Molecular and Stochastic Dynamics of Proteins

Walter Nadler; Axel T. Brunger; Klaus Schulten; Martin Karplus

Proceedings of the National Academy of Sciences of the United States of America,
Volume 84, Issue 22 (Nov. 15, 1987), 7933-7937.

Stable URL:
http://links.jstor.org/sici?sici=0027-8424%2819871115%2984%3A22%3C7933%3AMASDOP%3E2.0.CO%3B2-0

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Proceedings of the National Academy of Sciences of the United States of America is published by National Academy of Sciences. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/nas.html.

Proceedings of the National Academy of Sciences of the United States of America
©1987 National Academy of Sciences

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
Molecular and stochastic dynamics of proteins

(Brownian motion/computer simulations/correlation functions/intramolecular fluctuations/myoglobin/Mössbauer spectroscopy studies)

WALTER NADLER*, AXEL T. BRÜNGER†§, KLAUS SCHULTEN*, and MARTIN KARPLUS†

*Department of Physics, Technical University of Munich, 8046 Garching, Federal Republic of Germany; and †Department of Chemistry, Harvard University, Cambridge, MA 02138

Contributed by Martin Karplus, May 29, 1987

ABSTRACT The rapid fluctuations of protein atoms derived from molecular dynamics simulations can be extrapolated to longer-time motions by effective single-particle stochastic models. This is demonstrated by an analysis of velocity autocorrelation functions for the atoms of lysine side chains in the active site of RNase A. The atomic motions are described by a bounded stochastic model with the friction and noise parameters determined from a molecular dynamics simulation. The low-frequency relaxation behavior is shown to result from collisional damping rather than dephasing. Extrapolation of these results to the quasistochastic motion of the heme group in myoglobin provides an explanation for Mössbauer spectroscopic data.

The internal motions of protein molecules occur over a wide range of time scales (1). They extend from the vibrations of single atoms and torsional oscillations of side chains in the femtosecond to picosecond range, through motions of larger protein segments on the nanosecond to microsecond time scale, to global fluctuations, including the transition to the unfolded state, that require microseconds to seconds or longer. Some of these motions are relevant to protein function (1–4).

An understanding of protein dynamics is thus a prerequisite for a complete description of protein function. Simulation of the atomic motions by molecular dynamics provides a theoretical approach to this problem (1, 5). Due to the computer-time requirements of such simulations, most studies of proteins have been limited to subnanosecond periods. Since many processes that are of functional importance or experimentally measurable occur on longer time scales, it is desirable to extend the theoretical methodology to include such events. Vibrational analyses of proteins have demonstrated that the lowest-frequency normal modes are near 3 cm⁻¹ (6), corresponding to periods of 30 ps. Since the normal mode description is not limited in time scale, this requires that longer time phenomena result from anharmonic effects, including transitions among multiple minima (7, 8). Examples include reactions with an activation barrier and slow conformational changes, such as the quaternary transition in hemoglobin.

One approach to long-time phenomena is stochastic modeling of the internal motions (9, 10), in which only the relevant portion of the protein is explicitly included and the remainder of the molecule, as well as the solvent, serves to provide an effective potential, a frictional drag, and a heat bath (11). We here make use of picosecond molecular dynamics results to implement a stochastic model for atomic and group motions in proteins. The damping and diffusion coefficients required for the model are derived from a stochastic boundary simulation for RNase A (12) and the atoms examined are those of lysine side chains. The resulting stochastic parameters are extrapolated to the motion of the heme group, which can be compared with phenomenological descriptions of myoglobin ⁵⁷Fe Mössbauer experiments (13–16). It is found, in contrast to a published assertion (13), that the molecular dynamics simulations and their stochastic extensions, including barrier crossings, provide a consistent picture of protein dynamics.

