Anomalous Orientation-Dependent Effective Pair Interaction among Histidine and Other Amino Acid Residues in Metalloproteins: Breakdown of the Hydropathy Scale Index†

Arnab Mukherjee‡ and Biman Bagchi*,§

CNRS, UMR 8640, Paris, Université Pierre et Marie Curie, Paris6, and Chemistry Department, Ecole Normale Superieure, 24 rue Lhomond, Paris, F-75005 France, and Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore, India 560 012

Received November 8, 2005; Revised Manuscript Received February 15, 2006

ABSTRACT: Hydropathy scale is widely used to obtain a measure of the effective interaction between any two amino acid residues in proteins and is based on the assumption that attraction between two hydrophobic groups and repulsion between hydrophilic groups (in water) can be translated straightforwardly to protein environment. Here we employ a recently developed statistical mechanical approach combined with the Protein Data Bank to obtain both distance- and orientation-dependent potential of mean force (ODPMF). This allowed us to explore effective pair potential among many amino acid residues and to examine the validity of the hydropathy scale in modeling the interaction among amino acid residues. We find that in some cases, like Phe-Phe and Lys-Lys, the hydropathy scale approach is largely obeyed. However, we also observe many unexpected pair interactions which defy the trend given by published hydropathy scales. An example of the former is the arginine−arginine (Arg-Arg) pair interaction which is found to be strongly and surprisingly attractive at short separation, even though it is the most hydrophilic residue. Here the head-to-head (see text) interaction is also stabilized. Tryptophan residues also exhibit strong attractive interaction. Equally important, we find strong influence of metal in determining effective interaction among the amino acid residues. It is the behavior of the histidine (His) which is found to be the most unusual. It exhibits a strong attractive interaction with itself which gets significantly enhanced in metalloproteins. These results highlight the important (sometime hidden) role of metals in protein structure and folding.

It is generally believed that the functionality of a protein is closely related to its three-dimensional structure. To understand the folding dynamics and stability of a protein, it is important to understand the nature of interactions among various amino acid residues in the system. This is hard because the amino acid residues themselves are complex entities, consisting of many atoms of different size and shape and different polar and aromatic character. Several studies adopted various approaches to model the interactions in a protein that lead to its unique native structure. The most viable approach is the molecular dynamics simulation with all-atom representation of a protein (*1*). However, it is still such a computationally demanding technique that there exist only a few microsecond long folding trajectories (*2*), even after using the parallel and distributed computing techniques (*3*). In addition, it is not always easy to extract a physical picture of stability and dynamics from all-atom approaches. Almost on the other extreme are the coarse-grained implicit minimalistic models which try to retain only the essential details of the nature of interactions in proteins. Such coarse-grained models have contributed significantly to the general understanding of the protein folding problem (*4*−*11*). Nevertheless, the scope of these minimalistic models is limited and lacks many specific interaction details (*5*, *7*, *9*).

More often than not, interactions among the amino acid residues in a protein are analyzed in terms of simple chemical properties of the amino acids such as hydrophilicity and hydrophobicity (*12*−*16*). While this gross classification is helpful in understanding the nature of the interaction in the early stage (the collapse) of folding (*7*), it, however, may not be useful to understand the late stage of folding where the specific tertiary interactions build up (*17*). Several minimalistic models have used the hydropathy scale (*18*−*20*) to characterize the model interaction potential which is then used in computer simulations (*6*, *9*, *10*, *21*, *22*). Among several different procedures of constructing the hydropathy scale [SPLIT server considers 88 hydropathy scales for the prediction of transmembrane protein structure (*23*)], the most common scale of Tanford (*14*) gives a measure of the free energy cost of transfer of an amino acid from hexane to water. This scale is used to obtain the interaction energy among amino acid residues. Therefore, such a scale does not do full justice to the protein environment, which also needs to be considered in order to understand the role of the hydrophobic effect in protein folding and stability. Rose and co-workers (*24*) calculated the hydropathy scale of amino acid residues from mean fractional area loss while Sippl and co-workers (*25*) calculated the scale from the structure of proteins. Although these scales were shown to correlate well with the Tanford hydropathy scale (*14*), the mapping of interaction energy to hydropathy (using principle component

† This work was supported in part by IFCPAR, New Delhi, India, EGIDE, France, and DST, India.

* To whom correspondence should be addressed. Phone: +91 80 2293 2926. Fax: +91 80 2360 1310. E-mail: bbagchi@sscu.iisc.ernet.in.

‡ CNRS, UMR 8640, Université Pierre et Marie Curie, and Chemistry Department, Ecole Normale Superieure.

§ Indian Institute of Science.

10.1021/bi0522899 CCC: $33.50
Published on Web 04/01/2006

analysis) involves huge loss of information (*25*). Chandler has recently discussed the difficulty of quantifying hydrophobic effect due to the "multifaceted" nature of the phenomenon (*26*). Chandler et al. have also discussed the length scale dependence of the hydrophobic effect (*27*). In this work we address the issue of the hydrophobic effect in the context of the well-known hydropathy scale, with particular attention to metalloproteins.

One of the important aspects of the structure−function relationship in proteins is observed in the case of the metalloproteins, i.e., proteins containing metal atoms (*28*−*30*). Approximately one-third of all proteins are metalloproteins (*31*). Not only do the metal ions play an important role in defining the three-dimensional structure of the metalloproteins [such as in DNA-binding zinc finger proteins (*32*)], these ions are also found to serve as prosthetic groups in the active sites of proteins or as a cofactor in metal-activated proteins. Metalloproteins take part in biological reactions such as electron transfer, transportation of reactive metals through the circulatory system without any local reactions, and formation of ion channels in transmembranes. The metal ions are also responsible for the structure and stability of the metalloproteins. Different functional activities carried out by proteins are consequences of the widely differing structural arrangements of the amino acid residues in a protein. This arrangement of some of the amino acid residues in a protein is largely influenced by the presence of a particular metal. The potential of mean force (PMF) study presented in this paper brings out the aspects of this involvement of metals in the structural arrangement of a few specific amino acid residues in metalloproteins.

Evaluation of the potential of mean force is often based on a statistical mechanical approach which uses the information hidden in the three-dimensional structure of the folded protein to understand the nature of interactions. So PMF is kind of an inverse protein folding problem having a goal to solve the original folding problem.

Tanaka and Scheraga (*33*) pioneered the technique of potential of mean force calculation using available three-dimensional structures of proteins. This approach was subsequently followed by Miyazawa and Jernigan by explicitly considering the solvent effect (*34*). Sippl (*35*) and others (*36*, *37*) extended these methods to include the dependence of PMF on pairwise separation of residues in space and along the primary sequence. Sippl (*35*) contributed greatly to the advancement of the PMF approach by developing a systematic approach (dynamical averaging technique) to overcome the averaging problem of a smaller data set of protein structures.

