ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/231398537

# Cytochrome c: A molecular proving ground for computer simulations

**ARTICLE** *in* THE JOURNAL OF PHYSICAL CHEMISTRY · APRIL 1993

Impact Factor: 2.78 · DOI: 10.1021/j100115a008

CITATIONS
29

READS
23

**5 AUTHORS**, INCLUDING:

Chong Zheng
Northern Illinois University
**154** PUBLICATIONS **2,641** CITATIONS
SEE PROFILE

Peter Wolynes
Rice University
**359** PUBLICATIONS **28,710** CITATIONS
SEE PROFILE

Available from: Chong Zheng
Retrieved on: 17 January 2016

# FEATURE ARTICLE

## Cytochrome c: A Molecular Proving Ground for Computer Simulations

**Chung F. Wong***

*Department of Physiology and Biophysics, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, New York 10029-6574*

**Chong Zheng**

*Department of Chemistry, Northern Illinois University, DeKalb, Illinois 60115*

**Jian Shen and J. Andrew McCammon**

*Department of Chemistry, University of Houston, Houston, Texas 77204-5641*

**Peter G. Wolynes**

*Department of Chemistry, University of Illinois, Urbana, Illinois 61801*

*Received: May 22, 1992*

This article surveys molecular simulation methods in the context of their applications to the electron-transfer protein, cytochrome *c*. Both classical and quantum simulations are considered, and it is shown how these can provide a wide variety of information on the functional properties of the protein.

### Introduction

Computer simulations are an increasingly powerful tool for studying the correlations between dynamics and function in proteins. In this feature article, we attempt to illustrate the current status of this field by describing a variety of calculations we have carried out on a single protein, cytochrome *c*, which is involved in electron transfer. Among the specific roles of atomic motion in the case of cytochrome *c*, possible examples include conformational adjustments in the surface residues to allow optimal binding to redox partners, and atomic displacements that reduce the barrier for electron transfer and influence the redox potentials of the reduced and oxidized states of the protein.[1-3] Although the present focus is on cytochrome *c*, the studies described here yield insights to the behavior of proteins more generally, e.g., the coupled dynamics of surface residues and solvent, and the occurrence of glasslike transitions at low temperatures. Finally, because of the availability of high quality experimental data, cytochrome *c* represents an excellent system for testing simulation methods, although detailed quantitative comparisons between experimental and simulation results remain to be carried out.

The following three sections describe different simulation tools and the kinds of information that they can provide for a protein such as cytochrome *c*. The section on classical simulations begins with a brief introduction to the molecular dynamics method and its application to cytochrome *c* in water. Of particular interest are the functional roles of the low frequency motions in the protein, the effects of solvent on the structure and dynamics of the protein, and the influence of the protein on the structure and mobility of the surrounding solvent.

The section on quantum simulations provides a somewhat more detailed introduction to the relatively new area of Feynman path simulations of atomic fluctuations in proteins. These simulations in "imaginary time" can be carried out using straightforward modifications of classical molecular dynamics programs. Such simulations yield the quantum corrections to equilibrium properties quite directly. As expected, these corrections are small for the root-mean-square position fluctuations, which are governed by low-frequency, classical motions. But the corrections are larger for local motions involving light atoms, e.g., the rotation of amino groups. Such motions are of special importance in the nuclear reorganization associated with electron transfer, and these kinetic issues have been studied by specialized quantum simulations that make use of thermodynamic integration techniques.

The section on Green's functions briefly describes new tools that can be used to study systematically how a perturbation applied at one part of a protein affects the structure at other parts of the protein. These tools show promise for the understanding of protein structure and the design of new proteins and drugs.

### Classical Simulations

Molecular dynamics simulations have been used for about a decade to provide detailed information about the structural, thermodynamic, and dynamic behavior of bimolecular systems.[4,5] In a conventional molecular dynamics simulation, Newton's equations of motion for a system of particles are solved on a computer. The interaction forces among the atoms in the system are usually derived from empirical potentials obtained by fitting experimental results for small molecules, liquids, and solids. Some potential parameters have also been obtained from quantum calculations on small molecules. A variety of dynamical quantities (e.g., rate constants, diffusion constants) and thermodynamic quantities (e.g., free energies) can be obtained from the trajectories generated in a molecular dynamics simulation. By displaying trajectories on graphics systems, the motion of groups in biomolecular complexes can be visualized to suggest how local and collective motions of important functional groups and amino acid residues contribute to biomolecular activity. Mutagenesis studies can also be performed, since fragments of biomolecules can be modified easily in computer models.

Simulation methods and computer models have been improved substantially since the first molecular dynamics simulation of a small, isolated protein was done.[6] One can now simulate large

0022-3654/93/2097-3100$04.00/0 © 1993 American Chemical Society

biomolecules with explicit solvent molecules included. Advances in supercomputer technology have also helped to extend the time scale of computer simulations. As potential functions are being improved, more reliable predictions can be obtained from computer simulations.

We here present an application of molecular dynamics to study an electron-transfer protein, tuna ferrocytochrome *c*. In the first molecular dynamics simulation on cytochrome *c*, conducted several years ago,[7] solvent effects were only included in a very approximate way. The results presented here are from a simulation of cytochrome c in which explicit water molecules have been included. This should be a better model for the behavior of cytochrome *c* under physiological conditions. The simulation has also been carried out for a longer period of time than the previous simulation.

The simulation was done by starting with Takano and Dickerson's X-ray structure of tuna ferrocytochrome *c* at 1.8-Å resolution.[2,3] Polar hydrogens were introduced and nonpolar hydrogens were only included implicitly with the extended atom representation.[5] We put a single protein molecule in a truncated octahedral box filled with SPC water molecules.[8] Water molecules that had oxygens lying at 2.3 Å or less to a protein heavy atom were deleted. The system contains 3128 water molecules and the distance between the edge of the box and the closest protein atom is at least 10 Å. Nonbonded interactions were cut off at 9 Å without using a switching or shifting function. We equilibrated the solvent by first performing 100 steps of steepest descent energy minimization to relieve bad contacts and then running dynamics with frequent reassignment of Maxwellian velocities corresponding to 300 K to the atoms in the system. Protein atoms were held almost stationary by using harmonic constraints with very stiff force constants. Maxwellian velocities were reassigned every 0.5 ps and the equilibration was continued for 5 ps. SHAKE was used to constrain all the bond lengths,[9] and a time step of 2 fs was used during the calculation. The protein atoms were then allowed to move by removing the harmonic restraints and the equilibration was continued for an additional 10 ps with Maxwellian velocity reassignments at 0.5-ps intervals. The system was then coupled to a temperature bath with relaxation time of 0.1 ps.[10] The equilibration was continued for 5 more picoseconds before velocity reassignments were stopped. After this 20 ps of equilibration, we continued the dynamics for 105 ps and the last 100 ps were used for analysis. The simulations were carried out on the Cyber205 at the John von Neumann Supercomputing center. Each picosecond of simulation took about 1 CPU hour on the Cyber205.

We have calculated the root-mean-square deviation of 20-ps-averaged structures from the X-ray structure (Figure 1). One can see that the drift from the X-ray structure continues up to about 60 ps. This drift reflects both the continuing equilibration of the solvent (with concomitant adjustments of the protein) and the sampling of the large number of conformational substates available to the protein.[11-13] During this period, several water molecules moved into the protein, as can be seen from Figure 2. This may be a result of the opening up of small channels, perhaps associated with protein side chains being extended in an aqueous environment. Due to the extension of side chains into the solvent, the 100-ps-averaged structure has a radius of gyration of 13.2 Å, which is larger than the value of 12.7 Å obtained from the crystal structure. The interior of the protein is also slightly less well packed in the simulation. The number density calculated within 8 Å about the protein center of mass (0.0455 $Å^{-3}$) is smaller than that in the X-ray structure (0.0499 $Å^{-3}$). The root-mean-square (rms) positional fluctuations of the protein atoms do not change much with time, however (Figure 3).

Figure 4 compares the 100-ps-averaged structure with the X-ray structure as a function of residue number. One can see that the N-terminal shows the largest deviation from the X-ray structure,

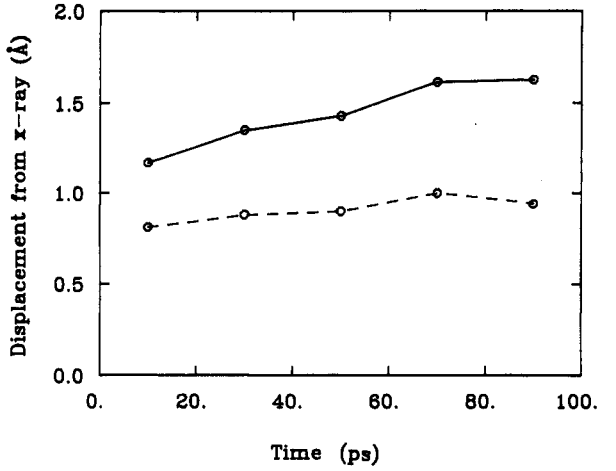


**Figure 1.** Root-mean-square atomic displacement of time-averaged structures from the X-ray structure. The time averages are over successive periods of 20 ps. The solid and dashed lines are for all and heme group atoms, respectively.