THEORETICAL FRAMEWORK

A molecular dynamics simulation determines the phase space trajectories of protein atoms by numerical integration of the Newtonian equations (1, 5). The motions of the densely packed protein atoms resemble the dynamics of classical fluid particles on a time scale during which diffusion transport does not exceed atomic diameters (1, 5). On longer time scales, account has to be taken of the fact that protein atoms are localized by their covalent interactions, while those in a fluid are not. The motion of individual fluid particles for times longer than the microscopic collision times can be modeled by use of effective single-particle equations of motion. The Langevin equation with a potential of mean force, $V(x)$,

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -m\gamma v - \nabla V(x) + A(t),$$

[1]

and the corresponding diffusion (Fokker–Planck) equation can be used (11, 17). In Eq. 1, $x$ and $v$ are the relevant particle coordinate and velocity, $\gamma$ is a friction coefficient, and $A(t)$ is a fluctuating force that can be approximated by Gaussian white noise on the time scale considered. One expects such a simplification to be applicable to the motion of atoms or groups of atoms in proteins for times long compared to the time of atomic vibrations (10). The protein and solvent environment exert the average potential $V(x)$ that confines the motion and determines the friction coefficient $\gamma$ and the amplitude of the noise term $A(t)$.

To investigate the stochastic character of internal protein motions one has to determine whether the results of molecular dynamics simulations can be reproduced by Eq. 1. Such studies, based on the displacement correlation function, have been made for tyrosine side chains (10) and other group motions (18) in proteins. Here we make use of the velocity autocorrelation function, $C_v(t) = \langle \langle v(t)v(0) \rangle \rangle$, which has been employed recently in determining the frictional coefficient for the Langevin buffer atoms in stochastic boundary simulations (12, 19). The velocity autocorrelation function is used because it relaxes rapidly (less than 1 ps), is dominated by frictional effects (i.e., the relaxation depends sensitively on $\gamma$), and is rather insensitive to the characteristics of the potential $V(x)$. This contrasts with the displacement autocorrelation function, which has a more complex behavior for protein atoms (18).

1Present address: Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, CA 91125.

4Present address: Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06472.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
METHOD

We have analyzed a stochastic boundary molecular simulation of a portion of RNase A that included the active site and the surrounding solvent molecules (12). The motions of the atoms of two active-site lysine residues (Lys-7 and Lys-41) were examined. Since a lysine side chain is long and flexible, the influence of fluctuations in the environment on the motion of its individual atoms should be significant. Both Lys-7 and Lys-41 are in the active site of RNase A, where they experience interactions with other protein atoms and with solvent molecules (see figure 1 of ref. 12). In the stochastic boundary simulation, all the atoms of the two lysines were treated by full molecular dynamics, except for the C\text{\textalpha} of Lys-41, which is in the buffer region and is, therefore, constrained and coupled to a random force (12, 19). A comparison of the behavior of these two side chains, which are otherwise in very similar environments, furnishes information on the range of the influence of the stochastic boundary.

The velocity autocorrelation functions were determined for a time interval between \( t = 0 \) and \( t = t_{\text{max}} \), where \( t_{\text{max}} \) is 1 ps, by averaging over a 30-ps molecular dynamics trajectory. The frequency spectrum \( C_\omega(\omega) = \int_0^{t_{\text{max}}} dt \cos(\omega t) C_\omega(t) \) was calculated by numerical integration over the time interval (0, \( t_{\text{max}} \)). Since \( C_\omega(t) \) is known only for a finite time interval, its spectrum is defined mathematically solely at the discrete values \( \omega = 2 \pi n / t_{\text{max}} \), \( n = 0, 1, 2, \ldots \). However, \( C_\omega(\omega) \) was evaluated also for other values of \( \omega \), which results in the presence of oscillations of frequency \( 2 \pi / t_{\text{max}} \), corresponding to the artificial period of \( t_{\text{max}} \) in \( C_\omega(t) \); these oscillations act to smooth the discrete spectrum, as an alternative to the introduction of a window function.

