

MATTERS OF TIME: Tick Tock Go the Simulations

By Katharine Miller

Time flows like a continuous, steady river. And it moves forward—never back. These facts create inherent challenges for computer simulations of biological molecules in motion.

It would be lovely if time could be efficiently simulated as a flowing variable. But time has to march in discrete steps for computers to handle the complex movements of molecules. And that matters: The length of the step (be it a femtosecond, a millisecond, a minute, or a year) affects the stability and accuracy of a simulation; limits the amount of total time that a simulation can reasonably cover; and generates error terms that must be accounted for. In addition, researchers add inaccuracies of their own by coarse-graining models, simplifying the simulations in space as well as time to improve efficiency and cover the longer time spans of biological interest. And then there's the fact that time has a directional arrow—one that's hard to untangle from the energy landscape at microscopic scales.

Despite the challenges of simulating time, researchers remain committed to molecular dynamics (MD) simulations—including coarse-graining—because they provide insight, says **William Noid, PhD**, assistant professor of chemistry at Penn State University. Indeed, he predicts coarse-grained models will always be useful because, as he puts it, “the human imagination and computational demands will always progress at a rate far exceeding Moore's law.” But it's important to keep in mind that these simulations are models, not

Discretizing Time

Researchers simulate biological molecules to gain an understanding of how they function in living systems. These molecules move

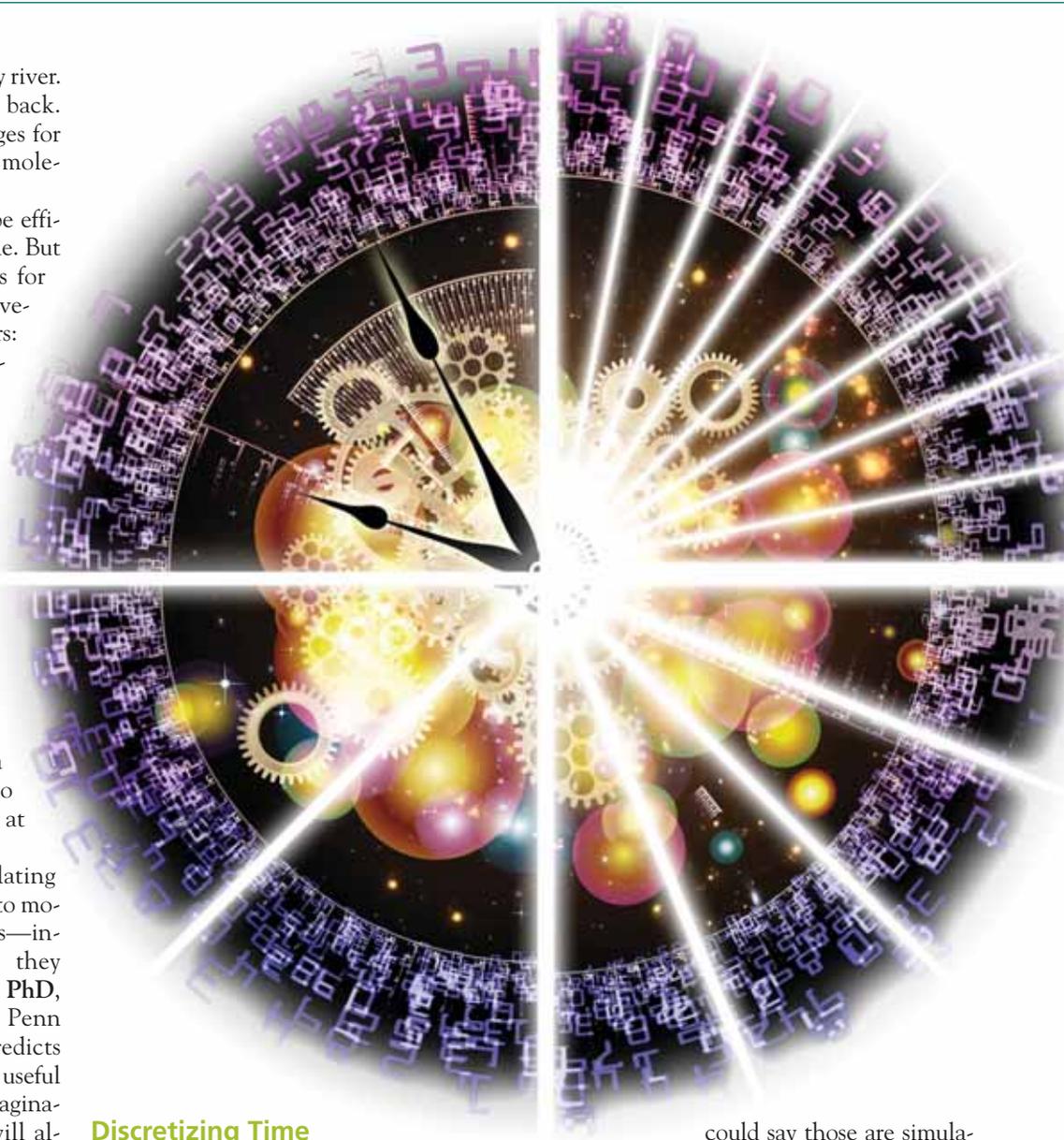
could say those are simulations with continuous time,” says **David Sivak, PhD**, a systems biology fellow at the University of California, San Francisco. But for more complicated molecular systems, the equations can't be solved exactly, even by a computer, he says. Breaking time into discrete steps becomes a way to make the calculations computationally tractable.

