

Editorial

## Molecular Modeling of Complex Chemical Systems

Donald G. Truhlar

*J. Am. Chem. Soc.*, **2008**, 130 (50), 16824-16827 • DOI: 10.1021/ja808927h • Publication Date (Web): 10 December 2008

Downloaded from <http://pubs.acs.org> on February 8, 2009

### More About This Article

---

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)

### Molecular Modeling of Complex Chemical Systems

This third *JACS* Select issue is devoted to molecular modeling of complex systems, and it showcases 23 Articles and Communications<sup>1–23</sup> that illustrate the breadth and power of computational chemistry as applied to such systems. Computational chemistry has developed into an important tool in almost all areas of chemistry. Although one of the first applications of quantum mechanics was to chemical bonding,<sup>24</sup> computational chemistry had a long growth period until theoretical methods and practical algorithms advanced to their present stage of usefulness for our field. In the early days only very basic systems could be tackled, and it took decades for useful results to emerge even on most simple systems.

Almost all modern theoretical chemistry is computational chemistry, because most of the progress that can be made with pencil and paper without a computer has been already made. Computations on complex systems are, in my opinion, the current frontier of theoretical chemistry. Let's look at progress. Pre-World War II theory developed the framework for modern physical chemistry in the works of London, Debye, Wigner, Eyring, Slater, Pauling, Wilson, Mulliken, and others. Pre-World War II computations "solved" the helium atom and the hydrogen molecule. Modern computational chemistry can probably be considered to have started in the 1960s<sup>25</sup> (based on theoretical groundwork laid down earlier). A classic publication of that era was the calculation of the barrier to internal rotation in ethane,<sup>26</sup> a complex problem for its day, and the 1960s also saw the first application of classical trajectory methods to calculate the equation of state of a fluid,<sup>27</sup> albeit a hard-sphere fluid, and the first Monte Carlo classical trajectory calculations of a chemical reaction.<sup>28</sup> The 1970s saw the first reliable calculation of a chemical reaction barrier height<sup>29</sup> and rate constant,<sup>30</sup> and condensed-phase simulations were extended to a dipeptide in aqueous solution.<sup>31</sup> The 1980s saw the publication of a general docking program for drug design,<sup>32</sup> the first systematic deployment (including a molecule as complex as methylamine) of the quantitative first generation of electronic structure wave function model chemistries that are still in wide use today,<sup>33</sup> the first converged quantum dynamics