Apart from residue-only based approaches, there have been efforts to understand the interaction of the protein residues through other kinds of PMF. Skolnick et al. proposed a residue triplet term (*38*). Nishikawa et al. (*39*) and Kocher et al. (*40*) proposed the PMF of dihedral angles, and Matsuo et al. proposed the solvent accessibility and hydrogen bonding (*39*). Recently, Lazaridis and co-workers calculated the potential of mean force of ionizable amino acid residues in explicit solvent using umbrella sampling (*41*). The PMF approach has also been used to calculate the protein−protein interaction (*42*).

Scheraga and co-workers (*43*) calculated the long-range and local interaction energy terms of proteins from the crystal structures, and then they optimized the relative weights of the energy terms so that the native structures of the selected proteins become the lowest energy structures (*43*). The potential generated by Scheraga et al. is known as united residue force field, where the coarse graining is performed on the lesser important degrees of freedom. Feytsman and co-workers approached the evaluation of PMF of atoms constituting the amino acids from the chemical point of view (*44*) by categorizing all of the 167 heavy atoms into 40 groups depending on the bond connectivity or the chemical environment of the atoms. Naturally, the statistics of PMF calculated by this procedure are much higher, and this kind of grouping is chemically more relevant than the grouping of atoms based on the names of the residues. The potential generated by Feytsman et al. showed a high selectivity toward the native topology compared to the misfolded decoys (*44*, *45*).

The above-mentioned approaches of calculating PMF (both all-atom or the coarse-grained model) considered primarily the distance dependence of the potential. Perhaps Scheraga et al. were the first to use the Gay−Berne potential (*46*) to mimic the orientation dependence of interaction among the amino acid side chains. Recently, Buchete, Straub, and Thirumalai (BST)[1] (*47*) calculated an orientation-dependent PMF from the crystal structures of proteins. They generated a right-handed coordinate system with C_α, C_β, and C_γ atoms for each residue and thus calculated a five-dimensional potential which they later split to a sum of three-dimensional potentials that depend on all of the three polar coordinates (*47*). In a more recent development of the model, the same authors showed that the backbone−backbone and the backbone−side chain interactions are important. They fit their potential to spherical harmonic functions after which the potential seems to be even better in recognizing the native structures of the proteins (*48*).

Here we use a newly developed orientation-dependent potential of mean force introduced by us (*49*) to analyze specific aspects of interactions among the amino acid residues, both in the presence and in the absence of the metal ions in proteins. Our ODPMF (*49*) was tested on several sets of decoys, and the calculated *Z*-scores were shown to be comparable to the ODPMF developed by BST (*47*). Once the effectiveness of the ODPMF was demonstrated, we have carried out further analysis of specific amino acid pair interactions in the absence and presence of metals. ODPMF allowed us to probe the interactions in certain ways which are otherwise difficult to explore. We analyzed the correlation of the effective pair interactions obtained from ODPMF with the published hydropathy scales and found significant deviations from established belief that hydrophilic residues repel each other. Our studies also bring forth the role of metals in fostering interaction among the amino acid residues.

The results show that there is a significant enhancement in the effective interaction of the histidine residues in the case of the metalloproteins. While one would expect such enhancement in metal-induced interaction among amino acid residues, the merit of the present study lies in the quantifica-

[1] Abbreviations: ODPMF, orientation-dependent potential of mean force; Phe, phenylalanine; Lys, lysine; His, histidine; Arg, arginine; Trp, tryptophan; Ala, alanine; Glu, glutamic acid; SASA, solvent-accessible surface area; RSA, relative solvent accessibility of a residue; BST, Buchete, Straub, and Thirumalai; dpxr, average side chain residue depth.

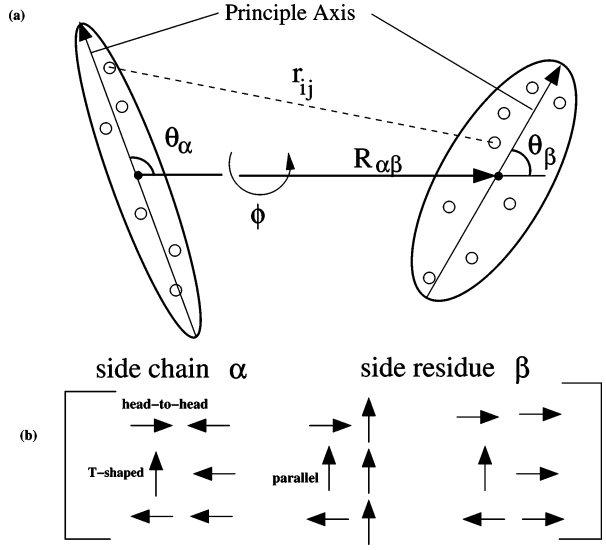

FIGURE 1: (a) The construction of four-dimensional ellipsoidal potential of mean force is shown schematically. $\mathbf{R}_{\alpha\beta}$ is the center of mass (COM) separation vector between the two ellipsoids. θ_α and θ_β are the corresponding angles between the principal axis of ellipsoid α and β and the COM separation vector $\mathbf{R}_{\alpha\beta}$ of the two ellipsoids, respectively. The circles on the ellipsoids denote heavy atom sites constituting the particular amino acid side chain represented by the ellipsoid. r_{ij} is the distance of separation between the sites i in ellipsoid α to another atom j in ellipsoid β. ϕ is the torsion angle between the principal axes of the two ellipsoids joined by their COM separation vector. (b) Arrows indicate the representative ellipsoids, approximately in line with the vector starting from the C_β atom to the center of mass of the particular side chain. Nine possible orientations among two interacting ellipsoids are shown (θ_α and θ_β change by 90° and $\phi = 0$). In the subsequent figures, PMF's are plotted only for these orientations. Some important orientations are referred by specific names.

tion of the effect of such interaction by using the Protein Data Bank (PDB) (*50*) directly. The present analysis also leads to many surprising results which are at complete odds with the trend given by published hydropathy scales. For example, arginine−arginine (Arg-Arg) interaction is found to be strong and surprisingly attractive at short separation, even though it is the most hydrophilic residue in the Kyte−Dolittle hydropathy scale (*20*). Here the head-to-head (see Figure 1) interaction is also a stable orientation. Tryptophan residues also exhibit strong attractive interaction in the protein environment. We find that metal ions can exert strong influences in determining effective interaction among the amino acid residues. In fact, the behavior of the histidine (His) residue is found to be the most unusual. It exhibits a strong attractive interaction with itself which gets significantly enhanced in metalloproteins. The results presented here highlight the important (sometimes even a hidden) role of metals in protein structure and folding.