For example, a system of atoms or molecules can be described by a series of differential equations that evaluate how the particles' positions and velocities change over time in accordance with Newton's second law of motion ($F=ma$, or force equals

“The human imagination and computational demands will always progress at a rate far exceeding Moore's law,” William Noid says.

reality, Noid says. And the ravages of time are likely to always play a role in keeping it that way.

according to the laws of physics. For very simple systems, Newton's equations of motion can be solved exactly. “In that sense you



mass times acceleration). An MD simulation might then numerically integrate these equations of motion over a series of time steps. The computer calculates the forces on each particle based upon their current positions and then assumes that the forces on the particles are constant for a short increment of time, Noid says. During this short increment, the particle positions and velocities change due to the forces on them. The computer then updates the forces on each particle based on their new positions and velocities, and the process is repeated. Although researchers use many more sophisticated ways to integrate the equations of motion to achieve greater accuracy and efficiency, “most are only slight variations on this simple mechanism,” Noid says.

The basic assumption here—that forces are constant during the time step—inaccurately represents reality. “It’s like a movie; It’s really a series of discrete snapshots played fast enough that it looks continuous,” says **Greg Bowman, PhD**, a research fellow at the University of California, Berkeley.

Longer time steps reduce a simulation’s stability as well as its accuracy. Indeed, if the steps are too long, the molecules being simulated start to take on unfavorable conformations. “You get a cascading problem.

comes time for publication, Bowman notes, reviewers don’t necessarily notice the details unless the paper’s results don’t make sense. “Only then will they look back and question the parameters.”

A better practice, Bowman says, is to take the time to find the right time scale for the problem. One approach is to find the largest time step where things don’t blow up. Another way to think about it, Noid says, is to find the largest time step over which one can reasonably approximate the forces as being constant—where the particles haven’t moved enough to alter the forces appreciably. In the case of MD simulations, the appropriate time step is determined by the interaction that changes most rapidly, Noid says. For simulations of atoms, this ends up being on the order of one femtosecond—the rate of jiggling and wiggling of bonds or water molecules.

When a continuous process is simulated using discrete time, there are always errors—discrepancies between the calculated results and the true underlying behavior. Errors pose another consideration for choosing the duration of the timestep, Sivak says. So, for example, researchers might observe the error at the largest time step where things don’t blow up and then do the same

additional work on the system,” Sivak says. This realization allows researchers to quantify how far out of equilibrium a simulation is simply due to the discretization of time—even when the system otherwise would be in equilibrium. It becomes possible to characterize this “shadow work” and correct for it, separating the physically realistic aspects of the simulation from the artifacts of the computer method, Sivak says.

Temporal Coarse Graining and the Time/Space Connection

To overcome the limits of computer power, researchers often create simplified models that allow for more efficient MD simulations over longer time scales. The simplifications can be spatial—e.g., treating a group of molecules as a single ball; or temporal—e.g., using longer time steps. In reality, says **Thomas Miller, PhD**, professor of chemistry at the California Institute of Technology, the two go hand in hand. “You can’t coarsen spatially without coarsening in time,” he says. If atoms are clumped into a ball, the corresponding time scale for the movement of the ball is slower than it was for the atoms. “That’s two halves of the benefit of the process,” he says. “As you eliminate unnecessary spatial motions,

“You can’t coarsen spatially without coarsening in time,” Thomas Miller says.

It’s not a subtle thing: Your simulations just blow up,” says Sivak.

But many biological events of interest take too long to simulate using small time steps, given the limits of computational power. For example, proteins take milliseconds to fold—a process that would take more than a trillion femtosecond timesteps to simulate—beyond the capacity of typical computational resources. On the other hand, it takes only a thousand microsecond timesteps to simulate a millisecond. Researchers have to balance their desire to integrate the equations of motion as accurately as possible, against their need to make the problem computationally manageable.

Picking a Timestep

In practice, many researchers don’t contemplate the size of the time step. They use the default settings or recommended time discretizations in readymade software packages, Sivak says. Or they copy the parameters used by others without necessarily evaluating where they came from or why they were chosen, Bowman adds. When it

at a smaller timestep to see how the error changes as the timestep shortens. If they know the level of error they are comfortable with, they can then pick a particular time step, he says.