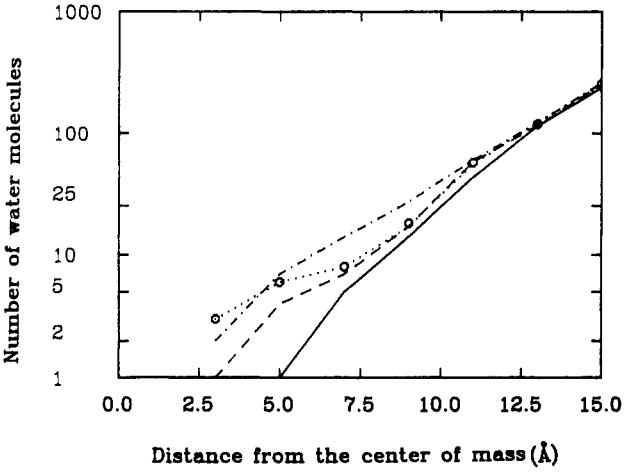


**Figure 2.** Number of waters inside spheres centered at the center of mass of the protein, against the radius of the sphere, averaged over each 20 ps. The solid, dashed, circle, dot, and dashed dot lines are for the first to the fifth 20 ps, respectively.

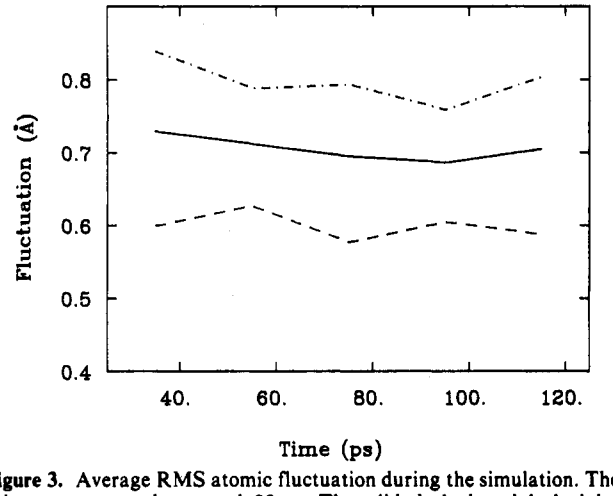


**Figure 3.** Average RMS atomic fluctuation during the simulation. The values are averaged over each 20 ps. The solid, dashed, and dashed dot lines are for all, backbone, and side-chain atoms, respectively.

probably because the N-terminal helix is rather exposed to the solvent and is mobile in solution. The C-terminal also has a large deviation from the X-ray structure. When averaged over all atoms, the deviation of the average structure from the X-ray structure is 1.34 Å. In viewing the deviation as a function of distance from the protein center of mass (Figure 5), one can see that the deviation increases with distance from this center. This

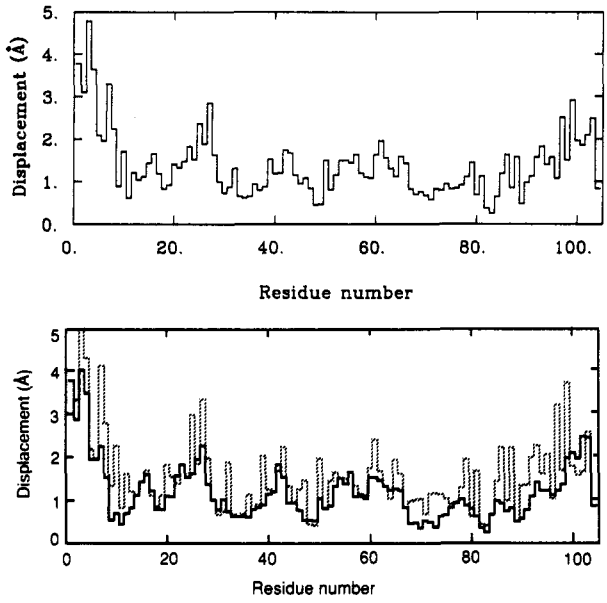


**Figure 4.** Mean atomic displacement of each residue in the time-averaged structure from the X-ray structure: (a) is for all atoms; in (b), solid lines are for backbone and broken lines for side chain.

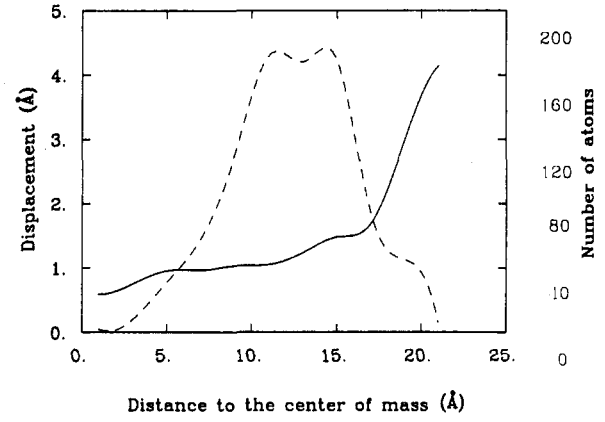


**Figure 5.** Root-mean-square atomic displacement from the X-ray structure vs the distance to the center of mass (solid line). The dashed line referred to the right vertical axis shows the number of atoms of the protein within a shell of 2-Å width at such distance.

is in qualitative agreement with a previous pseudo-solvated simulation of tuna ferrocytochrome $c$[7] but the quantitative deviation in the present results is somewhat smaller, especially near the surface of the protein (cf. 1.5 Å versus 1.7 Å previously for residues between 14 and 16 Å from the center of mass of the protein).

The rms fluctuations of different residues are shown in Figure 6 where they are also compared with results obtained from measured temperature factors. There is qualitative agreement between the simulated and experimental results but there are some quantitative discrepancies. One should note that the simulation was done to simulate an infinitely dilute solution whereas the experimental results were obtained in a crystal environment. Some of the discrepancies therefore arise from the elimination of contacts between neighboring protein molecules, which permits greater freedom of motion of the corresponding surface regions in the simulation. Additional differences can be expected due to approximations made in the computer model (such as approximate potentials and the limited period of averaging) as well as approximations made in extracting fluctuation information from crystallographic data (such as the assumption that each atom is moving in a harmonic and isotropic manner and the inclusion of static disorder contributions). The fluctuations as a function of distance from the protein center of mass are shown in Figure 7. One can see that the fluctuation

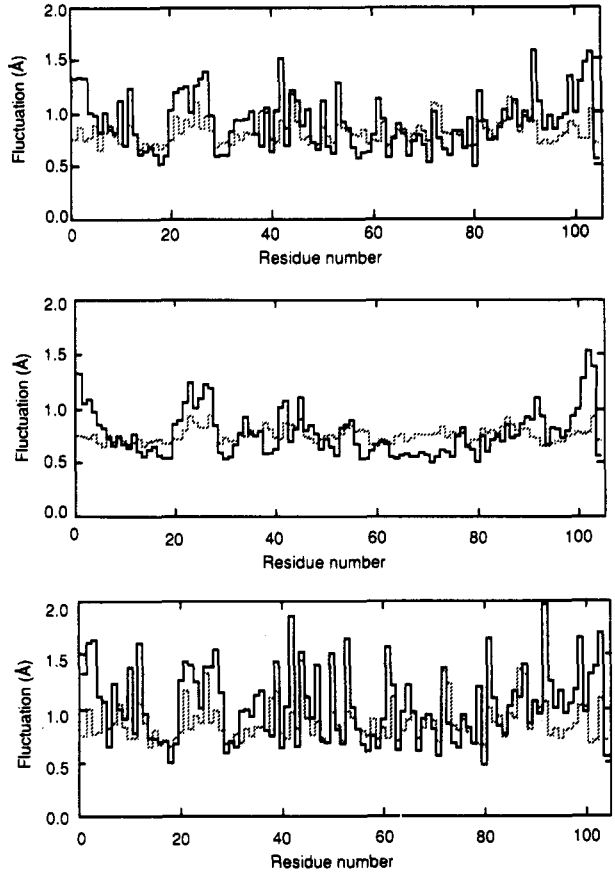


**Figure 6.** Average RMS atomic fluctuation of each residue during the 100 ps simulation. (a), (b), and (c) are for all atoms, backbone, and side chain, respectively. The broken lines correspond to MD and the solid lines to X-ray temperature factors.

