Polymer Modeling: Where Has it Been and Where is it Going?

David Rigby,[1] B.E. Eichinger[2]*

1. Accelrys, Inc., 9685 Scranton Road, San Diego, CA 92121-3752, USA
2. Department of Chemistry, University of Washington, Seattle, WA 98195, USA
Email: david@accelrys.com , eichinge@chem.washington.edu

Summary: The development of the area of polymer modeling often referred to as molecular modeling has been reviewed from its early beginnings to the present day. Key forces influencing the development include computational power, algorithmic advances and access to computational resources. The desire to apply modeling techniques to predict the properties of increasingly complex polymer-containing systems, taken in conjunction with a number of current limitations discussed in this brief review, is expected to define in part some essential future developments.

Keywords: equation of state, force field, molecular modeling, polymer conformation, simulations

Introduction

The term *Polymer Modeling* has been used to describe a number of techniques whose aim is to describe or predict experimentally measurable properties of polymers in a wide variety of situations including dilute and concentrated solutions, bulk systems, crosslinked materials, chains adsorbed at interfaces or restricted to confined geometries, and polymers in liquid crystalline or other types of mesophase. Typical methods used for modeling these systems range from quantum mechanical and force field based simulations, through various types of coarse-grained approaches, to semi-empirical correlation-based methods and pure thermodynamic methods of varying degrees of sophistication. For the purpose of the present review, based on work presented at the IUPAC-PC2002 conference on the Mission and Challenges of Polymer Science and Technology held in Kyoto Japan, we restrict our focus mostly to the area of polymer modeling often referred to as molecular modeling, involving construction, usually using a computer, of atomic-scale models representing the geometry, spatial configuration, topology and energetics of systems containing chain molecules. The closely related topic of mesoscale modeling, which directly or indirectly probes behavior associated with longer length scale aspects of structure, includes studies of morphology and

 DOI: 10.1002/masy.200351105

evolution of morphology in systems undergoing microphase separation, and studies of the various topological features found in elastomeric network forming systems and their consequences. Mesoscale modeling is discussed at length elsewhere in this volume and accordingly is not covered further here.

Early Polymer Models

The use of molecular models in polymer science emerged naturally from the early 1920s work of Staudinger and others aimed at establishing the validity of the macromolecular hypothesis, which is described in some detail in the historical review by Morawetz.[1] Following the general acceptance that high molecular compounds such as caoutchouc were indeed composed of long sequences of covalently bonded repeating units, the stage was set for a whole era of scientific enquiry into the spatial configurations of these flexible macromolecules. Thus, for example, we begin to see the development of expressions for describing characteristic spatial properties such as the moments of the end-to-end distance distribution and the radius of gyration, treated initially within the framework of the statistics of the random walk, which had been discussed much earlier by Pearson[2] and treated mathematically by Lord Rayleigh.[3] In an application to rubber elasticity, Treloar derived the exact expression for the probability distribution for the end-to-end distance of walks of any length,[4] which rapidly approaches the Gaussian form, and later constructed some of the first physical models of random walk chains using wooden sticks, wire and glue.[5]

Contemporaneous studies of chain molecules of different types, including natural polymers starch, cellulose and proteins together with a range of commercial synthetic polymers contributed to rapid increases in sophistication of what initially were mostly mathematical models. Theoretical studies of these models quickly established that any flexible chain molecule in which the range of interactions is local – as represented by chemical bond lengths, valence angles, and chain backbone torsional angles (possibly restricted by near-neighbour steric interactions along the backbone) – would exhibit universal behavior, for which the mean-squared end-to-end distance would always be given by the expression $\langle r^2_o \rangle = C_n n l^2$, where n denotes the number of backbone bonds of length l, and where the characteristic ratio C_n becomes a constant for chains of sufficient length. Although developed based on the treatment of isolated chain molecules, it was proposed by Flory that the same overall statistics would be followed by multichain systems in two situations, namely polymers in dilute solutions of certain

solvents (*theta* solvents, for which the second virial coefficient vanishes) and chains in the melt, a postulate which has subsequently been confirmed by experiment on many occasions. In other situations, chain molecules experience net non-local interactions (sometimes ambiguously termed long-range interactions) leading to mean-squared dimensions scaling with a power of n greater than unity (approximately 1.18 in good solvents). A description of the development of theories and related experimental work aimed at quantifying behavior of perturbed chains can be found in Yamakawa's comprehensive 1971 monograph.[6]

