

# The Looming Petascale

By Elizabeth K.  
Wilson

## Molecular Dynamics

Simulation of a villin headpiece folding on a seven-microsecond timescale, performed using NAMD.

Credit: Peter Freddolino

**"IN THE LATE** 1990s, all supercomputing machines looked alike," according to University of Utah chemistry professor Thomas E. Cheatham III. In general, they had similar structures—clusters of processors that communicate with fast networks.

Now, that's all about to change as a new generation of ultrafast supercomputers looms on the horizon. As a result, computational chemistry is facing the need to radically retool its methods.

The new petascale computers will be 1,000 times faster than the terascale supercomputers of today, performing more than 1,000 trillion operations per second. And instead of machines with thousands of processors, petascale machines will have many hundreds of thousands that simultaneously process streams of information.

Not everyone is building these computers in the same way. Some make use of the massively parallel graphical processing units found in high-end video game consoles, some are packing more computing core elements into single chips than are found in conventional chips. "The machines are becoming complex again," said Cheatham, coorganizer of a petascale computing symposium sponsored by the Computers in Chemistry Division at last month's American Chemical Society national meeting.

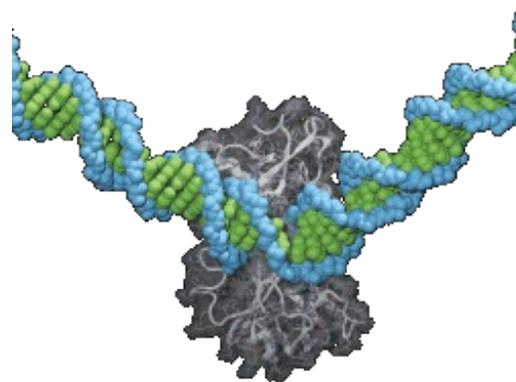
## Computing Action

Fpg, a DNA glycosylase, is locating and removing oxidative damage in DNA in this snapshot from a molecular dynamics simulation. Petascale computers will advance this computationally intensive modeling.

Credit: Carlos Simmerling

This technological sprint could be a great boon for chemists, allowing them to computationally explore the structure and behavior of bigger and more complex molecules. For example, molecular dynamics simulations of protein folding or interactions with solvents have been hampered by limitations in computational processing speeds. Simulations of processes that last at most a few hundred nanoseconds tax even the fastest terascale computers. Petascale computers offer the promise of simulations of processes that could last milliseconds, during which some of the most interesting chemistry happens.

But the massively parallel structure inherent to these machines often doesn't lend itself well to chemical problems. Tasks that involve repeating calculations with different starting conditions—in climate modeling, for example—are ideally suited for parallel computers. Some chemical calculations, however, such as those that predict electronic structures, require constant communication between processors, and this is more difficult to achieve on massively parallel computers.



Because of this parallel-computing problem, chemists "have been left behind" in software development, said Shawn T. Brown, a chemist at the [Pittsburgh Supercomputing Center](#) who was also a coorganizer of the symposium. "We have to change the way we do things," he said.

Apparently, there's no shortage of ideas. Several dozen chemists turned out to speak at the symposium, outlining new strategies for using statistical Monte Carlo methods and density functional theory and describing new ways to tackle problems ranging from electronic structure calculations to modeling biomass conversion to ethanol.

Last year, [Los Alamos National Laboratory](#) unveiled Roadrunner, the first supercomputer capable of sustaining a petaflop, or 1,000 trillion operations per second ([C&EN, June 16, 2008, page 12](#)). Much of Roadrunner's allotted time, however, will go to classified research.

Chemists are lining up in anticipation of the petascale computers that will soon be coming on-line. Klaus Schulten, director of the theoretical and computational biophysics group at the University of Illinois, Urbana-Champaign, has developed a code, [NAMD](#), for molecular dynamics simulations on parallel computing systems. His group has already been allotted time on Blue Waters, a petascale supercomputer under construction at the [National Center for Supercomputing Applications](#), run by UIUC, once it's operational in 2011.

The independent research company [D. E. Shaw Research](#) recently completed Anton, a massively parallel computer system built solely for molecular dynamics simulations. Scientists say Anton is capable of performing millisecond-long simulations of systems containing tens of thousands of atoms in a few months???100 times faster than current supercomputers can manage.

Using a different strategy, Vijay Pande, a chemistry professor at Stanford University, has been achieving petascale processing speeds for several years. His Folding@Home strategy takes advantage of thousands of processors from a network of home computers and even Sony PlayStations to perform protein-folding simulations.

Folding@Home's bottleneck is the speed of communication between processors, but Pande approaches the problem statistically by building simulations of protein folding from small snippets of the process. He likened it to charting out points on a road map. "Simulating a single, long trajectory is often neither necessary nor preferred," he said.

Simulation of a villin headpiece folding on a seven-microsecond timescale, performed using NAMD.

Chemical & Engineering News

ISSN 0009-2347

Copyright © 2014 American Chemical Society

Copyright ©2014 American Chemical Society