A Computational Method to Identify Residues Important in Creating a Protein Promoting Vibration in Enzymes

Joshua S. Mincer‡ and Steven D. Schwartz*,†,‡

Department of Physiology and Biophysics and Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received: September 17, 2002; In Final Form: October 30, 2002

In this paper, we present a computational method to screen a large set of protein residues to identify those residues the motions of which help create a protein promoting vibration and are therefore important for catalysis. The method is illustrated for the case of horse liver alcohol dehydrogenase (HLADH). In this system, the protein promoting vibration is the relative motion between hydride donor and acceptor, that is, the benzyl alcohol substrate and the nicotinamide adenine dinucleotide (NAD^+) cofactor, respectively. The resulting subset of screened residues compares favorably with existing experimental data and also suggests additional residues as objects for potential study. The method presented in this paper employs correlated motion as the basis for identifying residues important in catalysis. As such, the success of the method in the case of HLADH further supports the importance of protein dynamics in certain enzyme systems.

Introduction

The chemical dynamics that allow an enzyme to accelerate the rate of a reaction are still a topic of debate. Work in our group over the past few years has focused on investigating the possibility that in some enzymes protein dynamics plays a central role in the catalytic event.[1]

In some enzymes, we have shown that protein promoting vibrations, internal protein motions, are important to catalysis by modulating both the height and width of the barrier.[2−4] Such concepts have been used to explain anomalous experimental results.[2−6,8] A dynamic view of catalysis can help explain the paradoxical result found in the thermophilic alcohol dehydrogenase (ADH) from *Bacillus stearothermophilus*. Here, tunneling decreases with temperature. This has been rationalized using a dynamic view of catalysis as follows: at lower temperatures, the protein is less flexible and therefore the promoting vibration couples less strongly to the reaction.[8]

It is important to note that different views of catalysis are not mutually exclusive. In different types of enzymes, different mechanisms of catalysis may operate to a greater or lesser degree. One type of enzyme in which protein promoting vibrations may be employed to cause catalysis is found in the alcohol dehydrogenase family. In this enzyme type, a hydride is transferred from a donor to an acceptor; in horse liver alcohol dehydrogenase (HLADH), the hydride donor is the substrate alcohol and the acceptor is the NAD^+ cofactor.[17] Experimental evidence has pointed to quantum mechanical tunneling as the mode of this transfer in HLADH.[5,6] Because the probability of tunneling is dependent on the donor−acceptor distance, modulation of this distance would strongly impact the rate of catalysis. Other views of catalysis involving the creation of reactive species have maintained that the enzyme binds hydride donor and acceptor in such a way as to force them closer together to restrict their separation.[7] The theory of protein promoting vibrations, by contrast, maintains that the dynamics of the protein sets up an oscillation between donor and acceptor, termed the promoting vibration.

In previous papers published in our group, a computational method to identify the existence of a protein promoting vibration was developed[9] and successfully applied to HLADH.[10] This methodology enables one to determine whether a protein promoting vibration exists in any enzyme system through molecular dynamics modeling. In the present work, we have developed a computational method to identify the specific motions of the protein involved in creating the promoting vibration. This algorithm will identify specific residues the motions of which impact donor and acceptor in such a way as to create the promoting vibration. In this paper, we apply the method to HLADH because extensive experimental evidence allows confirmation of the validity of the algorithm.

Methods

Theoretical Background. For a residue to be involved in driving the protein promoting vibration (PPV), its motion must be strongly correlated with the PPV. The center of mass velocity of the residue is chosen to represent the motion of the residue. The PPV is, in the case of alcohol dehydrogenase, the relative motion between donor and acceptor. For a protein residue to help modulate the distance between donor and acceptor, this motion must have a component along the donor−acceptor axis. Therefore, we investigate this motion projected onto the unit vector connecting donor and acceptor.