Further development of actual physical models continued into the 1950s with noteworthy contributions from Corey and Pauling[7] in their development of space filling models, made from hardwood and plastic, for visualizing conformations of proteins. In the early 1960s, these models evolved into the well-known CPK molecular models following the redesign work of Koltun initiated at the U.S. National Institutes of Health Biophysical Laboratory. While such models, and their modern-day computer-generated counterparts, continue to be useful for gaining insight into various aspects of polymer conformation, especially in the biological field, the overwhelming number of conformations available to flexible chains inevitably necessitates the introduction of the machinery of statistical mechanics. Thus, given some model of a polymer chain – with specified bond lengths, valence angles, etc. – combined with a prescription for calculating energies of different conformers and for quantifying nonbonded interactions between remote segments of a chain, or perhaps with another chain or small molecule, it is straightforward to write expressions for the partition function and associated averages of properties over phase space. However until the advent of electronic computers, evaluation of such averages was not routinely feasible.

Beginnings of Computational Polymer Modeling

As already implied, much of the framework on which polymer modeling would eventually be based was established prior to the development of the first electronic computers, though numerical applications had yet to emerge. This situation began to change during the 1950s as a result of a number of important developments. First, increasing access to centralized computational facilities, coupled with standardization of programming languages, for the first time made it possible to undertake tasks such as Monte Carlo sampling or complete enumeration of all possible conformations accessible to simple self-avoiding lattice chain models[8] (of relevance to the so-called excluded volume problem), and to contemplate

evaluation of the integrals over phase space required to compute the properties of systems containing polymers. Secondly, this era witnessed the emergence of a number of simple and efficient algorithms destined to find widespread use in simulations, together with a comprehensive and rigorous formalism for treating the configurational statistics of unperturbed chains, as described briefly below.

One of the most significant algorithmic contributions to studies of equilibrium behavior is found in the so-called Metropolis Monte Carlo algorithm,[9] which for the first time made it possible to sample configurations of a model system in an efficient but unbiased manner, such that the probabilities of the resulting configurations would conform to canonical ensemble statistics. This algorithm was widely applied in polymer modeling, both in studying basic polymer models (e.g. lattice-based chains), and also in simulations of chemically more realistic models of chains either in isolation or in condensed phases.

The complementary area of molecular dynamics simulation, in which Newton's laws are solved iteratively for the atoms (or other subunits) comprising model polymers yielding information on both equilibrium and dynamical behavior, also benefited from key early algorithmic developments. Some of the most important include the simple integration algorithm introduced by Verlet,[10] and various prescriptions for controlling the temperature and pressure in simulations of condensed phases, as proposed by Berendsen,[11] Nose,[12] Hoover[13] and Parrinello and Rahman.[14] A summary of these algorithms may be found in the now-classical text of Allen and Tildesley.[15]

Another significant development of this early era was the establishment of the rotational isomeric state (RIS) scheme, which replaces the integration over all chain backbone dihedrals contained in the configurational partition function by a discrete sum over preferred, and optionally energetically-weighted states. The basic formalism for computing the second moment of the end-to-end distance distribution of a polymer chain was treated by a number of workers at the end of the 1950s, and summarized in the works of Volkenstein[16] and of Birshtein and Ptitsyn.[17] Shortly thereafter the application of the technique was greatly expanded by Flory and coworkers, beginning with the treatment of polymethylene chains subject to (realistic) interdependent torsional potentials by Abe et al.,[18] with subsequent extension to a wide variety of polymers and extensions to other properties as described in a

number of volumes including Flory's 1969 monograph[19] and subsequent work of Mattice and Suter.[20]

Other important polymer-specific algorithms which emerged during these early years include pivot algorithms as employed by Lal[21] and by Stellman and Gans[22] for implementing Monte Carlo moves aimed at rapid sampling of conformation space, as used subsequently in the RIS Metropolis Monte Carlo method.[23] Also at this time, we see the emergence of conformational sampling methods which change the coordinates of sequences of 3 or 4 bonds, leaving the rest of the chain fixed,[24] which were used for later modeling of condensed phases.