The decay of the velocity autocorrelation functions of the lysine atoms is so fast that during the relaxation process the atoms explore only the region in the neighborhood of the local potential minimum governing their average positions. Thus, the effective potentials are essentially harmonic and the velocity relaxation process can be represented by a Langevin harmonic oscillator model; i.e., in one dimension a potential of the form \( V(x) = k_B T_x x^2 \) is used in Eq. 1. For \( \omega_x > \gamma / 2 \), which corresponds to the present applications, the oscillations are underdamped. In this case the velocity autocorrelation function (11) is given by \( C_\omega(t) = (\gamma^2 - \omega^2)^{-1} \cos(\omega t) (\gamma^2 - \omega^2)^{-1} \cos(\omega t) \) where \( \omega_x = (k_B T_x / m)^{1/2} \gamma \). This expression has been fitted to the low-frequency portion (\( \omega < \approx 1.5 \text{ ps}^{-1} \)) of \( C_\omega(\omega) \) obtained from the molecular dynamics simulation; to yield agreement with the simulation results for certain atoms a superposition of two independent modes with different \( \omega_x \) values was used.

RESULTS

The velocity correlation function \( C_\omega(t) \) and the spectral densities \( C_\omega(\omega) \) obtained from the simulation are shown in Fig. 1 for the atoms of the Lys-7 side chain; the spectral densities for the atoms of Lys-41 are also shown for comparison. For every atom (see Fig. 1a), \( C_\omega(t) \) exhibits a decay with a relaxation time of less than 0.1 ps onto which fast oscillations are superimposed. In \( C_\omega(t) \) shown in Fig. 1b and c, the fast oscillations appear as discrete lines. The sharp lines near 220 ps\(^{-1} \) (\( \approx 1000 \text{ cm}^{-1} \)) correspond to skeletal vibrations of the lysine chain; this series of lines, associated with the motions of \( C^\beta \), \( C^\gamma \), and \( C^\delta \), leads to the beats modulating the rapidly oscillating component of \( C_\omega(t) \) in Fig. 1a. In addition, there is a line at about 300 ps\(^{-1} \) for the terminal \( N^\bullet \) atom that corresponds to a deformation mode of the \( NH_3 \) group.

The low-frequency region (0–100 cm\(^{-1} \)) of \( C_\omega(\omega) \) in Fig. 1b is congested due to a large number of lines. This feature, an essential property of all the spectral densities, is due to the interactions of the lysine atoms with their environment. The effect of the environment can be seen clearly from \( C_\omega(\omega) \) of a free lysine chain in the absence of solvent shown in Fig. 2; the free chain has been simulated with a harmonic constraint and a random force acting on the \( C^\alpha \) atom. The low-frequency part of \( C_\omega(\omega) \) for the atom \( N^\bullet \) of the free lysine consists of only a few resolved lines. This demonstrates that the broadening of the Lys-7 spectra is due to coupling to the immediate proteins and solvent environment. The \( NH_3 \) deformation mode for the free lysine chain lies at about 350 ps\(^{-1} \), in contrast to the value 300 ps\(^{-1} \) obtained for the hydrogen-bonded \( NH_3 \) group. The frequency lowering of the solvated \( NH_3 \) group is due to the fact that it is coupled to water molecules, thereby the effective mass of the vibration is increased. This interpretation is in accord with the fact that the line at 220 ps\(^{-1} \), which corresponds to a skeletal vibration of the neighboring \( C^\alpha \) atom, is not shifted by solvent though its intensity is increased.

