

Birth and Future of Multiscale Modeling for Macromolecular Systems (Nobel Lecture)**

Michael Levitt*

coarse-grained models · hybrid QM/MM models ·

molecular potential energy · Nobel lecture ·

protein dynamics in water

Introduction

Winning the Nobel Prize is a unique and marvelous experience that no one can prepare for or in any way know what to expect. The instantaneous transformation from an ordinary human, toiling away to solve the problems that come before us, into being a symbol, a celebrity, is a remarkable phenomenon. On the one hand, a mature person is likely to be pretty happy with the way they have been living until the moment of transformation and thus wants things to continue as they were before. On the other hand, any scientist appreciates just how important role models were for their entire career and thus wants to continue the tradition and be just such an example for future generations. This is a quandary that is with me now and is likely to require decades to solve.

The Nobel Lecture is different from other lectures in that it combines past, present, and future and is given to a diverse audience ranging from the interested school child to the expert colleague. Writing such a lecture down tends to follow the centuries-long tradition of scientific paper writing that can miss some of the freshness of the actual lecture. Faced with the challenge, I have decided to base this written lecture closely on my Nobel talk, using the slides as the Figures. The main text provides a simple narrative, while the figure legends facilitate more-specific comments and discussions.

Stand on the Shoulders of Giants

An obvious requirement for doing ground-breaking work that comes to fruition decades later—Nobel-Prize-winning research—is to start off on high ground and climb onto the shoulders of giants so as to see as far as possible into the future. In my case, these giants had discovered a new way to think about all of biology, a way that lent itself to computer modeling on many scales.

Francis Crick (Figure 1) was easy to appreciate as being a brilliant scientist with a passion for science and indeed life in general. Thinking back to my earliest memories of our encounters, I cannot help but be impressed by the fact that he owned a fancy sports car, a yellow Lotus Elan. What I think was most surprising about this is how it enabled me as a 21 year boy to relate to the obvious boy in him.

A few years after Crick and Watson solved the structure of DNA, John Kendrew (Figure 2) determined the three-dimen-

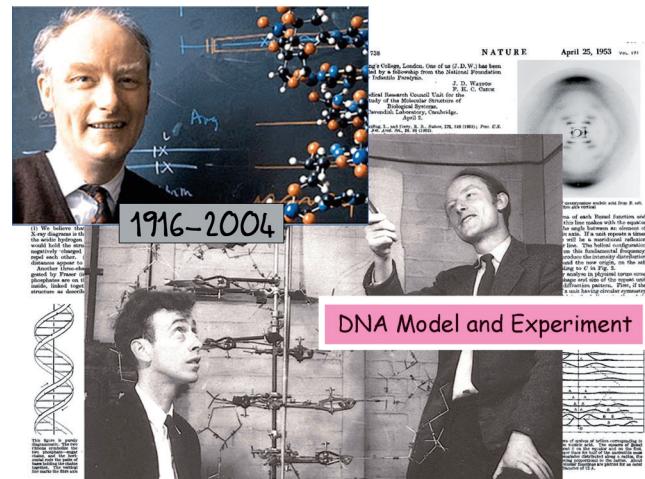


Figure 1. Francis H. C. Crick may well be one of the two or three best known scientists of the 20th century, a period that seems to have overflowed with great minds who changed the course of human thought. His contributions to our story are many and varied. I met Crick in 1968 aged 21 and worked with him closely for the following decade. He taught me to think carefully by asking and then trying to answer simple questions. With James Watson, Crick used molecular modeling to combine diverse data into a three-dimensional structure of DNA that proved to be sufficiently correct so as to explain how genetic information is kept error-free when copied. Their ability to combine partial data from many source to give a correct answer seemed like magic.^[1,2] It provided a paradigm that all non-experimental theoretical structural biologists would aim to imitate for the next 60 years.

sional structure of a protein, in this case myoglobin isolated from whale muscle, readily available back then. The approaches of Crick and Kendrew to determining the three-

[*] Prof. M. Levitt

Stanford University School of Medicine
Department of Structural Biology
299 West Campus Drive, Stanford, CA 94305 (USA)
E-mail: michael.levitt@stanford.edu

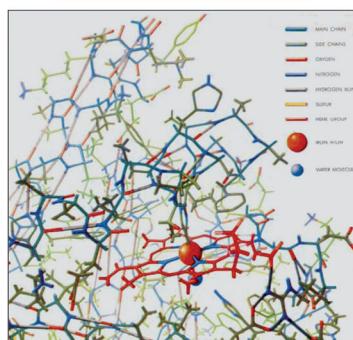
[**] Copyright © The Nobel Foundation 2013. We thank the Nobel Foundation, Stockholm, for permission to print this lecture.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201403691>.



1917–1997

Scientific American 1961

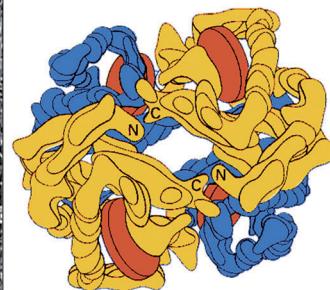


Painted by artist Irving Geis

First protein X-ray structure.



1914–2002



The REAL HERO of structural biology.

Figure 2. John C. Kendrew used X-ray crystallography to solve the three-dimensional structure of the protein myoglobin.^[3] This structure as present on the cover of *Scientific American* in 1961^[4] was drawn from a wire model hand-built to fit the electron density by the artist Irving Geis. It showed a complex structure built from 153 amino acids and over 2600 atoms with a precise three-dimensional shape that seemed to be determined by the forces between the atoms. This shape seemed to explain how the heme group shown in red could store oxygen in whale muscle setting the stage for molecular biology, where molecular function depends on structure in a precise manner. The fact that biology works like a clock, made our work possible.

dimensional shapes of biomolecules could not have been more different. Kendrew replaced Crick and Watson's brilliant inspiration with a painstaking method, which could be applied to any protein that could be crystallized. The method was invented by Kendrew's PhD supervisor, Max Perutz (Figure 3), who also supervised Francis Crick and was the leader of the lab where they all worked together in Cambridge. The method, known as Heavy Atom Replacement,^[5] is what made crystallographic protein structure determination possible and applicable broadly. For this Perutz shared the 1962 Nobel Prize in Chemistry with Kendrew and their work led to the explosive growth of protein three-dimensional structures from one structure in 1959 to almost one hundred thousand structures today, 55 years later.



Michael Levitt was born in Pretoria in 1947. He studied Physics at King's College London, and got his Ph.D. in Biophysics from the MRC Laboratory of Molecular Biology and Cambridge University. After having worked at the Weizmann Institute, the MRC Laboratory of Molecular Biology, and the Salk Institute, he became Professor of Structural Biology at the Department of Structural Biology, Stanford University School of Medicine in 1987. Today he is the Robert W. and Vivian K. Cahill Professor of Cancer Research. To the honors he had received before being awarded the Nobel prize belong the Federation of European Biochemical Societies Anniversary Prize, being elected as a Fellow of the Royal Society and becoming a member of the US National Academy of Science.

© Linda Cicero

Michael Levitt was born in Pretoria in 1947. He studied Physics at King's College London, and got his Ph.D. in Biophysics from the MRC Laboratory of Molecular Biology and Cambridge University. After having worked at the Weizmann Institute, the MRC Laboratory of Molecular Biology, and the Salk Institute, he became Professor of Structural Biology at the Department of Structural Biology, Stanford University School of Medicine in 1987. Today he is the Robert W. and Vivian K. Cahill Professor of Cancer Research. To the honors he had received before being awarded the Nobel prize belong the Federation of European Biochemical Societies Anniversary Prize, being elected as a Fellow of the Royal Society and becoming a member of the US National Academy of Science.

Figure 3. Max F. Perutz is shown working on his brass-wire model of hemoglobin, which is four times bigger than myoglobin; its structure was also much harder to solve as the crystals were not sufficiently ordered. Thus, at a time when Kendrew knew where every atom was in myoglobin, Max Perutz had to be content with the balsa wood model on the right showing the general shape of the four globin chains.^[6] With PhD students like Crick and Kendrew, it was Perutz who established the field of structural biology. He was a wonderful leader and a warm, clever human being who knew how to get the best out of all who worked with him.

