

Investigating vibrational energy relaxation and collective motions in proteins

Hiroshi FUJISAKI* and John E. STRAUB
(Chemistry Department, Boston University)

Abstract

Our recent studies on vibrational energy relaxation (VER) and collective motions in proteins are summarized. As a first topic, we discuss the validity of Fermi's golden rule for VER problems in proteins, and to get over this limitation, we apply the finite-time perturbation theory due to Okazaki's group to our VER problem of a localized mode in a protein. As a second topic, we discuss the dynamical effects of solvent water on the collective protein dynamics, employing several time series analysis methods such as principal component analysis and the power spectra of kinetic energies.

1 Vibrational energy relaxation in proteins

Vibrational energy relaxation (VER) and dephasing in large molecules including proteins are important and interesting issues in chemical physics because they characterize both time-dependent and time-independent spectroscopic properties. Even more significant is the challenge to relate VER and dephasing to fundamental reaction processes, such as conformational changes or electron-transfer processes in a protein, associated with protein functions.

Recently we have studied VER of cytochrome c (cyt c) using a reduced model approach [1]. Our Hamiltonian is decomposed as

$$\mathcal{H} = \mathcal{H}_S + \mathcal{H}_B + \mathcal{V}_3 + \mathcal{V}_4 + \cdots, \quad (1)$$

$$\mathcal{H}_S = \frac{p_S^2}{2} + \frac{\omega_S^2}{2} q_S^2, \quad (2)$$

$$\mathcal{H}_B = \sum_k \frac{p_k^2}{2} + \frac{\omega_k^2}{2} q_k^2, \quad (3)$$

with

$$\mathcal{V}_3 = \frac{1}{3} \sum_{k,l,m} C_{klm} q_k q_l q_m, \quad (4)$$

$$\mathcal{V}_4 = \frac{1}{4} \sum_{k,l,m,n} C_{klmn} q_k q_l q_m q_n. \quad (5)$$

*e-mail: fujisaki@bu.edu, <http://members.aol.com/Hiro2fuji/homepage/fujisaki.html>

where \mathcal{H}_S is the system Hamiltonian, \mathcal{H}_B is the bath Hamiltonian, both of which are modeled as harmonic oscillators, and the system-bath coupling is represented by \mathcal{V}_3 and \mathcal{V}_4 etc. By assuming the Markov property and using Fermi's golden rule, we have derived the following formula for a VER rate (we have only included the contribution from \mathcal{V}_3)

$$\frac{1}{T_1} \simeq \tanh\left(\frac{\beta\hbar\omega_S}{2}\right) \frac{\hbar\gamma}{2\omega_S} \sum_{k,l} \frac{(C_{Skl})^2}{\omega_k\omega_l} \left[\frac{1+n_k+n_l+2n_kn_l}{\gamma^2+(\omega_k+\omega_l-\omega_S)^2} + \frac{1+n_k+n_l+2n_kn_l}{\gamma^2+(\omega_k+\omega_l+\omega_S)^2} \right. \\ \left. + \frac{n_k+n_l+2n_kn_l}{\gamma^2+(\omega_k-\omega_l-\omega_S)^2} + \frac{n_k+n_l+2n_kn_l}{\gamma^2+(\omega_k-\omega_l+\omega_S)^2} \right] \quad (6)$$

where $n_k = 1/(e^{\beta\hbar\omega_k} - 1)$ is the thermal phonon number. This is actually equivalent to the Maradudin-Fein formula used by Leitner [2].

Using this formula, we found that the VER time scale T_1 is sub ps for the CD mode in cyt c, a localized mode of a methyl group of methionine 80 in the protein. The mechanism of the VER was found to be due to the Fermi resonance among three modes strongly coupled to one another, and localized in the methyl group [1]. The result is consistent with the absorption widths ($\sim 10 \text{ cm}^{-1}$) by Romesberg's group if the width is mainly determined by T_1 .

