# The Computational Chemistry "Toolbox"

BY SAREN JOHNSTON

## Turning sophisticated theory into scalable code

YOU NEED THE RIGHT TOOLS FOR THE job at hand. It's a familiar phrase. We hear some version of it every day in all forms of advertising, from newspaper ads to television commercials to unwanted e-mails. The sales pitch is a popular one, mostly because the reality is pretty much undeniable: using the appropriate tools for a specific task almost always results in time-, energy- and cost-savings.

No need to tell this to Mark Gordon – you'd be preaching to the choir. Gordon, director of Ames Laboratory's Applied Mathematics and Computational Sciences program and an Iowa State University Distinguished Professor of chemistry, has devoted much of his career to making sure the right tools are available in GAMESS, the "toolbox" that he and thousands of researchers worldwide use on a daily basis to learn more about the goings-on of molecules and atoms – how they act, interact and react.

GAMESS, or the General Atomic and Molecular Electronic Structure System, is a software suite chock-full of quantum chemistry codes and functionalities – the tools that significantly enhance efforts focused on finding solutions to problems involving molecular processes, such as designing new fuels and gaining better insights into combustion, catalysis and photochemical energy conversion.

Gordon fosters the development of GAMESS and its software tools, which continue to grow and diversify as a result of his contributions to the suite and those of many other "code crafters," including Gordon's co-worker, Ames Lab associate Mike Schmidt, various current and past members of the Gordon research group, and numerous professional colleagues.

The GAMESS software tools are designed to take advantage of parallel, high-performance supercomputing systems that are capable of doing trillions of calculations per second. On such systems, the versatile software suite can perform the very highest levels of theory for modeling complex physical, chemical and biological systems. The continued enhancements Gordon and his colleagues make to GAMESS dramatically improve the ability to do such high-end and accurate calculations, a feat that at first glance may seem to impact only a select few. However, Gordon promptly rejects this all-too-quick assumption.

"If you have something that's the highest level of theory you can do, and you make that better in some way – faster, more convenient, more accurate – that has an impact on your ability to say, 'We've nailed this now, and now other people can use this as a benchmark for simpler methods,' " explains Gordon.

### "Toning up" the toolbox

Helping advance the capability to do top-notch theory is an enhancement to the GAMESS toolbox that combines the Fragment Molecular Orbital, or FMO, method of handling large molecular systems quantum mechanically with the Polarizable Continuum Model, or PCM, for solvent computations. Now that's a lot of technical talk, so let's see if we can break it down.

Interestingly, breaking it down is exactly the idea behind the FMO portion of this combination tool. Gordon explains, "What you'd like to do, in principle, is take the highest level of theory you can do and apply it to a big molecule, such as a polymer, enzyme or protein. The problem is that to do quantum chemistry on something that big is hugely expensive computationally, especially if you want to do a level of theory that is very demanding. The FMO method offers a way of 'chopping' a very large molecule into relatively small pieces by doing explicit calculations on monomers, the repeat units in a polymer, and dimers, which are two monomers. Some very clever methods are then used to put all those fragments together to make the larger system."

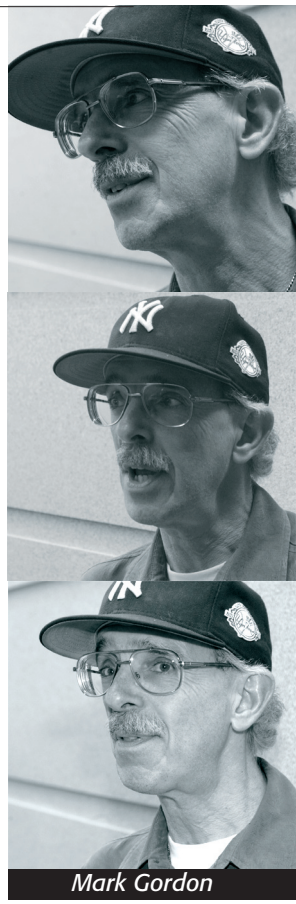
> "There is a great satisfaction in building good tools for other people to use."
>
> ~FREEMAN DYSON,
> English-born American physicist and mathematician

