



POWER PLAY

A German physicist and a hedge-fund magnate are competing to push protein simulations into the realm of the millisecond. **Brendan Borrell** finds out what is at stake.

For a while, Klaus Schulten did not mind the Godiva chocolates arriving in his team's mailboxes at the University of Illinois in Urbana-Champaign. Nor was Schulten, whose biophysics group boasted one of the fastest algorithms for simulating protein structures, much concerned when his programmers received e-mails heralding a job opportunity at an undisclosed Manhattan firm that aimed to “fundamentally transform the process of drug discovery”.

It was early 2004, and Schulten's 40-strong group was attracting close to \$2 million a year in grant money. Nearly 20,000 users had downloaded his software, called NAMD for Nanoscale Molecular Dynamics, for use on computers running hundreds of parallel

microprocessors to simulate how individual atoms behave in proteins and other large molecules. Schulten's group itself was working on a million-atom model of the satellite tobacco mosaic virus, which the researchers called “the first all-atom simulation of an entire life form”¹.

But the German-born physicist got his wake-up call in 2006, when he saw a table of computing benchmarks in a report from that year's supercomputing conference in Tampa, Florida. A new program called Desmond, he saw, could calculate each step of a standard molecular-dynamics simulation — the 23,558 atoms in a system involving the protein dihydrofolate reductase — in a little over a thousandth of a second. NAMD was ten times slower. “Suddenly,”

Schulten says, “we were not the best anymore.”

The title had passed to the sender of the chocolates — David Shaw, a hedge-fund magnate and computer expert who taught himself physical chemistry. Over the previous few years, he had recruited more than 50 scientists and engineers, including three former students from Schulten's group, and put them to work in his midtown Manhattan high-rise.

In the paper from the supercomputing conference, Shaw's team wrote that Desmond “is faster than NAMD at all levels of parallelism examined”². And the group noted that on one simulation Desmond ran faster on 1,024 processors than NAMD ran on the 16,384 processors of IBM's Blue Gene/L — the world's fastest supercomputer.



Model behaviour: Klaus Schulten is pursuing his dream of creating a 'computational microscope' to study complex molecular dynamics.

The numbers shocked Schulten, who believed his team was on course to simulate molecular dynamics on the scale of milliseconds — longer than anyone had previously achieved. Even with cutting-edge programs such as Desmond and NAMD, scientists have been able to glimpse only the fastest-folding proteins, such as the villin headpiece, which folds in about 10 microseconds. The number of possible configurations of atoms in larger molecules, over time and in three dimensions, is astronomical. If these kinds of simulation could be sped up 1,000-fold, which even then could take a month of computing time, the pay-off could be high. They might, for instance, reveal binding sites for new drugs to tackle a wide range of medical problems.

Shaw and Schulten are now spending millions of dollars each to break the millisecond barrier. But some in the field aren't sure what the all-out push will come to. As Ross Walker, a computational biologist at the San Diego Supercomputing Center in California, puts it: "A lot of what they are going to see are limitations on the underlying computational models."

Pushing the envelope

To make molecular-dynamics simulations feasible with today's computers, scientists have had to make a number of simplifying assumptions. Typical simulations calculate the forces acting on each atom from a century's worth of chemistry experiments on organic molecules much smaller than the proteins scientists wish to simulate. The simulated molecules are also pegged together like Tinkertoys; they can change shape during the simulation, but cannot react to form new molecules.

The first software that sought to capture this world was developed at Harvard University in the late 1970s. In a paper in *Nature*, a team led by Martin Karplus published its 458-atom simulation of a tiny protein on an IBM 370, a top-of-the-line supercomputer³. Today, development teams around the world continue to work on CHARMM, or Chemistry at Harvard Molecular Mechanics, even as other algorithms such as NAMD have risen to compete with it.

One of the biggest factors limiting the development of molecular dynamics has always been computational power — which is where Shaw comes in. Having stepped back from running his hedge fund around 2001 (see 'From science to finance and back again', overleaf), Shaw, who is also an adjunct professor of biomedical informatics at Columbia University in New York, returned to his first enthusiasm — the architecture of massively parallel supercomputers. Predicting the motions of large systems of atoms requires finding the best way to communicate particle positions and forces among multiple processors. And on a scorching afternoon in June 2003, Shaw holed himself up at a friend's house and found a way to speed things up.