- 
- (1) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 9296.
  - (2) Ess, D. H.; Houk, K. N. *J. Am. Chem. Soc.* **2007**, *129*, 10646.
  - (3) Hagberg, D.; Bednarz, E.; Edelstein, N. M.; Gagliardi, L. *J. Am. Chem. Soc.* **2007**, *129*, 14136.
  - (4) Shi, F.-Q.; Li, X.; Xia, Y.; Zhang, L.; Yu, Z.-X. *J. Am. Chem. Soc.* **2007**, *129*, 15503.
  - (5) Gokhale, A. A.; Dumesic, J. A.; Mavrikakis, M. *J. Am. Chem. Soc.* **2008**, *130*, 1402.
  - (6) Stoica, I.; Sadiq, S. K.; Coveney, P. V. *J. Am. Chem. Soc.* **2008**, *130*, 2639.
  - (7) Difley, S.; Beljonne, D.; Van Voorhis, T. *J. Am. Chem. Soc.* **2008**, *130*, 3420.
  - (8) Sproviero, E. M.; Gascón, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *J. Am. Chem. Soc.* **2008**, *130*, 3428.
  - (9) Kan, E.-j.; Li, Z.; Yang, J.; Hou, J. G. *J. Am. Chem. Soc.* **2008**, *130*, 4224.
  - (10) Datta, A.; Hrovat, D. A.; Borden, W. T. *J. Am. Chem. Soc.* **2008**, *130*, 6684.
  - (11) Barbatti, M.; Lischka, H. *J. Am. Chem. Soc.* **2008**, *130*, 6831.
  - (12) Migani, A.; Blancafort, L.; Robb, M. A.; DeBellis, A. D. *J. Am. Chem. Soc.* **2008**, *130*, 6932.
  - (13) Pang, J.; Hay, S.; Scrutton, N. S.; Sutcliffe, M. J. *J. Am. Chem. Soc.* **2008**, *130*, 7092.
  - (14) Pincus, D. L.; Hyeon, C.; Thirumalai, D. *J. Am. Chem. Soc.* **2008**, *130*, 7364.
  - (15) Arai, N.; Yasuoka, K.; Zeng, X. C. *J. Am. Chem. Soc.* **2008**, *130*, 7916.
  - (16) Yan, L.; López, X.; Carbó, J. J.; Sniatynsky, R.; Duncan, D. C.; Poblet, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 8223.
  - (17) Yudanov, I. V.; Matveev, A. V.; Neyman, K. M.; Rösch, N. *J. Am. Chem. Soc.* **2008**, *130*, 9342.
  - (18) Cheng, J.; Hu, P. *J. Am. Chem. Soc.* **2008**, *130*, 10868.
  - (19) Dubbeldam, D.; Galvin, C. J.; Walton, K. S.; Ellis, D. E.; Snurr, R. Q. *J. Am. Chem. Soc.* **2008**, *130*, 10884.
  - (20) De Vivo, M.; Dal Peraro, M.; Klein, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 10955.
  - (21) Réal, F.; Vallet, V.; Wahlgren, U.; Grenthe, I. *J. Am. Chem. Soc.* **2008**, *130*, 11742.
  - (22) Serr, A.; Horinek, D.; Netz, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 12408.
  - (23) Dastidar, S. G.; Lane, D. P.; Verma, C. S. *J. Am. Chem. Soc.* **2008**, *130*, 13514.
  - (24) Heitler, W.; London, F. Z. *Phys.* **1927**, *44*, 455.
  - (25) *Theory and Applications of Computational Chemistry: The First Forty Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005.
  - (26) Pitzer, R. M.; Lipscomb, W. N. *J. Chem. Phys.* **1963**, *39*, 1983.
  - (27) Alder, B. J.; Wainright, T. E. *J. Chem. Phys.* **1960**, *33*, 1439.
  - (28) Wall, F.; Hiller, L.; Mazur, J. *J. Chem. Phys.* **1961**, *35*, 1284. Bunker, D. L. *J. Chem. Phys.* **1962**, *37*, 393.
  - (29) Liu, B. *J. Chem. Phys.* **1973**, *58*, 1925.
  - (30) Garrret, B. C.; Truhlar, D. G. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 4755.
  - (31) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1913.

calculation for a reaction more complex than  $H + H_2$ ,<sup>34</sup> and the application of Monte Carlo simulation to polydisperse liquids (such as micellar solutions, but in 1988 modeled by simple potentials) with temperature- and pressure-dependent distributions.<sup>35</sup>

By the 1990s, applications to complex systems were in full swing—for example, sorting out complex reaction mechanisms in organic chemistry. A noteworthy example in medicinal chemistry was the use of molecular modeling in the design of HIV protease inhibitors.<sup>36</sup> A noteworthy example in statistical mechanics was the use of configurational bias Monte Carlo sampling in the Gibbs ensemble calculation of the critical constants for higher alkanes (up to  $C_{48}H_{98}$ ),<sup>37</sup> which are not accessible to experiment due to pyrolysis but are essential for the crude oil fractionation industry.