Construction of the Orientation-Dependent PMF (ODPMF). The construction of the ODPMF has been discussed in detail elsewhere (*49*). Here we mention the essential points and the modifications. The construction of ODPMF involves the following steps: (i) The construction of atomistic PMF is done following Feytsman's procedure (*44*). This provides only distance-dependent (one-dimensional) PMF among the atomistic "sites". This is the only step that involves statistics, details of which are given in the next subsection. (ii) The distance-dependent atomistic PMF's thus obtained are converted to four-dimensional ODPMF's using rigid ellipsoidal models of side chains with the help of eqs 4 and 5. This process is schematically shown in Figure 1. Each ellipsoid in this figure represents a rigid side chain amino acid residue, and the circles represent the constituting atoms. This step does not involve any probabilistic calculations. Rather, this is a coarse graining from atomistic to a residue-based potential.

To study the effects of metals, the nonhomologous protein data set is divided into two parts: metalloproteins and nonmetalloproteins. The potentials are calculated separately for the two data sets, as described below.

(a) Construction of Site−Site PMF. There are several procedures to construct atomistic PMF (*35*, *43*, *51*). Most of these atomistic PMF's are based on either amino acids or all-atom PMF. Here we have followed the procedure of Feytsman (*44*) and classified the total 167 heavy (nonhydrogen) atoms into 40 different sites depending on their chemical environment and bond connectivity (*44*). Of these 40 different sites, 35 sites are exclusively located in the side chain part of the amino acid residues, and the remaining 5 sites belong to backbone atoms of a protein (*44*) (please see ref *44* for details of this classification). For example, the alanine (Ala) residue has only a single terminal methyl group in its side chain. A similar terminal methyl group is present in many other amino acids such as valine, leucine, isoleucine, and threonine. All of these terminal methyl groups belong to a particular site (no. 6 according to ref *44*). This kind of classification not only has a firm chemical basis but also allows us to obtain better statistics for the evaluation of PMF between amino acid residues (*44*). For 40 different sites, 1600 site−site PMF's (U_{ss}) are generated using the dynamical averaging technique of Sippl (*35*). We describe the procedure of the dynamical averaging technique just for clarity. First we accumulate the occurrence of two sites in bins of 0.5 Å up to a separation of 15 Å for all of the available pairs. Let $h(i,j,k)$ denote the number of occurrences for a certain pair of sites i and j at the kth bin (where $i, j = 1, 40$ and $k = 1, 30$). Initially, the relative probability $f(k)$ is generated by summing over all of the sites as follows:

$$f(k) = \frac{\sum_{i=1}^{40}\sum_{j=1}^{40} h(i,j,k)}{\sum_{i=1}^{40}\sum_{j=1}^{40}\sum_{k=1}^{30} h(i,j,k)} \quad (1)$$

The relative probability $f_{ij}(k)$ of any two particular $i-j$ pairs is given by

$$f_{ij}(k) = h(i,j,k)/H_{ij} \quad (2)$$

where $H_{ij} = \sum_{k=1}^{30} f(i,j,k)$ denotes the total number of observations for a particular $i-j$ pair. Finally, the expression of PMF of two $i-j$ atom sites is given by

$$U_{ij}(k) = \ln[1 + H_{ij}\sigma] - \ln\left[1 + H_{ij}\sigma\frac{f_{ij}(k)}{f(k)}\right] \quad (3)$$

where the weight function $\sigma = 1/50$ as proposed by Sippl (*35*) and also used by Feytsman et al. (*44*). For the construction of the histogram $h(i,j,k)$, we have used a nonredundant set of protein structures which were used by

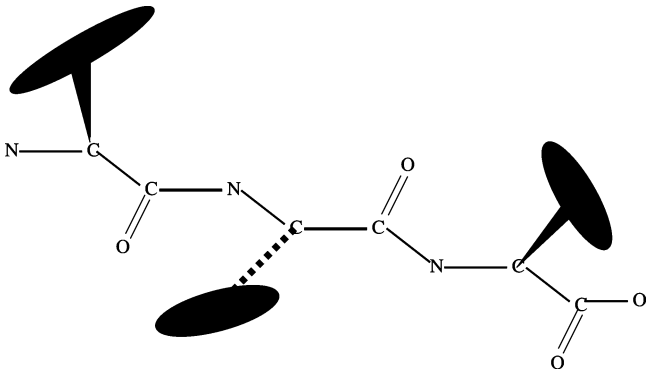

FIGURE 2: The basic model of a protein is shown. All non-hydrogen backbone atoms are considered. All of the atoms of the side chain are represented by a single ellipsoid.

Scheraga et al. (*43*) and BST (*47*). The names of the proteins are available in ref *43*.

Once these atomistic one-dimensional PMF's are generated, we used them in the construction of the ODPMF between ellipsoids. We calculate the effective PMF among the ellipsoids by varying the position and orientation of the two ellipsoids with respect to each other. For each separation and orientation of the two ellipsoids, the effective ODPMF is obtained from the total interaction of the atomistic sites as described below.

(b) Construction of Ellipsoidal Residue−Residue ODPMF. Most of the side chain residues can certainly be modeled better as ellipsoids than as spheres. Alanine and the backbone atoms are modeled as spheres because they comprise a single site. The basic construction of a protein using the coarse-grained model is schematically shown in Figure 2. The side chain residues are considered to be a rigid ellipsoid, with the center of mass denoting its position and the principal axis denoting its orientation.

Energy of interaction of a pair of residues depends on four variables: the distance between the center of masses of the two residues (R), the angles (θ_α, θ_β) which the principal axes of the two ellipsoids make with the interresidue vector, and the torsional angle between the two ellipsoids (ϕ) as shown in Figure 1. The principal axis is calculated by diagonalizing the moment of inertia matrix constructed from the atoms constituting a particular side chain residue, and the orientation of the axis is chosen to approximately match the vector starting from C_β to the center of mass of the side chain. This is done in order to remove the symmetry with respect to the minor axes. Next, the potential energy is calculated for all different R, θ_α, θ_β, and ϕ using the following equations. For a particular value of all of the above variables, the energy is calculated from a pairwise sum of site−sits potential as follows:

$$U_{\alpha\beta}(\mathbf{R}_{\alpha\beta},\theta_\alpha,\theta_\beta,\phi) = \sum_{i=1}^{N_\alpha}\sum_{j=1}^{N_\beta} U_{s_\alpha(i)s_\beta(j)}(|\boldsymbol{r}_\alpha(i) - \boldsymbol{r}_\beta(j)|) \quad (4)$$

where the subscript $s_\alpha(i)$ denotes site i of the α ellipsoid. Equation 4 sums all of the site−site PMF for particular values of R, θ_α, θ_β, and ϕ of two ellipsoids. The distance R is varied by 0.2 Å, and the angles θ_α, θ_β, and ϕ are varied each by 30° to get a four-dimensional potential ($U_{\alpha\beta}(\mathbf{R}_{\alpha\beta},\theta_\alpha,\theta_\beta,\phi)$) among any two particular amino acid residues. For all intermediate points, the values have been obtained by interpolation.