But how good is the piecing together of fragments relative to doing the full calculation if you could do the full calculation? Gordon says test calculations can be done on small molecules in which the full calculations can be accomplished without any approximation. These calculations can then be compared to those derived from the molecule's FMO. "The errors are within what most people call chemical accuracy," he notes. "The way to think about this is if an experimentalist had error bars, then the accuracy would be within those error bars and would never impact the experiment. And that's pretty good!"

Gordon credits his former graduate student, Dimitri Fedorov and his colleague Kazuo Kitaura, both of the Japanese National Computational Laboratory, AIST, for developing the FMO method for the GAMESS software suite. "The FMO method has been implemented for five different levels of theory, from the very lowest level of quantum chemistry theory that we do to the very highest," he says. "Fedorov and Kitaura are interested in biomolecules, so they've done calculations on systems with as many as 20,000 atoms, which is unheard of in quantum chemistry. You're lucky if you can do 100 atoms with quantum chemistry, so this is orders of magnitude larger than what people can usually do."

### Adding water to the tool "pool"

Making a good tool even better, Gordon led a team that included his postdoctoral research associate, Hui Li; former graduate student Jan Jenson, now at the University of Iowa; Fedorov; and Kitaura in developing a method of modeling the activity of molecules in solution. They created the new GAMESS tool by interfacing the Polarizable Continuum Model, PCM, with the FMO method. "If you're interested in polymers and enzymes, whatever they do, they do in solution, usually in aqueous solution," says Gordon. "That means just doing FMO calculations is not enough – you need to get the solvent in there." Expanding, he says, "If you imagine a solvent and throw in something like a protein, the protein sort of sweeps out a volume. We're not looking at individual water molecules now – we're representing the whole solvent sort of as a sea." Offering an analogy, he continues, "It's exactly as if you went swimming. You don't see the individual water molecules; you see the whole ocean. So that's the nature of this type of method, and we have a very nice one in GAMESS."


*Mark Gordon*

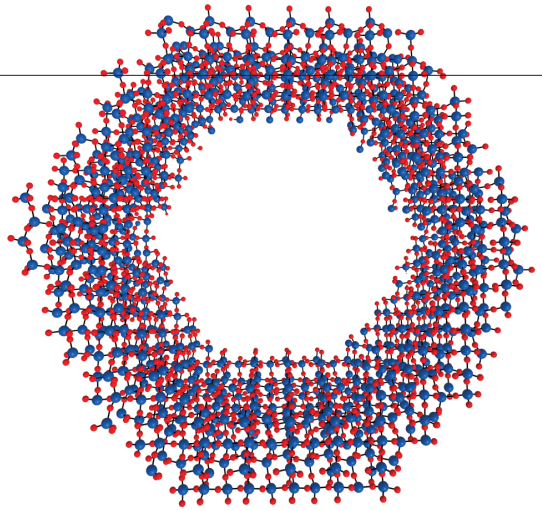
### A tool of prediction

A GAMESS tool that has been around since 1990 and continues to be improved upon is the Effective Fragment Potential, or EFP. When scientists deal with systems containing so many atoms that quantum mechanics can't be done, even at simple levels, the EFP method offers an alternative means. By combining the EFP with quantum mechanics in a way that a quantum mechanical description is produced only for that portion of a system that is going through a chemical change, scientists can then approximate the remainder of the system based on that fragment of information. The entire calculation takes orders of magnitude less time than a fully quantum calculation.

