial version of the WWW-based text on molecular simulation will be released for public access on July 1, 1998. One URL for accessing the text is <http://flory.engr.utk.edu/~kofke>; by July 1, a more intuitive domain name will be registered for the text and will be accessible from this site.

Bibliographical Information

1. Allen, M. P. and Tildesley, D. J., 1987, *Computer Simulation of Liquids*. Oxford: Oxford University Press.
2. Panagiotopoulos, A. Z., 1987a, Adsorption and capillary condensation of fluids in cylindrical pores by Monte Carlo simulation in the Gibbs ensemble. *Molec. Phys.*, **62**, 701.
3. Panagiotopoulos, A. Z., 1987b, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Molec. Phys.*, **61**, 813.
4. Frenkel, D. and Smit, B., 1996, *Understanding Molecular Simulations: From Algorithms to Applications*, Academic Press.

Biographical Information

PETER T. CUMMINGS holds a joint appointment at Distinguished Professor of Chemical Engineering at the University of Tennessee and Distinguished Scientist in the Chemical Technology Division of Oak Ridge National Laboratory. He is at the forefront of the study of aqueous phases by molecular simulation and the development and application of non-equilibrium molecular dynamics techniques to engineering problems.

HANK D. Cochran is a Group Leader in the Chemical Technology Division of Oak Ridge National Laboratory and an adjunct faculty member in the Department of Chemical Engineering at the University of Tennessee. He has served on numerous ABET curriculum review teams. He has performed extensive research on aqueous supercritical phases, and is now active in the molecular simulation studies of rheological properties of alkanes.

JUAN J. DE PABLO is professor of Chemical Engineering at the University of Wisconsin. He is one of the originators of the configurational-bias Monte Carlo method, which is the premier technique for study of chain molecules by molecular simulation. He is also very active in molecular dynamics studies of complex systems, and in the evaluation of phase equilibria by molecular simulation.

DENIS J. EVANS is professor and dean of the Research School of Chemistry at the Australian National University in Canberra, Australia. He is one of the founders of the field of nonequilibrium molecular dynamics and has published widely on this subject, including prime authorship of a highly-regarded research monograph on molecular simulation, *Statistical Mechanics of Nonequilibrium Liquids*.

DAVID A. KOFKE is professor of Chemical Engineering at the State University of New York at Buffalo. He introduced methods that are widely used for the simulation of mixtures, and is the inventor of the Gibbs-Duhem integration technique for evaluation of phase equilibria; with this method he has led the study of phase equilibria involving solids.

ATHANASSIOS Z. PANAGIOTOPOULOS holds joint professorship in the Department of Chemical Engineering and in the Institute for Physical Sciences and Technology at the University of Maryland. He is the inventor of the Gibbs ensemble, a technique that revolutionized the study of fluid-phase equilibria by molecular simulation; he is now directing the development of intermolecular potentials that are useful in engineering applications.

RICHARD L. ROWLEY is J. J. Christensen Professor of Chemical Engineering at Brigham Young University. He is accomplished not only in statistical mechanics and molecular simulation, but also in their translation to the

classroom. He is the author of *Statistical Mechanics for Thermophysical Property Calculations*, a new textbook that is being rapidly adopted nationwide for graduate instruction in applied statistical mechanics.