The Modern Era

The past twenty five years or so has seen a rapid acceleration in the amount of research activity involving molecular modeling of polymers, and as previously this development is linked to advances in the power and availability of computational resources. Thus at the beginning of this period, widespread access to powerful supercomputers from vendors such as CDC, Cray, Fujitsu and IBM led to comprehensive studies of hitherto inaccessible phenomena, increasingly focusing on dense polymer systems in contrast to the isolated chain models studied in the earlier work. A few examples taken from the 1991 book *Computer Simulations of Polymers*[25] include: PVT relations of melts and glassy polymers, dynamical motions of long chains in dense melts, diffusion studies of light gases in dense systems, polymer melting and crystallization, and mechanical behavior of amorphous polymers. By the middle of this period – around 1988-1992 – powerful UNIX workstations affordable by most laboratories began to shoulder much of the computational burden formerly offloaded to the supercomputers, in addition to providing a platform for high quality graphical systems, often empasized as part of commercial modeling software systems. During the past few years, fast microprocessors have found their way into desktop and laptop computers, and into high end massively parallel computers with phenomenal processing power.

The supercomputer and workstation era also saw the development of further ingenious Monte Carlo move and condensed phase model building algorithms. A number of noteworthy examples include the introduction of the slithering snake (sometimes called reptation) algorithm,[26] as well as a number of end and internal bridging algorithms,[27] and chain regrowth algorithms.[28] Other methods well suited to rapid equilibration of dense systems

include applications of the bond fluctuation method,[29] and use of high coordination lattices.[30] It should be remarked that these methods have led to an impressive increase in the range of systems and problems amenable to modeling, though there are frequently limitations in application to arbitrary chain topology and chemistry.

Force Field Based Simulations

From a survey of recent advances in polymer modeling,[31] as well as in a subsequent comprehensive review of the field,[32] it is apparent that while progress continues to be made in the area of algorithms and understanding the physics of high polymer systems, there is also an increasing emphasis on modeling the behavior of chemically diverse polymer systems and complex situations encountered in modern materials (mixtures, interfaces, confined systems, etc.). One of the essential requirements for such studies to be ultimately successful is the availability of accurate force fields. In this regard, the last two decades have witnessed

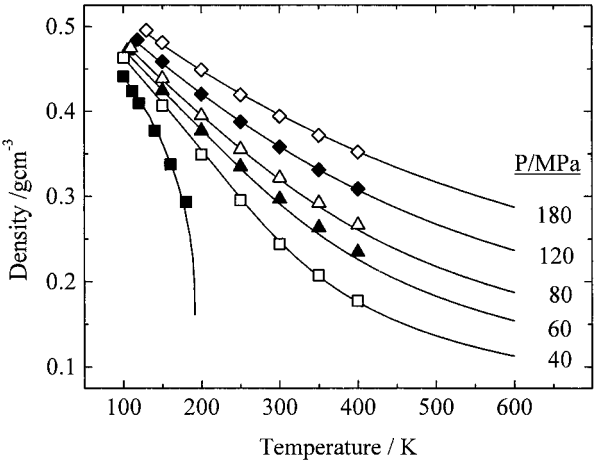

Fig. 1. Density of methane calculated from constant pressure molecular dynamics simulations using the COMPASS force field.[33,34] 100ps dynamics simulations used for each point. Solid curves denote experimental data.

significant advances in the sophistication and accuracy of force fields for modeling organic material, originating mostly from simulations of macromolecules conducted in the life sciences. In recent years, the methodologies developed for deriving force field parameters for these biologically-oriented force fields, usually involving high level ab initio quantum mechanical calculations, have been applied to polymers and small molecule organics (with

careful attention also given to parameterizing intermolecular nonbonded interactions using non quantum based methods). An example of what can now be achieved with small molecules is shown in Figure 1.

Here it is observed that PVT bevaviour can be accurately predicted over broad ranges of temperature and pressure. In some of our earlier work, we have also performed similar calculations on melts of systems containing oligomers and high polymer, where similarly good agreement with experiment can be obtained (with extrapolation of oligomer data to infinite molecular weight when studying high molecular weight polymer). A typical example is shown in Figure 2.

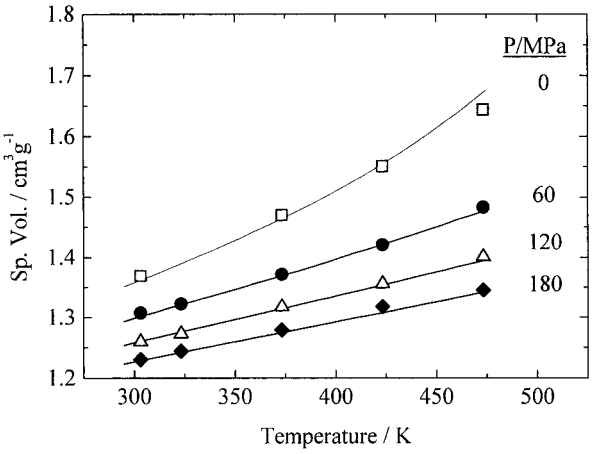

Fig. 2. Specific volume-temperature curves for $C_{11}H_{24}$ obtained as in Figure 1.