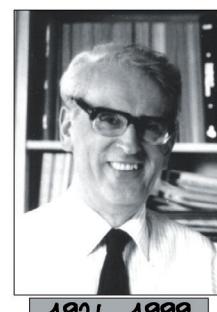
Another important influence on my career was the biophysicist David Phillips from Oxford. He solved the first enzyme structure, that of the protein lysozyme, in 1966, and like Kendrew published this in *Scientific American*, an accessible and highly desirable journal with its color figures (Figure 4). Lysozyme is an enzyme, a protein that can catalyze

SCIENTIFIC AMERICAN

ENZYME MOLECULE

SIXTY CYTOS

November 1966



1924–1999

Early supporter of Computational Biology

Figure 4. David C. Phillips used X-ray crystallography to solve the three-dimensional structure of the enzyme lysozyme.^[7] This structure allowed him to model the substrate in the active site of the enzyme and to speculate about the nature of enzyme action. He proposed that the six-membered sugar ring was distorted by its steric interaction with the active site.^[8] This was wrong but led to the development, with Arieh Warshel, of hybrid QM/MM quantum/classical models showing that the strain was electrostatic not steric.

a reaction, the cleavage of the sugar chains that provide the armor around bacteria. Together with myoglobin, lysozyme features prominently in setting the stage for the future of computation in structural biology (see below).

Another giant of that period on whose shoulders we stood and still stand is Linus Pauling, who in 1951 correctly predicted the structure of the α -helix and the β -sheet, the two major modules reused in many different protein structures. I did not know Pauling until much later, but in 1990 did have the pleasure and privilege of lecturing to him about the simulation of α -helix dynamics in water and showing him a movie of how the α -helix comes apart at high temperature.

The Birth of Computational Structural Biology

In 1967, there were two seemingly different raging torrents of scientific discovery and technological advances. Scientific discovery had revealed the X-ray structures of myoglobin (Figure 2) and lysozyme (Figure 4), which showed that the molecules carrying out all the key functions of living systems have incredibly complicated structures. Their detail is not something baroque or incidental, rather it is essential for carrying out crucial biological functions. Technological advances had shown that computers could be flexibly programmed to carry out all manner of calculations. These machines were just becoming commercial and developments were proceeding rapidly. Computational structural biology was born when these two torrents joined in a huge and powerful stream that is still propelling the field forward almost 50 years later.

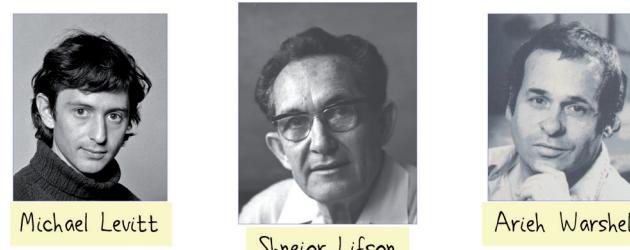
Like many interesting events in history this occurred by a rare coming together of three individuals with very different talents, backgrounds, and approaches. Even more remarkable, another individual was responsible for this meeting and planned it carefully. It started with a philosophical idea concerning the nature of the model used to represent a molecule. The man who had this idea was Shneior Lifson, a professor of Chemical Physics at the Weizmann Institute in Rehovot, Israel. He argued that the energy function and its first derivative, the force field, had to be consistent. This meant that there should be a small number of atom types for each element and that the energy parameters should not depend on the local environment of the atom. For example, there could be two types of carbon, aromatic and aliphatic, but once this distinction had been made, the same parameters should define the energy of the carbon atom. This consistency means that there is a small number of parameters and that these parameters are transferable from one situation to another.

Implementing this idea was not simple. One needed to first compute diverse properties of small molecules including their geometry, their strain energy, and their vibration frequencies, next compare these calculated values with the corresponding measured experimental values, and then finally change the parameters to get the best agreement between calculated and measured properties. The implementation was designed by the second person, Ariele Warshel, Lifson's PhD student, who also decided which systems to study and which properties to calculate. I arrived on the scene

in October 1967 aged 20 and just as this work was gearing up (Figure 5). My initial role as the third person was to be their computer programmer, writing a program to calculate the potential energy, its first derivative, the force vector, and its second derivative, the curvature of the energy surface.

CONSISTENT FORCE-FIELD 1968

Weizmann Institute 1967-1968

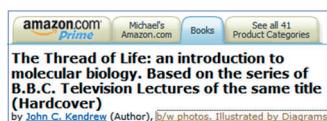


THE JOURNAL OF CHEMICAL PHYSICS VOLUME 49, NUMBER 11 1 DECEMBER 1968
Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules
S. LIFSON AND A. WARSHEL
Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel
(Received 13 May 1968)

Figure 5. In 1967, John Kendrew insisted that I spend a year with Shneior Lifson in Israel before I would be allowed to begin a PhD in Cambridge. Arriving in Israel in October 1967, I met Shneior Lifson and his PhD student Ariele Warshel and began a journey that would eventually bring me to Stockholm. The key to the work that led to the Nobel Prize was Lifson's philosophical concept that is known as the "Consistent Force Field". Energy calculations had been done on small molecules mainly for the purpose of calculating vibrational spectra. In these calculations there were energy parameters that described the force between atoms but the forces were not consistent in that different parameters were used for the same atom type, e.g. carbon, in different environments. Lifson wanted there to be very few atom types and to have the different energy terms define the influence of the environment. Using computer programs that I wrote, Ariele Warshel was able to define a consistent set of parameters for a series of small organic hydrocarbon molecules as published in their 1968 landmark paper.^[9]

This occurred remarkably quickly and within six months useful calculations were being run on the very powerful Golem A computer at the Weizmann Institute. Golem A was a home-built, second-generation machine that followed on from the Weizac built in the mid 1950s using the architecture developed by John von Neumann at the Institute for Advanced Study in Princeton. The Golem A was in operation from 1964 to 1974 and had a memory capacity of 32768 words of 75 bits (ca. 300000 bytes). It was programmed in the FORTRAN language with programs written on punched cards.

One man, John Kendrew brought this unlikely trio (Lifson, Warshel, and Levitt) together and he did it with remarkable foresight. As mentioned above and in Figure 2, Kendrew shared the 1962 Nobel Prize in Chemistry with Max Perutz. About a year later, Kendrew delivered a series of lectures on BBC television (Figure 6) that caught my attention as a 17 year old boy just arrived in London. The new discoveries in what was termed "molecular biology" were so exciting that I decided to study Physics at King's college in



Nobel Prize in 1962
Gave TV Series in 1964
Sent me to Israel in 1967

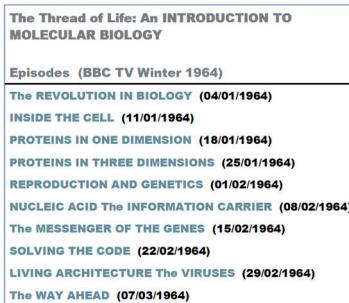


Figure 6. John Kendrew had a greater influence on my career than anyone else but this influence was indirect. One year after winning the 1962 Nobel Prize, Kendrew wrote and presented a BBC television program entitled the Thread of Life. I had arrived from South Africa two months before the program began to be aired on January 4th, 1964. I was living with my aunt and uncle, both scientists, in London and had never seen TV before. Although the screen was small, the resolution low and the color more black and yellow than black and white, I was immediately addicted. Thankfully, I got to watch Kendrew's program which no longer exists and got the most amazing introductory course in molecular biology imaginable. The topics dealt with could be the backbone of a modern course in molecular biology, starting with "The REVOLUTION IN BIOLOGY" (on January 4, 1964) and ending with "The WAY AHEAD" (on March 7, 1964).

London, home to Maurice Wilkins, who shared the 1962 Nobel Prize with Crick and Watson, and where there was a third-year biophysics option. In 1967, towards the end of my BSc degree I applied to Kendrew and Perutz to do a PhD at the Medical Research Council Laboratory of Molecular Biology in Cambridge but they turned me down for lack of space. Persuaded by friends (who went on to be very successful businessmen), I asked to be considered for 1968. This time they invited me for an interview but their decision to consider me in 1968 left me at a loose end. Again my friends worked on me and I drove up to Cambridge, accosted Max Perutz in the corridor, and when he agreed to discuss my case with Kendrew, I beat a hasty retreat. I was overjoyed when I heard few days later that I had definitely been accepted for 1968. Kendrew went on to insist that I spend the intervening year with Lifson at the Weizmann Institute, and made his suggestion very attractive by getting me, who had just finished his BSc, a Royal Society Exchange postdoctoral fellowship at the Institute.