It is interesting to compare the early prototype applications to the problems addressed in this issue of *JACS Select*. This collection encompasses drug design; heterogeneous and homogeneous catalysis; catalysis by enzymes and nanoparticles; photochemistry, photostability, and photosynthetic water splitting; ion coordination chemistry in the condensed phase; polyoxometalate anions; selectivity, kinetics, and tunneling in organic reactions; separations of mixtures; light-emitting devices and photovoltaics; RNA stability; nanostructured semiconducting materials; polymorphic transitions in nanotubes; and polypeptide friction and adhesion. Computational chemistry applied to complex systems like these examples is often more significant for the new insights that emerge than for any numerical result. The insights gained from simulation are synergistic with those that arise from new experiments, and sometimes they lead the way on problems where experiments are not available. Rather than give an example from the Articles and Communications in this issue of *JACS Select*, I will provide two examples from other recent *JACS* publications, to further expand this overview of the broad applicability of theory. In one Article,<sup>38</sup> the authors studied the question: how can one enhance the catalytic activity of nanostructured gold? On the basis of a combination of experimental evidence and calculations, they proposed a simple way to stabilize and chemically activate gold nanoclusters on MgO, and they suggested how their theoretical prediction can be validated experimentally. The authors of another Communication<sup>39</sup> reported both neutron scattering experiments and molecular dynamics simulations of the mean-square displacements of nonexchangeable H atoms of a hydrated protein. Because the simulation results are in qualitative agreement with the experiments, and because the simulation provides access to atomistic details not available experimentally, the authors were able to provide a detailed picture of the functionally relevant dynamical coupling of protein dynamics to hydration-water dynamics that would be impossible to obtain purely by experiment; this work also suggests additional experiments and simulations that can further increase our knowledge of this coupling in a realistic biological context.

What are the tools that are being used to make progress on these complex problems? Molecular computation usually has two components: (i) obtaining the potential energy surface and its gradient field, which provides the interatomic forces, and (ii) carrying out chemical dynamics or statistical mechanics or both with the chosen force field.

There are a wide variety of methods for carrying out step i, and they range in complexity from molecular mechanics, in which the conventional approach consists of parametrizing analytical potentials in terms of internuclear distances, bond angles, and torsion angles (dating in modern form to 1969,<sup>40</sup> at which time it was applied to a protein), to correlated wave function theory (WFT), exemplified by coupled cluster theory (with the first molecular application in 1972<sup>41</sup> to  $BH_3$ ). In between we find theories like semiempirical molecular orbital theory (developed in its most useful modern form in 1985,<sup>42</sup> with an application to a molecule as complex as tricyclo[3.3.1.1<sup>3,7</sup>]decane). A special place is occupied by density functional theory (DFT), where the first broadly useful functional for chemistry was published in 1994,<sup>43</sup> with application to 4-methyl-2-oxetanone. DFT and semiempirical molecular orbital theory are the methods currently in use for the most complex problems, and the higher accuracy achievable

(32) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269.

(33) DeFrees, D. J.; Raghavachari, K.; Schlegel, H. B.; Pople, J. A. *J. Am. Chem. Soc.* **1982**, *104*, 5576.

(34) Haug, K.; Schwenke, D. W.; Truhlar, D. G.; Zhang, Y.; Zhang, J. Z. H.; Truhlar, D. G. *J. Chem. Phys.* **1987**, *87*, 1892.

(35) Stapleton, M. R.; Tildesley, D. J.; Sluckin, T. J.; Quirke, N. *J. Phys. Chem.* **1988**, *92*, 4788.

(36) Dorsey, R. B.; et al. *J. Med. Chem.* **1994**, *37*, 3443.

(37) Siepmann, J. I.; Karaborni, S.; Smit, B. *Nature* **1993**, *365*, 330.

(38) Pacchioni, G.; Siculo, S.; Di Valentin, C.; Chiesa, M.; Giamello, E. *J. Am. Chem. Soc.* **2008**, *130*, 8690.

(39) Wood, K.; Frolich, A.; Paciaroni, A.; Moulin, M.; Hartlein, M.; Zaccai, G.; Tobias, D. J.; Weik, M. *J. Am. Chem. Soc.* **2008**, *130*, 4586.

(40) Levitt, M.; Lifson, S. *J. Mol. Biol.* **1969**, *46*, 269.

