Powering new discoveries
One of the model problems specified by the National Science Foundation in its solicitation for a sustained-petascale computing system was a very large protein simulation using NAMD, the molecular dynamics code developed by the team of University of Illinois researcher Klaus Schulten. Schulten was one of more than 20 scientists with whom NCSA collaborated in developing the proposal for Blue Waters the center’s winning response to that petascale solicitation.

As leader of the Theoretical and Computational Biophysics Group at the University of Illinois, Klaus Schulten has pioneered the development of tools and techniques that he collectively refers to as a “computational microscope.” The analogy is apt; just as light microscopes gave scientists the first glimpse of cells, today Schulten and others use computational methods to obtain an even more fine-grained look at the basics of life.

“We know the physical structure of the molecular machinery in cells, and we can now apply the laws of physics and our knowledge of chemistry to describe how these structures and materials evolve in biological functions,” Schulten says. “This microscope gives us an unprecedented view of the details that happen in living cells and make life possible.”

Researchers like Schulten typically begin with information obtained from X-ray crystallography, which can reveal the structure of the biomolecules within cells. But while crystallography provides a static view, the “computational microscope” lets Schulten observe how these biomolecules behave.
“We can put them in the environment in which they are found in the living cell…and just let them move and see what they do when they are out of the confinement under which they are observed,” he says.

Using a football game as an analogy, Schulten says that in crystallography the players are just standing in a line. The computational microscope shows how they play.

More computer power, more detail

As computers have become more powerful, Schulten and other researchers have been able to observe biomolecular processes in greater and greater detail. Twenty years ago, it was a triumph to simulate part of the cell wall with the correct physical properties, a simulation that involved tens of thousands of atoms.

“It was a very important result, because the lipids that make the walls of cells are disordered. There’s no way of capturing their exact composition in some kind of other microscopy, but we could describe it,” Schulten says.

Later, a simulation of about 100,000 atoms could describe proteins in the cell walls. “We could see, for example, a water channel that permits water to go through bacterial membranes,” Schulten explains.

Today’s high-performance computers enable scientists to simulate systems of several million atoms. Schulten and his team, for example, were the first researchers to simulate an entire lifeform—the satellite tobacco mosaic virus. Even for such a tiny form of life, the simulation involved 1 million atoms.

Simulations continue to grow in size as researchers seek to understand more complex processes in greater detail.

“The key processes of proteins and other biomolecules in cells are actually team sports,” Schulten says. “If you want to understand how life comes about, how it organizes itself, then we need to understand how teams of molecules work together. And that means we simulate 10 proteins at a time, 50 proteins at a time.”

The impact of Blue Waters

Blue Waters, which will be capable of sustained performance of a petaflop or more when it comes online in 2011, is squarely aimed at addressing this need for increased computing power to enable greater scientific insight.

Schulten anticipates gaining a 50 times speedup in his NAMD code thanks to the few hundred thousand cores offered by Blue Waters.

“We will use that increase in two ways. On one side we want to use it to compute structures that are more and more like the entire biological cell,” he says. “If we use current computers, we can’t show how the structures within cells are formed. ‘The factor of 50 will give us an increase that will be about one-tenth the size of the cell that we are going to describe.’

The increased power will also allow Schulten to look at biomolecular processes for longer time scales. His team has just recently been able to simulate processes for about 10 microseconds, but even that achievement is too short a window, since many processes in living cells take a full millisecond.

“This is a big goal for the field—the holy grail—to be able to simulate a millisecond,” Schulten says. Blue Waters could bring that holy grail within reach, enabling the simulation of many “almost a universal class of processes” that occur in a millisecond.
Cosmologists are always clamoring for more computing power, and with good reason—the object of their study is the universe, and it simply doesn’t get any bigger than that. Cosmologists seek to understand how our vast expanding universe sparked into existence and unspooled, how the Big Bang gave birth to particles and gases that coalesced into stars, planets, galaxies … and everything else.

According to Princeton’s Jeremiah Ostriker, the combination of recent advances in the cosmological model and in simulation codes, a growing influx of data from ground- and space-based observation, and petascale computing could make it possible to computationally construct a model universe.

“The goal is to see if you can simulate the universe. Can you compute it and can you make comparisons to observed data?” he explains.

The first simulations of the universe were done in a volume of 32 x 32 x 32. “You can barely make out a face in that!” Ostriker says. “Now we’re up to 2,048 x 2,048 x 2,048. And that’s a gigantic improvement. But still it’s pathetic compared to the real universe.”

“Right now, we can either simulate one or two galaxies well, or we can do many galaxies poorly,” he says. Petascale simulations could provide detailed images of many galaxies for comparison with observational data.

“With this extra horsepower, we should be able to make a huge stride forward,” he says. “Needless to say, we’re all drooling. It’s a very exciting prospect for us.”

For astrophysicist Paul Woodward, petascale computing could bring a solution to a problem that just a few years ago was considered impossible.

“The large problem I’m working on now has to do with what goes on inside the central region of a star near the end of its life,” he says. Woodward explains that in the center of the dying star, hotter gases are rising and cooler gases are descending. At the bottom of the central region, nuclear fusion is occurring and helium is burning and forming carbon; at the top is unburned hydrogen. If some of the unburned hydrogen could be pulled down to the shell of burning helium by the moving gas, it would set off multiple reactions affecting which heavy elements the star ultimately would produce.

“The problem is that the motions of this material are fast but not too fast, about Mach 1/30th,” he says. That means that it takes about 20 times as many time steps to simulate the same amount of fluid motion as can be more quickly modeled for other problems.

“It costs 20 times as much to do the calculation,” Woodward says. And calculating the reactions that occur when the hydrogen hits the helium shell is computationally expensive “because you have to follow all the reaction productions and make sure you know where they went and what they turned into, and then where that went, and so forth.”

Petascale computing is “just the kind of development that we need to make it possible at a reasonable cost to do simulations like this,” Woodward says. “If you go to the petascale level you can do the whole layer for a long time.

“It could simulate an hour of the whole star, of that whole shell, doing its thing at very high resolution, so that I should hope to get the answer kind of right. And it’s affordable. And I’ll just be able to solve it! That’s kind of a neat thing.”