The consistent force field description of the potential energy function of a molecule (Figure 7) is very powerful as it can be used to compute all the properties of any molecular system by a combination of the methods shown in Figure 8. Relying on the transferability of the energy parameters, I realized that although Lifson and Warshel had not included amino acids in their parameter determination, and indeed had determined energy parameters for only two (H and C) of the four (H, C, N, O) atom types that occur commonly in amino

MOLECULAR POTENTIAL ENERGY

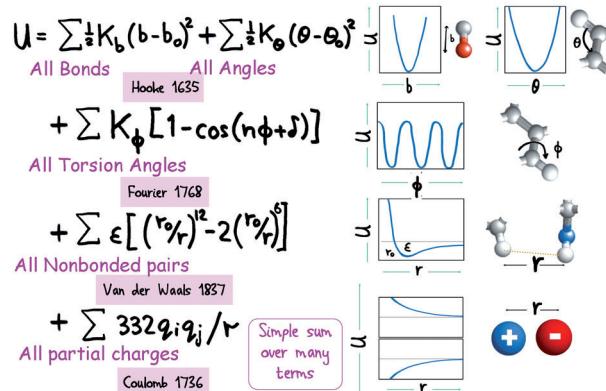


Figure 7. The energy function of any molecule is classical both in that it does not use quantum mechanics and also because it relies on a classical description of the molecule as a collection of balls connected by springs. The terms shown here have been used with little alteration since 1970. They account for bond stretching and bond angle bending as harmonic springs. Both degrees of freedom b and θ have an equilibrium value given by the energy parameters b_0 and θ_0 , respectively. The potential energy of a single bond or bond angle increases if the bond (or angle) is compressed or extended. The stiffness of the spring is given by other energy parameters, K_b and K_θ . The other energy terms are a little more complicated but they follow the simple bond and angle terms in that they depend on the types of interacting atoms and each interaction contributes to the total potential energy in a simple additive fashion. Different terms use different energy parameters, which must be determined by least-squares refinement of calculated molecular properties against those observed. Lifson and Warshel started this process in 1968 and it is still used to refine the most-modern classical molecular potential energy functions. The newest force fields are based on high-order quantum calculations^[10] rather than experimental data.

acids, their methods could be extended. This made me start to do calculations on protein molecules that had many hundreds of atoms compared to the few tens of atoms in the molecules studied by Warshel and Lifson.^[9] My idea was to energy-minimize the atomic structure of an entire protein by moving the atoms in Cartesian coordinates (x, y, z). Such a calculation was feasible even though the Golem A had so little memory because one did not require first derivatives for energy minimization: it was sufficient to follow the forces downhill by a method called steepest descents. Consider a small molecule with 30 atoms. Its second-derivative matrix requires $(3 \times 30)^2 / 2 = 4050$ memory words. This space suffices for the first-derivative vector of a protein with 1350 atoms, more than enough for lysozyme with 964 heavy atoms or myoglobin with 1120 heavy atoms.

The issue was where to get the X-ray-determined atomic coordinates for these two proteins. Fortunately, Prof. Nathan Sharon and his PhD student Yuval Eshdat had obtained printouts of the coordinates of these proteins from David Phillips and John Kendrew, respectively, so that they could build a brass-wire model with what are known as Watson-Kendrew components. I had volunteered to help Yuval build the model of lysozyme (Figure 9). This allowed me to get the printout typed onto punched cards and then run the first

MOVING OVER ENERGY SURFACE

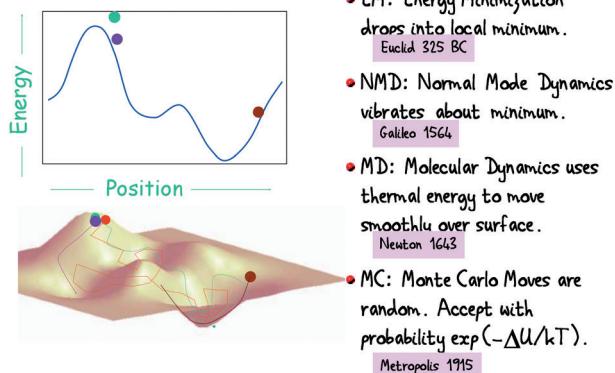


Figure 8. Given the potential energy function of any molecular system, all static, dynamic, and thermodynamic properties can be calculated by simple methods. Energy Minimization (EM) is simplest in that one moves over the energy surface (illustrated in one and two dimensions) to reach a local minimum, where all net forces on every atom are zero and the system is at equilibrium. Normal Mode Dynamics (NMD) focuses on the energy surface around the minimum, where the surface is basin-like and the system will vibrate about the equilibrium following an analytical path. Molecular Dynamics (MD) is a more general method for simulating molecular motion that does not depend on being in an energy basin. Algorithmically, it is a simple variant of energy minimization. The conformation is changed to follow the net forces towards a local minimum; the loss of potential energy is converted into kinetic energy, which gives every atom a velocity to allow it to move over energy barriers. While the three methods EM, NMD, and MD, all arose centuries ago, the forth method known as Monte Carlo (MC) is much more recent, originating as it did with the simulation of neutron diffusion in hydrogen bombs. It is the simplest but also the most generally applicable method (see Figure 14).

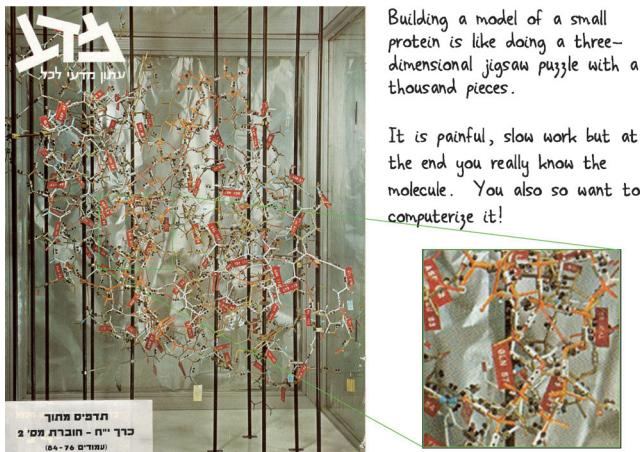


Figure 9. As seen in Figures 2 and 3, the physical models of the first protein structures were built from brass components, known as Kendrew Models. In 1968, together with Yuval Eshdat, I built such a model of hen egg white lysozyme using coordinates determined by David Phillips (Figure 4) and sent on a computer printout to Nathan Sharon, Yuval's PhD supervisor. Such manual modeling was slow and difficult but it provided me the impetus to do the first energy calculations on an entire protein (Figure 10).

MACROMOLECULAR ENERGY MINIMIZATION 1969

Refinement of Protein Conformations using a Macromolecular Energy Minimization Procedure

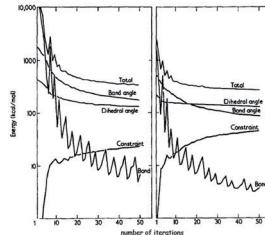
MICHAEL LEVITT[†] AND SIMEON LIPSON
Weizmann Institute of Science

J. Mol. Biol. (1969) **46**, 269–279

$$E = \sum_{\text{all bonds}} K_b (b - b_0)^2 + \sum_{\text{all bond angles}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{all dihedral angles}} K_\phi \{1 + \cos(\phi - \phi_0)\}$$

$$+ \sum_{\text{all non-bonded pairs}} \epsilon_{ij} \{(r_{ij}^0/r_{ij})^{12} - 2(r_{ij}^0/r_{ij})^6\}$$

$$+ \sum_{\text{all atomic co-ordinates}} \frac{Z}{w} (x_i - x_i^0)^2$$



First protein structure refinement



Figure 10. Steepest-descent energy minimization was used to move all non-hydrogen atoms of the two proteins myoglobin and lysozyme by changing their Cartesian coordinates. This reduced the net forces and moved the structure towards an equilibrium. Note how a restraint on atom positions was used to correct for limitations of the energy function, principally the omission of the Coloumbic electrostatic term. Our paper^[11] reports 50 steps of minimization, which is totally trivial by today's standards; these 50 steps took about 1000 seconds on the Golem A computer. The same calculation of forces used for energy minimization could also be used to simulate molecular dynamics (Figure 8), which had previously been applied by Annesur Rahman to liquid argon^[12] and then together with Frank Stillinger to much more complicated liquid water.^[13]

energy minimization on an entire protein structure (Figure 10).

This was the start of the multiscale modeling of complex macromolecules recognized by the Nobel Committee for Chemistry. The key problem was one of simplification as attributed to Einstein (Figure 11). Our calculations had to be

EINSTEIN* ON SIMPLIFICATION

"Everything Should Be Made As Simple

As It Can Be, But Not Simpler"

*Einstein may have crafted this aphorism, but there is no direct evidence in his writings. He did express a similar idea in a lecture but not concisely. Roger Sessions was a key figure in the propagation of the saying. In fact, he may have crafted it when he attempted to paraphrase an idea imparted by Einstein.