Fig. 1c presents \( C_\omega(\omega) \) for the Lys-41 side chain, whose \( C^\alpha \) atom is in the stochastic buffer zone and so has its motion described by a Langevin equation, instead of Newtonian mechanics. As expected, \( C_\omega(\omega) \) for \( C^\alpha \) of Lys-41 shows a strongly broadened spectrum, in contrast to that for \( C^\alpha \) of Lys-7. However, the \( C^\alpha \) atom of Lys-41 already shows low-frequency behavior that is similar to that of Lys-7, and the other side-chain atoms of the two lysines have essentially identical low-frequency spectra. This indicates that the range of influence of the stochastic boundary is very short. Further, because the Langevin equation applied to \( C^\alpha \) of Lys-41 essentially uncouples the side chain from the protein, the similarity of the low-frequency spectra of Lys-7 and Lys-41 demonstrates that the dominant broadening mechanism arises from collisions with the environment (particularly solvent in this case), rather than coupling to vibrational modes of the rest of the protein. This is in accord with the absence of broadening for atoms other than \( C^\alpha \) in the free lysine simulation.

The congested vibrational bands observed in the low-frequency range can be approximated by one or two stochastically damped harmonic oscillators. The damping simulates collisional effects and any coupling between the closely spaced vibrational modes. When \( C_\omega(t) \) and \( C_\omega(\omega) \) are fitted to the molecular dynamics results by adjusting the values of \( \omega_0 \) and \( \gamma \), the parameters shown in Table 1 are obtained; only for \( C^\alpha \) and \( N^\bullet \) are two oscillators required. The broken lines in Fig. 1b and c correspond to the fits and demonstrate that excellent agreement is obtained in the low-frequency range. This supports the supposition that a stochastic single-particle model is adequate for describing these motions. It is such low-frequency modes that dominate the amplitudes of the atomic fluctuations (6, 18).

The friction coefficients for the various atoms lie in the range 19–45 ps\(^{-1} \) (see Table 1). They show excellent correspondence for the two lysine residues, in accord with their comparable protein and solvent environments. Diffusion coefficients calculated by use of the Einstein relation, \( D = k_B T / \pi \eta_r \), for the bounded diffusive motion of the lysine side chain atoms are in the range 4–10 \times 10^{-5} \text{ cm}^2/\text{s}, which is somewhat larger than the self-diffusion coefficient (3 \times 10^{-5} \text{ cm}^2/\text{s}) obtained for the water model used in the simulations.

DISCUSSION

It has been shown that the motions of protein atoms calculated by molecular dynamics simulations that extend over 100 ps exhibit properties that can be modeled by single-particle...
stochastic equations. The comparison of the results obtained for lysine side chains in three different simulations demonstrates that the dominant factor inducing the stochastic behavior is the collisions with the local environment (protein...
Fig. 2. Spectrum of the velocity autocorrelation function for the atom Nε of Lys-41. Solid line, in solution as part of protein; dotted line, unsolvated free chain.

and solvent) rather than the coupling with the high density of low-frequency modes of the remainder of the protein.

The present results apply to the oscillations of individual atoms of a flexible side chain in a low-frequency potential well. We now consider the extension of these results to the displacement of larger protein segments over longer times. Although the lysine atoms were found to be in the intermediate damping limit, purely diffusive behavior is expected for the motion of larger, relatively rigid protein segments (e.g., the heme group). Using Stokes law to scale the lysine atom diffusion coefficient, we write $D$ as proportional to $(m R_m)^{-1}$, where $m$ is the mass and $R_m$ is the effective hydrodynamic radius of the segment (20). To take account of the fact that longer time motions (longer than $\sim 100$ ps) involve transitions over barriers, it is necessary to go beyond the harmonic oscillator approximation for the potential $V(x)$. For a static potential with a repeating multiminimum structure, the effective diffusion coefficient, $D_{eff}$, has the form

$$D_{eff} = D_0 \exp(-\Delta E/k_B T),$$

(2)

where $D_0$ is the diffusion coefficient for motion in a well and $\Delta E$ is the height of the barriers separating the wells. For barrier heights significantly greater than $k_B T$, as required for Eq. 2 to be valid, the value of $\Delta E$ clearly cannot be determined from standard picosecond simulations because times of nanoseconds or longer are involved. Thus, $\Delta E$ must be obtained in another way; adiabatic dynamics simulations (22), or experiment (13-16) can be used.