(41) Paldus, J.; Cizek, J.; Shavitt, I. *Phys. Rev. A* **1972**, *5*, 50.

(42) Dewar, M. J. S.; Zebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

(43) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

and anticipated to be achievable by density functional theory, especially with modern functionals,<sup>44–47</sup> is a key element in making the simulation of complex systems so promising.

Altogether 17 of the 23 publications employ some form of electronic structure theory, either WFT or DFT, and the other six articles and communications employ only all-atom molecular mechanics (five selections) or coarse-grained molecular mechanics (two selections; note that the numbers do not add up to 23 because some publications use more than one method). In coarse-grained molecular mechanics,<sup>48</sup> some functional groups or more broadly defined groups of atoms are treated as a “united atom”, which is more efficient for (for example) sampling conformational space in complex chain molecules. Several publications employ combined quantum mechanical and molecular mechanical methods in which an active site or primary subsystem is treated by a quantum mechanical method and the remainder of a large system is treated by all-atom molecular mechanics; thus, in total 10 publications involve all-atom molecular mechanics. An alternative to treating the secondary subsystems by molecular mechanics is to treat them by a continuum solvation model, which is employed in three of the publications.

A large variety of WFT methods are available. One key distinction is whether the zero-order description of the system (often called the reference wave function) is a single Slater determinant or more generally a single configuration state function (CSF; a CSF may be a single determinant or a linear combination of determinantal wave functions with fixed coefficients determined by symmetry or by a valence bond scheme). Such methods are called single-reference methods. In order to treat many kinds of complex systems, such as open-shell systems, systems with half-broken bonds, systems with nearly degenerate HOMO and LUMO, and electronically excited states, a multiconfiguration reference function is required, and methods with this complication are called multireference methods. Multireference WFT methods are said to include near-degeneracy (or “static” or “nondynamical”) correlation energy; an advantage of DFT is that it includes some static correlation effects even in a formally single-determinantal framework, although these effects are present (at the current level of theory development) in a somewhat uncontrolled way.

It is interesting to note that the technique of fitting potential energy surfaces for specific systems to physically chosen analytic forms has never advanced to systems more complex than about nine atoms. Instead, almost all dynamics and statistical mechanics calculations for complex systems are carried out either by molecular mechanics with transferable parameters or by the method of direct dynamics, in which “all required energies and forces for each geometry that is important for evaluating dynamical properties are obtained directly from electronic structure calculations.”<sup>49</sup> An important area of current research is making direct dynamics more efficient, by better algorithms or by using interpolation at some of the geometries.

With these classification schemes in hand, we can give a final summary of the methods used to obtain potential energy surfaces and force fields in the 23 selections: 10 use all-atom molecular mechanics, nine use DFT, three use semiempirical single-reference WFT (in particular, two employ AM1, and one employs CBS-QB3), two use ab initio single-reference WFT, four use ab initio multireference WFT, three use continuum solvation models, and two use coarse-grained molecular mechanics. Several important tools for attacking complex chemical systems are unfortunately not represented in this *JACS* Select collection but will certainly see increasing use for complex problems in the future. One example is reactive molecular mechanics,<sup>50,51</sup> in which analytical functions with transferable parameters are used in such a way that bond breaking, chemical reactions, and centers with multiple coordination numbers (e.g., transition metals) can be modeled. Another example is polarizable molecular mechanics<sup>52,53</sup> (or a more quantum mechanical polarizable force field), in which the secondary subsystem is allowed to self-consistently polarize. A third example is the use of specific reaction parameters<sup>54,55</sup> in which semiempirical parameters are introduced for a specific reaction or specific range of systems rather than as transferable parameters. Another strategy to look for in future work is multiscale and multilayer modeling; the examples in this issue of *JACS* Select involve at most two layers

(44) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(45) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207.

(46) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2005**, *123*, 121103.

(47) Zhao, M.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(48) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469.