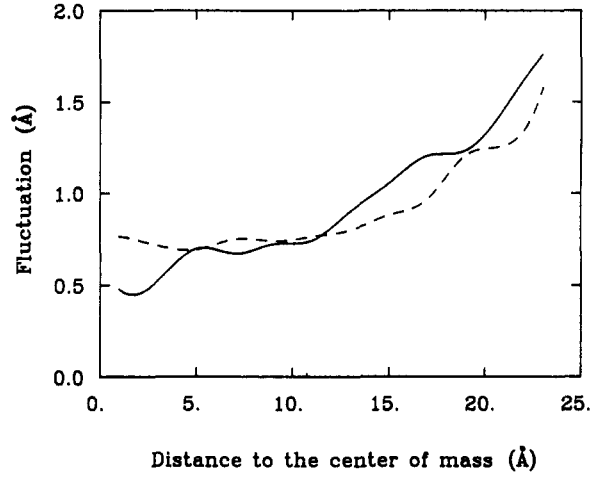


**Figure 7.** Average RMS fluctuation in MD vs the distance to the center of mass (thick solid line). The dashed line is based on the X-ray temperature factors.

increases as a function of distance from the protein center of mass, also in qualitative agreement with previous simulations.[7]

The solvent-accessible surface area of proteins has an important influence on the stability of protein folding and complexation. It is also important in determining the rates of reactions that are preceded by protein–protein or protein–ligand precursors, such as electron transfer reactions in cytochromes.[14,15] Previous analyses of protein surface areas have been based on static molecular structures, usually those obtained by X-ray crystallography. The surface features are, however, likely to be the most sensitive to the environment and it is useful to compare results obtained from solution simulations and from crystallographic measurements. The algorithm we used to calculate the

**TABLE I: Surface Areas of Residues Accessible to a 1.4-Å-Radius Probe (in Å²)**

| | | X-ray | ⟨A⟩ | $\langle\Delta A^2\rangle^{1/2}$ | | | X-ray | ⟨A⟩ | $\langle\Delta A^2\rangle^{1/2}$ |
|---|---|---|---|---|---|---|---|---|---|
| GLY | 1 | 21.3 | 61.9 | 29.8 | LYS | 53 | 85.1 | 58.9 | 22.2 |
| ASP | 2 | 145.4 | 130.2 | 52.5 | SER | 54 | 92.8 | 94.1 | 13.5 |
| VAL | 3 | 49.1 | 139.3 | 45.8 | LYS | 55 | 93.7 | 68.8 | 32.4 |
| ALA | 4 | 74.9 | 79.1 | 8.4 | GLY | 56 | 65.7 | 46.8 | 11.3 |
| LYS | 5 | 85.6 | 229.9 | 91.7 | ILE | 57 | 34.5 | 32.5 | 8.0 |
| GLY | 6 | 0.7 | 5.2 | 5.5 | VAL | 58 | 62.0 | 68.5 | 17.2 |
| LYS | 7 | 108.0 | 217.1 | 87.4 | TRP | 59 | 0.4 | 2.8 | 3.4 |
| LYS | 8 | 83.9 | 130.4 | 63.8 | ASN | 60 | 80.4 | 68.5 | 27.5 |
| THR | 9 | 11.6 | 12.4 | 6.1 | ASN | 61 | 44.1 | 52.3 | 26.7 |
| PHE | 10 | 12.0 | 21.8 | 13.9 | ASP | 62 | 156.7 | 167.4 | 61.3 |
| VAL | 11 | 160.8 | 93.2 | 17.6 | THR | 63 | 25.6 | 19.5 | 5.3 |
| GLN | 12 | 99.8 | 132.5 | 42.0 | LEU | 64 | 0.8 | 4.4 | 5.2 |
| LYS | 13 | 70.4 | 84.2 | 45.4 | MET | 65 | 42.0 | 52.1 | 28.1 |
| CYS | 14 | 0.5 | 4.5 | 4.3 | GLU | 66 | 151.4 | 91.3 | 22.3 |
| ALA | 15 | 34.0 | 43.3 | 11.1 | TYR | 67 | 0.0 | 1.9 | 2.6 |
| GLN | 16 | 76.8 | 139.7 | 33.6 | LEU | 68 | 0.0 | 3.4 | 4.9 |
| CYS | 17 | 1.5 | 26.8 | 17.9 | GLU | 69 | 55.9 | 61.1 | 16.6 |
| HIS | 18 | 4.0 | 11.6 | 9.7 | ASN | 70 | 2.6 | 39.3 | 28.8 |
| THR | 19 | 20.0 | 19.3 | 10.5 | PRO | 71 | 0.0 | 2.2 | 3.1 |
| VAL | 20 | 16.9 | 57.3 | 32.9 | LYS | 72 | 157.3 | 92.9 | 30.5 |
| GLU | 21 | 184.7 | 97.8 | 24.6 | LYS | 73 | 84.4 | 117.1 | 36.5 |
| ASN | 22 | 72.2 | 108.5 | 50.1 | TYR | 74 | 24.8 | 8.7 | 5.3 |
| GLY | 23 | 69.9 | 55.6 | 15.9 | ILE | 75 | 3.0 | 2.1 | 1.9 |
| GLY | 24 | 27.3 | 3.5 | 3.1 | PRO | 76 | 101.7 | 95.0 | 11.9 |
| LYS | 25 | 207.8 | 134.3 | 44.3 | GLY | 77 | 48.1 | 51.4 | 10.0 |
| HIS | 26 | 102.8 | 60.5 | 20.0 | THR | 78 | 12.2 | 24.6 | 11.6 |
| LYS | 27 | 39.4 | 138.3 | 28.8 | LYS | 79 | 98.3 | 89.5 | 21.8 |
| VAL | 28 | 87.9 | 92.1 | 22.5 | MET | 80 | 1.0 | 1.6 | 1.5 |
| GLY | 29 | 0.0 | 1.3 | 1.5 | ILE | 81 | 142.3 | 138.7 | 19.8 |
| PRO | 30 | 0.0 | 3.0 | 3.2 | PHE | 82 | 31.9 | 33.5 | 16.2 |
| ASN | 31 | 2.4 | 10.0 | 6.2 | ALA | 83 | 114.5 | 82.8 | 13.2 |
| LEU | 32 | 1.1 | 6.9 | 6.6 | GLY | 84 | 9.8 | 11.4 | 5.2 |
| TRP | 33 | 65.5 | 110.3 | 35.6 | ILE | 85 | 3.0 | 7.2 | 6.5 |
| GLY | 34 | 29.5 | 64.8 | 22.6 | LYS | 86 | 90.3 | 106.4 | 35.5 |
| LEU | 35 | 0.4 | 10.8 | 9.9 | LYS | 87 | 131.9 | 94.3 | 36.3 |
| PHE | 36 | 42.7 | 91.8 | 33.1 | LYS | 88 | 237.9 | 132.2 | 47.9 |
| GLY | 37 | 54.8 | 81.3 | 22.3 | GLY | 89 | 77.2 | 53.1 | 22.3 |
| ARG | 38 | 31.1 | 30.7 | 20.4 | GLU | 90 | 50.6 | 72.2 | 26.2 |
| LYS | 39 | 137.7 | 129.1 | 40.1 | ARG | 91 | 0.0 | 9.5 | 11.6 |
| THR | 40 | 2.6 | 6.3 | 4.4 | GLN | 92 | 28.5 | 40.5 | 22.3 |
| GLY | 41 | 10.4 | 20.5 | 11.7 | ASP | 93 | 25.1 | 53.8 | 27.4 |
| GLN | 42 | 56.0 | 121.0 | 56.3 | LEU | 94 | 4.4 | 6.8 | 4.9 |
| ALA | 43 | 16.1 | 22.6 | 15.0 | VAL | 95 | 3.1 | 29.5 | 14.2 |
| GLU | 44 | 192.2 | 228.3 | 63.8 | ALA | 96 | 17.6 | 53.1 | 14.5 |
| GLY | 45 | 64.8 | 57.2 | 30.1 | TYR | 97 | 11.4 | 31.2 | 10.0 |
| TYR | 46 | 32.5 | 40.0 | 12.1 | LEU | 98 | 0.5 | 6.2 | 6.0 |
| SER | 47 | 52.1 | 56.6 | 23.8 | LYS | 99 | 65.5 | 60.1 | 16.3 |
| TYR | 48 | 13.2 | 11.1 | 8.3 | SER | 100 | 50.1 | 50.6 | 20.6 |
| THR | 49 | 29.3 | 14.2 | 9.3 | ALA | 101 | 21.0 | 36.7 | 10.1 |
| ASP | 50 | 125.1 | 111.9 | 13.8 | THR | 102 | 0.0 | 11.9 | 8.8 |
| ALA | 51 | 37.7 | 22.6 | 12.4 | SER | 103 | 101.0 | 106.5 | 17.0 |
| ASN | 52 | 0.0 | 1.1 | 1.1 | HEM | 104 | 38.9 | 51.4 | 13.0 |
| | | | | | total | | 5667.1 | 6285.5 | 228.1 |

solvent-accessible surface area is that by Lee and Richards.[16] The area is defined as that swept out by the center of a sphere of 1.4-Å radius (modeling a water molecule) when the sphere rolls over the van der Waals surface of the protein. The average surface area and its fluctuation are calculated from the last 95 ps of the MD simulation. Table I lists the results for the average surface area and its fluctuations in comparison with the area of the X-ray crystal structure. Although several residues are completely inaccessible in the X-ray structure (Gly29, Pro30, Asn52, Tyr67, Leu68, Arg91, and Thr102), these at least become transiently exposed to the solvent in the simulation. Two of these residues, Asn52 and Tyr67, are considered to play an important role in conformational switching during the electron transfer reaction.[2,3,17] In general, polar residues have slightly larger average surface areas in the simulation than in the crystal structure. The average exposed surface area per polar residue is 83 Å² with an average fluctuation of 30 Å², compared to the average surface area per polar residue of 78 Å² in the crystal structure. Slightly larger increases are found for nonpolar residues, with per residue average exposed areas of 59 Å² and fluctuations of 21 Å² in the simulation, versus an average exposed area of 53 Å² in the X-ray structure. The total surface area is increased by 11% compared to the X-ray structure. The total surface area of the X-ray structure is equal to that of a sphere of radius 21.2 Å, while the corresponding result for the simulation is 23.4 Å. These radii are close to the effective Stokes–Einstein radius of 21 Å from diffusion constant measurement at 25 °C.[18] The surface areas of some of the long side chains (e.g., Lys, Arg and Met) have larger relative fluctuations than do those of the short residues. The average surface area and fluctuation per residue of Lys, Arg, and Met are 99 ± 37 Å² while those for all residues are 60 ± 2 Å.² The magnitude of the fluctuation of the total surface area is small, about 4%. The fluctuation is most pronounced at very low frequencies. Figure 8 shows the total surface area fluctuation during the simulation time (only a 60-ps window is plotted for clarity) and its Fourier transform. As can be seen from the Fourier spectrum, there is a high density of fluctuation below 0.01 ps⁻¹. However, the frequency of 0.01 ps⁻¹ is only an upper limit because of our limited simulation time and thus restricted Fourier window.