Finally, the ellipsoidal potential $U^{e}_{\alpha\beta}$ is calculated by averaging the energies with Boltzmann factor weight for all possible orientations obtained by rotation with respect to the principal axis; i.e.

$$U^{e}_{\alpha\beta}(\mathbf{R}_{\alpha\beta},\theta_\alpha,\theta_\beta,\phi) = \frac{\sum U_{\alpha\beta}(\mathbf{R}_{\alpha\beta},\theta_\alpha,\theta_\beta,\phi)\exp[-U_{\alpha\beta}(\mathbf{R}_{\alpha\beta},\theta_\alpha,\theta_\beta,\phi)]}{\sum \exp[-U_{\alpha\beta}(\mathbf{R}_{\alpha\beta},\theta_\alpha,\theta_\beta,\phi)]} \quad (5)$$

where $\sum$ indicates the summing over all possible rotations about the principal axes. This ensures that chemical properties of the side chains become symmetrical about the principal axes and that the calculated $U^{e}_{\alpha\beta}$ is pertinent to the present ellipsoidal model.

For some amino acids, side chains cannot be represented as ellipsoids. For example, glycine does not have a side chain. Alanine has only a C_β atom in the side chain, so it is modeled as a spherical atom. Serine and cysteine have only two atoms in the side chain, and their principal axes are taken to be $C_\beta \rightarrow$ terminal atom vector.

(c) Choice of Protein Structures. For the construction of atomistic site−site PMF, we have used the same nonhomologous set of protein structure as used by Scheraga (*43*) and BST (*47*). Some of the PDB ID's are changed from that of Scheraga's due to PDB errors. The general interaction of the amino acid residues is shown using the total data set. However, to show the metal-influenced interaction among a few specific amino acid residues i.e., histidine, we have separated the proteins into the sets of metallo- and non-metalloproteins. The ODPMF is calculated separately from each set. The effective interactions among the amino acid residues obtained from two different data sets are compared. This is discussed in detail in the subsequent sections.

Pairwise Interactions in Protein. A residue-based potential is useful to project out the various forms of interactions present among the residues in a protein environment. The advantage of the ODPMF constructed here is that it retains many details of information about the interaction while it brings out the more general aspect such as hydrophobic and hydrophilic interaction. In this section we discuss both types of interactions. We show in the following subsections that the hydrophobic residues such as phenylalanine−phenylalanine (Phe-Phe) interact strongly and attractively with each other whereas hydrophilic residues such as lysine−lysine (Lys-Lys) have rather weak repulsive interactions. The interactions are strongly orientation dependent. Later, we show that a hydrophilic residue such as His has strong interaction with itself, and the interaction enhances considerably in the presence of metal. One weakly hydrophilic residue such as tryptophan (Trp), on the other hand, has also very strong interaction.

(a) Phe-Phe Pair Interaction. Phenylalanine is characterized as both a hydrophobic and an aromatic amino acid residue. The interaction potential of this pair in a protein as a function of distance is shown in the Figure 3 for a range of orientations and has several interesting features. Except for a single arrangement where the C_β's are packed too closely, at short separation the remaining arrangements have

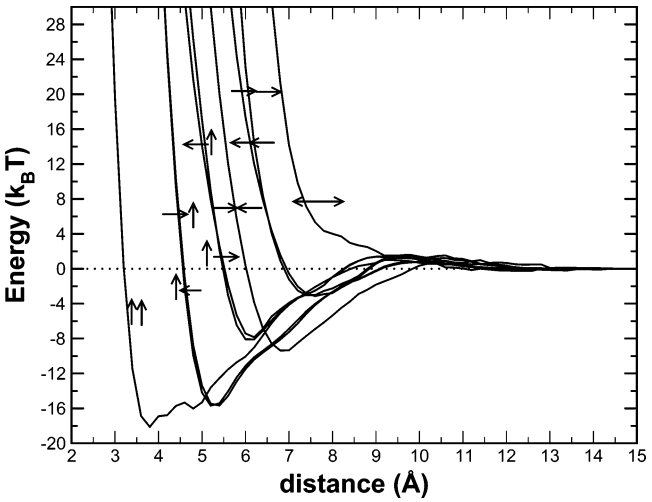

FIGURE 3: The interaction of the phenylalanine−phenylalanine pair is plotted as a function of the distance for different θ_α and θ_β orientations. See Figure 1. The dotted line in this figure and subsequent figures is just to guide the eye.

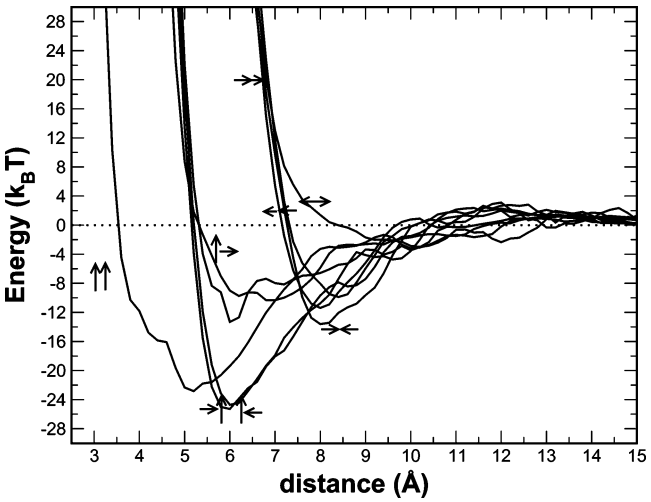

FIGURE 4: Energy of interaction between two tryptophan groups. Note the huge stabilization even being weakly hydrophilic in the Kyte−Doolittle scale. See Figure 1 for other details.

a pronounced minimum and hence a large stabilization. Most of the considered arrangements show a large stabilization, indicating that both the aromatic interaction (*52*) and the hydrophobic interaction join together to provide such stabilization. The interaction of phenylalanine has been studied by several groups (*53*, *54*), and it mainly belongs to the category of $\delta^+-\pi$ face-to-edge interaction (*53*). In this paper, the face-to-face π stacking interaction and the face-to-edge $\delta^+-\pi$ interaction are averaged out due to the axial symmetry of the ellipsoids. This reflects also in the width of the potential in parallel orientation. The distance dependence of the interaction is shown to have quadrupolar contribution (*55*). An ab initio study on the interaction of benzene revealed that a slipped parallel orientation is the maximum stabilized one (more than face-to-edge interaction) (*56*). Although these subtle geometrical arrangements are out of the scope of chosen orientational degrees of freedom of ellipsoidal residues, the constructed PMF gives a physically intuitive understanding of the packing and its energetics for pairwise interactions between hydrophobic aromatic side chains in the protein.