<http://quoteinvestigator.com/2011/05/13/einstein-simple/>

Figure 11. Key to useful multiscale models is proper simplification of the complex chemical systems under study. In our work, simplicity was needed for three reasons. Firstly, the calculations had to be feasible with the very limited computational resources available to us on the Golem A, one of the most powerful computers in the world in 1967. Secondly, parameterization had to be feasible with a small number of parameters and transferable atom types (see text). Thirdly, the conformational space associated with the model needed to be simple enough so as to allow adequate exploration of different structures.

COARSE-GRAINED SIMULATION OF FOLDING 1975

Michael Levitt* & Arieh Warshel* Nature Vol. 253 February 27 1975

Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel

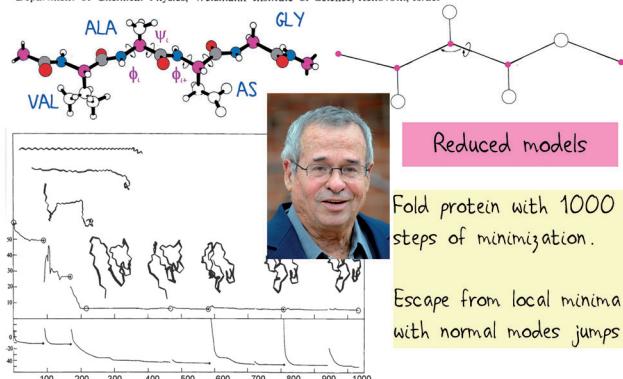


Figure 12. The first application of energy calculations to protein folding required a drastic simplification through the use of what are now known as coarse-grained energy functions. In protein folding, we aim to explore conformation space thoroughly so as to find the low-energy conformations that are not just local energy minima. We did this by simplifying the polypeptide chain by collapsing all the side-chain atoms into a single interaction center and by collapsing all the main-chain atoms into a second interaction center. We sometimes used a simpler model that had one interaction center per amino acid residue. Torsion angles were varied to reduce the number of degrees of freedom by about 30-fold and cut the time to compute a single energy value about 100-fold. Energy minimization converged to a true local minimum. The trajectory was then continued by fitting the local minimum energy basin by an analytical function and using it to predict how to jump out of the minimum with least increase in energy. 1000 cycles took 600 s on an IBM 370/165 computer.

simple if they were to run in reasonable time but they had to still provide useful results. The first energy minimization of a protein with all heavy atoms published in 1969 was followed in 1975 by a model that simplified the structure to have just one interaction center per residue (Figure 12). This enabled us to fold up an extended polypeptide chain in the first simulation of protein folding.^[14,15] The methods used on these simpler systems were actually more complicated changing as they did the torsion angles as Scheraga and Gibson had pioneered^[16] and also using normal modes to calculate low-energy paths out of the local minima. This enabled energy minimization to change conformation a lot (Figure 12).

The next use of multiscale models depended on Arieh Warshel's knowledge of quantum mechanics (Figure 13) and led to the QM/MM method that Arieh has continued to improve. Over the next decade, together with Ruth Sharon, we developed a model for a protein with all atoms in a box of explicit water molecules (Figure 14). This greater realism allowed the simulation to remain much closer to the known X-ray structure than had earlier in vacuo simulations. With this greater realism Dr. Valerie Daggett and I were able to simulate α -helix unfolding (Figure 15).

The period from 1967 to 1976 were my "golden years" with my first 13 papers, six that were sole-author and five more that were co-authored with Arieh Warshel and in two cases Shneior Lifson. Although focused on multiscale models, this body of work also dealt with tRNA structure, folding of

QUANTUM MECHANICS OF ENZYMIC REACTIONS 1976

J. Mol. Biol. (1976) 103, 227–249

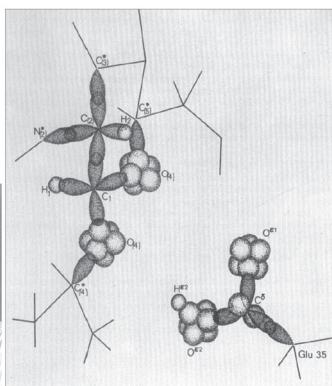
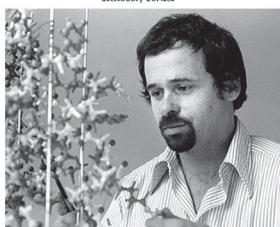
A. WARSHEL AND M. LEVITT
Medical Research Council Laboratory of Molecular Biology
Hills Road, Cambridge CB2 2QH, Englandand
Department of Chemical Physics
The Weizmann Institute of Science
Rehovot, Israel

Figure 13. When Phillips solved the X-ray structure of lysozyme, he proposed that its catalytic action is due to using the binding energy to distort the substrate. Specifically, the six-membered sugar ring adjacent to the bond to be cleaved was thought to be deformed from a chair to a half-boat. Calculations done in my thesis^[17] and published in the proceedings of a conference^[18] showed that the enzyme was too soft to cause such a deformation and led us to propose electrostatic rather than steric strain. With Arieh Warshel, we added quantum mechanical orbitals to a small part of the system, while the rest was still treated classically in what has become known as QM/MM. The calculations now possible showed that the substrate is indeed electrostatically strained.^[19]

RNA, secondary-structure prediction and analysis of structural patterns in globular proteins.

Present: Multiscale Dynamics of Huge Systems

Much of biology is now seen to be driven by large molecular machines consisting of hundreds of thousands of atoms. Unlike smaller globular proteins, these machines are complexes of many different protein chains and have moving parts and fixed parts just like the machines we are familiar with from the world around us. Studying these systems by the same sort of atom-based molecular dynamics is impractical as 100 000 atoms are defined by 300 000 Cartesian coordinates and 1 000 000 000 iterations would be needed to simulate just 1 microsecond (simulation time-steps are typical 1 femtosecond apart). Even if the calculations could be done, analysis would mandate some sort of simplification. Simplification can be done in two ways. Firstly, keep the same degrees of freedom but reduce the number of interacting centers. This is like what we did for our coarse-grained model (Figure 12). Secondly, keep the same interaction centers—the atoms—but move them with collective degrees of freedom rather than atomic Cartesian coordinates. Both tricks can be combined as we did for the simulation of protein folding (Figure 12). The same sort of shortcuts are used in modern studies of the dynamics of large molecular machines. Herein we illustrate this with three examples.

RNA polymerase II is an essential macromolecular machine transcribing the library copy of DNA in the cell nucleus to a working copy of RNA to be used for protein

SIMULATION OF PROTEIN DYNAMICS IN SOLUTION 1988

MICHAEL LEVITT* AND RUTH SHARON Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7557–7561, October 1988
Department of Chemical Physics, Weizmann Institute of Science, Rehovot 76100

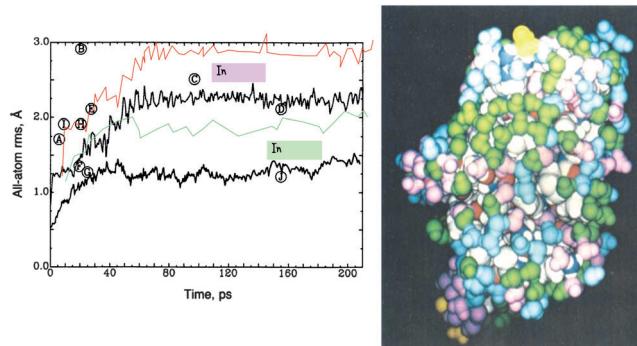


Figure 14. The first molecular dynamics simulation of a protein^[20] was done in a vacuum. While this simplification greatly speeded the simulation, it omitted a very important part of the system, namely the solvent. Running simulation of proteins in a periodic box of explicit water molecules is much more difficult as the force field used for the protein must match that used for the water. Efficiency is paramount as each energy evaluation is some 10 to 20 times slower. The first simulation of the small protein BPTI in water showed that the protein remained much closer to the known X-ray structure than in a comparable simulation in vacuo.^[21] As a result, almost all current simulations use this protocol and include thousands of water molecules. The first glimpse of protein dynamics in vacuum is illustrated in a video in the Supporting Information.

