Recipes for Replication

Novel supercomputing methods on Ranger help University of California-Irvine researchers understand DNA

At this very moment, cells in your body are replicating. Supercoils of DNA that code for mucus or muscles are unraveling to allow new strands to assemble from nucleotides.

But to do so, the DNA must first relax.

This is just one of the intriguing biochemical processes that Ioan Andricioaei, assistant professor of chemistry at the University of California-Irvine, studies, and that can only be revealed through high-resolution computational simulations. Using Ranger — the most powerful computing system in academia and among the most powerful supercomputers in the world — at the Texas Advanced Computing Center (TACC), Andricioaei performs virtual molecular experiments that recreate the conditions inside the nucleus and reveal the workings of the body.

Andricioaei’s interest in topoisomerase is far from academic. “A cell that cannot loosen its supercoils cannot replicate, and if you think about what are the fastest replicating cells, they are the cancerous cells. So if you can come up with a certain way to stop topoisomerase from relaxing supercoils, that particular cell will not be able to replicate,” Andricioaei explained. “That’s a very important anticancer strategy.” In fact, topotecan, an anticancer drug that uses just that technique to stop tumor growth, received Food and Drug Administration approval in 2007 and is currently used for cancer treatments.

Since DNA has a hard time unwinding on its own, a small enzyme, topoisomerase, wraps around the strands of DNA and makes a cut, permitting the helix to spin and relieve its tension. Then, once the DNA is relaxed, the same enzyme reconnects the broken strands and the genetic material can be replicated.

In the last two decades, high performance computing (HPC) has become an invaluable tool for biological research, helping scientists understand the underlying mechanisms of life’s processes, and using that knowledge to develop drugs and delivery systems to combat disease. As scientific interest has moved from bulk experiments to single-molecule, nanoscale studies, computational science has edged from the periphery of biological research to a more fundamental position in the field. Working hand-in-hand, experimentation and simulation are finding answers to important biological problems with greater speed than ever before.

Marcos Sotomayor and Klaus Schulten, in their recent review of laboratory and computational single-

For more information, please contact:
Faith Singer-Villalobos, Public Relations, faith@tacc.utexas.edu, 512.232.5771
Recipes for Replication

molecule experiments (Science Magazine, May 2007), noted the differences between the two forms of scientific discovery as they relate to biological research. “In vitro [experiments] enable the characterization of the mechanical response of biological matter at the nanometer scale. However, they do not reveal the molecular mechanisms underlying mechanical function,” they wrote. “These can only be readily studied through molecular dynamics simulations of atomic structural models: ‘in silico’ (by computer analysis) single-molecule experiments.”

Laboratory experiments can detect the dynamics of systems at about the millisecond (10^-3) timescale, but by harnessing the power of thousands of parallel processors to simulate the motions of each atom in a carefully modeled system, computational researchers can follow the dynamics of biological systems down to the femtosecond (10^-15 seconds) scale, showing the underlying molecular mechanisms.

Similarly, single molecule experiments can only show one or two parameters, like the end-to-end extension, at a time, “while in the simulation, we can see all the atoms at all the times,” Andricioaei explained.

In the complex simulations that he produces, that can mean tracking millions of interacting atoms — representing all the components of a system: water, solvent, ions, proteins, and DNA — over thousands of time-steps. “This requires very large scale computing where the molecular dynamics simulations are distributed among a large numbers of parallel processors.”

But large numbers of processors alone cannot solve the problem, and Andricioaei faces the same dilemma all molecular dynamics researchers confront: How to extend computer simulations to time-scales that are biologically relevant without compromising resolution?

Andricioaei developed an approach to this problem using stochastic path integrals — a method derived from probability theory — that he believes may help researchers create longer time-scale simulations in biology, as well as in other fields.

To explain the approach, Andricioaei offered an analogy. “Let’s say you want to search a range of mountains by foot. You’ll need a lot of energy to do this in the summertime, because you have to go over every little hill. But in wintertime, if you transform the landscape by letting snow accumulate, then you can just ski around and find the low conformations relatively easily,” he said.

This technique smooths out the details of the potential energy surface to let the supercomputer analyze...
Recipes for Replication

the entire problem more completely. “You lose the physicality, and the surface is different, but there are tricks that you can play to see through the snow and figure out what the kinetics should’ve been if you examined the ski traces on the original surface,” Andricioaei said.

By adding equations to solve the stochastic path integrals, he is able to reproduce the dynamics of the biomolecular interactions with particular effectiveness, producing all-atom simulations of biological lengths.

In Andricioaei’s case, this means watching topoisomerases snip the twisted DNA strands “like little pac-men.” Or, in another promising research project, learning how dendrimers — branched nanoparticles that may be the delivery system for future gene therapies — poke holes in membranes.

“These are things that the scientific community understands in broad strokes, but the more specific aspects are still a mystery,” he said. “We want to understand the fundamental mechanism that may aid in developing this understanding at a larger scale. So what we’re doing is vital to understanding and optimizing future methods and tools.”

By anticipating health breakthroughs, and helping fine-tune cancer drugs and treatment methods, Andricioaei’s research has the potential for significant biomedical impact. But the implications of his research may go even further.

“This work spurs new theoretical ways of thinking about things at the nanoscale, where the regimes are different from those in the bulk,” Andricioaei said. “Additionally, the path integral studies can draw in scientists working in unrelated areas. Stock markets are modeled with the same kinds of stochastic prob-

ability methods that we’re using to understand long-time dynamics, so, across disciplines, it can galvanize novel research.”

###

The Ranger supercomputer is funded through the National Science Foundation (NSF) Office of Cyberinfrastructure “Path to Petascale” program. The system is a collaboration among the Texas Advanced Computing Center, The University of Texas at Austin’s Institute for Computational Engineering and Science, Sun Microsystems, Advanced Micro Devices, Arizona State University, and Cornell University.

Ranger is a key resource of the NSF TeraGrid (www.teragrid.org), a nationwide network of people, resources and services, also sponsored by the NSF Office of Cyberinfrastructure, which enables discovery in U.S. science and engineering. The TeraGrid provides scientists and researchers expertise in and access to large-scale computing power, networking, data-analysis, and visualization systems.

Publications:


Recipes for Replication


Aaron Dubrow
Texas Advanced Computing Center
Science and Technology Writer
July 2, 2008