Other behavior which must be accurately predictable if modeling is ever to be used to make quantitative calculations of *excess* mixture properties such as the enthalpy of mixing includes cohesive properties. Here also, it is found that heats of vaporization or, alternatively, solubility parameters, of organic solvents can be accurately predicted when a carefully parameterized force field is used. Moreover, although the range of experimentally-reported solubility parameters for polymers is often quite broad, there are good indications that molecular modeling can also be used to make reliable and accurate predictions of this important quantity. As an example, Figure 3 illustrates the method of estimating the solubility parameter at 298K for high molar mass poly(ethylene oxide), which was shown in a previous publication to be in excellent agreement with experiment.[35]

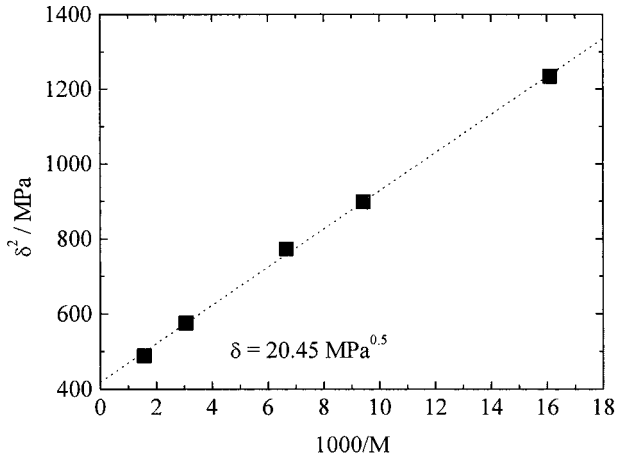

Fig. 3. Calculation of solubility parameter δ of high molar mass poly(ethylene oxide) by extrapolation of oligomer data vs. the reciprocal of the molar mass M.

Concluding Remarks

Polymer modeling has evolved dramatically since the advent of the first electronic computers, with early applications using simple, often lattice-based, models of isolated chains to investigate the factors which govern the spatial configuration of chain molecules upon which many important phenomena such as solution viscosity depend. Widespread access to supercomputers later effected a shift in focus of polymer modeling activity to a broader range of properties, mostly pertaining to bulk melt, glassy and crystalline systems. This trend has continued through the era of the laboratory graphical workstation and most recently into a new era of desktop computers and large scale parallel systems, while the breadth of applications continues to expand, ranging from studies of fundamentals of high polymer physics, through property prediction for bulk polymer and solvent, to studies of chemical phenomena (such as degradation), and increasingly involving studies of complex systems encountered in materials research.[36] Parallel developments in the field involving longer length scale simulations have fueled an interest in generating the molecular level properties required as input to these mesoscale simulations using data obtained purely from atomistic simulation.

Although it is difficult to speculate on the future direction of polymer modeling, it is nonetheless possible to identify some current limitations which in principle should drive future developments.

In the little-explored area of polymer chemistry there are a few needs:

1. Improved understanding of polymer degradation, perhaps involving reactive force fields.
2. More extensive, quantum mechanics based, studies of polymerization behavior, including ring formation, transesterification, etc.

In the area of force field based simulations, some limitations are as follows:

1. The accuracy of current force fields is impressive. However work is still required to establish whether the current level of accuracy is adequate for direct prediction of excess thermodynamic properties of mixtures, for example.
2. Alternative (e.g. Gibbs-ensemble) methods for studying phase behavior[37], while successful, are hampered by limited coverage of the united atom force fields currently used, while usage of widely available and more extensively parameterized all atom force fields is limited by low efficiency of the associated simulation algorithms.
3. Force field-based methods for accurately and routinely predicting properties of glassy polymers (e.g. volume-temperature behavior below and around T_g) are still in need of further development.
4. Although promising work involving use of atomistic methods to study aspects of viscoelastic behavior has recently been reported[38], this has so far been rather limited.
5. Methods for prediction of diffusivities and solubility of molecules of moderate size (e.g. of the size of drugs with 30-50 heavy atoms) in amorphous polymer have yet to be perfected.