(b) Trp-Trp Pair Interaction. Tryptophan (Trp) is a weakly hydrophilic and aromatic amino acid residue. It has the largest surface area (*57*) and has been shown to impart thermal stability to a protein (*58*). The environment of the constituting atoms of Trp is discussed in detail by Chakrabarti et al. (*59*). It has been shown that a Trp residue is in the most favorable position around another Trp residue (*59*). The nature of the interaction involving the Trp residues is $\pi-\pi$ stacking, hydrogen bonding, van der Waals interaction, and NH$-\pi$ interaction (*59*). Here, we have plotted the PMF of two Trp residues in Figure 5. The strong attractive interaction asserts the strong preference of a Trp residue around another. Moreover, the ODPMF here points toward the preferable orientation of the residues. In Figure 5, we have shown two representative orientations of two interacting Trp residues obtained from the PDB (*50*).

(c) Lys-Lys Pair Interaction. Figure 6 presents the results of the PMF of Lys-Lys arrangements. Note that lysine is a positively charged hydrophilic residue. However, the charge is only at the end of a hydrophobic tail part. For all of the

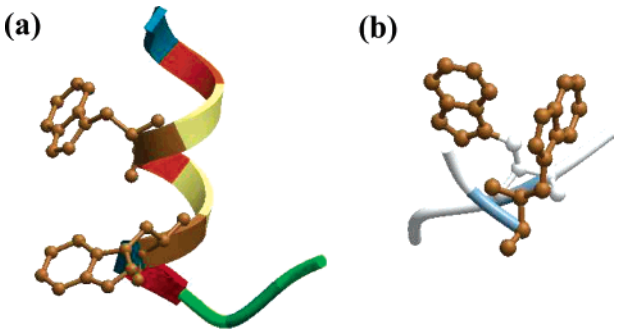

FIGURE 5: Two representative orientational arrangements of two Trp residues are shown. (a) and (b) are taken from proteins with PDB ID 5RUB and 2MSB, respectively.

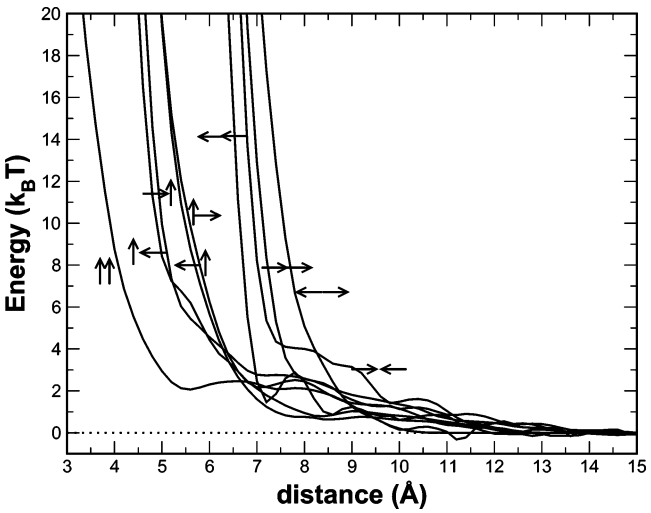

FIGURE 6: Energy of interaction of the lysine−lysine pair. The interaction is hardly attractive. Other details are the same as in Figure 1.

considered arrangements for the two residues, the interaction is largely repulsive. The long-range repulsion shows a stronger than $1/r$ Coulombic interaction but significantly weaker than van der Waals $1/r^{12}$ repulsive distance dependence, indicating a screening of charge by surrounding residues in the protein environment. This kind of electrostatic interaction is often treated with the Poisson−Boltzmann equation, which takes into account the shape of the molecule

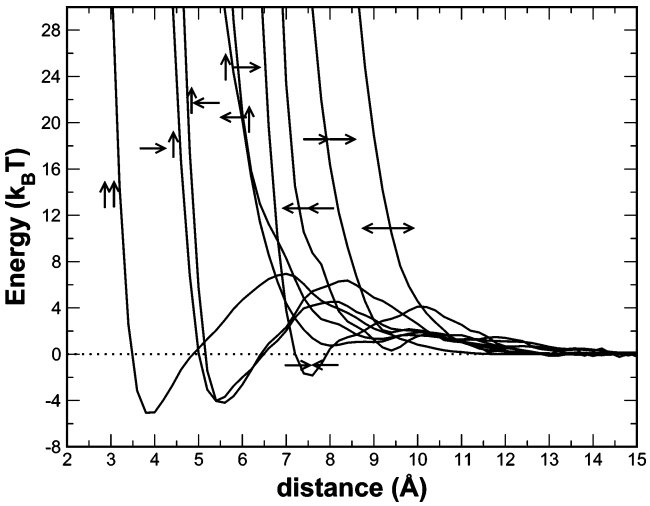

FIGURE 7: Energy of interaction between arginine groups. See Figure 1 for other details. Note that even though arginine is hydrophilic, it has considerable stabilization in parallel orientation. Some stabilization could also be observed for head-to-head orientation. Note also the intermediate maxima.

and its charge distribution and the distance-dependent dielectric constant of the medium (*60*). The special case where the residues maximize the contact area (parallel configuration) has a smaller and a softer core and, more interestingly, shows a plateau at intermediate distances before showing the typical long-ranged repulsion.

(d) Arg-Arg Pair Interaction. Arg is also a positively charged residue, and it is the most hydrophilic residue in the Kyte−Dolittle hydropathy scale. It has a long hydrophobic tail. However, Arg is different from Lys because the charge in Arg is more diffused, which allows it to have strong interaction with itself in the protein environment due to its hydrophobic tail. The parallel orientation among the two Arg residues is the most stable configuration, as shown in Figure 7. However, a somewhat surprising observation is the stable head-to-head orientation (see Figure 7) which is normally expected to have a repulsive effect (see the difference with Lys in Figure 6). Therefore, the possible stabilization of the head-to-head orientation could come from a dimer formation among two arginine groups. A similar observation for glutamate residues is mentioned elsewhere (*49*).