In traditional parallel approaches, each processor calculates forces to update the position of all the particles in its own small box of simulated space. But to do so, it must import positional data from neighbouring boxes within a certain radius. Shaw's strategy, implemented in Desmond, changes the geometry of this import region from a hemisphere to a semi-circular plate and a rectangular tower. As the number of processors available to Desmond

grows, the volume of this import region shrinks more quickly than in the approaches used by NAMD and CHARMM. In one of the first studies to use Desmond, this speed-up gave Shaw and his collaborators an unprecedented view of the workings of an ion transporter that the bacterium *Escherichia coli* uses to maintain its salt and pH balance⁴.

But Shaw knew that software alone could not obtain millisecond-long molecular simulations. His plan has been to build a supercomputer so dumb, he says, that it can do nothing except molecular dynamics. "But," he beams, "it's really fast at that." He calls it a computational microscope and has named it Anton, after Anton von Leeuwenhoek, the seventeenth-century Dutch scientist and builder of microscopes. The first segment of Anton is due to arrive in Shaw's lab at the end of the year.

Need for speed

Anton uses a high-speed task pipeline to accelerate the most computationally intensive tasks of molecular dynamics — modelling certain long-range interactions among atoms. But the chip does not have the ability to speed up software-based operations to the same extent, and the hard-wired pipeline may not be flexible enough to efficiently incorporate advances in the field. "At this point, though, we placed our bets," Shaw says.

When Shaw began the work, he estimated that Anton would run molecular-dynamics simulations 1,000 times faster than previous parallel supercomputers. In recent months, he has stopped presenting the 1,000-fold estimate in talks, although he still believes Anton will run more than 100 times faster than today's machines. But with general-purpose hardware doubling in speed about every two years,

many wonder how long Anton might maintain a lead. "If you are a little bit of a sceptic," says Schulten, "you would say it is another attempt for a special-purpose processor that will be overrun by market forces."

The field is littered with what Gregory Voth, a computational chemist at the University of Utah in Salt Lake City, calls

"dead bodies". In 1984, the late biochemist Cyrus Levinthal designed a molecular-dynamics computer called FASTRUN, but it took his group six years to get it running. During the past ten years, IBM and RIKEN, Japan's main research institute, have collaborated on several generations of chips intended for molecular-dynamics simulations, called MD-GRAPE, without producing any major breakthroughs in the field. At the National Institutes of Health in Bethesda,

"I wouldn't have told them about a great solution I had developed, and they wouldn't tell me their solution."

— Klaus Schulten

Maryland, in the late 1980s, Bernard Brooks abandoned his effort, dubbed Gemstar, when Hewlett-Packard announced its blazingly fast 9000 series — which could be had for as little as \$12,000. Scientists are racing not just against each other, but against Silicon Valley.

Schulten has played that game before. In Munich in the late 1980s, he built his own parallel supercomputer out of 60 processors mail-ordered from England. He carried his computer in a backpack to his new laboratory in Illinois, where he ran a 30,000-atom simulation of the bacteriorhodopsin protein, which drives the photosynthetic reaction that turns light into an electric charge. His simulation lasted 263 picoseconds — less than a millionth of a millisecond — and required more than two years of continuous computation⁵. By then, his machine was obsolete.

Thinking big

In the past 15 years, Schulten's ambitions have grown: from 100,000 atoms in 1999, to 300,000 in 2003, and culminating with his million-atom simulation of the tobacco mosaic virus published in 2006. To match his models, Schulten developed software that could scale with advances in parallel computers, something CHARMM could not do at the time. Chemist Richard Hilderbrandt, who supported the early development of NAMD at the computing directorate of the US National Science Foundation, says that the idea “was to take a large molecule and break it up into patches to distribute to processors. It was quite a bold step”.

The drawback of Schulten's strategy was that it could not simulate the behaviour of smaller molecules significantly faster than it could large ones. “If you have a protein of 500 atoms,” he says, “it's very difficult to put it on a parallel computer with 5,000 processors.”

Schulten emphasizes that his publicly funded group had to focus on ensuring that NAMD, which is freely distributed, would run on a wide range of platforms. Shaw's team, in contrast, could tune Desmond for its state-of-the-art computing cluster, about a year before similar clusters were available at National Science Foundation computing centres.