However, there are deficiencies in this method. We need to include an empirical parameter γ , and the result can depend on γ significantly. Furthermore, this is based on the Markov property assumed in Fermi's golden rule. As shown before [3], this might not be the case for high frequency modes in large molecules like proteins. Hence, we are now pursuing more advanced methods to address this issue. One such method has been presented by Okazaki's group [4]. Using the path integral method and the perturbation theory, they derived an appealing formula for a density matrix, which contains the effects of both VER and dephasing. Here we use the following expression for the diagonal density matrix:

$$\rho_{11}(t) \simeq \exp\{-W_{0\leftarrow 1}(t) - W_{2\leftarrow 1}(t)\} \quad (7)$$

where

$$W_{0\leftarrow 1}(t) = \frac{\hbar}{2\omega_S} \sum_{k,l} \frac{(C_{Skl})^2}{\omega_k\omega_l} [(1+n_k)(1+n_l)f_+(t) + n_kn_lg_+(t) \\ + n_l(1+n_k)f_-(t) + n_k(1+n_l)g_-(t)] \quad (8)$$

$$W_{2\leftarrow 1}(t) = \frac{\hbar}{\omega_S} \sum_{k,l} \frac{(C_{Skl})^2}{\omega_k\omega_l} [(1+n_k)(1+n_l)g_+(t) + n_kn_lf_+(t) \\ + n_l(1+n_k)g_-(t) + n_k(1+n_l)f_-(t)] \quad (9)$$

with

$$f_{\pm}(t) = \frac{1 - \cos(\omega_S - \omega_k \mp \omega_l)t}{(\omega_S - \omega_k \mp \omega_l)^2}, \quad g_{\pm}(t) = \frac{1 - \cos(\omega_S + \omega_k \pm \omega_l)t}{(\omega_S + \omega_k \pm \omega_l)^2}. \quad (10)$$

Since it is shown that $\left. \frac{dW_{1\leftarrow 0}(t)}{dt} \right|_{t \rightarrow \infty} = \frac{1}{1 - e^{-\beta\hbar\omega_S}} \frac{1}{T_1} \Big|_{\gamma \rightarrow 0}$, Eq. (7) can be regarded as a generalization of Eq. (6) with important improvements. (1) We do not need to introduce

the annoying factor γ , (2) we can use Eq. (7) for much shorter time scales where the Markov property does not hold, and (3) we can study time-dependent properties of the system. The restrictions are (1) we have used the lowest order perturbation theory, and (2) we have treated the system-bath coupling perturbatively.

The concrete procedure is as follows: 100 ps simulations for (a) methionine in water, and (b) cyt c in water are carried out, and 100 sample trajectories from such runs are collected. All the atoms except the active region around methionine for (a) and around methionine 80 for (b) are deleted. Instantaneous normal mode analysis is carried out to calculate the anharmonic coefficients C_{Slm} for the CD mode [1]. The resulting data are plugged into Eq. (7). Finally an ensemble average is taken for the density matrix.

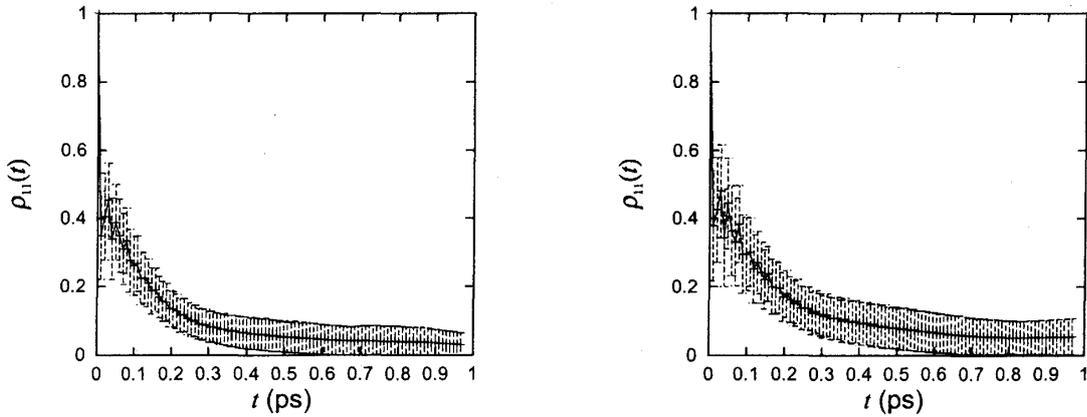


Figure 1: Time evolution of the diagonal density matrix for methionine in water (left) and in cyt c (right) at 300K.

