Simbios: an NIH national center for physics-based simulation of biological structures

Scott L Delp,[1] Joy P Ku,[2] Vijay S Pande,[3] Michael A Sherman,[2] Russ B Altman[4]

[1]Departments of Bioengineering, Mechanical Engineering, and Orthopaedic Surgery, Stanford University, Stanford, California, USA
[2]Department of Bioengineering, Stanford University, Stanford, California, USA
[3]Department of Chemistry, Stanford University, Stanford, California, USA
[4]Departments of Bioengineering, Genetics, and Medicine, Stanford University, Stanford, California, USA

Correspondence to
Professor Russ B Altman, Department of Bioengineering, 318 Campus Drive, MC 5444, Stanford University, Stanford, CA 94305-5444, USA; russ.altman@stanford.edu

Received 14 July 2011
Accepted 26 July 2011

ABSTRACT

Physics-based simulation provides a powerful framework for understanding biological form and function. Simulations can be used by biologists to study macromolecular assemblies and by clinicians to design treatments for diseases. Simulations help biomedical researchers understand the physical constraints on biological systems as they engineer novel drugs, synthetic tissues, medical devices, and surgical interventions. Although individual biomedical investigators make outstanding contributions to physics-based simulation, the field has been fragmented. Applications are typically limited to a single physical scale, and individual investigators usually must create their own software. These conditions created a major barrier to advancing simulation capabilities. In 2004, we established a National Center for Physics-Based Simulation of Biological Structures (Simbios) to help integrate the field and accelerate biomedical research. In 6 years, Simbios has become a vibrant national center, with collaborators in 16 states and eight countries. Simbios focuses on problems at both the molecular scale and the organismal level, with a long-term goal of uniting these in accurate multiscale simulations.

MISSION

The mission of Simbios is to develop, disseminate, and support a simulation toolkit (SIMTK) that enables biomedical scientists to develop and share accurate models and simulations of biological structures—from molecules to organisms. We have developed, tested, and released multiple versions of each of the components within SIMTK. The core of SIMTK includes high performance algorithms for performing matrix operations, generating and integrating equations of motion, performing linear and non-linear optimization, modeling contact between bodies, and calculating molecular interaction forces (figure 1).[1]

SIMTK has enabled the development of powerful graphics-based applications. For example, OPENSIM,[2 3] built from SIMTK, is focused on the simulation of human biomechanics, using LAPACK for linear algebra, SIMBODY for multibody dynamics,[4] IPOPT for optimization,[5] and other SIMTK components. OPENMM ZEPHYR, an easy-to-use interface for atomistic molecular dynamics, builds on SIMTK using GPU-accelerated molecular force-field calculations.[6–9] RNABUILDER[10–12] simulates coarse-grained models of large complexes of RNAs and proteins, making extensive use of LAPACK, SIMBODY, MOLMODEL, and other SIMTK components. Through SIMTK and the applications that use it, we have enabled thousands of researchers to perform accurate physics-based simulations of many different biological structures.

SIMTK has been developed and tested in close collaboration with hundreds of biomedical scientists to ensure its accuracy and utility. Our driving biological problems have included research projects in RNA folding,[13 14] protein folding,[15–18] myosin dynamics,[19] cardiovascular mechanics,[20–24] and neuromuscular biomechanics.[25–29] By choosing driving biological problems that represent important areas of research, our software innovations find broad applications.

There are two major complementary systems that make up SIMTK: a sophisticated open-source multibody mechanics code, SIMBODY, that forms the basis for modeling applications in biomechanics and molecular mechanics, and an interacting particle open-source code, OPENMM, that provides extremely fast force-field computations for large numbers of interacting components. These codes are based on state-of-the-art research innovations and are built and documented by experienced software engineering professionals who have developed and delivered complex software packages to thousands of users.