The above considerations can be applied to the $^{57}$Fe Mössbauer effect in myoglobin and its phenomenological analyses (13-16). Several aspects of the experimental observations are important. The Lamb–Mössbauer factor, $-\ln f_0$, increases linearly with temperature between 0 and 170 K and then rises much more rapidly between 200 and 300 K. Since the Lamb–Mössbauer factor is related to the mean-square displacement that occurs in times shorter than the lifetime of the $^{57}$Fe excited state (0.14 μs), this suggests that there is a change in the protein dynamics between 180 and 220 K. Further, a broad line appears in the Mössbauer spectrum at about 180 K, and both its width and its intensity increase strongly with increasing temperature; the linewidths reflect relaxation processes on the time scale $1 \leq t < 100$ ns.

To model these results, the iron motion has been represented in one dimension (13-15), as well as in two and three dimensions (16). The strong temperature dependence of the width $\Gamma$ of the broad line (it increases by approximately a factor of 4 between 200 and 300 K) has been fitted by assuming that the iron moves in a temperature-independent multiminium potential with a harmonic envelope (15, 16). Use of Eq. 2 leads to $\Delta E$ equal to 1.8–2.1 kcal/mol (1 kcal = 4.18 kJ). $D_{eff}$ can be estimated from the magnitude of $\Gamma$, since it depends on the time scale of the Fe motion. With experimental estimates for $\Delta E$ and $D_{eff}$, Eq. 2 can be used to determine $D_0$. The resulting value for $D_0$ is $0.8-1.1 \times 10^{-7}$ cm$^2$/s (16), approximately 50 times larger than $D_{eff}$ at 250 K.

If $-\ln f_0$ is related to the mean-square displacement ($\delta x^2$) by the Gaussian approximation (see below), the temperature value of $\delta x^2$ is $0.065 \AA^2$. This leads to a constant $k/\delta x^2$ for the envelope potential equal to $k/\delta x^2 = 4.6 \times 10^2$ K/Å$^2$ from the relation $k/\delta x^2 = k_B T$. However, between 300 and 225 K, $-\ln f_0$ decreases much more rapidly than expected for a harmonic potential. To explain this temperature dependence, phenomenological models have been introduced in which there are one or more very narrow, deeper traps [\(\Delta E_{trap} \approx 4\) kcal, with $\delta x^2 = 0.1-0.2 \AA$ for one- and three-dimensional models, respectively (16)]. The observed temperature dependence of $-\ln f_0$ is then governed by the change in the probability of being excited to the envelope potential versus that of being in the deeper well.

The above description of the iron motion is generally consistent with simulation results for the dynamics of myoglobin (8), which demonstrate that the underlying potential surface is characterized by a large number of thermally accessible minima in the neighborhood of the native structure; the mean-square fluctuations are a superposition of oscillations within a well and transitions among wells. Computer simulations of myoglobin (K. Kuczera, J. Kuriyan, and M.K., unpublished results) have demonstrated that the iron is tightly coupled to the heme and that the longer-time, larger-scale iron displacements involve motions of the entire heme group. For the heme group, one estimates a diffusion coefficient $D_0$ of $2-5 \times 10^{-2}$ cm$^2$/s by Stokes law scaling of the lysine atom values. This is in accord with the estimate of $D_0$ from Mössbauer spectroscopy cited above. However, the extreme narrowness of the deep traps in the phenomenological models (13-16) is rather surprising. A possible interpretation is that the effective potential in which the iron moves is itself temperature dependent. It has been shown by crystallographic studies that myoglobin contracts as the temperature is lowered (23). Further, the onset of the increase in $-\ln f_0$ and the appearance of the broad line occur at a temperature in the neighborhood of 200 K, which coincides with the melting of the aqueous crystal environment of the protein (24). Normal mode and molecular dynamics simulations have shown that the dominant contribution to mean square displacements came from larger-scale correlated motions of many atoms (6, 18). Since these involve the protein surface, it is likely that they are quenched when the solvent freezes. The effect of this on the iron motion could be that it is trapped in one of a series of local minima, whose effective barriers are significantly higher in the low-temperature system. One possibility is a minimum in which the heme itself is trapped; from the simulation the root-mean-square