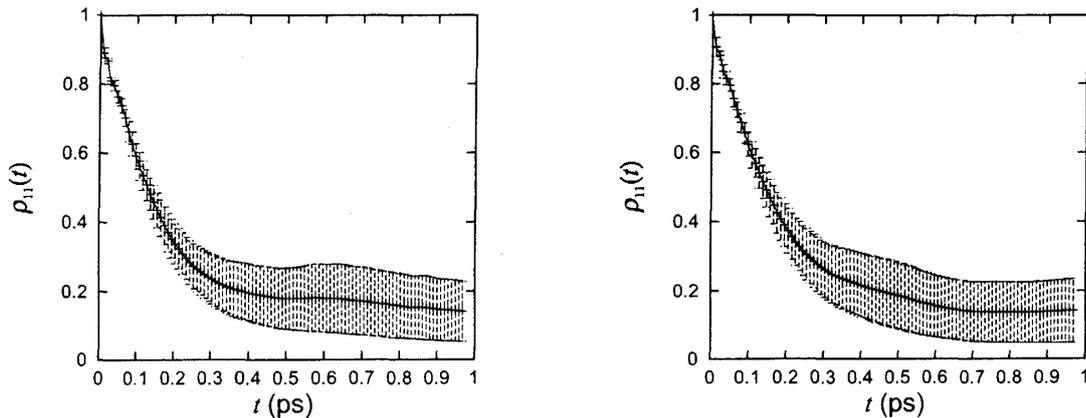


Figure 2: Time evolution of the diagonal density matrix for methionine in water (left) and in cyt c (right) at 0K.

The results are shown in Figs. 1 (at 300K) and 2 (at 0K), which indicate that the VER time scale is sub ps as shown before [1]. However, for the 300K case (Fig. 1), it

is difficult to fit the time dependence with a single exponential. Actually there is strong sample dependence for this case as indicated by the error bars. In some cases, the density matrix becomes zero within 20 fs, due to the significant contribution from low frequencies with the large thermal phonon number n_k .

For the 0K case, we can nicely fit the data with a single exponential (up to 0.2 ps) and the VER time turns out to be 0.2 ps for both (a) and (b). (In any case, we should not trust the time dependence after the VER time, because this is a just perturbative treatment.) In this case, the contribution from low frequencies becomes minimum, because there is no contribution from the thermal phonon number n_k .

Thus we observe no difference between (a) methionine in water and (b) cyt c in water in terms of VER of the CD mode. Recent experimental data show that there could be a difference of a factor of four for the absorption width [5], which also depends on the deuterium position. To solve these problems, we need to take into account the dipole functions (oscillator strength), and the effects of dephasing with QM/MM-MD type methods.

2 Collective motions in proteins

Many experimental studies have indicated that functionality of proteins is often related to collective motions in them, which accompany structural changes. Such large amplitude and low frequency motions can be well described by normal mode analysis or principal component analysis (PCA) [6]. Since proteins usually function when they are in solvents, the collective motions are naturally coupled to solvents, leading to the importance of studying solvent effects for collective motions in proteins. The first approximation of the solvent effects might be friction effects: Such a picture actually holds for principal components (PCs) with the use of Langevin analysis [7], which implies that PCs normally diffuse in configuration space. However, recent studies clarified that the diffusion of PCs can be abnormal (subdiffusive) [8]. Those results stimulated us to investigate more detailed aspects of solvent effects in proteins.

One important observation is that there exists a correlation between hydrated structure of water and the collective motions in proteins [9, 10]. In addition, the dynamics of hydrated water around a protein surface have been studied by many researchers [11], and a correlation between hydrated water and a protein was also found by Bizzarri and Canistraro in terms of $1/f$ fluctuations. However, the dynamical coupling between collective motions in proteins and solvents has not been clearly understood.

