All-atom Simulations of Biological Function: the impossible dream

Bob Eisenberg
October 30, 2013

Finite numbers are sometimes so large that they seem infinite. But they are not. Computers are so powerful nowadays[§], that their power seems infinite too. It is not.[2, 3]

Computers can calculate motions of all the atoms of proteins, a capability celebrated recently in a Nobel Prize to Warshel, Levitt, and Karplus. But many important biological functions (produced by proteins) cannot be computed in an all-atom simulation. We need to remember that, even as we celebrate what can be done so well.

The nerve signal, for example, is an important biological function called the action potential. The action potential carries information as an all-or-none binary electrical signal over long distances (say 0.001 m to 10 m) in nerve cells, mostly nerve axons. The electrical potential is a waveform in time and distance. It is a propagating wave that forms the neuronal signal. Currents drive the action potential. The electric currents are carried by two types of ions—sodium and potassium ions—as they move through two types of channel proteins that sit in nerve membranes. The channels provide the currents that generate the nerve signal. The electrical voltage that is the nerve signal helps control channels, including sodium and potassium channels. The channels generate the voltage; the voltage helps control the channels. Macroscopic electric fields interact with atomic structures to make the nerve signal.

Sodium is abundant outside cells. Potassium is abundant inside cells. One channel type selectively catalyzes[4] the flow of sodium ions across the otherwise insulating lipid membrane. The other channel type selectively catalyzes the flow of potassium ions. These channels open in response to changes in the electrical potential across the voltage sensors of the channels, one after the other, loosely speaking. The gradients of concentration, electrical, and electrochemical potential for sodium drive sodium ions (and their electrical current) inward through the sodium channel, when it opens. The inward sodium current provides the energy for propagation of the action potential waveform. The different gradients of concentration, electrical and electrochemical potential for potassium drive potassium ions outwards when the potassium channel opens. The outward potassium currents helps promptly terminate the nerve signal so another signal can soon follow.

The electric current carried across the membrane by these channels spreads down the length of axons the way electric current flows down transmission lines. Both follow Kelvin's sub-marine cable equations,[5, 6] worked out to describe current flow in telegraph cables under the ocean. These equations show that the potential at one location depends on the current flow from other locations millimeters to centimeters away. The potential at one location depends on the potential

[§] Computer power is some 10^9 greater than in 1965, when Moore[1] feared it could not continue to increase very much longer.

at other locations, a fact expected in theory and easy to verify in experiments. Some 10^{15} ions interact this way, in a 0.3 mm diameter nerve fiber (from the squid), with action potential wave length of 20 mm, if the action potential duration is 1 msec with velocity 20 m/sec.

An all-atom computation of 10^{15} ions is not possible, if each ion interacts with all the others through the electric field, even if the induced polarization charge of water is neglected. The electric field induces (voltage-dependent) polarization charge in the lipid membrane. That charge produces forces that act on every ion. In that way, every ion is coupled to every other ion, producing a staggering number of interactions, far beyond the number of pairwise interactions. Everything interacts with everything else, through the induced charge on the lipid membrane and induced charge at any other place where polarization charge varies with location and electrical potential. The number of computations needed to compute all the interactions—not just pairwise—of 10^{15} interacting ions is beyond astronomical. Indeed, computing just the pairwise interactions of 10^{15} ions every 10^{-15} seconds is daunting.

An all-atom simulation of a nerve signal is not needed, which is just as well, since it seems impossible. A multiscale analysis will do all we need. The multiscale description of the action potential is known, thanks to the work of Hodgkin, Huxley, and Cole,[7, 8, 9] also recognized in a Nobel Prize (to two of them). The atomic description of nerve propagation is not needed. The atomic scale description of current through single protein molecules—individual sodium and potassium channels—is nearly enough, because the interaction of channels is described accurately[7, 8, 9] by Kelvin's cable equation, usually called the transmission line equation in the mathematics literature today.

Current through channels is measured every day in hundreds or thousands of laboratories. The voltage clamp technique of Cole and Hodgkin measures the properties of ensembles of noninteracting channels. That is why the technique was invented. Single channel recording allows measurement of current through one channel protein at a time, as celebrated by the Nobel Prize to Neher[10] and Sakmann.[11] The voltage and time dependence of ensembles of noninteracting channels is simple. An action potential can be reconstructed by summing the currents through single channels themselves[11] and then using the summed current in an intermediate scale equation,[7, 8, 9] the transmission line equation.

A multiscale analysis succeeds in this way. It links atomic scale currents through single channels with the biological reality of the action potential (waveform), the nerve signal. The multiscale equation of transmission lines is all that is needed to go from ions moving through single protein molecules to the macroscopic biological function of the action potential waveform in space and time.