SIMBIOS ACHIEVEMENTS

Simbios has had a major impact on biomedical research by bringing physics-based simulation software to researchers and hospitals across the nation and the world. The highlights of our achievements include:

1. Becoming an international hub for physics-based simulation, with over 30 funded collaborations, and deep connections to other centers, the Physiome project,[30] and many individual NIH grantees (figure 2).
2. Producing a powerful simulation toolkit, SIMTK, which enables dynamic simulation of biological structures across a broad range of scales, from molecular to whole organism.
3. Enabling thousands of researchers and dozens of hospitals that are using application programs built from SIMTK.
4. Making new discoveries across a broad spectrum of biomedical science, developing valuable new computational methods, and publishing more than 150 articles.

Just as simulation has revolutionized other areas of science and engineering, Simbios aims to transform biomedical research by enabling advanced simulations of complex biological structures. We provide examples in three areas:

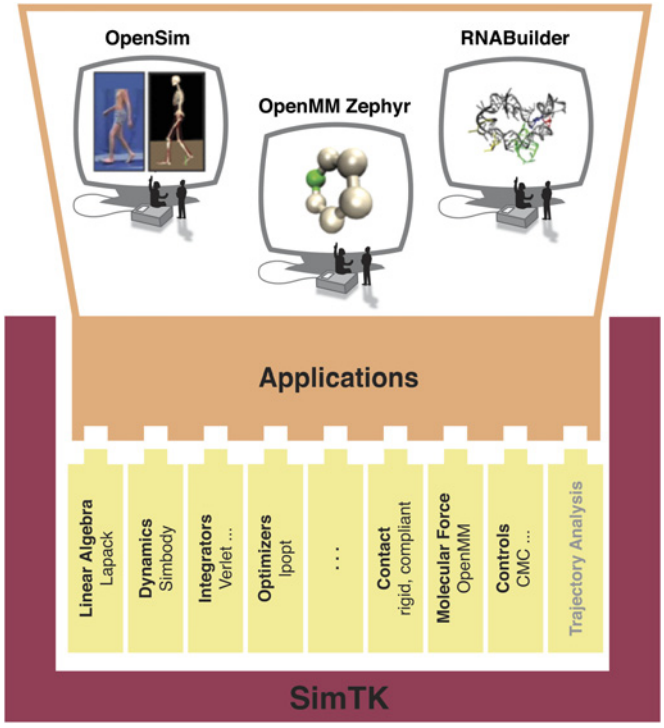

Figure 1 SIMTK is an open-source simulation toolkit serving a broad range of biosimulation domains. Shown are the major computational components (yellow) and some applications that are built from SIMTK components (OPENSIM, OPENMM ZEPHYR, and RNABUILDER). SIMTK and these applications are available for download at our http://simtk.org website as easy-to-install packages. Over 6000 unique users have downloaded these three applications, and over 3000 users have downloaded SIMTK components for use in their own biosimulation projects. There is a high degree of reuse of SIMTK components in the application software (eg, they all use the same linear algebra, dynamics engine, integrators). Over the next project period, we will extend SIMTK with new tools motivated by our driving biological projects, as suggested by the grayed-out 'Trajectory Analysis' component, and provide major advances of the other computational components.

1. **Infrastructure for simulating biological structural dynamics**: Simbios has created important simulation infrastructure, including (1) an open-source dynamics code, SIMBODY, that forms the basis for applications in biomechanical and molecular modeling, (2) an accelerated molecular dynamics code, OPENMM, that provides 100–1000× speed up in molecular force calculations, (3) a finite element package for simulating blood flow, SIMVASCULAR, and (4) a set of focused applications for molecular and macroscale simulation now used by thousands of biomedical researchers, including applications for RNA 3-D structure modeling (NUCLEIC ACID SIMULATION TOOL and RNABUILDER) and the simulation of movement (OPENSIM).
2. **Simtk.org website, infrastructure for software development and dissemination:** We established the site with 10 members in 2005. http://simtk.org has now grown to over 16 000 members and hosts more than 500 projects. Discussion forums, bug trackers, wiki support, mailing lists, daily backups, and dashboard reports on nightly builds are available for every project.
3. **Biomedical computation education:** Simbios has provided intense research training for 25 graduate students and 13 post-doctoral fellows at Stanford University. Simbios publishes a magazine, *Biomedical Computation Review*,[31] that has become a valuable vehicle to communicate the achievements and promise of biomedical computation to the broader scientific community. More than 600 researchers have benefited from over 20 Simbios workshops that provide hands-on training for SIMTK-based applications and SIMTK components.