To attack this issue, we have conducted numerical simulations using cyt c in water. We used a standard and simple procedure to simulate the system for 1ns.¹ Figure 3 shows the result for the variance $\langle(\Delta x_i)^2\rangle$ of each α carbon (there are 104 α carbons), where the bracket indicates a long time average up to 1 ns. From the left figure, we can see that there is a transition at around 200K, and the fluctuation becomes large above such a temperature. The most fluctuating part is the loop region of cyt c (around 20th ~ 30th residues).

We also calculated the variance-covariance matrix $\langle\Delta x_i\Delta x_j\rangle$; we can extract principal components by diagonalization of this matrix. The projection of the most fluctuating PC

¹The Verlet algorithm with SHAKE and the periodic boundary condition with an octahedral box was used, the time step was 1.0 fs, and 13 Å cut off was applied. We used the CHARMM software.

is shown in the right of Fig. 3. We can confirm that the most fluctuating PC represents the fluctuations in the system quite well.

We also calculated the diffusive motion of the most fluctuating PC, $PC_1(t)$, using $D(t) = \langle [PC_1(t) - PC_1(0)]^2 \rangle$ in the left of Fig. 4. There are three time regions: (1) normal diffusive region ($t < \sim 1$ ps), (2) sub diffusive region ($1 \text{ ps} < t < 10$ ps), and (3) super diffusive region ($t > \sim 10$ ps). It is very likely that the super diffusive behavior is an artifact of PCA, and we might expect normal diffusion again for such longer time scales. We thus only consider the sub diffusive behavior at the intermediate time scale.

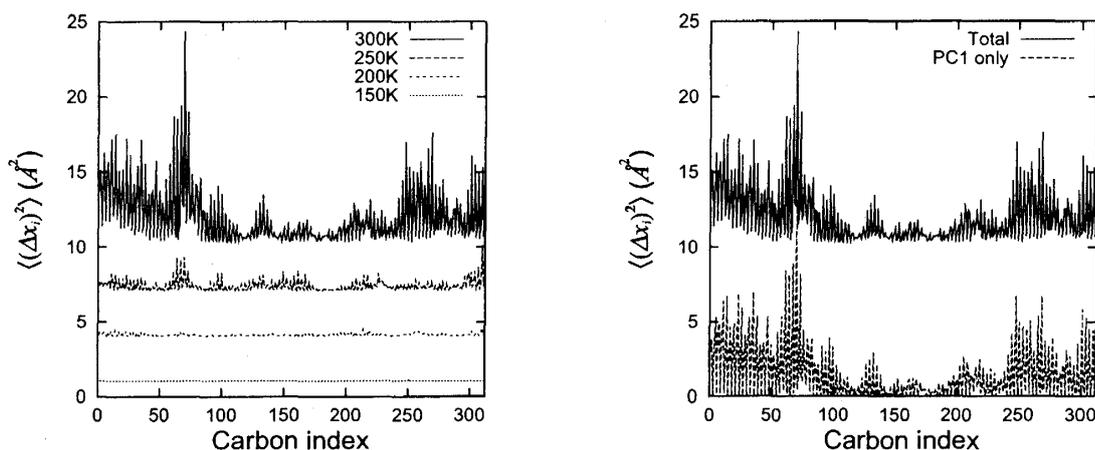


Figure 3: Left: Variance of α carbons at four different temperatures. Right: Comparison between the variance of α carbons at 300K and that of the most fluctuating principal component at 300K.