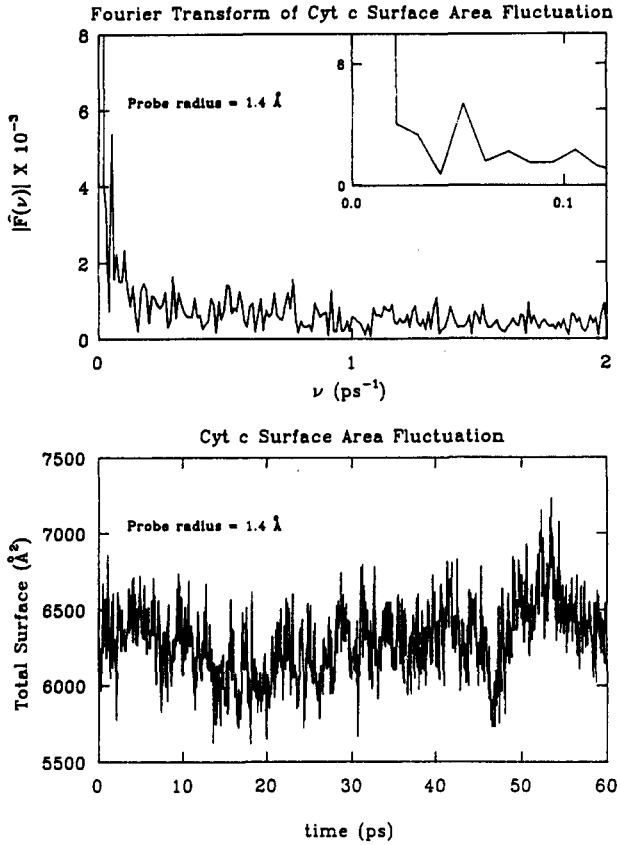


**Figure 8.** Total surface area accessible to a 1.4-Å radius probe as a function of time (lower plot) and the absolute value of its Fourier transform (upper plot). Only a 60-ps window from the full simulation is shown in the lower plot.

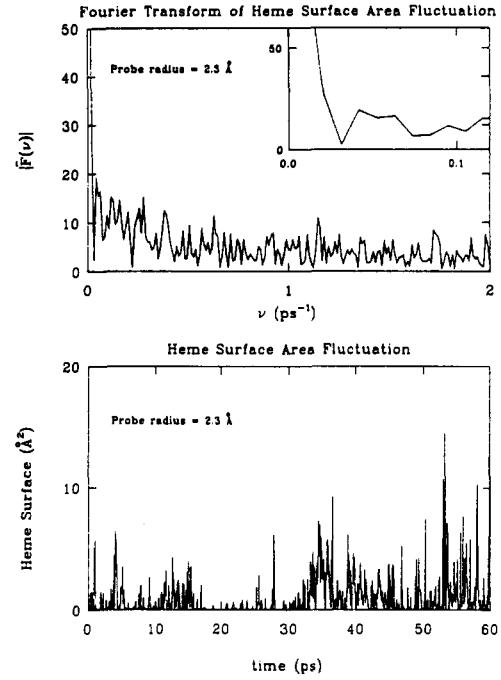


**Figure 9.** Heme surface area accessible to a 2.3-Å radius probe as a function of time (lower plot) and the absolute value of its Fourier transform (upper plot). Only a 60-ps window from the full simulation is shown in the lower plot.

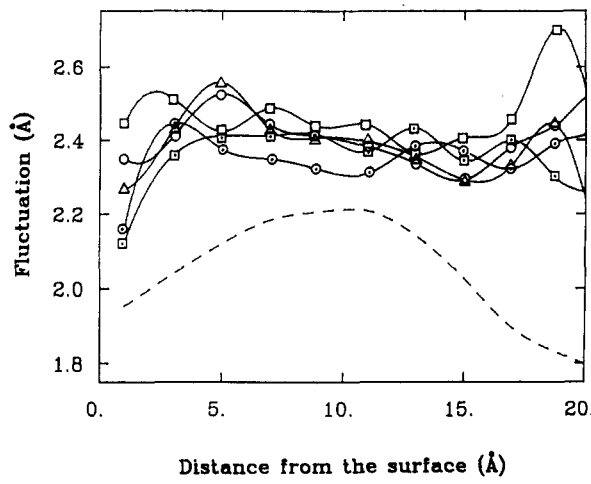


**Figure 10.** Average root-mean-square atomic fluctuation of water oxygens in MD vs the distance to the surface of the protein. The ○, □, △, ⊙, and ⊡ stand for the first, ... to the fifth 20 ps, respectively. The dashed line shows the distribution of the waters.

The extent of the heme exposure is crucial to the electron transfer process involving a redox partner such as cytochrome peroxidase or inorganic oxidant $[Fe(CN)_6]^{3-}$. It is seen from Table I that the fluctuation of the solvent-accessible area of the heme (HEM) is about 25% of the average area. The accessible area decreases with the increasing radius of the probe sphere. To a probe of radius 2.3 Å, the heme is inaccessible in the average MD structure. To see how the heme crevice opens and closes, we plot the time evolution of the heme surface area accessible to a 2.3 Å probe and the Fourier transform in Figure 9. From the Fourier spectrum, we can see that the opening frequency is also low (<0.01 $ps^{-1}$), but the high-frequency intensity is significant, suggesting that the large opening of the heme crevice at low frequencies is accompanied by small openings at higher frequencies. The average heme area accessible to the 2.3-Å probe is 1.1 $Å^2$ with a fluctuation of 1.8 $Å^2$. The ratio of the time that the accessible area of the heme fluctuates to more than 1.8 $Å^2$ versus the total time of the simulation is 0.2.

We have also looked at the mobility of water as a function of distance from the surface of the protein (Figure 10). The distance of a water molecule to the surface of the protein was calculated as the distance between the average atomic position of the water oxygen and that of the closest protein atom. The mobility of water increases as one moves away from the surface of the protein. In contrast to previous results for trypsin, we did not observe a clear intermediate region with higher mobility.[19] We have calculated temperatures of water in different spherical shells about the protein center of mass and found that the temperature at the protein-water interface was slightly higher (about 5–10 K) than in the rest of the system. Preliminary analysis suggests that the slightly elevated temperature at the protein–solvent interface is partly due to the slow equilibration of this region of the system. Although the structure of the solvent (initially in a configuration representative of pure liquid water) relaxes substantially during the 20-ps equilibration period, some additional relaxation occurs during the subsequent 105 ps. As shown in Figure 2, the number of water molecules inside the protein was increasing during the 105-ps simulation. The use of a 9 Å cutoff for the electrostatic interactions also appears to contribute to the slight warming of water near the protein surface (Shen et al., unpublished). A preliminary simulation using a 16 Å cutoff appears to reduce the temperature elevation near the protein–water interface. Our tentative explanation is that water molecules close to the protein experience discontinuous changes in their interactions with charged protein atoms more frequently than do water molecules far from the protein.

Temperature-dependent measurements have shown that proteins can behave like glass at low temperatures.[5,20–27] It is usually believed that the glasslike property is the consequence of multiple conformational states that can occur in protein molecule. Recent experimental and theoretical studies suggest that the glasslike behavior observed in proteins could be controlled by the surrounding solvent.[28,29] In our theoretical studies, glassy states

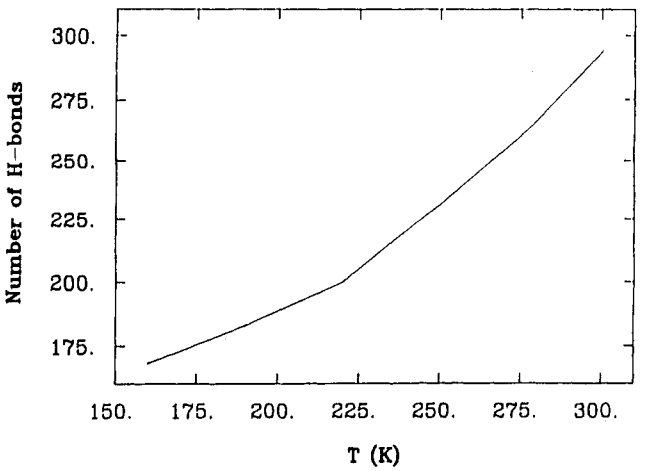


**Figure 11.** Plot of the number of different hydrogen bonds observed in the protein as a function of temperature. A hydrogen bond is said to exist if the distance between the hydrogen acceptor and the hydrogen is less than 0.25 nm and the donor–hydrogen–acceptor angle is greater than 135°.