ACTIVITIES AND GOALS

Biocomputational research

Physical simulation is one of the most computationally intensive activities in biocomputing, and therefore is highly dependent on advances in hardware technology. Recently, there has been a shift in hardware toward complex heterogeneous multicore architectures. This is not simply computing with graphics cards, but is a much more fundamental shift in how Moore's law of computing power will advance: clock rates have stopped improving, but transistors continue to get smaller and will be arranged in massively parallel arrays on special purpose hardware. We will take the lead in ensuring that biophysical simulation develops appropriately to use these new architectures and have engaged in a collaboration with the Stanford Pervasive Parallelism Laboratory to design 'domain specific languages' that will provide an application programmer interface that hides the complexity of programming these complex new architectures (figure 3).[32]

Our research is focused on defining fundamental data structures for simulation (particles, multibodies, and trajectories) and creating application programmer interfaces for manipulating these data structures on next-generation hardware. Our activities in core biocomputational research can therefore be summarized in four critical areas.

1. **Design and implement a family of domain-specific languages** with a common syntax that implement physical simulation operations as primitives and are implemented for maximal computational speed and development ease.
2. **Create advanced methods for multibody mechanical simulation**. We will extend the open source multibody code to provide high performance to a wider range of applications and include advanced contact modeling and controls.
3. **Create advanced methods for molecular dynamics**. We will extend the OPENMM framework to include support for symbolic representations of potentials, Markov state models to scale to billions of cores, enhanced free energy computations, and support for coarse-grain simulation.
4. **Create new methods for simulation trajectory analysis**. We will provide a set of tools for finding patterns and features of biological relevance in the trajectories that result from simulations of biological structures over relevant time frames. These will include new methods for clustering, classification, visualization, and modeling of simulation output.

Driving biological projects

There are two biological applications currently driving our biocomputational research: a neuroprosthetic dynamics project and a drug target dynamics project.

Neuroprosthetic dynamics

The long-term goal of this project is to develop arm prostheses for amputees that can be directly controlled by recordings of brain activity. Achieving this goal requires decoding motor intention from recordings of brain activity during complex movement patterns. Although most motor control research is done in highly constrained laboratory environments, understanding the control of complex movement in natural settings

Figure 2 The Simbios network of collaborators, including (yellow) institutions with collaborating R01 grants, (orange) institutions with Simbios seed grants, (red) co-authors on Simbios papers, contributors to Simbios software, co-investigators on related grants, and members of the Simbios Science Advisory Board. Collaborations are supported on http://simtk.org, a development and dissemination website that hosts more than 500 projects and supports over 16 000 members.

is critical for controlling prostheses in natural settings. Working with Dr Krishna Shenoy at Stanford, we will establish a freely moving animal model to directly measure the context-dependency of motor cortical activity.[33 34] Our model will include wireless transmission of neural data from electrode arrays chronically implanted in the brains of monkeys, and computer-vision algorithms and biomechanical models to automatically determine body and limb orientations during free movement over long periods of time. We will develop new mathematical and computational methods for extracting information from the high-dimensional neural and behavioral activity acquired to compare and contrast neural firing properties under different conditions. The integration of neuroscience, computer vision, and biomechanical modeling will enable the unprecedented study of motor control during natural behavior. This new paradigm will greatly enhance neuroscience investigations of motor control, advance neuroengineering studies aimed at designing high-performance neural prostheses, and improve the quality of life for individuals with physical disabilities by restoring lost motor function.