Shaw says that profits are a long way off, and that he is working to share his team's technology

From science to finance and back again

For a man whom *Fortune* magazine once named King Quant, David Shaw does not come across as particularly regal. “I never understood why, if you want to be accepted in the business community, you have to wear something that restricts blood flow,” he says, tieless and with his top shirt button undone, in his office at the Columbia University Medical Center in New York.

Shaw's integration of science and trading may have been predestined: his stepfather was a professor of finance at the University of California, and his biological father was a plasma physicist who worked in the defence industry. For his part, the younger Shaw founded a technology consulting company while an undergraduate at Stanford University and, even after he joined Columbia University with a generous faculty package, he approached venture capitalists in the hope of getting \$10 million to develop his own parallel computing venture.

Instead of bringing money into his lab, Shaw got sucked into quantitative finance, where investors study the mathematical behaviour

of the markets to plan their strategies. In 1986, Morgan Stanley hired him to run its technology team, where he stayed for two years before founding his quantitative hedge fund, D. E. Shaw, one of the first to use sophisticated algorithms to exploit inefficiencies in the marketplace. “It was a field in its relative infancy and what that means to an academic type is there's low-hanging fruit,” he says. “This was part of the attraction: we could discover things no one else had found.”

His company became widely known for its 18% annual returns and a highly selective recruiting process in which only 1 in every 250 candidates got selected, many of them PhDs in the hard sciences or winners of prestigious maths competitions.

Even on Wall Street, Shaw never strayed too far from science. Always a major contributor to the Democratic party, in 1994 Bill Clinton appointed him to the President's Council of Advisors on Science and Technology, where he served for seven years and was charged with improving the use of technology in schools.

But as Shaw's 50th birthday neared in 2001, he began to look for an exit. His company had more than 1,000 employees, and Shaw was no longer engaged in the quantitative problem-solving that fascinated him. His sister was battling breast cancer — she died in 2003 — and Shaw believed he could contribute to medicine, not just financially but intellectually. In his spare time, he had been reading up on the computational puzzles of molecular dynamics and talking with academic friends.

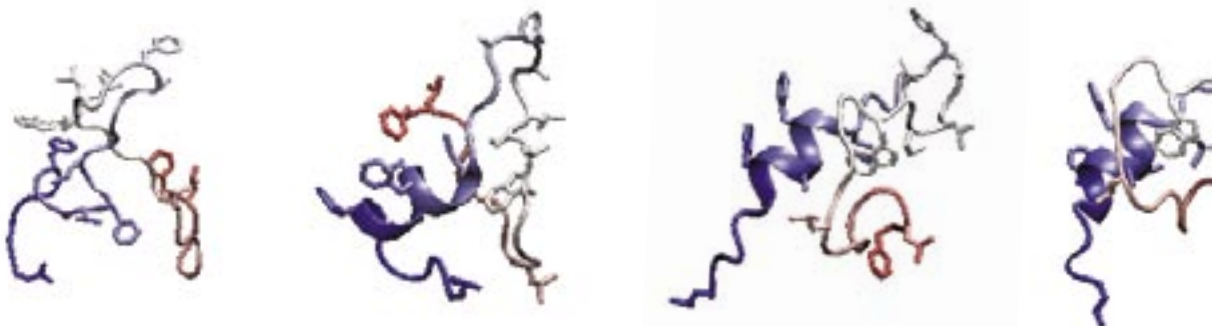
In October 2002, Shaw hired his first computer scientist, trained at the Massachusetts Institute of Technology, to manage operations at D. E. Shaw Research. The venture, which Shaw compares to a tiny and highly focused Bell Labs, now has nearly 60 staff. Some researchers in the field are still getting used to the newcomer, but Shaw does not see science as a contest. “Maybe it was because I was in the financial field,” he says. “It's a zero sum game — you cannot make money unless someone else is losing money. That's one of the reasons why I like science: it doesn't work that way.” **B.B.**

as much as possible. But his proprietary algorithm will ultimately be sold to industry through an agreement with Schrödinger, a biotechnology company founded by chemist Richard Friesner, a colleague of Shaw's at Columbia. Schulten had only inklings of Shaw's ambitions when he gave a seminar at D. E. Shaw Research in October 2004. “At that time it was clear that there was a competition,” he says, “but in a very civilized way.” Even so, he

says, “I wouldn't have told them about a great solution I had developed, and they wouldn't tell me their solution.”