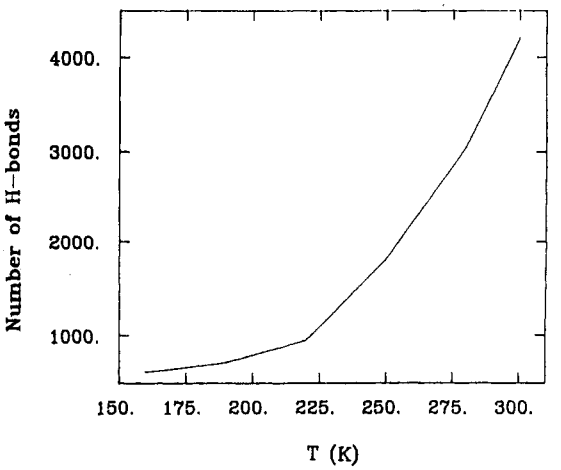


**Figure 12.** Plot of the number of different hydrogen bonds observed in liquid water as a function of temperature. The number of different hydrogen bonds reflects the mobility of water at different temperatures. That is, this number refers to the number of hydrogen bonds observed between distinct donors and acceptors, and will increase monotonically with the period of simulation if the molecules are mobile.

were formed by quenching the cytochrome *c*–water system down to a low temperature of 160 K and then reheating the system to 300 K. [The cooling and heating rates in a computer simulation are much larger than those in actual experiments so that the transition temperatures obtained in simulations may not be exactly the same as those obtained experimentally.] By looking at the temperature dependence of different quantities as the protein is reheated (one example is shown in Figure 11), one can see a transition at about 220 K. A parallel simulation of pure water with similar cooling and heating rates shows a similar transition temperature (an example is shown in Figure 12). These results suggest that glass transitions observed in proteins could be influenced by the surrounding solvent. This argument is reinforced by the fact that the same kind of transition has not been observed in a dry protein.[26]

These results have demonstrated the potential of using molecular dynamics to understand the behavior and function of cytochrome *c*. We have seen that simulating the fluctuations of cytochrome *c* can help to relate the flexibility of the protein to its function. The opening of the heme crevice, for instance, has been found to fluctuate at a low frequency and this opening frequency can determine how easily an inorganic oxidant can approach the heme group and subsequently react with the heme iron. More insights to the function of cytochrome *c* can be obtained as the simulation model is refined, which includes the incorporation of quantum nuclear tunneling effects as described in the next section, and the improvement of the potential model. The simulation needs to be run longer to allow sufficient time for equilibration, especially near the protein-water interface. The availability of fast supercomputers and high-speed workstations should make it easier to simulate this protein in an aqueous environment in the nanosecond timescale, with ~1000 CPU hours on the Cray-YMP. Extended analyses, including the ones described earlier and others such as the generalized Green's function approach described later, can then help to understand the properties and function of cytochrome *c* further.

## Quantum Simulations

Although classical mechanics is expected to provide an accurate description of many types of fluctuations in proteins, there are other phenomena where quantum effects will be important. Nuclear tunneling effects have been observed in the binding reaction of heme proteins[30] and in electron-transfer reactions.[31-33] The fact that some biological electron transfer reactions become temperature independent at temperatures lower than ~150 K suggests that there are substantial quantum effects even under physiological conditions.[31,34] Quantum mechanical nuclear tunneling may also contribute to the rate of an electron transfer that is conformationally gated.[34,35] Large isotope effects suggest a major role of quantum effects for proton transfer reactions, as well.[36-38] Tunneling between conformational substates may contribute to the thermal properties of proteins at low temperatures.[20,23,39,40] Thermal factors, measured by X-ray diffraction and the Mössbauer effect, are influenced by quantum effects, especially at low temperature.[26,41] Finally, because many modes of protein motion are high in frequency, zero-point energies have nonnegligible contributions to the overall thermodynamics.[42] Therefore in some respects biomolecules are at the borders of the domain of applicability of classical mechanics. The role of quantum mechanical effects in biomolecular structure and function is thus worth investigating.

To account for such quantum contributions, various semiclassical and quantum theoretical models have been formulated, such as those by Hopfield[43] and Jortner[44] in the theory of electron transfer. These theories, however, do not deal with the detailed correlation of structure and activity in proteins. With the assistance of fast computers, simulations of these detailed quantum aspects are becoming possible.

One approach, and perhaps the easiest to implement in a molecular dynamics (MD) or Monte Carlo (MC) scheme, is to make use of the isomorphism between the Feynman path integral formalism and polyatomic MD.[45,46] This approach yields time-averaged properties (e.g., root-mean-square fluctuations of atomic positions) quite readily; time-dependent properties are more difficult to calculate. The isomorphism, described in detail by Chandler and Wolynes,[47] is briefly reviewed in the following. The transition amplitude that a particle travels from A to B in space is equal to the transition amplitude that it travels from A to an intermediate point C, times the transition amplitude that it travels from C to the final destination B, with the summation over all possible intermediate points C. In quantum mechanical propagator notation, the above statement reads

$$\langle x_A|e^{-itH/\hbar}|x_B\rangle = \int dx_C \langle x_A|e^{-itH/2\hbar}|x_C\rangle\langle x_C|e^{-itH/2\hbar}|x_B\rangle \quad (1)$$

where $\hbar$ is the Planck's constant over $2\pi$, $t$ the time that the particle takes to travel from A to B, and $H$ the Hamiltonian. $\langle x_A|e^{-itH/\hbar}|x_B\rangle$ describes the transition amplitude of the particle traveling from A to B under the influence of the time evolution operator $e^{-itH/\hbar}$. One can insert more intermediate points D, E, .... Similarly, the canonical partition function in quantum statistical mechanics can be written as the products of two or

more factors over which intermediate variables are integrated:

$$Q(\beta) = \mathrm{tr}\, e^{-\beta H} = \int dx_1 \langle x_1|e^{-\beta H}|x_1\rangle = \int dx_1\, dx_2 \langle x_1|e^{-\beta H/2}|x_2\rangle\langle x_2|e^{-\beta H/2}|x_1\rangle = \ldots = \int dx_1 \ldots dx_P \langle x_1|e^{-\beta H/P}|x_2\rangle\langle x_2|e^{-\beta H/P}|x_3\rangle\langle x_3|e^{-\beta H/P}|x_4\rangle \ldots \langle x_{P-1}|e^{-\beta H/P}|x_P\rangle\langle x_P|e^{-\beta H/P}|x_1\rangle \quad (2)$$

where $\beta^{-1} = k_B T$, $k_B$ is the Boltzmann's constant, and $T$ the absolute temperature of the system. The transition amplitude $\langle x_i|e^{-\beta H/P}|x_{i+1}\rangle$ can be rewritten as $\langle x_i|\exp(-(i\tau/P\hbar)H)|x_{i+1}\rangle$, where $\tau = \beta\hbar/i$ is an imaginary time. Thus, a quantum statistical density matrix element can be obtained from transition amplitudes involving imaginary rather than real-time evolution operators.

When $P$ in eq 2 is large enough so that $\beta H/P$ is very small, $\langle x_i|e^{-\beta H/P}|x_{i+1}\rangle$ is essentially the propagator for a free particle in a small perturbation potential where only the first order correction needs to be considered. It can be shown that this requirement is met when

$$P \gg \beta\hbar^2/m\sigma^2 \quad (3)$$

where $\sigma$ is the characteristic length scale over which the potential function in the Hamiltonian changes.[48] For smooth potentials (such as of bond angles and dihedral angles) and heavy nuclear mass, small P will suffice. Whereas in the case of light mass such as of an electron, a large $P$ is needed.[49] When the first-order correction is added to the easily solvable equation for a free particle, the propagator is reduced to

$$\langle x_i|e^{-\beta H/P}|x_{i+1}\rangle \cong \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{3/2} \exp\left\{-\beta\left[\frac{1}{2}k_P (x_i - x_{i+1})^2 + \frac{1}{P}V(x_i)\right]\right\} \quad (4)$$

where

$$k_P = mP/\beta^2\hbar^2$$

This is a Boltzmann factor for a classical harmonic oscillator with force constant $k_P$ in a potential $V(x_i)/P$. Substituting each of the propagators in eq 2 with eq 4, we can rewrite the quantum canonical partition function as

$$Q(\beta) \cong \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{3P/2} \int dx_1 \ldots dx_P\, e^{-\beta\Phi_P(x_1 \ldots x_P;\beta)} \quad (5)$$

where

$$\Phi_P(x_1 \ldots x_P;\beta) = \frac{1}{2}k_P\sum_{i=1}^{P}(x_i - x_{i+1})^2 + \frac{1}{P}\sum_{i=1}^{P}V(x_i) \quad (6)$$

or

$$Q(\beta) \cong \int dp_1 \ldots dp_P \int dx_1 \ldots dx_P \exp[-\beta H_{\mathrm{eff}}] \quad (7)$$

where

$$H_{\mathrm{eff}} = \sum_{i=1}^{P}\frac{p_i^2}{2m'} + \Phi_P(x_1 \ldots x_P;\beta) \quad (8)$$

$m'$ is chosen to recover the preintegral factor in eq 5 when the integration over the momentum space is carried out. The quantity $p_i$ is a fictitious momentum conjugate to $x_i$.