Interestingly, we observe intermediate maxima in the interaction potential of Arg in Figure 7. This kind of maxima arises probably due to the competition between the dispersion (which has $1/r^6$ dependence) and electrostatic potential (which has $1/r$ dependence). A representative figure of the unusual orientation is shown in Figure 8.

Lazaridis and co-workers (*41*) recently carried out extensive molecular dynamics simulations with an umbrella sampling technique in order to calculate the PMF of the charged amino acid residues. Here we would like to compare some of our statistical PMF with their physical PMF results. They have shown that the distance dependence of the PMF of two Arg residues in head-to-head orientation has a huge barrier at intermediate distance (*41*). Our statistical PMF of two Arg residues as shown in Figure 8 portrays a similar behavior. Now we compare the interaction of two oppositely charged residues such as glutamate (Glu) and lysine, or glutamate and arginine. Figure 9 shows the PMF of Glu with Arg, Lys, and Glu in head-to-head orientation. Glu-Arg and

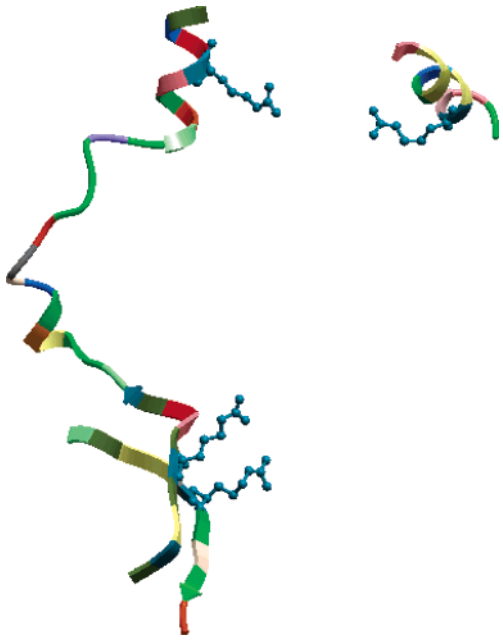

FIGURE 8: Two different orientations of two Arg residues are shown. The figure is obtained from the protein with PDB ID 3RUB.

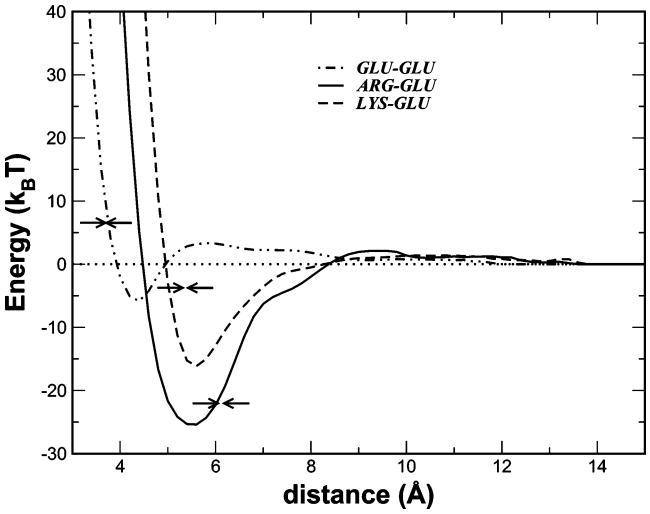

FIGURE 9: Energy of interaction of glutamic acid with lysine, arginine, and glutamic acid in head-to-head orientation. See Figure 1 for other details.

Glu-Lys are salt bridge-type interactions, and naturally they are stable. However, Glu-Arg interaction is more stable than Glu-Lys. This is in line with the observation of Lazaridis and co-workers (*41*) for the same interactions. The stability of Glu-Glu can be explained from the possibility of a double hydrogen-bonding (acetic acid dimer) type of arrangement (*49*). All of the interactions show a long-range repulsive tail before decaying to zero.

(e) Histidine−Histidine (His−His) Pair Interaction. The interaction among the histidine residues is perhaps the most surprising of all. This uniqueness can be attributed to the unique chemical nature of the histidine. Histidine is generally a hydrophilic residue in the Kyte−Doolittle scale. Due to the imidazole side chain and relatively neutral pK_a (6.0) value, the charge of the histidine is changed by a relatively small change in cellular pH. Therefore, histidine can exist in both neutral and positively charged forms at physiological pH. Histidine interacts strongly with aromatic residues and

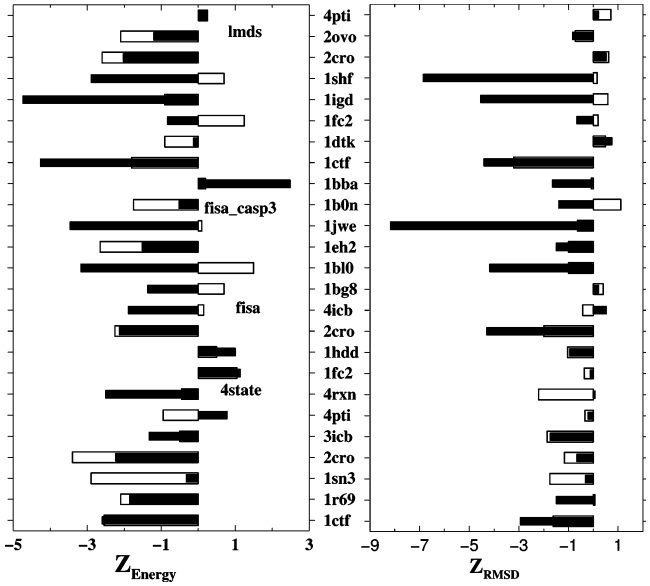

FIGURE 10: Comparison of the *Z*-score between BST potential (*47*) (unfilled bars) and this work (filled bars).

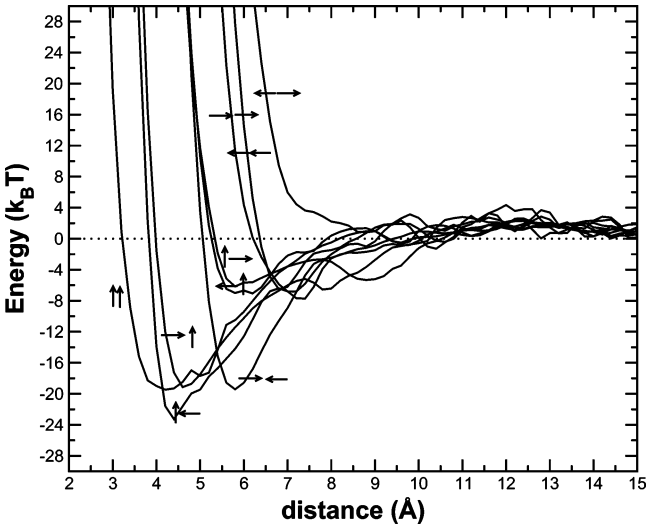

FIGURE 11: The interaction of the histidine−histidine pair plotted as a function of the distance for different θ_α and θ_β orientations. Also see caption for Figure 1. Note the strong attractive potential at short distances.