More specifically, we are interested in two functions of time (Figure 1). The first, $A(t)$, is the relative velocity between donor and acceptor, v_{DA}, projected onto the unit vector between them, ur_{DA}. The second, $B(t)$, involves two vector projections. First, the residue center of mass velocity, v_R, is projected onto the unit vector connecting the residue center of mass and the donor or acceptor, ur_{RD} or ur_{RA} (for a given residue, the choice of donor or acceptor depends on whether the residue lies behind donor or acceptor, as discussed below). This vector is then projected onto the unit vector connecting donor and acceptor, ur_{DA}.

* Corresponding author E-mail: sschwartz@aecom.yu.edu.
‡ Department of Physiology and Biophysics.
† Department of Biochemistry.

10.1021/jp027017j CCC: $25.00
Published on Web 12/07/2002

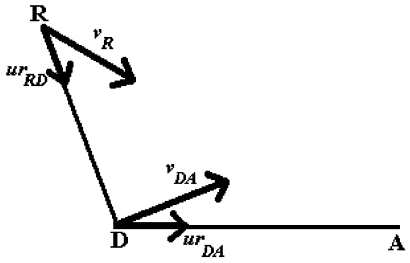

Figure 1. Motions of interest: R = residue center of mass; D = donor; A = acceptor. Vectors are defined in the text.

$A(t)$ represents the PPV. The significance of $B(t)$ is that it addresses two criteria that a residue must satisfy to possibly drive the PPV motion. The first vector projection determines to what extent the residue could drive donor or acceptor motion. The second projection determines whether this driving is in the right direction to create the PPV. The following definitions are written for $A(t)$ and $B(t)$:

$$A(t) = \vec{v}_{DA} \cdot \vec{ur}_{DA}$$

$$B(t) = (\vec{v}_R \cdot \vec{ur}_{RD})(\vec{ur}_{RD} \cdot \vec{ur}_{DA})$$

Note that $B(t)$ is defined for the case in which the residue lies behind the donor. Were it behind the acceptor, ur_{RA} would be used in place of ur_{RD}.

One might think to define $B(t)$ more simply as the residue center of mass velocity projected onto the unit vector connecting donor and acceptor:

$$B(t) = \vec{v}_R \cdot \vec{ur}_{DA}$$

That this is not sufficient is seen in the following case. Imagine a residue that lies directly above the donor. That is, the vector connecting the residue center of mass and the donor is perpendicular to the vector connecting donor and acceptor. Imagine that the center of mass velocity has a component parallel to the donor−acceptor axis. In this case, the simpler definition of $B(t)$ would yield a strong correlation between $B(t)$ and $A(t)$, and one would conclude that the residue could be driving the PPV. But this would be false. Because the residue lies directly above the donor, it could not possibly drive the donor toward the acceptor and so could not drive the PPV. The first projection in $B(t)$ rectifies this; in this hypothetical case, this first projection would be zero.

To determine whether $A(t)$ and $B(t)$ are correlated, we employ a time-correlation function and its associated spectral decomposition.[11] The time-correlation function is defined as

$$C_{AB}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} A(t+\tau)B(t)\, dt$$

Note that $C_{AB}(\tau)$ is a time average; it is equivalent, assuming the system is ergodic, to the ensemble average:[11]

$$C_{AB}(\tau) = \langle A(t+\tau)B(t) \rangle$$

Both $C_{AB}(\tau)$ and $C_{BA}(\tau)$ are calculated and averaged to give $C(\tau)$. A PPV, confirmed through the study of the dynamics of the quantity $A(t)$, was found in previous studies to be a function with one dominant frequency.[10] That is, the donor−acceptor oscillation occurs at a dominant frequency, but because the motion is not purely harmonic, in actuality the motion is a superposition of other frequencies as well. If a given residue drives the PPV, then its motion should be similar. $B(t)$ should also have a similar dominant frequency oscillation.