Drug target dynamics

The long-term goal of this project is to accelerate drug discovery and our understanding of drug side effects. Achieving this goal requires two major breakthroughs: improvement in our

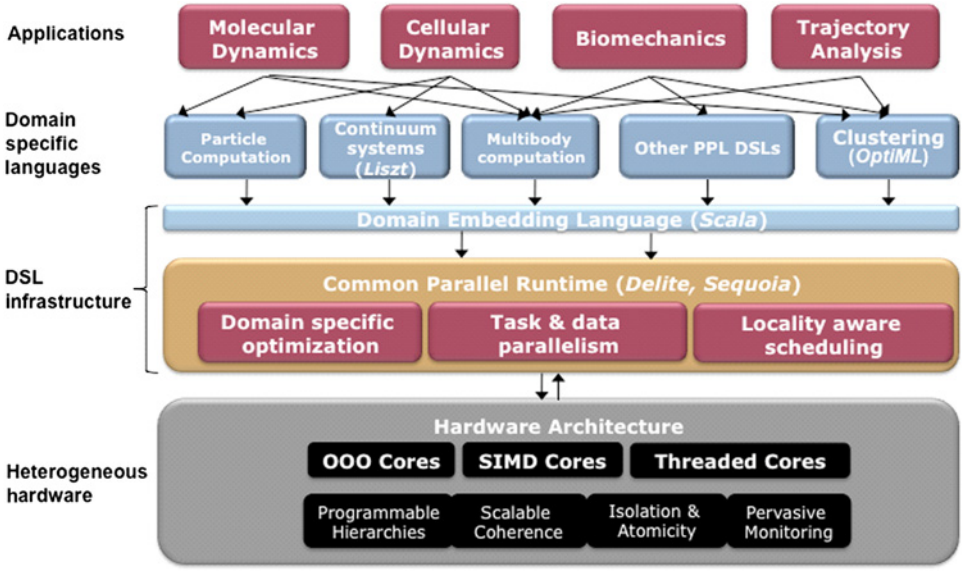

Figure 3 The Pervasive Parallelism Laboratory (PPL)—Simbios collaboration will define a family of domain specific languages (DSLs) related to physics-based simulation. Simbios programmers will build applications (top level) using the DSLs (second level). The DSLs provide an application programmer interface with a small set of primitives that are specific to the domain. DSLs are implemented in the PPL DSL infrastructure (third/fourth levels), which allows compilation and optimization for particular hardware (bottom level) so that the code runs efficiently. Application programmers currently program hardware directly, which is extremely difficult, expensive, and one-off. DSLs will provide an interoperable syntax, ease of programming, and high performance.

ability to model the interaction of a small molecule with a target and improvement in our ability to identify secondary targets that may mediate unexpected side effects. This project contributes to both problems by developing physics-based methods to improve drug-docking and modeling of unexpected drug–target interactions. Working with computational chemist Dr Brian Shoichet at U.C. San Francisco, Simbios collaborators will create physical models of the structure, function, and dynamics of G-protein coupled receptor proteins.[35–37] These proteins constitute approximately 50% of all drug targets and include recently solved 3-D structures for the β-2 adrenergic receptor, the adenosine A2A receptor, and others. These large, membrane-bound proteins are a critical molecular family. Understanding their structural dynamics should be key for understanding (1) how to inhibit or augment their function through small molecule interactions, and (2) how these molecules interact with intracellular G-proteins to trigger cell signaling cascades.

Infrastructure for physical simulation in biology

The rapid growth of our http://simtk.org user community requires that in the next 4 years we enhance the site's capabilities and create a sustainable model for support. Some of our software applications have achieved a critical mass of users who want to interact with and support one another. Thus, we will implement tools based on social networking to enable peer support. The development of virtual machines provides an opportunity for users of http://simtk.org to reproduce the results of others. Virtual machines provide a mechanism for an operating system, executable codes, and data sets to be preloaded into a binary format that represents an 'image' of a fully functional machine.[38] Our pilot studies have demonstrated the ability to deliver binaries, documentation, models, and full simulation trajectories.