Molecular Dynamics Simulation of Helix Denaturation

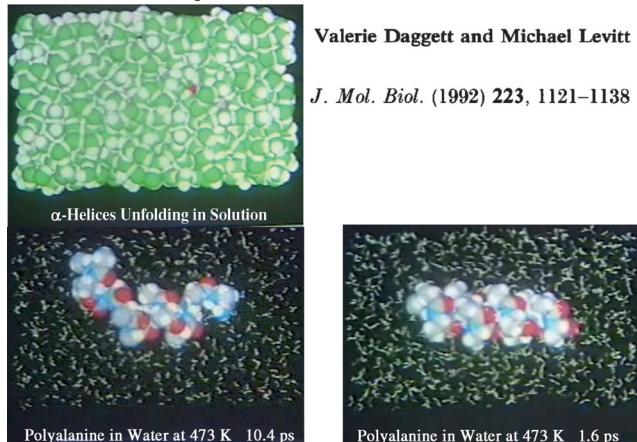


Figure 15. Simulation of temperature unfolding: By 1992, computer power had advanced sufficiently to enable simulation of the unfolding of a short α -helix of 13 alanine residues in a large box of water molecules.^[22] At room temperature, the α -helix is perfectly stable whereas as the temperature increases it becomes progressively less stable. We also showed that in vacuum the α -helix is very stable. This may be expected but such common-sense tests were essential in the early days of simulation. In the two decades since then computers have become much more powerful and simulations of much larger systems are possible with social computing^[23] or special-purpose hardware^[24] (see the video in the Supporting Information).

synthesis and in its own right as functional RNA of different types. This enzyme has been studied extensively by my close friend and colleague, Prof. Roger Kornberg, who character-

ized the system, purified it and solved the detailed three-dimensional structure of the complex in action.^[25] After he won the Nobel Prize for Chemistry in 2006, many in my group wanted to collaborate with him and his group (we are in the same tiny department at Stanford). For me the attraction was that this is a large molecular complex but also one where a close colleague has immense knowledge about all aspects of the system. RNA PolII is a large system with 10 protein

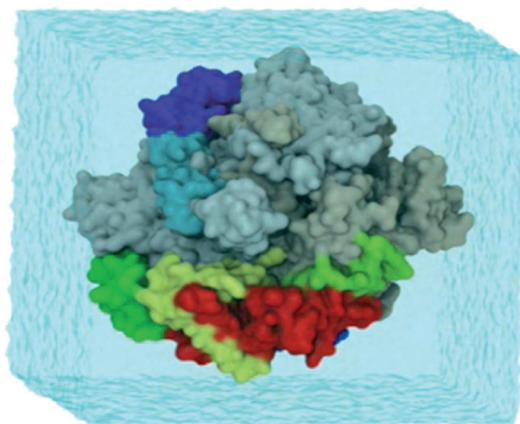


Figure 16. State dynamics of RNA polymerase II: a long simulation of the molecular dynamics of a large system in water with a Markov state model^[26] is shown for the action of the large molecular machine RNA PolII as it moves one base of the template DNA strand over the bridging helix so that it can be recognized by the correct incoming nucleoside triphosphate. Simulations lasting microseconds are easily achieved for a system with almost 500 000 atoms as illustrated in the video in the Supporting Information.

chains, the DNA template strand, and the growing RNA chain. It is also a machine with fixed and moving parts.

Working with Prof. Xuhui Huang, then a postdoc and now faculty at Hong Kong University, we set up the system in a huge box of explicit water molecules (Figure 16). We then ran many independent relatively short molecular dynamics simulations starting from conformations generated by morphing the structure along a path between end-points^[27] that characterize its biological function. Then we used the Markov State Model or MSM model^[26] to cluster the conformations along the trajectories into “states”. If we observed a transition between two states they were linked to form a graph of states. Long-time-scale motion is then simulated by randomly jumping from one connected state to the next. This is beautifully illustrated in the movie made by Dr. Daniel Silva working with Prof. Huang and is from a paper now published.^[28]

The second project involved an even larger system, the complete ribosome (Figure 17), whose structure won the 2010 Nobel Prize in Chemistry for Ramakrishnan, Steitz, and Yonath.^[30–32] The as yet unpublished work was done with Jenelle Bray and Junjie Zhang, two recent postdocs at LinkedIn and on the faculty at Texas A&M University, respectively. We used torsion angle normal modes to calculate how the system would move. This was done with two different

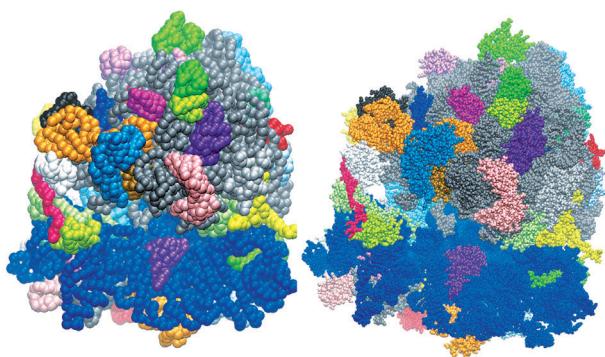


Figure 17. Coarse-grained and all-atom normal mode dynamics of the entire ribosome: Together with postdocs Jenelle Bray and Junjie Zhang, the torsional angle normal mode method we developed in 1985^[29] has been improved so that it can handle any number of independent bodies each with its own rotational and translational degrees of freedom. Although the entire ribosome is large with 4500 nucleotides in 7 RNA chains and 6000 amino acids in 49 protein chains, we can represent its low-frequency motion by just 538 degrees of freedom, 6 for each of the 56 chains and an additional 202 for internal degrees of freedom. The motion is simulated with all 167 000 atoms as well as with 11 062 interaction centers in a coarse-grained representation like that we introduced.^[14] The motions of the four lowest-frequency modes are very similar for the two models. The video of these modes in the Supporting Information shows a functionally suggestive relative motion of the heavy (30S) and light (16S) particles that includes jaw closing, rotational grinding, and rocking.

models of atomic interaction: a) a coarse-grained model termed, 1pt, which used one point-of-interaction center per amino acid or nucleotide, and b) a model that took all atoms except for non-polar hydrogen atoms into account. The calculations were very quick, taking no more than one day on a laptop. This speedup resulted from using Monique Tirion's trick,^[33] in which an artificial energy function is used to ensure that the starting X-ray conformation is indeed a local minimum. This approach, also known a quasi-elastic model, treats all pairwise interactions as springs whose equilibrium distance is the actual distance in the starting structure. Our programs can use any energy function and minimize it in torsion angle space; this work awaits publication.

The degrees of freedom we use are special in that every protein or RNA chain moves as a rigid body with a few additional internal degrees of freedom. The choice of these degrees of freedom is arbitrary but we used the simplest possible allowing an additional torsion-angle degree of freedom for every stretch of 50 amino acids or nucleotides along each chain. In spite of this simplicity, the movie showed in its four lowest-frequency mode motion that may help explain how the ribosome moves as it functions.

The third project involved another of the methods to simulate motion shown in Figure 8, namely, Monte Carlo random moves. Because of its simplicity, this method can be used to rapidly prototype energy functions without needing the cumbersome analytical derivatives I programmed as a 20 year old (Figure 5). It can also be used with any set of degrees of freedom, which can perturb the system in a totally arbitrary manner. Key is to find degrees of freedom that allow the conformation to change a good deal without increasing

the total energy so much as to make the proposed move totally unacceptable. For this, Dr. Peter Minary, then a post-doc with me and now a faculty member at Oxford, UK, developed a new method called Natural Move Monte Carlo or NM-MC,^[34] which is an extension of another pioneering study.^[35] The idea was to allow a degree of freedom to deform the structure in any way. This deformation could include breaking of bonds, which normally carries with it a huge energy penalty. Minary's new algorithm called Recursive Stochastic Chain Closure would then correct the broken bond locally while leaving the natural move perturbation in effect.