Ions moving through single channels are an obvious candidate for all-atoms simulations if induced polarization charges (at dielectric boundaries) are not too important. Such simulations may be possible, but are beyond our immediate reach[12] if we wish to compute single channel currents as they actually have been studied in hundreds or thousands of laboratories.[13]

Experiments show that properties of individual sodium and potassium channels depend sensitively on the solutions around them. These solutions are invariably ionic mixtures (made

mostly of Na, K, Ca, and Cl ions) derived from roughly half molar seawater. Seawater is highly concentrated (599 mM) compared to the dilute solutions (50 mM or much less depending on the criteria) in which either NaCl or KCl have the ideal properties of an infinitely dilute perfect gas.[14, 15, 16]

Atomic detail computations of ionic mixtures containing calcium ions pose certain challenges for all-atom simulations. All-atom calculation of trace concentrations of intracellular calcium ions is a particular challenge since ~55 moles of water must be calculated for each 10^{-7} moles of calcium ion in a 10^{-7} molar calcium solution. Calibration against the main physical properties of these mixtures (the free energy per mole, i.e., the electrochemical potential) is of course part of the calculation.

The calibration of the free energy (per mole) of calcium must be accurate because changes in this concentration serve as control signals in many biological systems. The resting concentration of calcium inside cells is typically 0.0000001 M. Four to ten-fold increases in calcium concentration control many biological systems. The calcium concentration determines the speed of many chemical reactions in cells the way a gas pedal controls the speed of a car. The trace concentration of calcium must be simulated accurately in biology, just as changes in the location of the gas pedal must be known in simulations of cars, to understand their speed.[13, 17] In addition, many channels become 'sick' (i.e., change properties pathologically sometimes irreversibly) if the concentration of calcium ions is changed too much, particularly on the inside of the channel. If the concentration of calcium is changed too much, many channels disappear from experiments measuring current, and are dead to the world, as far as those measurements are concerned, and cannot be studied. Calibrated simulations should reproduce all the properties of the systems they simulate. So channels should 'disappear' in simulations in those conditions in which they disappear in experiments. Uncalibrated simulations are likely to have 'invisible' (nonfunctional) channels or sick channels because simulations are likely to reproduce calcium concentrations incorrectly. Incorrect calcium concentrations would make the simulated channel nonfunctional, or sick, as they would in experiment. Uncalibrated simulations are likely to have different (effective) calcium concentrations each time the simulation is computed. Uncalibrated simulations are likely to be irreproducible, and give different results in different laboratories. Much confusion would result from these different results and progress would be slow.

Computations in atomic detail must also face the well-known problem of time scales. Time scales of action potentials range from milliseconds (for nerve) to nearly seconds (for the heart) and so all-atom simulations must be computed for milliseconds to seconds. The time step of all-atom simulations is typically 10^{-15} sec. Differential equations must be integrated for something like 10^{13} steps, producing certain difficulties of accuracy and reproducibility. Reproducibility is difficult in simulations like these of chaotic systems. Individual trajectories are extremely sensitive to initial conditions in chaotic systems. Chaotic systems have different types of behavior depending on where the system starts. Simulations starting with one set of conditions miss some of the behaviors. Convergence of simulations to stable results is not guaranteed. These problems limit simulations no matter how quickly they can be done.[2, 3]

The spatial scale of the action potential is millimeters to centimeters or more. All-atom simulations must calculate atoms from 10^{-10} m to action potentials that spread millimeters—and

propagate meters—if the simulation is to generate the nerve signal from an all atom simulation of ions moving through channels

All-atom simulations of biological systems have to be done away from equilibrium, in systems in which flows are involved in biological function. Nerve signals do not occur at equilibrium because nerve signals use flows of current through single channels to generate the propagating action potential.[7, 8, 9] These flows cannot occur in equilibrium systems (by definition). Equilibrium analysis (with a single solution on both sides of a channel) can only reproduce a dead system without a nerve signal at all. Many biological systems are at equilibrium only when they are dead.

We need simulations of live systems with different concentrations of ions inside and outside cells. We need to calculate ion flows driven by the different electrical, chemical, and electrochemical potentials found inside and outside cells. Simulations of 'live' diodes and 'triodes' (both bipolar and field effect transistors) are done every day in hundreds or thousands of laboratories. Techniques used in these simulations of computational electronics may be useful in all-atom simulations of biological systems.[18]

All-atom simulations of flows through ion channels may be impossible. They may not be necessary, either. A multiscale analysis of ion flows through channels may be enough and much simpler. Multiscale analysis can use the laws of electricity, and of electrodiffusion, to simplify the simulation and handle induced polarization charge at dielectric boundaries.