also forms hydrogen bonds with polar or charged (both positive and negative) residues (*61*). The recent study of Chakrabarti et al. showed that the preferred orientation of two histidine residues is parallel [i.e., the normals of the two planar rings have an angle of 0−30° contrary to the normal notion of a 90° interplanar angle in other aromatic interactions (*61*)]. Indeed, we find that the strongest interaction between two histidine residues comes from a parallel orientation as shown in Figure 11. However, the more important point to note is that the minimum in the interaction energy is deeper than that found between two most hydrophobic Phe residues. There is also a stabilizing head-to-head interaction among two histidine residues as shown in Figure 11. This may be due to the fact that, among two nitrogen atoms in the imidazole ring, one is slightly basic and other is slightly acidic, resulting in a stable dimer formation.

Because of the unique chemical nature of the histidine residue, it has a strong metal affinity where histidine residues act as ligands to a metal center, thereby playing an important role in the catalysis, oxidation reduction, and transport processes (*62*). In the following section, we discuss the effect of metal in the interaction of histidine residues.

(f) Test of the Potential. We have shown that the ODPMF described here can bring out several physical aspects of interaction among amino acid residues. It has been shown to correctly reproduce the hydrophobic, hydrophilic, charged interaction, etc. However, to establish the overall effectiveness of the potential, we have carried out *Z*-score calculations on several proteins and their decoys using our ODPMF (*49*). The *Z*-score for a particular quantity is defined as

$$Z_x = (x - \langle x \rangle)/\sigma_x \tag{6}$$

where $\langle x \rangle$ is the mean value and σ_x is the standard deviation for a certain distribution of x. Z_E is defined for the energy of the native structure (from the crystal structure or the NMR structure). Z_{RMSD} is defined with respect to the RMSD of the lowest energy structure obtained from the PMF calculation. If the *Z*-score (as defined above) is more negative, the potential is better in selecting out the native protein among its decoys. From the plot of the *Z*-score of energy and RMSD in Figure 10, we see that most of the *Z*-scores are negative. Our results have been compared to the recent work of BST (*47*). The filled bars show our results whereas the open bars show the results from BST (*47*). It can be noted that *Z*-score values for both energy and RMSD are quite comparable with those of BST (*47*).

Effect of Metals in the Interactions of Amino Acid Residues. In a detailed study of the geometry of metal ion and histidine interaction, it was found that the cations normally lie in the imidazole plane along the lone pair of the nitrogen atom (*63*). There are two tautomeric forms of the metal ion and histidine interaction; in one form, metal is coordinated to ND1 atom while in other it is coordinated to the NE2 atom of the imidazole ring of His with a proportion of 8.2 in basic aqueous solution (*64*). Some representative pictures of the histidine−metal interaction are shown in Figure 12.

The typical distance between the metal ion and the coordinating atom of His is ∼2 Å. Thus, when more than one His residue is coordinated to the metal ion, there is a spatial correlation between them. This effect of metal-influenced interaction is considered in this section with the help of our ODPMF approach. We have separated the set of proteins into metallo- and nonmetalloproteins. Then the ODPMF is calculated separately for the set of metallo- and nonmetalloproteins. Figures 13 and 14 show the orientation-dependent pair interaction of His with itself in the presence and absence of metal, respectively. There are several interesting points to note here. (a) First, there is a huge increase in the strength of interaction among the histidine residues in the presence of metal. (b) Second, the presence of metal does not alter the position of the minima in the PMF. (c) Third, note the geometry of the interaction. As pointed out by Chakrabarti and co-workers (*61*) that the interplanar angle between two histidine residues is around 90° when they interact through a metal ion, we also find in Figure 13 that the strongest interaction between two histidine residues in the presence of metal has indeed an angle of 90° (T-shaped as defined in Figure 1). Since we have an average over the principal axis of the ellipsoid, we cannot compare

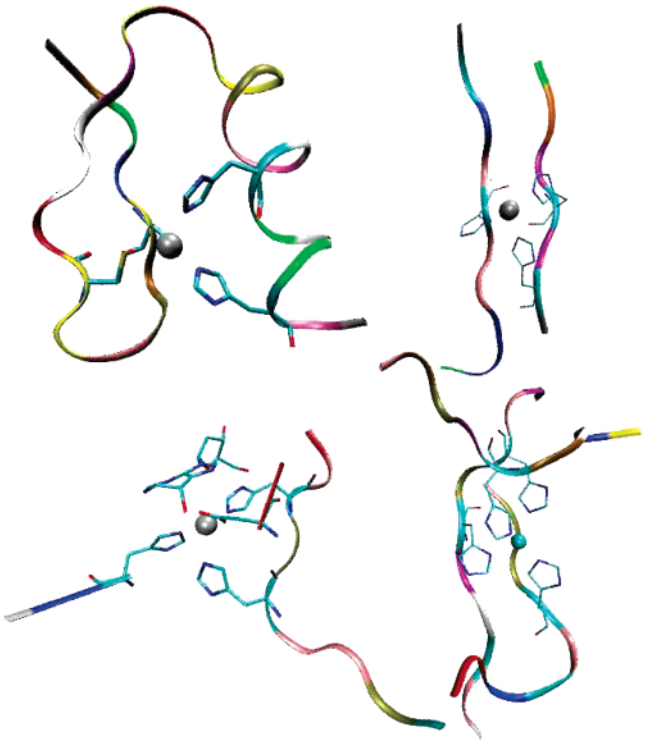

FIGURE 12: A few typical arrangements of histidine residues around metal. The pictures are made with the help of the metalloprotein database (*31*) and VMD (*65*).

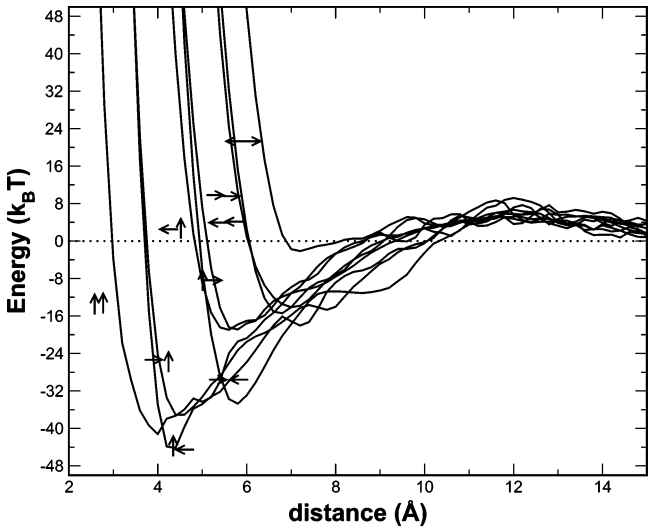

FIGURE 13: The interaction between histidine residues in the presence of metal. Compared to Figure 11, the potential is much more attractive. Also see caption for Figure 1.

the geometry of interaction along the θ angle mentioned in ref *61*. However, it is interesting to observe that although these PMF's have been constructed from atomistic PMF's, yet it retains the geometrical features when converted to a residue-based PMF.