Together with Adelene Sim my then PhD student and now a postdoc at the Bioinformatics Institute in Singapore, Minary and I showed that carefully chosen "Natural Moves" allow the Monte Carlo method to sample the conformational space of large RNA hairpins very efficiently (Figure 18). This work has

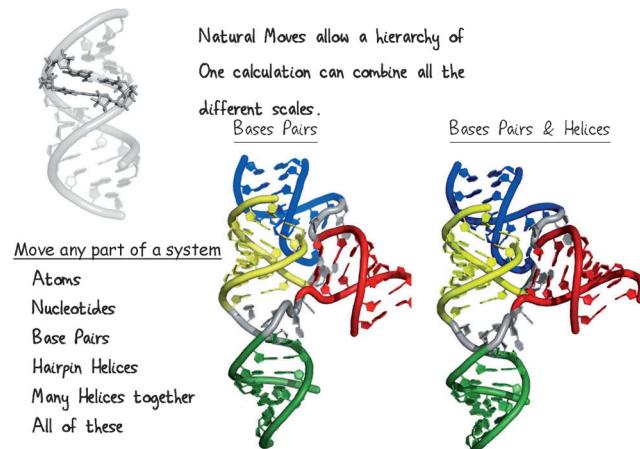


Figure 18. NM-MC calculation for RNA: This new method, developed with Peter Minary^[34] and tested by Adelene Sim,^[36] allows one to move a molecular system through any arbitrary degrees of freedom. Unlike torsion-angle variable (Figure 15), these "natural moves" break bonded chains which would normally cause unacceptably high energy values leading to rejection of all moves. We use stochastic chain closure to quickly close chain breaks and then proceed to accept or reject the move by the normal Monte Carlo criterion (see Figure 7). Our scheme can be used to combine any set of natural moves leading to very rapid sampling of conformational space. Here it is tested on an RNA, a class of molecules with a conformational space that is difficult to sample.

many future applications including the prediction of the location of nucleosomes by calculating the DNA deformation energy from first principles, namely the same consistent force field used for much of our work. This approximation to what localizes the nucleosome on DNA ignores the interaction of the DNA with the nucleosome but does as well as predicting nucleosome location with knowledge-based methods. In this study, the bent DNA is relaxed by NM-MC before determining its average deformation energy.^[37]

Future: Diverse Studies in Computational Biology

Although my group of four is much smaller than its normal size, this is deliberately intended to help more ational Institutes of Health (NIH) funding go to younger scientists. It also allows me to focus on my diverse interests as I did in those golden years between 1967 and 1977. There are four projects encompassing aspects of computational biology.

Dr. Andrea Scaiewicz is working on a project that is involved with genomics and protein function without concern for the detailed three-dimensional protein structure. She classifies all sequences of a genome by recognizing function motifs and then uses this to compare all known genomes. The method scales well allowing tens of thousands of complete genomes to be compared.

Dr. Ivan Ufimtsev is applying his PhD-derived expertise on density functional theory to a long-standing, very difficult problem, namely, determining the crystal structure of macromolecules from the scattered X-ray intensities. Obviating the needs for phases normally still generally determined by Perutz's heavy-atom method would dramatically speed up structure determination especially when used with the super-intense X-ray beams created by free-electron lasers.

Dr. Yana Gofman is developing methods to solve and refine membrane-protein structures by cryo-electron microscopy. She is working independently with co-workers with experimental expertise in a project that will benefit from the new generation of microscopes having higher resolution.

Dr. Nir Kalisman (now a young faculty member at the Hebrew University, Jerusalem) is using chemical cross-linking and mass spectrometric data combined with low-resolution X-ray and cryo-EM structural data to determine the structures of large complexes with less data. He has published studies on eukaryote chaperonin (CCT)^[38] as well as on the eukaryote transcript pre-imitation complex (PIC).^[39] In both cases his methods were able to fix the incorrect chain assignment of previous studies and gave models that explained the molecular function.

Applications to Biomedicine

Moving experimental chemistry into cyberspace should be of clear importance to biomedical science, as it allows one to accelerate the testing of hypotheses. Of course, this is only useful if the calculation is an accurate prediction of what an experiment is likely to show. The required level of accuracy is very problem-dependent. One of the most obvious applications of computational methods to biomedicine is the design of better binding drugs that are more specific for a particular therapeutic target protein. This task is actually very difficult for three independent reasons: a) empirical energy functions do not include all the atom types encountered in drug molecules, b) binding strength depends on the Gibbs energy of interaction of drug and protein compared to the energy of each alone in solution requiring broad conformational sampling, and c) a small change in the Gibbs energy can have a large effect on the binding energy (1 kcal results in a 5-

fold change in affinity). New quantum mechanical force fields^[40] offer hope of more accurate energies.

Fortunately, some problems need less computational accuracy. Thus, in 1987 I was asked to consult for a startup company, Protein Design Labs (PDL), and help them engineer better antibodies. Specifically, they wanted me to make a three-dimensional model of an arbitrary antibody sequence so that they could visualize which amino acids were most important (Figure 19). The task at hand was to design an

A humanized antibody that binds to the interleukin 2 receptor

(chimeric antibody/antibody affinity/autimmune disease)

CARY QUEEN*, WILLIAM P. SCHNEIDER*, HAROLD E. SELICK*, PHILIP W. PAYNE*, NICHOLAS F. LANDOLFI*, JAMES F. DUNCAN*, NEVENKA M. AVDALOVIC*, MICHAEL LEVITT\$, RICHARD P. JUNGHANS\$, AND THOMAS A. WALDMANN*

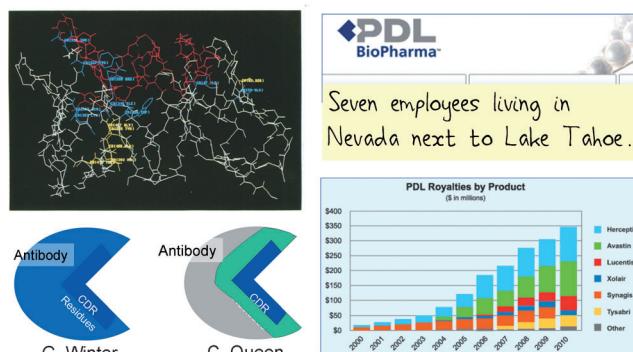


Figure 19. Computer modeling humanizes antibodies: Antibodies are the bodies' defense force but they sometimes need help recognizing treats. Work that started out as an academic exercise^[41,42] led to an automatic method for modeling the structure of any antibody sequence.^[43] More than two decades later, this work, when combined with genetic engineering, thorough patenting, marketing prowess, and massive investment in manufacturing, led via a tortuous path to one of the most successful anticancer therapies. More details are given in the text but this goes to show the potential power of computer methods in medicine. The example also shows how long is the road from basic research to practical treatment.

antibody drug against a natural receptor involved in cancer. Antibodies could be easily raised in mice inoculated with the particular target molecule but these antibodies were then unsuitable as they were deemed foreign by human cells and caused a severe immune reaction. What needed to be done was obvious: take the mouse antibody sequence as a starting point and modify its sequence so that it is not foreign to human cells but still maintains its ability to recognize and destroy the cancer cells. This had been pioneered by the group of Winter, who grafted the parts of the mice antibody recognizing the cancer onto a human antibody framework.^[44] Sadly, the resulting "humanized" antibodies were not as potent as the original mouse antibodies. Cary Queen at PDL used the computer models I built for them, to decide which additional framework residues to change (Figure 19). This eventually led to a series of successful anticancer drugs, made with the PDL patent, the most well-known of which are Herceptin and Avastin, but it took decades and tens of billions of dollars to follow a tortuous path from pure research to a clinically useful drug.

Some General Thoughts

Soon after the good news woke me in California at 2:16 AM on October 9, I mentioned in an interview that had the prize been awarded to four rather than three, the forth recipient should be the computer industry whose massive research and development efforts led to unimaginable gains in computer time (Figure 20). This growth in power, which has

YEAR	COST	SPEED	MEMORY	SIZE
1967	\$40×10 ⁶	0.1 MHz	1 MB	HALL
2013	\$4000	1 GHz	10 GB	LAPTOP
CHANGE	10000	10000	10000	10000

If cars were like computers then a new Volvo would cost \$3, would have a top speed of 1000000 Km/h, would carry 50000 adults and would park in a shoebox.

Figure 20. Pushed ahead by technology: It is difficult to imagine how much computers have developed since our first calculations in 1967. Surprisingly, there has been a 10000-fold improvement in each of four aspects: cost, speed, memory size, physical size. This means that the cost of a particular calculation is 100000000 less. The car analogy has been used before but not at this level of detail.

been so important in giving value to the multiscale models pioneered 45 years ago, was fueled by popular demand for computer power and not by scientific needs. The Cray X-MP supercomputer was essential for the first simulation of protein molecular dynamics in water in 1986 (Figure 14), but a decade later, Linus Torvald's Linux operating system opened up the power of home and office computers for science. This dropped prices as chip development is hugely expensive and needs to be offset by making huge numbers of computers. In some ways, the steady drop in efficiency with successive releases of the Windows operating system forced Intel to make faster and faster hardware, an unexpected bonanza for research computing.