In the area of long length scale simulations, perhaps the weakest link currently is the difficulty in obtaining accurate values of the necessary input parameters, especially thermodynamic interaction parameters, directly from simulations. In this regard, investigations of the sensitivity of the characteristic morphologies and their rate of development to variation/errors in input parameters would also be of interest.

[1] H. Morawetz, *Polymers. The Origins and Growth of a Science*, J. Wiley & Sons, New York 1985, Ch 10.
[2] K. Pearson, *Nature* **1905**, 77, 294.
[3] Lord Rayleigh, *Phil. Mag.* **1919**, 37, 321.
[4] L. R. G. Treloar, *Trans. Faraday Soc.* **1946**, 42, 77.
[5] see L. R. G. Treloar, *The Physics of Rubber Elasticity*, 3rd ed., Oxford Univ. Press 1975.
[6] H. Yamakawa, *Modern Theory of Polymer Solutions*, Harper & Row, New York 1971.
[7] R. B. Corey, L. Pauling, *Rev. Sci. Instr.* **1953**, 24, 621.
[8] see, for example, F. T. Wall, J. J. Erpenbeck, *J. Chem. Phys.* **1959**, 30, 634; M. E. Fisher, B. J. Hiley, *J. Chem. Phys.* **1961**, 34, 1253.
[9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **1953**, 21, 1087.
[10] L. Verlet, *Phys. Rev.* **1967**, 159, 98.
[11] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, 81, 3684.
[12] S. Nose, *J. Chem. Phys.* **1984**, 81, 511.
[13] W. G. Hoover, *Phys. Rev. A.* **1985**, 31, 1695.
[14] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, 52, 7182.
[15] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*. Oxford Univ. Press 1987.
[16] M. V. Volkenstein, *Configurational Statistics of Polymeric Chains*, Interscience, New York 1963.
[17] T. M. Birshtein, O. B. Ptitsyn, *Conformations of Macromolecules*, transl. S. N. Timasheff and M. J. Timasheff, Interscience, New York 1966.
[18] A. Abe, R. L. Jernigan, P. J. Flory, *J. Am. Chem. Soc.* **1966**, 88, 631.
[19] P. J. Flory, *Statistical Mechanics of Chain Molecules*, Interscience, New York 1969.
[20] W. L. Mattice, U. W. Suter, *Conformational Theory of Large Molecules*, Interscience, New York 1994.
[21] M. Lal, *Mol. Phys.*, **1969**, 17, 57.
[22] S. D. Stellman, P. J. Gans, *Macromolecules* **1972**, 5, 516.
[23] J. D. Honeycutt, *Comp. Theor. Polym. Sci.* **1998**, 8, 1.
[24] A. T. Clark , M. Lal, *British Polymer Journal* **1977**, 9, 92.
[25] R. J. Roe (Ed.), *Computer Simulations of Polymers*, Prentice-Hall, New Jersey 1991.
[26] M. Vacatello, G. Avitabile, P. Corradini, A. Tuzi, *J. Chem. Phys.* **1980**, 73, 548.
[27] P. V. K. Pant, D. N. Theodorou, *Macromolecules* **1995**, 28, 7224.
[28] J. J. de Pablo, M. Laso, U. W. Suter, *J. Chem. Phys.* **1992**, 96, 2395.
[29] I. Carmesin, K. Kremer, *Macromolecules* **1988**, 21, 2819.
[30] P. Doruker, W. L. Mattice, *Macromol. Theor. Sim.* **1999**, 8, 463.
[31] B. E. Eichinger, D. Rigby, *Current Opinion in Solid State and Materials Science* **2001**, 5, 445.
[32] B. E. Eichinger, R. Khare, *Molecular Modeling* in: *Encyclopedia of Polymer Science and Technology*, Interscience, New York 2003.
[33] H. Sun, *J. Phys. Chem. B* **1998**, 102, 7338.
[34] D. Rigby, revised COMPASS parameters for methane, unpublished.
[35] D. Rigby, H. Sun, B. E. Eichinger, *Polymer International* **1997**, 44, 311.
[36] see, for example, R. Khare, J. J. de Pablo, A. Yethiraj, *Macromolecules* **1996**, 29, 7910.
[37] S. K. Nath, F. A. Escobedo, J. J. de Pablo, *J. Chem. Phys.* **1998**, 108, 9905.
[38] O. Byutner, G. D. Smith, *Macromolecules* **2002**, 35, 3769.