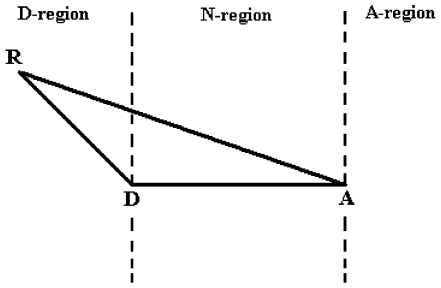

Figure 2. Angles RDA and RAD, measured to determine where the residue is in space in relation to donor and acceptor: R = residue center of mass; D = donor; A = acceptor.

To analyze $C(\tau)$, we calculate its Fourier transform, the spectral decomposition,[11] $G(\omega)$:

$$G(\omega) = \int_{-\infty}^{\infty} C(\tau)\, e^{-i\omega\tau}\, d\tau$$

The strength of the correlation at each frequency is given by the amplitude of the peak. If a residue is involved in driving the PPV, then we expect to see a strong peak in the spectral decomposition at the dominant frequency of the PPV.

Molecular Dynamics. The structure used for MD simulation is that of wild-type HLADH with benzyl alcohol substrate. Its preparation is described in ref 10 (the relevant structure is the reactants configuration). Minimization and dynamics were carried out using CHARMM[12] (MSI, San Diego, CA). The structure was minimized first for 1000 steps using steepest descent with a force criterion of 0.001 kcal per 10 steps, followed by 8000 steps using adopted-basis Newton Raphson with tolerance 1×10^{-9} per 10 steps. The nonbonding interactions cutoff was set at 14 Å. The dynamics protocol employed Verlet integration with time step 1 fs. The structure was heated for 2 ps with the temperature being raised 3 K every 20 steps to a final 300 K. Heating was followed by equilibration for 8 ps, followed by an observation period of 5 ps. Data collection took place over 30 ps; coordinates and velocities were recorded every five steps for a total of 6000 structures. (Note: we have performed 100 ps runs as well with no marked change in results.) All bonds involving hydrogen atoms were constrained with SHAKE[13] during dynamics. The stored velocities and coordinates were used to calculate the relevant time-correlation functions for each residue of interest.

The Algorithm. We seek to screen a large set of residues to determine which residues have motion that is correlated with the PPV. We employed the following algorithm: (1) From the crystal structure of the enzyme, identify the set of residues that lie at least partially within 10 Å of either donor or acceptor carbon. This forms our test set. Ten angstroms is chosen as an arbitrary cutoff. (For HLADH, the donor carbon is C7 of benzyl alcohol; the acceptor carbon is C4N of the nicotinamide ring of NAD^+). (2) Place each residue into one of three sets, depending on whether it lies behind donor or acceptor or between them. To do this, one measures two angles (Figure 2). If angle RDA is greater than 90°, then the residue lies behind the donor; the same goes for RAD and the acceptor side. If both angles are less than or equal to 90°, then the residue lies between donor and acceptor. The donor-side set will be termed D-region, the acceptor side A-region, and the intermediate region N-region (N for neither). (3) Using the trajectory from the MD simulation, calculate the appropriate time-correlation functions. For D-region, the unit vector ur_{RD} (the unit vector connecting

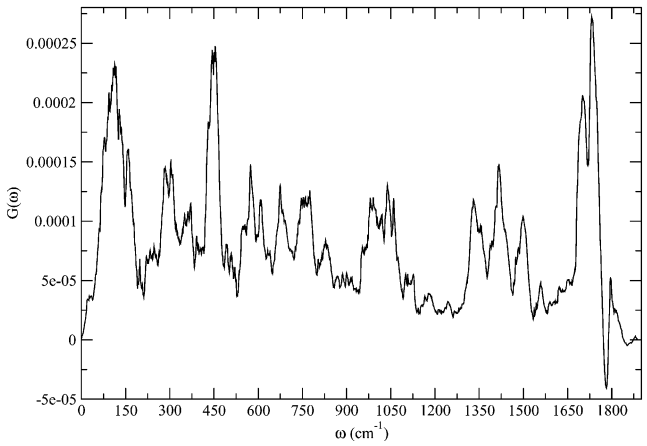

Figure 3. The spectral decomposition, $G(\omega)$, for the relative motion between donor and acceptor carbons. The spectrum is reported in CHARMM units.