VISION OF THE FUTURE

Although Simbios is broad in terms of the potential biological applications, it is not overly broad on the technical side. Our focus is on the biophysical dynamics of the structures we study and the physical context of biological processes. Our investigators start with Newton's 'F=ma' and work from there to generate models, degrees of freedom, equations of motion, integrators, trajectory analysis tools, and conclusions. This is the simple concept that unifies the center.

Over the initial 10-year life of Simbios, a worldwide community will generate an array of software, data, models, and simulations. These research resources will be available through http://simtk.org. SIMTK software will be embedded in software used by the biomedical research community. Physics-based modeling, simulation and data analysis is becoming fully integrated into the educational and professional activities of scientists, engineers, and clinicians. We will better understand the relationship between biological form and function, and we will know how to translate this understanding into improved human health.

Funding This work was supported by the National Institutes of Health through the NIH Roadmap for Medical Research grant U54 GM072970.

Competing interests None.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES

1. **Schmidt JP,** Delp SL, Sherman MA, *et al*. The simbios national center: systems biology in motion. *Proc IEEE Inst Electr Electron Eng* 2008;**96**:1266–80.
2. **Delp SL,** Anderson FC, Arnold AS, *et al*. OpenSim: open-source software to create and analyze dynamic simulations of movement. *IEEE Trans Biomed Eng* 2007;**54**:1940–50.
3. **Seth A,** Sherman MA, Reinbolt JA, *et al*. OpenSim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange. *Procedia IUTAM* 2011;**2**:212–32.
4. **Sherman MA,** Seth A, Delp SL. Simbody: multibody dynamics for biomedical research. *Procedia IUTAM* 2011;**2**:241–61.
5. **Wächter A,** Biegler LT. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math Program* 2006;**106**:25–57.
6. **Eastman P,** Pande VS. OpenMM: a hardware-independent framework for molecular simulations. *Comput Sci Eng* 2010;**12**:34–9.
7. **Eastman P,** Pande VS. Constant constraint matrix approximation: a robust, parallelizable constraint method for molecular simulations. *J Chem Theory Comput* 2010;**6**:434–7.
8. **Eastman P,** Pande VS. Efficient nonbonded interactions for molecular dynamics on a graphics processing unit. *J Comput Chem* 2010;**31**:1268–72.
9. **Friedrichs MS,** Eastman P, Vaidyanathan V, *et al*. Accelerating molecular dynamic simulation on graphics processing units. *J Comput Chem* 2009;**30**:864–72.
10. **Flores SC,** Altman R. Structural insights into pre-translocation ribosome motions. *Pac Symp Biocomput* 2011;**16**:205–16.
11. **Flores SC,** Altman R. Turning limited experimental information into 3D models of RNA. *RNA* 2010;**16**:1769–78.
12. **Flores SC,** Wan Y, Russell R, *et al*. Predicting RNA structure of multiple template homology modeling. *Pac Symp Biocomput* 2010;**15**:216–27.
13. **Jonikas MA,** Radmer RJ, Laederach A, *et al*. Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA* 2009;**15**:189–99.
14. **Jonikas MA,** Radmer RJ, Altman RB. Knowledge-based instantiation of full atomic detail into coarse grain RNA 3D structural models. *Bioinformatics* 2009;**25**:3259–66.
15. **Bowman GR,** Huang X, Pande VS. Network models for molecular kinetics and their initial applications to human health. *Cell Res* 2010;**20**:622–30.
16. **Voelz VA,** Bowman GR, Beauchamp K, *et al*. Molecular simulation of ab initio protein folding for a millisecond folder NTL9(1–39). *J Am Chem Soc* 2010;**132**:1526–8.