<table>
<thead>
<tr>
<th>Atom</th>
<th>$\omega_0$, ps$^{-1}$</th>
<th>$\gamma$, ps$^{-1}$</th>
<th>$\omega_0$, ps$^{-1}$</th>
<th>$\gamma$, ps$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$^\alpha$</td>
<td>28</td>
<td>45</td>
<td>27</td>
<td>40</td>
</tr>
<tr>
<td>C$^\beta$</td>
<td>28</td>
<td>19</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td>C$^\gamma$</td>
<td>31</td>
<td>26</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>C$^\delta$</td>
<td>13</td>
<td>22</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>N$^\varepsilon$</td>
<td>45</td>
<td>19</td>
<td>47</td>
<td>19</td>
</tr>
<tr>
<td>N$^\varepsilon$</td>
<td>14</td>
<td>40</td>
<td>24</td>
<td>40</td>
</tr>
<tr>
<td>N$^\varepsilon$</td>
<td>50</td>
<td>30</td>
<td>72</td>
<td>30</td>
</tr>
</tbody>
</table>

The quantities $\omega_0$ and $\gamma$ are the effective oscillator frequency and friction coefficient, respectively (see text).
fluctuation of the iron relative to the heme is \(7 \times 10^{-2} \text{Å}^2\) at 200 K, on the order of the estimate from the phenomenological model (16).

The Mössbauer and x-ray estimates of (\(\Delta x^2\)) at 300 K are 0.065 and 0.11 Å², respectively (13). A 120-ps simulation of carboxymyoglobin at 300 K (K. Kucecza, J. Kuriyan, and M.K., unpublished results) yields (\(\Delta x^2\)) = 0.057 Å² with the dominant component of the motion in the heme plane. Additional motions on the Mössbauer time scale between 1 and 100 ns clearly cannot be sampled by such a subnanosecond simulation and must involve crossing of higher barriers (see Eq. 2). The slower motions have been suggested to contribute about 40% of the Mössbauer value of (\(\Delta x^2\)) (13). This would mean that the simulation result is too large by 30%. Such an error in a vacuum simulation would not be surprising, and there is no reason to propose, as done in an analysis (13) of the Mössbauer data, that simulations yield motions that are 100 times too fast. An additional point is that the experimental estimates of (\(\Delta x^2\)) from Mössbauer or x-ray scattering may be in error. Only in the Gaussian approximation, which corresponds to harmonic motion, can the structure factor be written as \(f = \exp(-k^2(\Delta x^2))\). Thus, the observed "values" (13) of (\(\Delta x^2\)) refer to \(-\ln f/k^2\), where \(k\) is the momentum of the \(\gamma\)-quantum in Mössbauer spectrometry and the scattering vector in x-ray scattering. Molecular dynamics simulations have demonstrated that the local effective potentials for the atomic motions tend to be anistropic and anharmonic (1). Consequently, the distributions monitored by Mössbauer spectroscopy may be non-Gaussian, so that the structure factor itself must be used in a comparison of theory and experiment (25, 26). Also, it has been demonstrated by simulations that the magnitudes of the fluctuations are underestimated by x-ray refinement (26).

The present analysis demonstrates that molecular dynamics simulations and their stochastic extensions, including barrier crossing, provide a consistent picture of protein dynamics over a wide range of time scales. Additional experimental data, as well as further theoretical analyses, are needed, however, for a full understanding of the internal motions of proteins.

The authors thank C. L. Brooks, III, P. Tavan, and H. Treutlein for helpful discussions and assistance with some of the computations. This work was supported in part by the National Institutes of Health.