the residue center of mass and the donor carbon) is used to calculate $B(t)$. For A-region, ur_{RA} is used. For those residues in N-region, calculate both functions. Those in D-region could conceivably contribute to creating the PPV by driving the donor motion; those in A-region could could drive the acceptor motion. Those in N-region could drive either but only to disrupt the PPV (by driving donor and acceptor not toward each other but away from each other). (4) Generate the spectral decompositions for each time-correlation function. The spectrum is the tool actually used to screen residues. For each residue, record the height of the peak (if it exists) that falls in the range of the dominant PPV frequency. This is the first part of the screen; only residues with correlated motion of the correct frequency (the PPV frequency) are selected. Note that the PPV itself is not purely harmonic and thus correlations may exist at other frequencies. We employ the dominant frequency of the PPV as the diagnostic in the spectral decomposition $G(\omega)$. (5) Generate a histogram for the amplitudes of peaks for residues in the D- and A-regions. Select residues with the highest amplitudes. This represents the second part of the screen, that is, selecting strongly correlated residues.

Results and Discussion

Sixty-seven residues were found to lie at least partially within 10 Å of either donor (benzyl alcohol C7) or acceptor carbon (C4N of the nicotinamide ring). Of these, 25 lie in D-region, 31 in A-region, and 11 intermediate to donor and acceptor (N-region).

Figure 3 is the spectral decomposition of the autocorrelation function of the donor−acceptor relative velocity projected onto the unit vector connecting them. Thus, this spectrum corresponds to the autocorrelation function of the function $A(t)$ described above. This function is diagnostic of a PPV[10] with dominant frequency peak around 110 cm^{-1}. A closer view shows its breadth to be from about 80 to 150 cm^{-1}. Therefore, in recording the peak amplitude for each residue's spectral density, a peak in this range was chosen. For most residues, this was the dominant peak in the spectrum.

Figure 4 is an amplitude histogram for D- and A-region residues. Amplitudes range from 0.5 units to around 8.9 units (as seen in figures of the spectra, the actual values are number of units times 10^{-5}; for brevity this factor has been omitted). Only eight residues have peak amplitude in the upper third of this range, and these eight residues are therefore selected as having motion strongly correlated with the PPV. Spectral decompositions for three of these eight are shown in Figure 5 (the others may be found in Figures 1S and 2S of the Supporting Information). Figure 6 illustrates these selected residues in relation to NAD^+ and the benzyl alcohol substrate. Note that of these residues, only one, Ser144, lies on the donor side. The rest lie on the acceptor side. Note that out of all residues in the enzyme the screen has selected only eight. These eight residues and their peak amplitudes are listed in Table 1.

The results of the screen compare favorably with existing experimental data. Klinman and co-workers have measured indicators of tunneling, as well as rate of catalysis, in a number of mutant forms of HLADH. These experiments serve to identify residues important for catalysis. Namely, if mutation affects tunneling and rate of catalysis, then the mutated residue is thought to be important in the enzyme's catalytic function.

On the basis of their experiments, Klinman and co-workers have concluded that residue Val203 is very important for