17. **Voelz VA,** Singh VR, Wedemeyer WJ, *et al*. Unfolded-state dynamics and structure of protein L characterized by simulation and experiment. *J Am Chem Soc* 2010;**132**:4702–9.
18. **Ensign DL,** Pande VS. The Fip35 WW domain folds with structural and mechanistic heterogeneity in molecular dynamics simulations. *Biophys J* 2009;**96**:L53–5.
19. **Liao JC,** Elting MW, Delp SL, *et al*. Engineered myosin VI motors reveal minimal structural determinants of directionality and processivity. *J Mol Biol* 2009;**392**:862–7.
20. **Les AS,** Shadden SC, Figueroa CA, *et al*. Quantification of hemodynamics in abdominal aortic aneurysms during rest and exercise using magnetic resonance imaging and computational fluid dynamics. *Ann Biomed Eng* 2010;**38**:1288–313.
21. **Figueroa CA,** Baek S, Taylor CA, *et al*. A computational framework for fluid–solid-growth modeling in cardiovascular simulations. *Comput Methods Appl Mech Eng* 2009;**198**:3583–602.
22. **Taylor CA,** Figueroa CA. Patient-specific modeling of cardiovascular mechanics. *Annu Rev Biomed Eng* 2009;**11**:109–34.
23. **Tang BT,** Fonte TA, Chan FP, *et al*. Three-dimensional hemodynamics in the human pulmonary arteries under resting and exercise conditions. *Ann Biomed Eng* 2011;**39**:347–58.
24. **Kung EO,** Les AS, Medina F, *et al*. In vitro validation of finite-element model of AAA hemodynamics incorporating realistic outlet boundary conditions. *J Biomech Eng* 2011;**133**:41003.
25. **Arnold EM,** Delp SL. Fibre operating lengths of human lower limb muscles during walking. *Philos Trans R Soc Lond B Biol Sci* 2011;**366**:1530–9.
26. **Fox MD,** Delp SL. Contributions of muscles and passive dynamics to swing initiation over a range of walking speeds. *J Biomech* 2010;**43**:1450–5.
27. **Steele KM,** Seth A, Hicks JL, *et al*. Muscle contributions to support and progression during single-limb stance in crouch gait. *J Biomech* 2010;**43**:2099–105.
28. **Reinbolt JA,** Fox MD, Schwartz MH, *et al*. Predicting outcomes of rectus femoris transfer surgery. *Gait Posture* 2009;**30**:100–5.
29. **Liu MQ,** Anderson FC, Schwartz MH, *et al*. Muscle contributions to support and progression over a range of walking speeds. *J Biomech* 2008;**41**:3243–52.
30. **Hunter P,** Robbins P, Noble D. The IUPS human physiome project. *Pflugers Arch* 2002;**445**:1–9.
31. *Biomedical Computation Review Magazine*. http://biomedicalcomputationreview.org/ (accessed 13 Jul 2011).
32. **Gosh D.** DSL for the uninitiated. *Commun ACM* 2011;**54**:44–50.
33. **Gilja V,** Chestek CA, Nuyujukian P, *et al*. Autonomous head-mounted electrophysiology systems for freely-behaving primates. *Curr Opin Neurobiol* 2010;**20**:676–86.
34. **Foster JD,** Freifeld O, Nuyujukian P, *et al*. *Combining wireless Neural Recording and Video Capture for the Analysis of Natural Gait, in IEEE EMBS Conference on Neural Engineering*. Cancun, Mexico: IEEE, 2011:613–15.
35. **Carlsson J,** Yoo L, Gao ZG, *et al*. Structure-based discovery of A2A adenosine receptor ligands. *J Med Chem* 2010;**53**:3748–55.
36. **Kolb P,** Ferreira RS, Irwin JJ, *et al*. Docking and chemoinformatic screens for new ligands and targets. *Curr Opin Biotechnol* 2009;**20**:429–36.
37. **Kolb P,** Rosenbaum DM, Irwin JJ, *et al*. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc Natl Acad Sci U S A* 2009;**106**:6843–8.
38. **Dudley JT,** Butte AJ. In silico research in the era of cloud computing. *Nat Biotechnol* 2010;**28**:1181–5.

Simbios: an NIH national center for physics-based simulation of biological structures

Scott L Delp, Joy P Ku, Vijay S Pande, et al.

JAMIA published online November 10, 2011
doi: 10.1136/amiajnl-2011-000488