In this formalism, a quantum atom is equivalent to a cyclic ring (or "necklace" of $P$ classical particles (or "beads"). The $i$th bead is connected to two adjacent beads by a quantum spring with force constant $k_P$ and is subject to an external potential $V(x_i)/P$. The quantum spring reflects the causality of the wave propagation of the atom in the path integral formulation. The cyclic topology is due to the fact that the quantum partition function in eq 1 is the integral of an imaginary time propagator leaving and returning to the same point in space. Because of the uncertainty principle, the position of a quantum particle can not be sharply defined at thermal equilibrium. The breadth of the particle's thermal wavepacket is reflected in the average width of the necklace. The isomorphism can be extended to many-particle systems.[47] In that case, if the potential is from another quantized atom, i.e., a necklace, then only the beads in the same position (same $i$) interact with each other as a result of causality. Through this isomorphism a quantum system can be simulated by the MD method on a classical system characterized by an effective Hamiltonian $H_{\mathrm{eff}}$ in eq 8. This procedure has been used to simulate a wide range of systems ranging from the simple[48] to the complex.[49] One of the powerful attributes of this approach to quantum simulations is that it can deal with complicated anharmonic Hamiltonians. Most treatments of quantum effects on nuclear motion in proteins have been tied to harmonic approximations. An important contribution of this type is the study of the photosynthetic reaction center by Warshel and co-workers[50] which used classical molecular dynamics to infer an effective harmonic oscillator model for use in estimating electron transfer rates. Although these approximations may be valid for many of the high frequency modes of motion with the largest quantum effects, they cannot deal well with tunneling motions, as in rotations of side chains, and may not do well when quantum harmonic modes are strongly coupled to anharmonic low-frequency modes. One issue, of course, is whether these anharmonic effects are quantitatively important, however.

**TABLE II: Average Deviations from the X-ray Structure for Various Classes of Atoms[a]**

| atom class | classical | quantum | classical-quantum |
|---|---|---|---|
| all atoms | 1.71 | 1.71 | 0.63 |
| interior | 1.26 | 1.26 | 0.42 |
| exterior | 2.11 | 2.12 | 0.81 |
| non-haem | 1.76 | 1.76 | 0.65 |
| haem | 0.53 | 0.62 | 0.28 |
| backbone | 1.27 | 1.32 | 0.49 |
| side chains | 1.87 | 1.87 | 0.68 |
| $\alpha$-carbon | 1.30 | 1.36 | 0.50 |
| $\beta$-carbon | 1.39 | 1.42 | 0.53 |
| $\gamma$-atoms | 1.69 | 1.73 | 0.58 |
| $\delta$-atoms | 2.06 | 2.05 | 0.72 |
| $\epsilon$-atoms | 2.42 | 2.36 | 0.97 |
| atoms quantized in QD | 0.81 | 0.83 | 0.36 |

[a] All deviations are in Å. The average deviation of an atom is the distance of the average position of the atom in the simulation relative to the position of the same atom in the X-ray structure.

Because of the many degrees of freedom in proteins, quantum simulations demand large computational resources and are yet to be widely undertaken. With the investigation of quantum effects in electron-transfer reactions in mind, we have performed quantum simulations on tuna ferrocytochrome $c$ in which only residues that are thought to be involved in redox conformational changes were quantized.[51] Explicit water molecules have not yet been included in the quantum simulations. Rather, a distance dependent dielectric and a scaling of atomic partial charges according to their distances from the center of mass of the protein[7] were used to model the screening effect of water.

Many of the results of the quantum simulations turn out to be similar to what is found classically. As can be seen from Table II, parallel runs of classical MD and quantum simulation indicate that the positional fluctuations of the atoms are smaller in the interior than in the exterior, and smaller in the backbone than in the side chains. The further away an atom in the side chain is from the backbone, the larger the fluctuation. The heme group, which has a relatively rigid structure compared with the polypeptide and which is in the interior of the protein, shows relatively small fluctuations. The similarity of the quantum and classical results for the atomic positional fluctuations is expected,

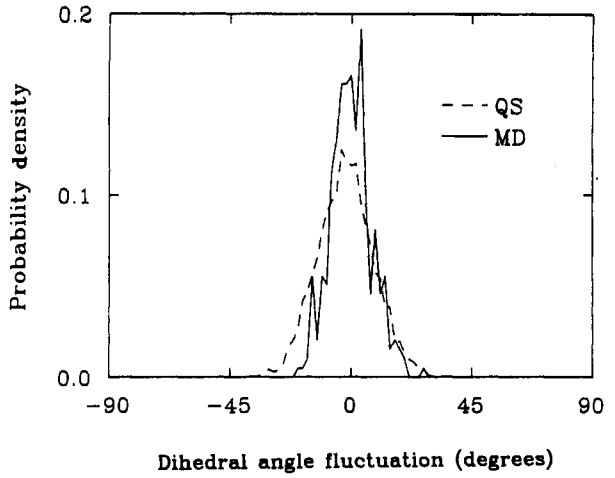


**Figure 13.** Asn 52 $NH_2$ dihedral angle distributions from the classical (—) and quantum (- - -) simulations. The distribution is of the deviations from the average dihedral angle in the fluctuating protein.

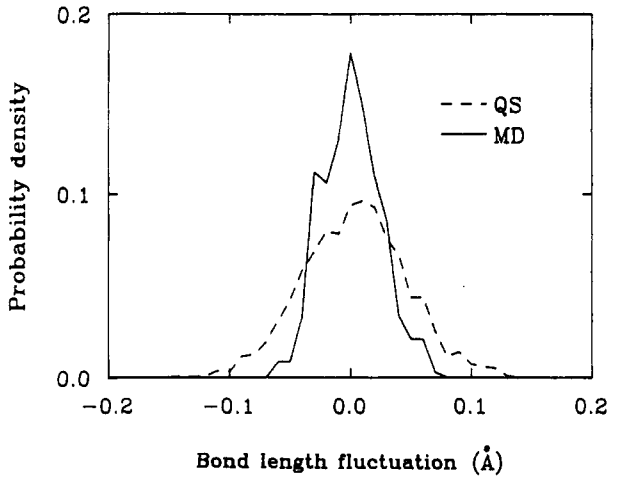


**Figure 14.** Fe–N(His 18) bond length distributions from the classical (—) and quantum (- - -) simulations. The distribution is of the deviations from the average bond length in the fluctuating protein.

because these fluctuations are dominated by low-frequency motions[5] and quantum corrections decrease with decreasing frequency.

There are significant differences in the classical and quantum distributions of other properties, however. Figure 13 shows the $NH_2$ dihedral angle distributions for Asn 52 from the two simulations. The width of the quantum distribution is larger, which is typical of quantum distributions. As in a harmonic oscillator,[43] the quantum effect should be more pronounced for larger ratios of $(\hbar\omega)/(k_BT)$. Figure 14 shows the Fe–N (His 18) bond-length distributions. The quantum effect is more apparent because the bond stretching has higher frequency, whereas the quantum signature of low-frequency modes in the protein is weak and can be buried in the noise in the simulation.

Both the classical and the quantum simulations give similar descriptions of the hydrogen bonding scheme. However, there is a difference in the Trp 59 NH–heme propionate 1O hydrogen bond. Because of the rotation of the propionate group at the simulation temperature (300 K), Trp 59 NH formed hydrogen bonds with 1O1 and 1O2 of the propionate group with nearly equal probability in the classical simulation (47% vs 49%). The occurrence of the two hydrogen bonds in the quantum simulation was less symmetrical, 41% for 1O1 and 29% for 1O2. Whether the breaking of the symmetry is of quantum nature or due to the limited simulation time is worth further investigation.

Although quantum simulations of large biomolecules are still at a very early stage of development, they are nevertheless feasible. The technique can be used to study proton transfer and certain conformational transitions, for example, $NH_2$ rotation, where quantum dispersion and tunneling effects are expected to be significant.

In biological reactions where conformational changes control the rate, such as in the conformationally gated electron transfer reactions between cytochrome *c* and cytochrome *b5*,[35] nuclear tunneling effects should play an important role. In fact, it has been shown for a similar heme protein system that classical theories do not adequately describe the kinetics even at physiological temperatures.[52] Thus, quantum simulation techniques should be used for such systems.

The quantum simulation technique outlined above can be used not only to sample configurational space in proteins, but also to estimate reorganizational activation energies and the quantum contributions. This requires a combination of the path integral quantum simulation technique[47] and the thermodynamic integration method.[5] This simulation scheme was initially applied successfully to small inorganic systems such as $Co(NH_3)_6$ in aqueous solution.[53,54] Recently, it was also used to calculate the activation energy for self-exchange electron transfer reactions in cytochrome *c*.[55] The quantum correction was estimated to be as significant as 20%. With this scheme, one can also estimate the effects of mutations on the rate of electron or proton reactions, which are of quantum nature.

### Generalized Green's Function Approach

Methods to help predict how the structure of a biomolecule may be altered when disturbances are introduced in different parts of the molecule would be useful for the design of novel therapeutic agents. For instance, one might wish to design a ligand which binds to a protein and changes the conformation of the protein so as to enhance or diminish the ability of the protein to bind another ligand or to bind to a receptor. The systematic study of the structural responses of a protein when each of its amino acids is successively perturbed by small applied forces may also help to identify key residues that play significant roles in determining the folded conformation of the protein. This is because the amino acids that play important structural roles may be those that interact strongly with the other parts of the protein and the disturbances of these residues may give rise to large structural responses of the protein molecule.