(49) Liu, Y.-P.; Lu, D.-h.; González-Lafont, A.; Truhlar, D. G.; Garrett, B. C. *J. Am. Chem. Soc.* **1993**, *115*, 7806.

(50) Brenner, D. W. *Phys. Rev. B* **1990**, *42*, 9458.

(51) Chenoweth, K.; Cheung, S.; Van Duin, A. C. T.; Goddard, W. A., III.; Kober, E. M. *J. Am. Chem. Soc.* **2005**, *127*, 7192.

(52) Stillinger, F. H.; David, C. W. *J. Chem. Phys.* **1978**, *69*, 1473.

(53) Ren, P.; Ponder, J. W. *J. Phys. Chem.* **2003**, *107*, 5933.

(54) González-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618.

(55) Tresadern, G.; Wang, H.; Faulder, P. F.; Burton, N. A.; Hillier, I. H. *Mol. Phys.* **2003**, *101*, 2775.

---

(a small or mid-sized subsystem treated at a higher level and its surroundings treated by a lower level) or two scales (atomistic and continuum), but the use of three or more layers<sup>56,57</sup> is becoming more common. Finally, we expect to see more use of adaptive multilayer and multiscale methods where the boundary between layers or scales adapts to the evolving dynamics.<sup>58,59</sup>

The second component in applications to complex systems is the post-potential-energy-surface step, involving dynamics, statistical thermodynamics, or both. In the present issue of *JACS Select*, we find a broad distribution of approaches for this step as well. Six publications involve all-atom classical dynamics (usually called molecular dynamics), six involve conventional transition state analysis (which might also be called conventional thermochemical kinetics), two involve generalized transition state theory with multidimensional tunneling, two involve reaction-path analyses, two involve course-grained classical mechanical dissipative dynamics, two involve free energy calculations, and one each involve spectral simulation, trajectory surface hopping, vibronic analysis, and configurational-bias Monte Carlo sampling.

Without the continuing evolution of these tools, applications to more challenging problems would come to a standstill. One often hears the statement that computational chemistry has advanced because of the increase in the power of computers. This is certainly a key factor. However, my late colleague Jan Almlöf used to pose a thought-provoking question (typically in those late-night sessions when conferees at an international meeting gather for discussion in a bar or other surroundings conducive to illuminating discussion): If you were going to an isolated island to do computational chemistry and you could take this year's computers and 10-year-old algorithms or this year's algorithms and 10-year-old computers, which would choose to take? Most computational chemists chose the latter. So I dedicate this virtual issue to the chemists who develop the powerful new methods that allow us to tackle more complex problems rather than dedicating it to our new computers.

This third issue of *JACS Select* contains a new feature called Spotlights. This content is included on an experimental basis, in keeping with the experimental nature of the *JACS Beta* Web site (<http://pubs.acs.org/JACSbeta>). If successful, these features might occur in future issues of *JACS Select*. The Spotlights were written in some cases by computational chemists and in others by experimentalists. In each case I invited a nonauthor who could provide a broad perspective on an Article or Communication I selected (with, in some cases, advice from colleagues, whom I thank). The willingness of the community to write these Spotlights on very short notice is greatly appreciated. I encouraged the Spotlight writers to develop their own style for these contributions; there were no precedents—the only rule was that each piece should be between 150 and 400 words, preferably under 300. Some Spotlights comment mainly on the significance of the publication, both in terms of what is already accomplished and in terms of what the work means for the future. Others add thoughts about what promise the field of the publication holds for the future. I hope our community finds these Spotlights interesting and will join me in thanking the writers for their selfless service to our community in starting a dialogue.

*Donald G. Truhlar, Associate Editor*  
*December 10, 2008*

JA808927H

---

(56) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419.

(57) Rudd, R. E.; Broughton, J. Q. *Phys. Status Solidi (b)* **2000**, *217*, 251.

(58) Praprotnik, M.; Delle Site, L.; Kremer, K. *J. Chem. Phys.* **2007**, *1269*, 134902.

(59) Heyden, A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 217.