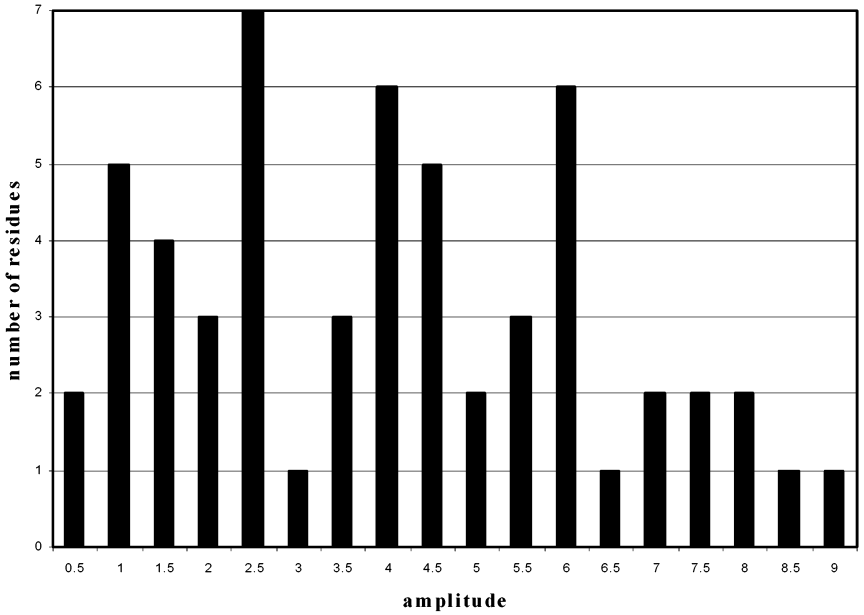

Figure 4. Histogram of peak amplitudes for the residues that lie in either D- or A-regions.

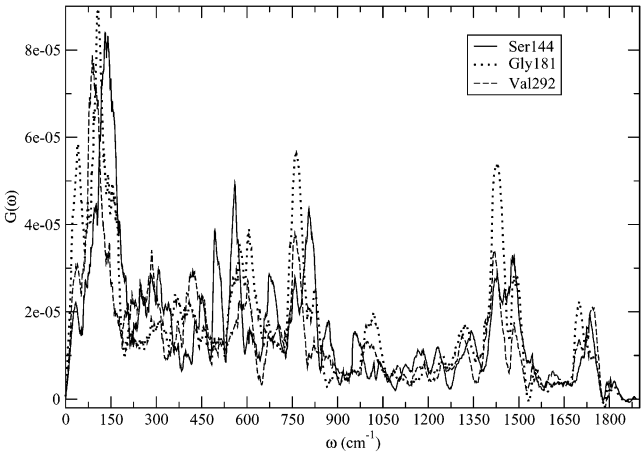

Figure 5. Spectral decompositions, $G(\omega)$, for three of the eight selected residues: Ser144, Gly181, and Val292. The spectra are reported in CHARMM units.

catalysis,[6] while Leu57,[5] Phe93,[5] and Val294[14] are not. Figure 7 and Figure 3S of the Supporting information show the spectral decompositions for these residues obtained in our analysis. Val203 has significant amplitude and is one of the eight selected residues. The others are seen to have little amplitude. Note that because Val294 lies in N-region, both donor-side and acceptor-side spectra are shown in Figure 3S. We conclude that the motion of Val203 is strongly correlated with the PPV, while the motions of the other three residues are not. As such, by driving the PPV, Val203 could be very important to catalysis, while the others are not. Our results are at variance with those of the Klinman group with respect to one residue, Ile269. We have selected Ile269 as one of the eight residues important in creating the PPV, while the Klinman group has concluded that mutation at that position does not signifiantly affect tunneling or rate.[14] It is possible that while motion of Ile269 is strongly correlated with the PPV, it is in fact not central to creating the

TABLE 1: The Eight Selected Residues, Their Peak Amplitudes as Measured in the Spectra, and the Region in Which Each Is Found

residue	amplitude ($\times 10^{-5}$ units)	region[a]
Ser144	8.4	D
Gly181	8.9	A
Val292	7.9	A
Glu267	7.6	A
Ile269	7.1	A
Val203	6.6	A
Gly204	7.3	A
Val207	6.6	A

[a] D = D-region; A = A-region.