Susnow et al.[56,57] recently introduced a molecular mechanics Green's function approach to study the structural responses of a molecule to small external forces applied to different parts of the molecule. The advantage of the method is that one can perform a single molecular mechanics calculation to study the structural responses of a molecule at a local energy minimum to disturbances applied to each atom of the molecule. However, there are several disadvantages of the method: (1) The molecular mechanics approach only consider a molecule at a local minimum. For studying biomolecules under physiological conditions, it is important to consider ensembles of structures that are accessible at finite temperatures. (2) Since the calculation of a molecular mechanics Green's function involves an inversion of a Hessian matrix of order $3N \times 3N$ where $N$ is the number of atoms in a molecule, it is difficult to apply the method to large biomolecules. (3) It is hard to include solvent effects in the molecular mechanics method.

To circumvent these problems, we have developed a generalized Green's function approach for studying structural relaxations of biomolecules due to forces applied to atoms (or groups of atoms) of the biomolecules. This approach is more appropriate for studying molecules at finite temperatures and can be more easily applied to study large biomolecules since no inversions of large matrices are involved. It is also easier to treat solvent effects with the new approach.

A generalized Green's function measures the extent to which the ensemble-averaged structure of a molecule is changed when

small forces are applied to atoms of the molecule. One can easily derive an expression for calculating the elements of a Green's function matrix $G_{ij}$ in the following manner:

Suppose a small perturbation potential $V_P$ of the form

$$V_P = \sum_i k_i(x_i - \langle x_i\rangle_0) \tag{9}$$

is applied to a system in which the atoms are interacting through a potential $V$. In eq 9, $x_i$ is a coordinate associated with atom $i$, $k_i$ is a constant, and $\langle ...\rangle_0$ denotes an average over an ensemble of an unperturbed system, i.e., a system with no perturbation potential $V_P$ applied. The perturbation potential $V_P$ imposes a force of $-k_i$ on atom $i$. An ensemble-averaged coordinate of atom $j$, $\langle x_j\rangle$, of the system with an applied perturbation potential $V_P$ can be written as

$$\langle x_j\rangle = \frac{\int x_j \exp[-\beta V - \beta\sum_i k_i(x_i - \langle x_i\rangle_0)]\, d\Gamma}{\int \exp[-\beta V - \beta\sum_i k_i(x_i - \langle x_i\rangle_0)]\, d\Gamma} \tag{10}$$

where $\int ...d\Gamma$ denotes an integration over all the atomic coordinates. For small $k_i$, eq 10 can be approximated by

$$\langle x_j\rangle = \frac{\int x_j[\exp(-\beta V)][1 - \beta\sum_i k_i(x_i - \langle x_i\rangle_0)]\, d\Gamma}{\int \exp(-\beta V)\exp[-\beta\sum_i k_i(x_i - \langle x_i\rangle_0)]\, d\Gamma} \tag{11}$$

Equation 11 can also be written in the form

$$\langle x_j\rangle = [\langle x_j\rangle_0 - \beta\sum_i k_i\langle \Delta x_i \Delta x_j\rangle_0] \times \langle \exp[-\beta\sum_i k_i(x_i - \langle x_i\rangle_0)]\rangle_0^{-1} \tag{12}$$

where $\Delta x_i = x_i - \langle x_i\rangle_0$. If one expands $\langle x_j\rangle$ on the left-hand side of eq 12 as

$$\langle x_j\rangle = \langle x_j\rangle_0 + \sum_i \frac{\partial\langle x_j\rangle_0}{\partial k_i} k_i + \mathcal{O}(\{k_i^2\}) \tag{13}$$

and comparing the coefficients of the linear terms in $k_i$ on both sides of eq 12, noting that

$$\langle \exp[-\beta\sum_i k_i (x_i - \langle x_i\rangle_0)]\rangle_0^{-1} = \frac{1}{1 + \mathcal{O}(\{k_i^2\})} \tag{14}$$

and assuming terms of order $k_i^2$ and higher are negligible, one can easily show that the elements $G_{ij}$ of a generalized Green's function matrix $G$ can be obtained from the familiar displacement correlation function[5]

$$C_{ij} = \langle \Delta x_i \Delta x_j\rangle / ((\Delta x_i^2)^{1/2}(\Delta x_j^2)^{1/2})$$

$$G_{ij} = d\langle x_i\rangle / df_j = \beta\langle \Delta x_i \Delta x_j\rangle = \beta C_{ij}((\Delta x_i^2)^{1/2}(\Delta x_j^2)^{1/2}) \tag{15}$$

where $f_j = -k_i$ is a component of a force applied to atom $j$. Each element of the Green's function matrix $G_{ij}$ measures the extent to which the averaged coordinate $\langle x_i\rangle$ is changed as a result of a small force $j_j$ applied to atom $j$. In its most general form, the Green's function matrix is of order $3N \times 3N$ where $N$ is the number of atoms in the system. However, rigid-body translation is usually not of interest and its contribution to the Green's function matrix can be easily removed. The overall tumbling motion of a molecule, on the other hand, is typically coupled to the intramolecular vibrations of the molecule. However, since the coupling is usually weak, it is often a good approximation to neglect the coupling between the overall tumbling motion and the intramolecular vibrations of the molecule. Accordingly, the contribution of the tumbling motion to the Green's function matrix can also be removed. The resulting Green's function matrix thus only accounts for structural deformations when small forces are applied to different parts of the molecule and is useful for identifying key amino acids that determine the conformation of a protein or for guiding the design of novel therapeutic agents which function by changing the conformations of their target biomolecules. The essence of eq 9 is that it quantitatively predicts the extent to which the average position of an atom is perturbed when a small force is applied to another atom. The extent of the positional perturbation not only depends on the degree to which the two atoms are correlated, but also depends on the positional fluctuations of the two atoms about their means.

Once a Green's function matrix is calculated, one can study how and to what extent the structure of a protein molecule can be altered when small forces are applied to different parts of the molecule. The correlation function in eq 9 has in fact been calculated for cytochrome $c$.[5] It shows that perturbations due to the binding of a ligand at certain parts of the protein could produce conformational changes at distant sites. The ability of a protein molecule to resist structural changes can also be measured by the Green's function.

The Green's function matrix can be manipulated in different ways to gain different insights into structure and function of a biomolecule. For example, one can calculate $G_j^r = \sum_i (G_{ij}^2)^{1/2}$. Amino acid residues with large $G_j^r$'s may play important roles in determining the structure of a protein molecule. One can also study collective structural responses of a protein molecule by diagonalizing the Green's function matrix. The analysis is very similar to a quasi-normal mode analysis.[58-61] In a quasi-normal mode analysis, one focuses on obtaining the effective normal modes of a molecule by using a potential of mean force derived from a molecular dynamics simulation. The eigenvalues of the effective normal modes can then be used to provide estimates of thermodynamic quantities, and the eigenvectors can give useful insights into the local and collective motions of the molecule. In a Green's function analysis, on the other hand, one can directly identify groups of atoms that cooperate to produce large (or small) local or global structural responses of a molecule by diagonalizing a Green's function matrix. For large biomolecules, the diagonalization of a Green's function matrix of order $3N \times 3N$ is difficult. However, the principal component analysis technique[62] provides a useful alternative. As discussed earlier,[63] a principal component analysis can be achieved through a singular value decomposition of an $m \times n$ matrix, where $m$ and $n$ are integers and need not be equal. When a singular value decomposition is applied to a Green's function submatrix of dimension $m \times n$, one can study how small forces applied to the part of the protein defined by n propagate to affect the part of the molecule defined by $m$.

There are other useful manipulations of the Green's function matrix and they will be discussed in future publications. Here, we use the Green's function method to identify the amino acids of cytochrome c that may play important roles in determining the structure of the protein. Although we have not yet calculated the Green's function matrix of cytochrome $c$ using the trajectories presented in this paper, we have used the displacement correlation matrix obtained earlier[5] from a simulation of an isolated cytochrome $c$ molecule to gain some preliminary insights (Figure 15). From the displacement correlation matrix, one finds that each residue or group of residues close to Asn22, Leu32, Arg38, Trp59, Leu68, Gly77, and Ser100 is strongly coupled to many other parts of the protein. Equation 9 then predicts that the structure of the protein is relatively sensitive to disturbances of these residues, and these residues may play important roles in determining the folded conformation of the protein. Mutations

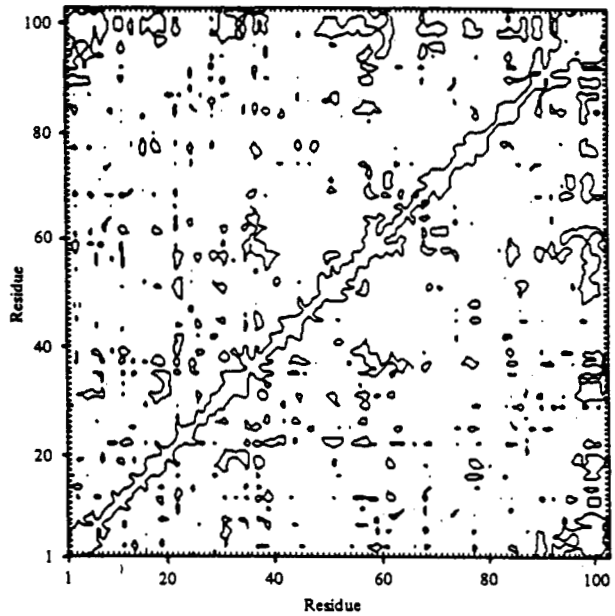


**Figure 15.** Displacement correlation matrix of cytochrome $c^{5,64}$ obtained from a 32-ps dynamical simulation. The correlation of residues i and j is computed as the normalized quantity $C_{ij} = \langle R_i R_j \rangle / \langle R_i^2 \rangle^{1/2} \langle R_j^2 \rangle^{1/2}$, where $R_i$ is the instantaneous displacement of the centroid of residue i from the mean position of the centroid. The values of $C_{ij}$ range from 0.6 to 1.0. A single contour is drawn at the value of 0.87.

of one or more of these residues may lead to conformations differ from the native protein, or may cause the protein to unfold. The hypothesis that these residues may play important structural roles is supported by the fact that all the residues except Asn22 and Ser100 are highly conserved residues in the family of cytochrome c proteins. Based on an analysis of the crystal structure of tuna cytochrome *c*, Takano and Dickerson have also suggested the possible roles of Leu32, Arg38, Trp59, Leu68, Gly77 in stabilizing the structure of the native protein. This preliminary analysis suggests that the Green's function approach may be a useful tool for identifying important components that play significant roles in determining the structure of a protein molecule. It will be of interest to repeat this analysis by using the results of a simulation of cytochrome c in which solvent effects are modelled more accurately, such as the simulation presented in this article.