I started the work honored by the Nobel Committee when I was 20 years old having been put in the right place at the right time by John Kendrew. Ten years later the work was essentially done but I have remained an active researcher and mentor who is proud to be a computer programmer.^[45] I have also been blessed by a wonderful wife, Rina, who gave me three sons and kept home-life steady during those very rocky early years. This makes me feel the need to try to influence the young by four simple pieces of advice (Figure 21). Clearly advice is cheap and I hope to help more by making sure that young scientists have the same remarkable opportunities afforded to me by my many mentors.

One area of advice concerns the need to move out of your comfort zone and take risks (Figure 22). I suppose I also need

- BE PASSIONATE
- BE PERSISTENT
- BE ORIGINAL
- BE KIND & GOOD

Figure 21. My advice to the young: Adults tend to give too much advice, so this is given in the expectation that it will be ignored. These four points are rather obvious but they certainly worked for me. Passion is needed for any endeavor. Being persistent means you believe in yourself; if you do not, why should anyone else? By being original, competition is less of a concern. By being kind and good, you make friends and not enemies.

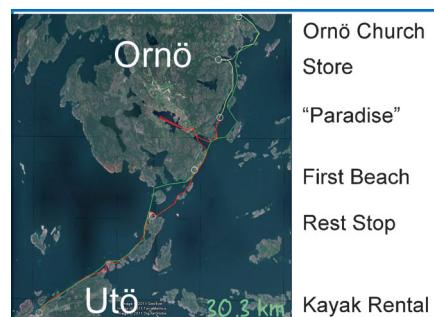
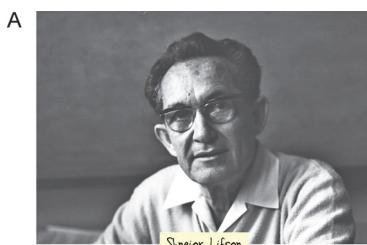


Figure 22. Take risks: It is difficult to predict the outcome of most actions. Taking some risks can lead you to wonderful places that would have been missed otherwise. This is true in science as it is in life. When a meeting I was attending in Sweden was held in Uppsala and not on the Stockholm archipelago, I decided to travel there alone. Advised against hiking as the islands are small and flat, I rented a sea kayak online. As a complete novice, I found a short movie and set out myself on the weekend before mid-summer's day. I was completely alone on the water but the sea was calm and the swans comforting until the wind hit (continued in Figure 23).



Figure 23. But do not be too stupid: The water was cold at 12°C so I stayed close to shore as I learned to balance. After a scary encounter with a Visby-class missile boat that passed as I beached my kayak, I proceeded up the coast to Ornö Church with the wind coming from behind. I headed back south to find a tiny island on the way where I camped for the night. I had a Swedish SIM card and felt comforted by e-mail and internet. Still it was hard to sleep without a facemask, something essential with such short nights. Next morning I headed back and had a hard time crossing about 1 km of open water against a head wind. My island was paradise, but perhaps it was a bit too risky?



Shneior Lifson

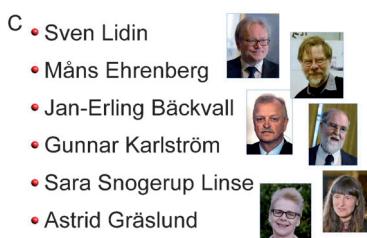


Figure 24. Special thanks to: A) Shneior Lifson, my mentor at the Weizmann Institute. B) John Kendrew, Max Perutz, Bob Diamond, Francis Crick, and Aaron Klug, my mentors in Cambridge. Bob Diamond was my actual PhD supervisor but independence was expected: I never wrote a paper with Diamond but we did write related papers adjacent to one another in the same journal. C) The 2013 Nobel Committee in Chemistry. This may seem obvious as they awarded me a share of the Nobel Prize for 2013. No, I thank them for their courage to recognize the role that computers have played in taking chemistry of complex biological systems from the experimental lab into cyberspace. Given the incredible increase in computer power, there is no doubt that their recognition of a field of increasing importance in biomedical science, will, itself, be recognized as formalizing the establishment of a new field.



Figure 25. With this recognition, the field of computational structural biology and indeed the broader field of computational biology, all those who have worked away in the belief that computers and biology belong together are winners. This photo was taken on Stanford's American Football field during the Homecoming game with UCLA on October 19, just 10 days after the Chemistry Prize announcement. Hearing 50 000 people screaming "Nobel Prize, Nobel Prize" is an indelible, treasured memory.

PhD Students

Miriam Hirshberg
Chris Lee
Britt Park
Dave Hinds
Enoch Huang
Jerry Tsai
Yu Xia
Michael Sykes
Rachel Kolodny
Nizar Batada
Sergio Moreno

Dahlia Weiss

Gaurav Chopra
Adelene Sim

Post Docs
S. Subbiah
Valerie Daggett
Peter David
Mark Gerstein
Steven Brenner
Boris Fain
Chen Keasar

Golan Yona

Ram Samudrala
Tanya Raschke
Patrice Koehl
Erik Sandelin
Chris Summa
Erik Lindahl
Julian Gough
Peter Minary
Gunnar Schroeder
Karine Bastard
Avraham Samson

Xuhui Huang

Julie Bernauer
Alena Shmygelska
Mitul Saha
Nir Kalisman
Andrea Scaiewicz
Jenelle Bray
Marie Brut
Junjie Zhang
Leonid Pereyaslavets
Yana Gofman
Ivan Ufimtsev



Figure 26. Past and present members of my group: Since 1987, I have had the privilege to mentor 14 PhD students and 29 postdoctoral fellows. They are all part of my family and a majority have followed my example and established independent academic careers. In my first 20 years as an independent scientist I worked with collaborators or alone, not trusting myself to direct others.

to mention that some things may be too risky (Figure 23) but what does not kill you may make you stronger.

As this unusual account comes to a close, I need to thank Shneior Lifson (Figure 24A), my earliest mentor at the Weizmann Institute, as well as John Kendrew, Max Perutz, Francis Crick, Bob Diamond, and Aaron Klug (Figure 24B), my mentors in Cambridge. Sadly, only Diamond and Klug are still among us to read these words. As a group, these are my towering heroes of science.^[46]

I also thank the 2013 Nobel Committee for Chemistry (Figure 24C) for daring to recognize the role that computers have played in multiscale modeling of the complex chemical systems so important in biology. This work is intrinsically multidisciplinary extending from the math and physics of atomic interactions to chemical reactions in biology to biomedical therapeutics. As a result of being honored with the Nobel prize the entire field of computational biology has become a big winner (Figure 25). Since moving to Stanford in 1987, I have been blessed by an exceptional group of PhD students and postdoctoral fellows (Figure 26) and I thank them all profusely for teaching me so much. My work is supported by the NIH R01 award GM063817.