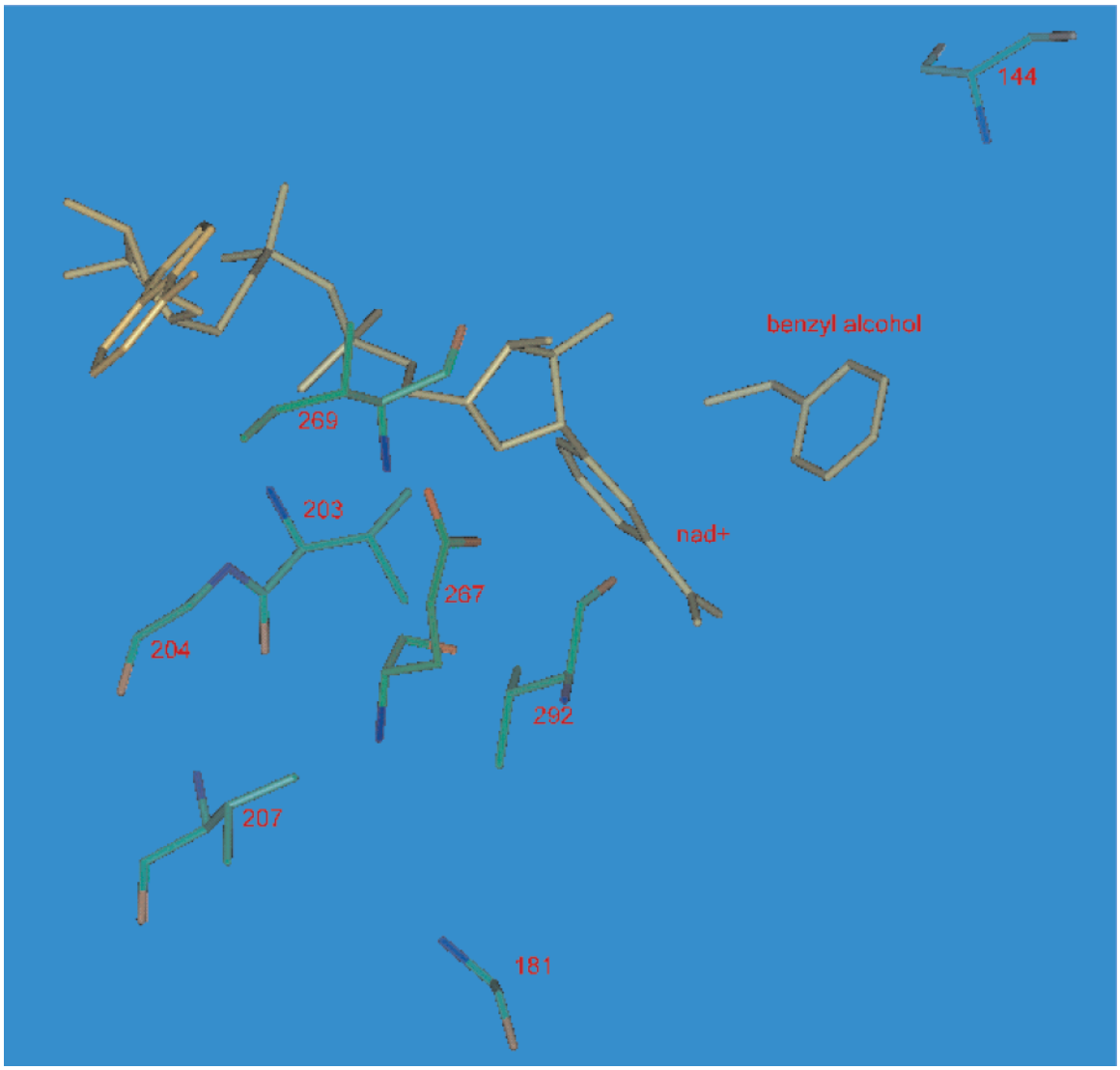

Figure 6. The eight selected residues in relation to the benzyl alcohol substrate and the NAD^+ cofactor.