Although Schulten's software has been a boon to many researchers, with a development cost of \$20 million it might also be considered a drain on their resource pool. Some scientists contend that the pursuit of speed has hindered alternative modes of inquiry. “I think it's unfortunate that some of the researchers who use

Twist in the tale: a simulation of some steps in the folding of the villin headpiece, one of the fastest-folding proteins.



R. LEMOINE

more established codes with a broader range of functionality are not getting the same access to national resources,” says computational chemist Charles Brooks of the Scripps Research Institute in La Jolla, California.

Tough decisions

Some participants at a 2001 supercomputing conference recall Hilderbrandt telling the audience that users should switch from older programs such as CHARMM to modern parallelized packages, such as NAMD. Hilderbrandt, who is now at the Department of Energy, does not recall being so specific, but says he still believes NAMD is “the program of choice” for most applications.

Michael Crowley at the National Renewable Energy Laboratory in Golden, Colorado, doesn't buy that. He uses CHARMM to study biofuels and says: “CHARMM has functionality that as far as I know, no other program comes near.” He says that when he has applied for supercomputing time from allocating agencies, “you can almost expect that somebody is going to suggest you use NAMD”.

There are deeper questions about the pursuit of ever-longer timescales. “It's clear to me that what's emerging out of both Schulten's and Shaw's efforts are technological advances that are going to affect the entire community,” says Brooks. “But whether an individual achievement of a millisecond timescale for any particular simulation is of great significance, I'm not entirely sure.”

Vijay Pande, at Stanford University in California, has pioneered the folding@home distributed-computing project, which uses the personal computers and Sony PlayStations of more than 250,000 volunteers to study protein folding. “The revolution that's going on,” he says, “is people are now treating molecular dynamics in a much more sophisticated way, where they are running hundreds or thousands or millions of simulations and then data-mining those simulations.” Because a simulation may take a slightly different course each time, he notes, a single long simulation cannot provide the statistical information that

“Researchers who use more established codes are not getting the same access to resources.”

— Charles Brooks

must be gathered over many runs, such as the affinities for binding to a drug.

Schulten and Shaw may also be pushing current models to their breaking point. Neither group is investing significant resources in improving fixed-charge force fields, which might turn out not to be accurate enough for lengthy simulations. For instance, when two atoms approach one another, the electron orbits of one can get sucked towards the positive charge generated by the other. This phenomenon, called polarizability, is cumbersome to model and slow to compute. Shaw estimates that it would slow down computation by roughly a factor of ten; Schulten thinks it may be only a factor of two.

Yet these difficulties may be a reason for moving forward, not calling a halt. Longer simulations can show where the models are failing, and they can guide the distributed-computing approach. Shaw believes his group can make a meaningful contribution to the field, but he is well aware of the problems ahead. “If you have something you're sure is going to work,” he says, “you're not being ambitious enough.”

Last year, Schulten's group started running a new version of NAMD that can handle smaller

molecules faster. His team has also started programming the graphics accelerator chips prized by PC gamers — an economical solution to the hardware problem that could further shrink Anton's expected lead. And, now that the team is up to speed with the University of Illinois's cluster, Abe, it has tailored a special version of NAMD to compete on equal terms with Desmond.

Two months ago, Schulten was delighted to tell Shaw about a simulation of a 38,000-atom protein, in which NAMD had set a new personal best, computing a 0.1-microsecond simulation in the course of a day. “We agreed, now the programs are pretty equal,” says Schulten. And for his part, Shaw may be starting to concede that each algorithm has its benefits. “Schulten has made extraordinary strides in his NAMD code,” he says, “so it's not obvious to me that Desmond will be significantly faster for all applications.”

Brendan Borrell is a freelance science writer in New York City.

1. Freddolino, P. L. *et al. Structure* **14**, 437–449 (2006).
2. Bowers, K. J. *et al. Proc. ACM/IEEE Conf. on Supercomputing (SC06)*, Tampa, Florida, 2006.
3. McCammon, J. A., Gelin, B. R. & Karplus, M. *Nature* **267**, 585–590 (1977).
4. Arkin, I. T. *et al. Science* **317**, 799–803 (2007).
5. Heller, H., Schaefer, M. & Schulten, K. *J. Phys. Chem.* **97**, 8343–8360 (1993).



Number cruncher: David Shaw has used his computer skills to make money and model proteins.