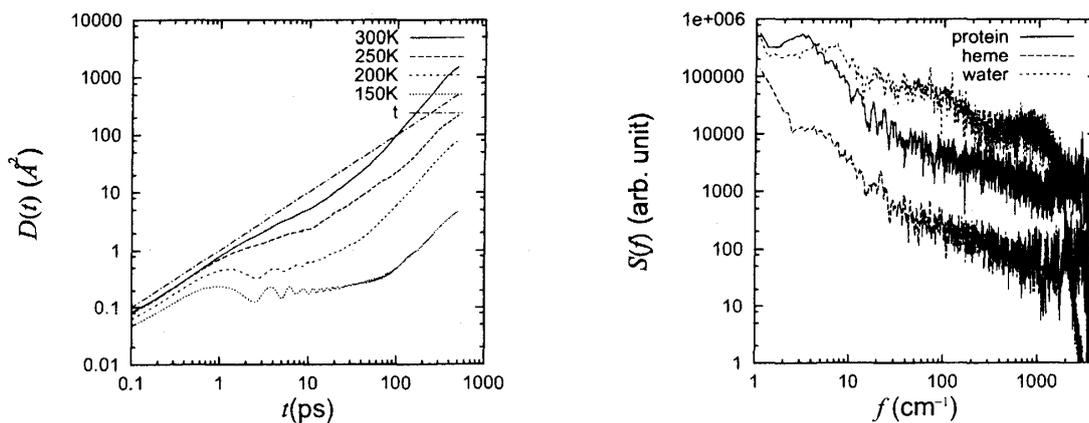


Figure 4: Left: Diffusive motion of the most fluctuating PC at four different temperatures. Right: Power spectra for the kinetic energies of the protein, water, and heme at 300K.

We expect that this sub diffusive behavior is transient as in the case of a tagged particle in glassy material [12]. This is supported by the power spectra for the protein dynamics

as shown in the right of Fig. 4. The protein exhibits $1/f$ like fluctuations as prominently as water does, and there might be a crossover to white noise at a low frequency. In the case of water, such a critical frequency is $\sim 1.0 \text{ cm}^{-1}$ [13], which means that the memory is lost after $\sim 30 \text{ ps}$. If this is also true for the protein, there should be some memory effects in the protein at least up to 30 ps , and there may be a relation between the sub diffusive behavior and the $1/f$ fluctuations.

To justify (or refute) this hypothesis, we need to study the details of the protein and water fluctuations at the molecular level. Some molecular insights have been given by Kitao *et al.* [9] and Koyama and Takano [14] for the coupling between a protein and water. Bagchi *et al.* [15] also gave a dynamical picture, from which one can derive the sub diffusion of the hydrated water as Brownian motion in a double well potential. The similar strategy should be pursued to understand the sub diffusive behavior of cyt c, especially around the most fluctuating loop region, and we must clarify whether we need to include the memory effects as mentioned above or if it is enough to consider Brownian motion in a general potential.

References

- [1] H. Fujisaki, L. Bu, and J.E. Straub, *Adv. Chem. Phys.* **130B**, 179-203 (2005); H. Fujisaki and J.E. Straub, *Proc. Natl. Acad. Sci. USA* **102**, 6726-6731 (2005).
- [2] D. Leitner, *Adv. Chem. Phys.* **130B**, 205-256 (2005).
- [3] H. Fujisaki, L. Bu, and J.E. Straub, in *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*, edited by Q. Cui and I. Bahar; e-print q-bio.BM/0408023.
- [4] T. Mikami and S. Okazaki, *J. Chem. Phys.* **121** (2004) 10052-10064; T. Mikami, M. Shiga, and S. Okazaki, *ibid.* **115** (2001) 9797-9807.
- [5] Matt Cremeens, private communication.
- [6] A. Kitao and N. Go, *Curr. Opin. Struc. Biol.* **9**, 164-169 (1999).
- [7] S. Hayward, A. Kitao, F. Hirata, and N. Go, *J. Mol. Biol.* **234**, 1207-1217 (1993).
- [8] A.E. Garcia and G. Hummer, *Proteins* **36**, 175-191 (1999); P. Carlini, A.R. Bizzarri, and S. Cannistraro, *Physica D* **165**, 242-250 (2002).
- [9] A. Kitao, F. Hirata, and N. Go, *J. Phys. Chem.* **97**, 10223-10230 (1993).
- [10] M. Tarek and D.J. Tobias, *Phys. Rev. Lett.* **88**, 138101-1-4 (2002).
- [11] A.R. Bizzarri and S. Cannistraro, *J. Phys. Chem. B* **106**, 6617-6633 (2002).
- [12] J. Habasaki, K.L. Ngai, and Y. Hiwatari, *J. Chem. Phys.* **122**, 054507-1-10 (2005).
- [13] I. Ohmine, *J. Phys. Chem.* **99**, 6767-6776 (1995).
- [14] Y. Koyama and M. Takano, private communication.
- [15] S.K. Pal, J. Peon, B. Bagchi, and A.H. Zewail, *J. Phys. Chem. B* **106** (2002) 12376-12395.