## Concluding Remarks

The simulation studies described above, although far from complete, illustrate the kinds of insights to cytochrome properties and activity that computer models can provide. Findings that may be worth further consideration are that the polar side chains at the surface may be more extended in solution than in the X-ray structure, that water molecules can penetrate deeply into the protein, that there are substantial fluctuations in the exposure of the heme edge, that nuclear tunneling may significantly influence the electron-transfer kinetics, and that certain residues may play key roles in determining the native structure of the protein.

**Acknowledgment.** This work has been supported in part by the National Science Foundation, the National Institutes of Health, the Robert A. Welch Foundation, the Petroleum Research Fund, the National Center for Supercomputer Applications, the Pittsburgh Supercomputing Center, and the San Diego Supercomputer Center. C.Z. is the recipient of a Presidential Young Investigator Award from NSF. J.A.M. is the recipient of the George H. Hitchings Award from the Burroughs Wellcome Fund.

## References and Notes

(1) Feng, Y.; Roder, H.; Englander, S. W. *Biochemistry* **1990**, *29*, 3494.
(2) Takano, T.; Dickerson, R. E. *J. Mol. Biol.* **1981a**, *153*, 79.
(3) Takano, T.; Dickerson, R. E. *J. Mol. Biol.* **1981b**, *153*, 95.
(4) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Adv. Chem. Phys.* **1988**, *71*, 1.
(5) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge 1987.
(6) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature (London)* **1977**, *267*, 585.
(7) Northrup, S. H.; Pear, M. R.; Morgan, J. D.; McCammon, J. A.; Karplus, M. *J. Mol. Biol.* **1981**, *153*, 1087.
(8) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(9) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(10) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haakk, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(11) Ansari, A.; Berendzen, J.; Browne, S. F.; Frauenfelder, H.; Iben, I. E. T.; Sauke, T. B.; Shyamsunder, E.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5000.
(12) Elber, R.; Karplus, M. *Science* **1987**, *235*, 318.
(13) Stein, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3670.
(14) Hasinoff, B. B.; Licht, A.; Pecht, I. *Biochim. Biophys. Acta* **1984**, *767*, 627.
(15) Wendoloski, J. J.; Mathew, J. B.; Weber, P. C.; Saleme, F. R. *Science* **1987**, *238*, 794.
(16) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379.
(17) Williams, G.; Moore, G. R.; Williams, R. J. P. *Commun. Inorg. Chem.* **1985**, *4*, 55.
(18) Koller, K. B.; Hawkridge, F. M. *J. Electroanal. Chem.* **1988**, *239*, 291.
(19) Wong, C. F.; McCammon, J. A. *Isr. J. Chem.* **1986**, *27*, 211.
(20) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C. *Biochemistry* **1975**, *14*, 5355.
(21) Bauminger, E. R.; Cohen, S. G.; Nowik, I.; Ofer, S.; Yariv, J. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 736.
(22) Frauenfelder, H.; Petsko, G. A.; Tsernoglou, F. *Nature* **1979**, *280*, 558.
(23) Goldanskii, V. I.; Fleurov, V. N. *Sov. Sci. Rev. B. Chem.* **1987**, *9*, 59.
(24) Hartmann, H.; Parak, F.; Steigemann, W.; Petsko, G. A.; Ponzi, D. R.; Frauenfelder, H. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4967.
(25) Knapp, E. W.; Fisher, S. F.; Parak, F. *J. Phys. Chem.* **1982**, *86*, 5042.
(26) Parak, F.; Frolov, E. N.; Mossbauer, R. L.; Goldanskii, V. I. *J. Mol. Biol.* **1981**, *145*, 825.
(27) Parak, F.; Knapp, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 7088.
(28) Ansari, A.; Berendzen, J.; Braunstein, D.; Cowen, B. R.; Frauenfelder, H.; Hong, M. K.; Iben, I. E. T.; Johnson, J. B.; Ormos, P.; Sauke, T. B.; Scholl, R.; Schulte, A.; Steinbach, P. J.; Vittitow, J.; Young, R. D. *Biophys. Chem.* **1987**, *26*, 337.
(29) Wong, C. F.; Zheng, C.; McCammon, J. A. *Chem. Phys. Lett.* **1989**, *154*, 151.
(30) Frauenfelder, H.; Wolynes, P. G. *Science* **1985**, *229*, 337.
(31) Chance, B., et al., Eds.; Academic Press: New York, 1979.
(32) DeVault, D. *Quantum-Mechanical Tunnelling in Biological Systems*, 2nd ed.; Cambridge University Press: New York, 1979.
(33) Mayo, S. L.; Ellis, W. R., Jr.; Crutchley, R. J.; Gray, H. B. *Science* **1986**, *233*, 948.
(34) Kihara, T.; McCray, J. A. *Biochem. Biophys. Acta* **1973**, *292*, 297.
(35) McLendon, G.; Pardue, K.; Bak, P. *J. Am. Chem. Soc.* **1987**, *109*, 7540.
(36) Albery, W. J.; Knowles, J. R. *J. Am. Chem. Soc.* **1973**, *292*, 297.
(37) Kreevoy, M. M.; Ostovic, D.; Lee, I. S. H.; Binder, D. A.; King, G. W. *J. Am. Chem. Soc.* **1988**, *110*, 524.
(38) Kreevoy, M. M.; Ostovic, D.; Truhlar, D. G.; Garrett, B. C. *J. Phys. Chem.* **1986**, *90*, 3766.
(39) Goldanskii, V. I.; Krupyanskii, Y. F.; Fleurov, V. N. In *Protein structure: molecular and electronic reactivity*; Austin, R., et al., Eds.; Springer: Berlin 1987.
(40) Goldanskii, V. I.; Krupyanskii, Yu. F.; Flerov, V. N. *Dokl. Biophys.* **1984**, *272*, 209.
(41) Ringe, D.; Petsko, G. A. *Prog. Biophys. Mol. Biol.* **1985**, *45*, 197.
(42) Friedman, H. L. *A Course in Statistical Mechanics*; Prentice-Hall: Englewood Cliffs, NJ, 1985.
(43) Hopfield, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3640.
(44) Jortner, J. *J. Chem. Phys.* **1976**, *64*, 4860.
(45) Feynman, R. P. *Statistical Mechanics*; Benjamin: London, 1982.
(46) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.
(47) Chandler, D.; Wolynes, P. G. *J. Chem. Phys.* **1981**, *74*, 4078.
(48) Berne, B. J.; Thirumalai, D. *Annu. Rev. Phys. Chem.* **1986**, *37*, 401.
(49) Kuki, A.; Wolynes, P. G. *Science* **1987**, *236*, 1647.
(50) Warshel, A.; Chu, Z. T.; Parson, W. W. *Science* **1989**, *246*, 112.
(51) Zheng, C.; Wong, C. F.; McCammon, J. A.; Wolynes, P. G. *Nature* **1988**, *334*, 726.
(52) Peterson-Kennedy, S. E.; McGourty, J. L.; Kalweit, J. A.; Hoffman, B. M. *J. Am. Chem. Soc.* **1986**, *108*, 1739.
(53) Bader, J. S.; Kuharski, R.; Chandler, D. *J. Chem. Phys.* **1990**, *93*, 230.
(54) Zheng, C.; McCammon, J. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6441.

(55) Zheng, C.; McCammon, J. A.; Wolynes, P. G. *Chem. Phys.* **1991**, *158*, 261.
(56) Susnow, R.; Nachbar, R. B., Jr.; Schutt, C.; Rabitz, H. *J. Phys. Chem.* **1991**, *95*, 8585.
(57) Susnow, R.; Nachbar, R. B., Jr.; Schutt, C.; Rabitz, H. *J. Phys. Chem.* **1991**, *95*, 10662.
(58) Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices*; Clarenden: Oxford, 1954.
(59) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325.
(60) Levy, R.; Karplus, M.; Kushick, J. N.; Perahia, D. *Macromolecules* **1984**, *17*, 1370.
(61) Edholm, O.; Berendsen, H. J. C. *Mol. Phys.* **1984**, *51*, 1011.
(62) Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*; Wiley: New York, 1958.
(63) Wong, C. F.; Rabitz, H. *J. Phys. Chem.* **1991**, *95*, 9628.
(64) McCammon, J. A. *Rep. Prog. Phys.* **1984**, *47*, 1.