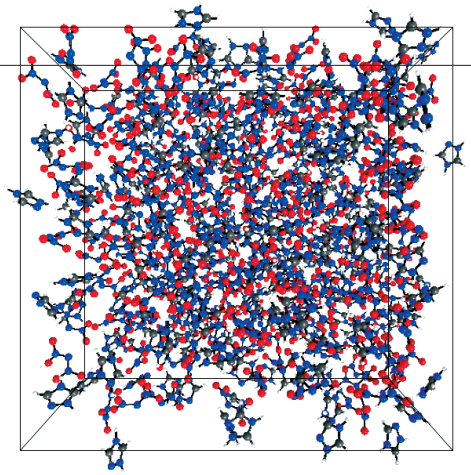
"The effective fragment potential is becoming more and more exciting as we try to model the interactions between molecules," says Gordon of the method that he and Jan Jenson developed in collaboration with Dr. Walter Stevens at the National Institute of Standards and Technology.

"This time we do care about each water molecule or each solvent molecule," Gordon notes. "Molecules interact with each other in a variety of different ways, and some of those interactions are easier to derive models for than others."

Some molecular interactions that are not so easy to model are those associated with dispersion and charged



*MacMolPlt, a GAMESS tool developed by Ames Lab associate scientist Brett Bode, makes it possible to visualize the results of GAMESS calculations on Macintosh computers. The above MacMolPlt image is of a mesoporous silica nanopore. Mesoporous silica materials hold great promise for the design of a new generation of highly efficient and selective catalysts.*



*This image is a molecular dynamics simulation of an ionic liquid composed of 512 ions. There are 256 positively charged ions (1,2,4-triazolium) and 256 negatively charged ions (dinitramide). The simulation was done on a 16-CPU parallel machine using the new GAMESS effective fragment potential code. The CPU time efficiency was 96 percent – testimony to the efficiency of the innovative GAMESS tool.*

species. Nonetheless, Gordon and his group members have added those capabilities to GAMESS in the past year.

## Dispersion

Dispersion is one means by which molecules exert forces on one another. Dispersion forces are randomly and spontaneously generated continuously over the entire molecule, making them extremely difficult to model. However, Ivana Adamovic, a former Gordon group member now at the Harvard School of Health in Boston, took on the challenge, developing a model potential for GAMESS that handles this fleeting molecular ebb and flow. Thanks to her contribution, the Gordon group has been able to use the new GAMESS tool to do calculations on DNA base pairs, and they are currently working with the ISU chemical engineering department to study the aggregation of polymers, such as latex.

## Charge transfer interactions

Charge transfer interactions are those that occur between ions. Ionic bonds are formed when the relationship between two atoms or molecules is such that one of them has a far greater ability to attract electrons and "steals" an electron from the other so that the now oppositely charged ions are attracted.

Hui Li, who will be an assistant professor at the University of Nebraska in the fall of 2006, and Heather Netzloff, a former graduate student of Gordon's now at Australian National University, enhanced the EFP method for GAMESS by respectively creating a molecular dynamics code to address the interactions that occur between ions and optimizing the code to take advantage of parallel computers.

The addition of the charge transfer code to the GAMESS toolbox is of particular importance to Gordon who is interested in ionic liquids in the design of high-energy fuels — products that bring together positively and negatively charged species. "You can't simulate those relationships adequately unless you have something that accounts for the charge transfer interactions, and we now have in GAMESS a molecular dynamics code that allows us to do that thanks to Hui and Heather," says Gordon.

## The "tool" de force

"GAMESS runs on every kind of computer platform and compiler you can imagine," says Gordon. "It just passed 20,000 registered user groups in over 100 countries. This translates into an estimated 100,000 to 150,000 individual users. They range from people in academics to industry to government labs."

GAMESS is distributed at no cost to users by accessing *www.msg.ameslab.gov* and signing a license agreement.

**For more information:**
Mark Gordon, (515) 294-0425
gordon@ameslab.gov

**Research funded by:**
U.S. Department of Defense
U.S. Department of Energy, Scientific Discovery through Advanced Computing