Received: March 25, 2014

Published online: August 5, 2014

- [1] "Molecular Structure for Nucleic Acids": J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.
- [2] "Molecular Structure of Deoxypentose Nucleic Acids": M. H. F. Wilkins, A. R. Stokes, H. R. Wilson, *Nature* **1953**, *171*, 738–740.
- [3] "Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å Resolution": J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore, *Nature* **1960**, *185*, 422–427.
- [4] "The Three-Dimensional Structure of a Protein Molecule": J. C. Kendrew, *Sci. Am.* **1961**, *205*, 96–110.
- [5] "The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method": D. W. Green, V. M. Ingram, M. F. Perutz, *Proc. R. Soc. London Ser. A* **1954**, *225*, 287–307.
- [6] "Structure of Haemoglobin: A Three-Dimensional Fourier synthesis at 5.5 Å Resolution, Obtained by X-ray Analysis": M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, A. C. T. North, *Nature* **1960**, *185*, 3416–3422.
- [7] "The Hen Egg-White Lysozyme Molecule": D. C. Phillips, *Proc. Natl. Acad. Sci. USA* **1967**, *57*, 483–495.
- [8] "The Three-Dimensional Strcuture of an Enzyme Molecule": D. C. Phillips, *Sci. Am.* **1966**, *205*, 96–110.
- [9] "A Consistent Force Field for Calculation on Conformations, Vibrational Spectra and Enthalpies of Cycloalkanes and n-Alkane Molecules": S. Lifson, A. Warshel, *J. Phys. Chem.* **1968**, *49*, 5116–5129.
- [10] "A Quantum Mechanical Polarizable Force Field for Biomolecular Interactions": A. G. Donchev, V. D. Ozrin, M. V. Subbotin, O. V. Tarasov, V. I. Tarasov, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7829–7834.
- [11] "Refinement of Protein Conformations Using a Macromolecular Energy Minimization Procedure": M. Levitt, S. Lifson, *J. Mol. Biol.* **1969**, *46*, 269–279.
- [12] "Correlations in the Motion of Atoms in Liquid Argon": A. Rahman, *Phys. Rev.* **1964**, *136*, A405–A411.
- [13] "Molecular Dynamics Study of Liquid Water": A. Rahman, F. H. Stillinger, *J. Chem. Phys.* **1971**, *55*, 3336–3359.
- [14] "Computer Simulation of Protein Folding": M. Levitt, A. Warshel, *Nature* **1975**, *253*, 694–698.
- [15] "A Simplified Representation of Protein Conformations for Rapid Simulation of Protein Folding": M. Levitt, *J. Mol. Biol.* **1976**, *104*, 59–107.
- [16] "Minimization of Polypeptide Energy. I. Preliminary Structures of Bovine Pancreatic Ribonuclease S-peptide": K. D. Gibson, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 420–427.
- [17] "Conformation Analysis of Proteins": M. Levitt, Ph.D. Thesis, Cambridge University, Cambridge, UK, **1971**. http://csb.stanford.edu/levitt/Levitt_Thesis_1971/Levitt_Thesis_1971.html.
- [18] "On the Nature of the Binding of Hexa-N-Acetyl Glucosamine Substrate to Lysozyme": M. Levitt, *Peptides, Polypeptides and Proteins*, Wiley, New York, **1974**, pp. 99–113.
- [19] "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme": A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227–249.
- [20] "Dynamics of Folded Proteins": J. A. McCammon, B. R. Gelbin, M. Karplus, *Nature* **1977**, *267*, 585–590.
- [21] "Accurate Simulation of Protein Dynamics in Solution": M. Levitt, R. Sharon, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7557–7561.
- [22] "Molecular Dynamics Simulation of Helix Denaturation": V. Daggett, M. Levitt, *J. Mol. Biol.* **1992**, *223*, 1121–1138.
- [23] "Atomistic Protein Folding Simulations on the Submillisecond Time Scale Using Worldwide Distributed Computing": V. S. Pande, I. Baker, J. Chapman, S. P. Elmer, S. Khalil, S. M. Larson, M. R. Shirts, C. D. Snow, E. J. Sorin, B. Zagrovic, *Biopolymers* **2003**, *68*, 91–109.
- [24] "Anton: A Special-Purpose Machine for Molecular Dynamics Simulation": D. E. Shaw, M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, K. J. Bowers, J. C. Chao, M. P. Eastwood, J. Gagliardo, J. P. Grossman, C. R. Ho, D. J. Ierardi, I. Kolossaváry, J. L. Klepeis, T. Layman, C. McLeavey, M. A. Moraes, R. Mueller, E. C. Priest, Y. Shan, J. Spengler, M. Theobald, B. Towles, S. C. Wang, Proceedings of the 34rd Annual International Symposium on Computer Architecture (ISCA), San Diego, June **2007**.
- [25] "The Molecular Basis of Eukaryotic Transcription (Nobel Lecture)": R. Kornberg, *Angew. Chem.* **2007**, *119*, 7082–7092; *Angew. Chem. Int. Ed.* **2007**, *46*, 6956–6965.
- [26] "Automatic Discovery of Metastable States for the Construction of Markov Models Of Macromolecular Conformational Dynamics": J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, W. C. Swope, *J. Chem. Phys.* **2007**, *126*, 155101–17.
- [27] "Can Morphing Methods Predict Intermediate Structures?": D. Weiss, M. Levitt, *J. Mol. Biol.* **2009**, *385*, 665–674.
- [28] "Millisecond Dynamics of RNA Polymerase II Translocation at Atomic Resolution": D.-A. Silva, D. Weiss, F. Pardo, L.-T. Da, M. Levitt, D. Wang, X. Huang, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6293–6298.
- [29] "Protein Normal-Mode Dynamics: Trypsin Inhibitor, Crambin, Ribonuclease and Lysozyme": M. Levitt, C. Sander, P. S. Stern, *J. Mol. Biol.* **1985**, *181*, 423–447.
- [30] "Unraveling the Structure of the Ribosome (Nobel Lecture)": V. Ramakrishnan, *Angew. Chem.* **2010**, *122*, 4454–4481; *Angew. Chem. Int. Ed.* **2010**, *49*, 4355–4380.
- [31] "From the Structure and Function of the Ribosome to New Antibiotics (Nobel Lecture)": T. A. Steitz, *Angew. Chem.* **2010**, *122*, 4482–4500; *Angew. Chem. Int. Ed.* **2010**, *49*, 4381–4398.
- [32] "Hibernating Bears, Antibiotics, and the Evolving Ribosome (Nobel Lecture)": A. Yonath, *Angew. Chem.* **2010**, *122*, 4438–4453; *Angew. Chem. Int. Ed.* **2010**, *49*, 4340–4354.

- [33] "Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis": M. Tirion, *Phys. Rev. Lett.* **1996**, 77, 1905–1908.
- [34] "Conformational Optimization with Natural Degrees of Freedom: A Novel Stochastic Chain Closure Algorithm": P. Minary, M. Levitt, *J. Comput. Chem.* **2010**, 31, 993–1010.
- [35] "Molecular Flexibility in ab initio Drug Docking to DNA: Binding-Site and Binding-Mode Transitions in All-Atom Monte Carlo Simulations": R. Rohs, I. Bloch, H. Sklenar, Z. Shakkeb, *Nucleic Acids Res.* **2005**, 33, 7048–7057.
- [36] "Modeling and Design by Hierarchical Natural Moves": A. Y. L. Sim, M. Levitt, P. Minary, *Proc. Natl. Acad. Sci. USA* **2012**, 109, 2890–2895.
- [37] "Training-free atomistic prediction of nucleosome occupancy": P. Minary, M. Levitt, *Proc. Natl. Acad. Sci. USA* **2014**, 111, 7665–7670.
- [38] "Subunit Order of Eukaryotic TRiC/CCT Chaperonin by Cross-linking, Mass Spectrometry and Combinatorial Homology Modeling": N. Kalisman, C. M. Adams, M. Levitt, *Proc. Natl. Acad. Sci. USA* **2012**, 109, 2884–2889.
- [39] "Architecture of an RNA polymerase II transcription pre-initiation complex": K. Murakami, H. Elmlund, N. Kalisman, D. A. Bushnell, C. M. Adams, M. Azubel, D. Elmlund, Y. Levi-Kalisman, X. Liu, B. J. Gibbons, M. Levitt, R. D. Kornberg, *Science* **2013**, 342, 1238724.
- [40] "Application of a polarizable force field to calculations of relative protein–ligand binding affinities": O. Khoruzhii, A. G. Donchev, G. Nikolay, A. Illarionov, M. Olevanov, V. Ozrin, C. Queen, V. Tarasov, *Proc. Natl. Acad. Sci. USA* **2008**, 105, 10378–10383.
- [41] "The Predicted Structure of Immunoglobulin D1.3 and its Comparison with the Crystal Structure": C. Chothia, A. M. Lesk, M. Levitt, A. G. Amit, R. A. Mariuzza, S. E. V. Phillips, R. J. Poljak, *Science* **1986**, 233, 755–758.
- [42] "The Conformations of Immunoglobulin Hypervariable Regions": C. Chothia, A. M. Lesk, M. Levitt, A. Tramontano, S. J. Smith-Gill, G. Air, S. Sheriff, E. A. Padlan, D. Davies, W. R. Tulip, P. M. Colman, *Nature* **1989**, 342, 877–883.
- [43] "A Humanized Antibody that Binds to the IL-2 Receptor": C. Queen, W. P. Schneider, H. E. Selick, P. W. Payne, N. F. Landolfi, J. F. Duncan, A. M. Avdalovic, M. Levitt, R. P. Junghans, T. A. Waldmann, *Proc. Natl. Acad. Sci. USA* **1989**, 86, 10029–10033.
- [44] "Replacing the Complementarity-Determining Regions in a Human Antibody with those from a Mouse": P. T. Jones, P. H. Dear, J. Foote, M. S. Neuberger, G. Winter, *Nature* **1986**, 321, 522–525.
- [45] "Michael Levitt National Academy of Science USA Member Details": http://nrc88.nas.edu/pnas_search/memberDetails.aspx?ctID=3012570.
- [46] "Life-Long Influence of Max Perutz and the Laboratory of Molecular Biology": M. Levitt in *Memories and Consequences* (Ed.: H. Huxley), **2013**. http://www2.mrc-lmb.cam.ac.uk/ebooks/Memories_and_consequences-Hugh_Huxley.epub.