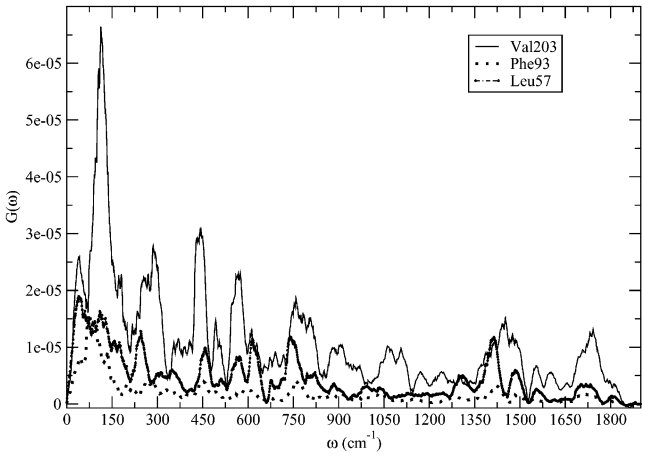

Figure 7. The spectral decompositions, $G(\omega)$, for residues for which experimental data exists: Val203, Phe93, and Leu57. The spectra are reported in CHARMM units.

PPV and therefore to catalysis. A look at Figure 6 shows that Ile269 lies close to the body of NAD^+ and not to the nicotinamide ring. It may be that while Ile269 contacts NAD^+, it is contacting the more massive body of the cofactor (as opposed to the nicotinamide ring), resulting in little displacement of the acceptor carbon.

None of the 11 residues that lie between donor and acceptor (N-region) have peaks with significant amplitude. All amplitudes are less than 4 units. We noted above the possibility that residues in N-region could be important to catalysis by disrupting the enzyme's functions. By driving donor or acceptor in the wrong direction (away from one another), these residues could disrupt the PPV. In the present case of HLADH, it seems unlikely that residues in N-region have much of an effect on the PPV.

Conclusions

We have presented a computational method to identify protein residues important in driving the protein promoting vibration and therefore important in catalysis. This work adds another computational tool to enable theoretical characterization of enzyme systems. Previously, Caratzoulas and Schwartz developed a method to determine whether a given enzyme utilized the mode of catalysis that we refer to as protein promoting vibrations.[9] With the method outlined in this paper, one may further characterize the dynamics of the enzyme by employing computer modeling. Thus, having established that a protein promoting vibration is present, we may now determine which protein residues are important in creating the promoting vibration.

We have applied our new algorithm to HLADH.[10] Out of the initial set of all protein residues, this method has isolated a small subset of eight residues. Comparison with the experimental studies of Klinman and co-workers demonstrates good agreement where data are available.[5,6,14]

The described method isolates residues on the basis of correlated motion, the correlation being defined in a physically motivated manner: which residues have motion that could drive the PPV. In this regard, the success of the method in HLADH further supports the theory of protein promoting vibrations. We have not isolated residues on the basis of, for example, geometric considerations of where they lie relative to donor and acceptor. Rather, we have asked the question which residues have motion that is in the right direction and has the right frequency to drive the protein promoting vibration, the oscillation between donor and acceptor.

The body of experimental evidence supporting the role of quantum mechanical tunneling in enzyme catalysis is growing.[15,18] With the method presented in this paper, one may now more fully study an enzyme system theoretically. Such study can help clarify the role of tunneling in enzymes for which there is currently experimental data. Furthermore, one can use the methods developed in this and previous papers to study systems as yet uncharacterized. If appropriately accurate potential surfaces are available, before a single experiment is done, one may determine the likelihood that tunneling is operating in a given enzyme system. Furthermore, one may determine which residues are important in effecting catalysis. This determination can guide mutational studies such as those carried out by Klinman and co-workers.[5,6,14]

The method presented here may also have implications for the design of de novo proteins. At present, while it is possible to define regions of sequence space that yield stable, water-soluble, α-helical proteins, designing proteins with high levels of enzymatic activity remains an elusive goal.[16] Study of enzyme systems in which protein promoting vibrations are operative may lead to identification of amino acid sequence motifs that are essential in general for creating the promoting vibration.

Acknowledgment. We gratefully acknowledge the support of the Office of Naval Research and the National Science Foundation through Grants CHE-9972864 and CHE-0139752. J.S.M acknowledges fellowship support from the Medical Scientist Training Program (MSTP). Finally, J.S.M. is happy to acknowledge helpful conversations with Dr. Dimitri Antoniou, Dr. C. K. Kalyanaraman, and Jodi Basner.