Certainly, the laws of electrodiffusion and electricity are not enough by themselves. Atomic detail will be crucial in some parts of the system, and atomic scale simulation of those parts of the system will be needed. Changes in a handful of atoms are enough to dramatically change the properties of single channels. The magnificent cathedrals of channel structure—recognized in the Nobel Prize to Mackinnon—have architectural detail that controls biological function. But full atomic detail is not needed everywhere to calculate biological function, in all likelihood.[19, 20]

All-atom simulations of ion channels remain an admirable dream, whether impossible or not. They form a productive, inspiring goal, ***so long as the simulations are calibrated.*** Simulations must actually reproduce the experimental properties of ions in mixtures, in bulk and in channels, as measured in the laboratory, if they are to help biologists and physicians in their everyday work. Uncalibrated, simulations can turn into dreadful nightmares, producing results irrelevant to biological function, wasting lifetimes of scientists trying to compute the uncomputable.

I believe multi-scale analysis[19, 20, 21] of simplified models will be needed, motivated and focused by what all-atom simulations cannot do.

References

1. Moore, G. E. Cramming more components onto integrated circuits. *Electronics Magazine.* **1965,** *38*, 114–117.
2. Maginn, E. J. From discovery to data: What must happen for molecular simulation to become a mainstream chemical engineering tool. *AIChE Journal* **2009,** *55* (6), 1304-1310.
3. Post, D. E.; Votta, L. G. Computational Science Demands a New Paradigm. *Physics Today* **2005,** *58*, 35-41.
4. Eisenberg, R. S. Channels as enzymes: Oxymoron and Tautology. *Journal of Membrane Biology* **1990,** *115*, 1–12. Available on arXiv as http://arxiv.org/abs/1112.2363.
5. Jack, J. J. B.; Noble, D.; Tsien, R. W. *Electric Current Flow in Excitable Cells*; Oxford, Clarendon Press.: New York, 1975.
6. Gabbiani, F.; Cox, S. J. *Mathematics for Neuroscientists*; Academic Press: New York, 2010.
7. Huxley, A. Sir Alan Lloyd Hodgkin, O. M., K. B. E. 5 February 1914-20 December 1998. *Biographical Memoirs of Fellows of the Royal Society* **2000,** *46*, 221-241.
8. Huxley, A. F. Kenneth Stewart Cole. *Biographical Memoirs of Fellows of the Royal Society* **1992,** *38*, 98-110 , see http://books.nap.edu/html/biomems/kcole.pdf
9. Huxley, A. F. From overshoot to voltage clamp. *Trends in neurosciences* **2002,** *25* (11), 553-558.
10. Neher, E. Ion channels for communication between and within cells Nobel Lecture, December 9, 1991. In *Nobel Lectures, Physiology or Medicine 1991-1995*, Ringertz, N., Ed.; World Scientific Publishing Co: Singapore, 1997, pp 10-25.
11. Sakmann, B.; Neher, E. *Single Channel Recording.*; Second ed.; Plenum: New York, 1995. p 700.
12. Eisenberg, B. Multiple Scales in the Simulation of Ion Channels and Proteins. *The Journal of Physical Chemistry C* **2010,** *114* (48), 20719-20733.
13. Hille, B. *Ionic Channels of Excitable Membranes*; 3rd ed.; Sinauer Associates Inc.: Sunderland, 2001. p 1-814.
14. Fawcett, W. R. *Liquids, Solutions, and Interfaces: From Classical Macroscopic Descriptions to Modern Microscopic Details*; Oxford University Press: New York, 2004. p 621.
15. Laidler, K. J.; Meiser, J. H.; Sanctuary, B. C. *Physical Chemistry*; Fourth ed.; BrooksCole, Belmont CA2003. p 1060.
16. Eisenberg, B. Interacting ions in Biophysics: Real is not ideal. . *Biophysical Journal* **2013,** *104*, 1849-1866.
17. Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular Biology of the Cell*; Third ed.; Garland: New York, 1994. p 1294.
18. Vasileska, D.; Goodnick, S. M.; Klimeck, G. *Computational Electronics: Semiclassical and Quantum Device Modeling and Simulation*; CRC Press: New York, 2010. p 764.
19. Gillespie, D. Energetics of divalent selectivity in a calcium channel: the ryanodine receptor case study. *Biophys J* **2008,** *94* (4), 1169-84.
20. Roux, B. Ion binding sites and their representations by reduced models. *The journal of physical chemistry. B* **2012,** *116* (23), 6966-79.

21. Eisenberg, B.; Hyon, Y.; Liu, C. Energy Variational Analysis EnVarA of Ions in Water and Channels: Field Theory for Primitive Models of Complex Ionic Fluids. *Journal of Chemical Physics* **2010,** *133,* 104104