Time At Work: An Intuitive Understanding of Timestep Error

Having selected a timestep and performed a simulation, researchers also have to correct for the errors the timestep creates. Recently, Sivak and his colleagues took a hard look at these errors and came up with an intuitive, physical way of thinking about them. The work was published in *Physical Review X* in January 2013.

Errors caused by time discretization turn out to be particularly important in so-called nonequilibrium simulations where the conditions are changing fast, such as where a protein is being stretched. Sivak and his colleagues found that just as you can mechanically put energy into a protein—by stretching it, for example—the discretization of time also puts energy into the protein. “The error arises because the simulation does

what’s left over moves more slowly so you can take bigger timesteps.”

In October 2012, Miller published in



the journal *Cell Reports* a coarse-grained simulation of the Sec translocon, a channel that allows proteins to pass through cell membranes. The feat required his team to coarse-grain out lots of faster molecular movements—from femtoseconds to hundreds of nanoseconds—in order to focus on the slower movements—from hundreds of nanoseconds to the full minutes it takes for a protein to pass through the channel. But before doing that, they had to determine the average effect of the faster motions. “We had millions of hours of underlying computer simulation time based on high-resolution models,” he says.

The Sec translocon paper demonstrates the degree to which complex biological machinery can be simplified while still capturing a wide array of experimentally observed phenomena in the system, Miller says.

Markov State Models: A Knob for Controlling Time and Space Resolution

Markov State Models (MSMs) offer another way to achieve longer time scales for MD simulations such as protein folding. An MSM can merge variations from thousands of successive protein-folding simulations and identify a set of relatively stable conformations along the protein’s many folding pathways. By choosing a timestep for the model as well as how many states to identify, whether 15 or 100,000, researchers can dial in the degree of complexity they seek.

The idea is that you’re removing the intermediate steps between these stable conformations, sort of like reducing the frame rate in a movie, Bowman explains. “We can use this time and space resolution basically as a knob to control how detailed our models are,” he says. The approach allows the simulation of

larger proteins for longer periods of time, permitting insight into how they function.

Tomorrow Differs from Today: Time’s Irreversibility and Biological Molecules

At the macroscopic scale, we have no doubt that time moves inexorably forward. A glass can fall off a table and smash to smithereens, but cannot jump back onto the

table in one piece. And we know instinctively when a movie of human-scale events is run in reverse.

But at the molecular scale, discerning forward from backward is much harder. That’s partially because everything is stochastic—tiny molecular machines fire randomly; they are not like steady car engines. Yet time’s forward arrow does exist at the molecular level thanks to the second law of thermodynamics which states that isolated systems spontaneously evolve toward maximum entropy. (All other laws of thermodynamics are equations that don’t care about time.)

It’s just that spotting entropy’s signature is tough at the molecular scale because the energy required to break time asymmetry—to move toward maximum entropy—is close to the entire local energy budget, says **Gavin Crooks, PhD**, senior scientist at Lawrence Berkeley National Lab. For example, an important molecule like ATP synthase—a tiny little molecular engine—functions at an energy level that is not much greater than the scale of energy fluctua-

tions in the environment.

Over the last ten years, Crooks and others have made progress toward spotting entropy’s signature against the fluctuating energy background in single-molecule experiments. It turns out that accounting for time asymmetry matters greatly in MD simulations of systems that are out of equilibrium—just the kinds of systems that interest Crooks. He has a grand vision of

[A]ccounting for time asymmetry matters greatly in molecular dynamics simulations of systems that are out of equilibrium—just the kinds of systems that interest Gavin Crooks.

thermodynamically realistic simulations of walking molecules, such as myosin stepping along an actin strand—a very non-equilibrium process. Such systems have their own intrinsic time asymmetry that needs to be untangled from the rest of the thermodynamics. “In the long run, I would like to do simulations of relevant biological systems that are active, that aren’t just at equilibrium. And I want to get the thermodynamics right,” he says.

Bridging Time Scales

The intrinsically molecular processes that govern our physiology include chemical reactions faster than a picosecond; bond rearrangements that take picoseconds to nanoseconds; changes in protein conformations that happen in microseconds; protein folding that occurs in milliseconds; and barrier-crossing events that take seconds to minutes, Miller says.

In biological systems, the separation of these time scales is not always clear. One process with higher time resolution may feed into a process with lower time resolution. “That complexity is potentially interesting but very challenging for the person doing the modeling,” says **Gerhard Hummer, PhD**, chief of the theoretical biophysics division at the National Institute of Diabetes and Digestive and Kidney Diseases at the National Institutes of Health. “To a large degree it’s an active area of research where there are no generally accepted and generally applicable solutions.”

Miller agrees. “Spanning these big ranges of time in biological systems is the big challenge of the field,” he says. “A whole lot of people with a whole lot of good ideas are trying to address that challenge.” □