Supporting Information Available: Spectral decompositions, $G(\omega)$, for five of the eight selected residues (Figures 1S and 2S) and spectra for residue Val294, another residue for which there exist experimental data (Figure 3S). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Antoniou, D.; Caratzoulas, S.; Kalyanaraman, C.; Mincer, J. S.; Schwartz, S. D. *Eur. J. Biochem.* **2002**, *269*, 3103−3112.

(2) Antoniou, D.; Schwartz, S. D. *J. Chem. Phys*. **1998**, *109*, 2287−2293.

(3) Antoniou, D.; Schwartz, S. D. *J. Chem. Phys*. **1998**, *108*, 3620−3625.

(4) Antoniou, D.; Schwartz, S. D. *J. Phys. Chem. B* **2001**, *105*, 5553−5558.

(5) Bahnson, B.; Park, D.-H.; Kim, K.; Plapp, B. V.; Klinman, J. P. *Biochemistry* **1993**, *32*, 5503−5507.

(6) Bahnson, B. J.; Colby, T. D.; Chin, J. K.; Goldstein, B. M.; Klinman, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12797−12802.

(7) Luo, J.; Kahn, K.; Bruice, T. C. *Bioorg. Chem*. **1999**, *27*, 289−296.

(8) Kohen, A.; Cannio, R.; Bartolucci, S.; Klinman, J. *Nature* **1999**, *399*, 496−499.

(9) Caratzoulas, S.; Schwartz, S. D. *J. Chem. Phys*. **2001**, *114*, 2910−2918.

(10) Caratzoulas, S.; Mincer, J. S.; Schwartz, S. D. *J. Am. Chem. Soc.* **2002**, *124*, 3270−3276.

(11) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA. 2000.

(12) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187−217.

(13) Allen, M. P.; Tildesley, D. J. *Computer Similation of Liquids*; Clarendon: Oxford, U.K., 1987.

(14) Chin, J. K.; Klinman, J. P. *Biochemistry* **2000**, *39*, 1278−1284.

(15) In addition to articles already cited, see, for example: (a) Sutcliffe, M. J.; Scrutton, N. S. *Eur. J. Biochem.* **2002**, *269*, 3096−3102. (b) Knapp, M. J.; Klinman, J. P. *Eur. J. Biochem.* **2002**, *269*, 3113−3121.

(16) Beasley, J. R.; Hecht, M. H. *J. Biol. Chem.* **1997**, *272*, 2031−2034.

(17) Theoretical studies of hydride transfer in LADH include the following: (a) Agarwal, P. K.; Webb, S. P.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2000**, *122*, 4803−4812. (b) Webb, S. P.; Agarwal, P. K.; Hammes-Schiffer, S. *J. Phys. Chem. B* **2000**, *104*, 8884−8894. (c) Billeter, S. R.; Webb, S. P.; Agarwal, P. K.; Iordanov, T.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2001**, *123*, 11262−11272. (d) Alhambra, C.; Corchado, J. C.; Sanchez, M. L.; Gao, J.; Truhlar, D. G. *J. Am. Chem. Soc.* **2000**, *122*, 8197−8203. (e) Alhambra, C.; Corchado, J.; Sanchez, M. L.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2001**, *105*, 11326−11340. (f) Truhlar, D. G.; Gao, J.; Alhambra, C.; Garcia-Viloca, M.; Corchado, J.; Sanchez, M. L.; Villa, J. *Acc. Chem. Res.* **2002**, *35*, 341−349.

(18) A recent theoretical study of tunneling in methylamine dehydrogenase: Faulder, P. F.; Tresadern, G.; Chohan, K. K.; Scrutton, N. S.; Sutcliffe, M. J.; Hillier, I. H., Burton, N. A. *J. Am. Chem. Soc.* **2001**, *123*, 8604−8605.