To further investigate the influence of metal ion in the mutual interaction of the amino acid residues, we have calculated the effective interaction by Boltzmann averaging over all of the calculated state points of all four variables, θ_α, θ_β, ϕ, and $\mathbf{R}_{\alpha\beta}$. Figure 15 shows the effective interaction among all of the residues (both the similar and dissimilar pairs) calculated from metalloproteins and nonmetalloproteins separately. The figure shows that there are considerable

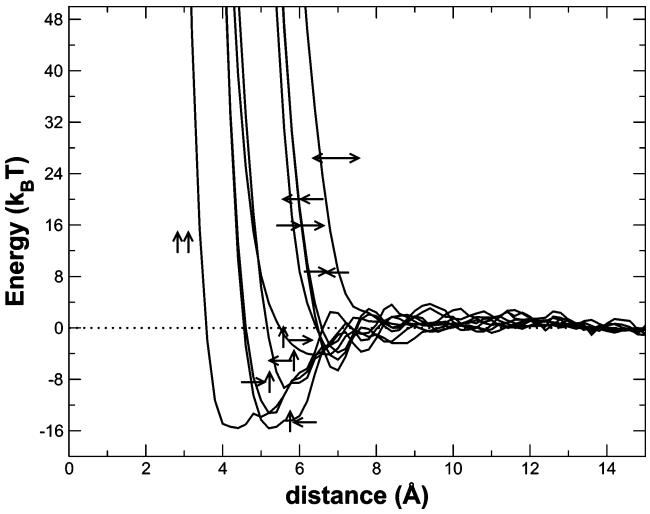

FIGURE 14: The interaction of histidine residues in the absence of metal. Note the dramatic decrease in stabilization compared to Figure 13. Also see caption for Figure 1.

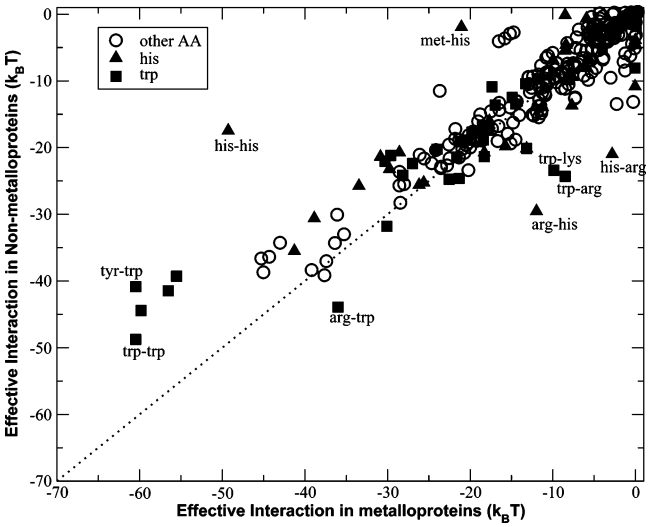

FIGURE 15: Comparison of the effective interactions of amino acid residues in the presence and absence of metal. Triangles and squares denote interaction of amino acid residues with histidine and tryptophan, respectively. All other interactions are shown by circles. Note that the deviations from the diagonal line are mainly for the histidine and tryptophan residues.

discrepancies (in terms of the departure from the diagonal line which depicts the indifference to a metal environment) in the effective interaction in the presence and absence of metal. The effect is most pronounced in case of His and in case of Trp residues, which are mainly heterocyclic compounds. These heterocyclic compounds are well-known to have strong interaction with transition metals due the exchange of electrons. Therefore, the presence of metal significantly influences the effective interaction among these residues. The points below the diagonal line signify a negative influence of the metal (i.e., which causes a decrease in the effective interaction strength) observed mostly in case of Arg and Lys, which are positively charged residues. Being positively charged species, Arg and Lys have stronger interaction with His in the absence of any metal. However, in metalloproteins, histidine residues interact more with metal, thus decreasing the interaction strength with positively charged amino acid residues.

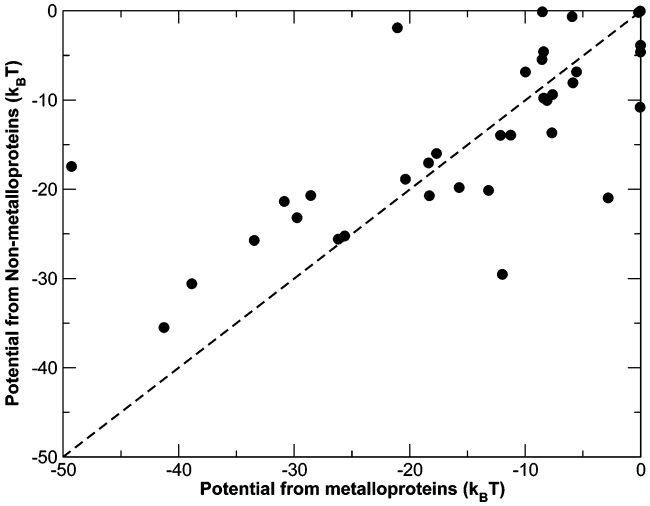

FIGURE 16: Similar figure as Figure 15, but only for histidine residues. This shows that most of the interactions are away from the diagonal line, which denotes an indifference toward the metal environment.

The comparison of effective interaction of His with all other residues in metal and nonmetal proteins is shown in Figure 16. Note the strong influence of metal, both positive and negative.

Comparison with Solvent Accessibility and Atom Depths. Solvent-accessible surface area (SASA) (*66*) and atom depth are two important parameters to study the structure of a protein. SASA is directly related to the hydropathy of the amino acid residues. However, SASA cannot provide sufficient insight to the stability of the protein in the interior of the proteins (*67*). The sum of atom depths, on the other hand, is shown to correlate much better with the stability of the protein (*67*), whereas the average residue depth again correlates well with the hydropathy scale (*68*). In light of the above facts, we have calculated relative residue accessibility (RSA) and the atom depths in both sets of metallo- and nonmetalloproteins.

In Figure 17, we have plotted four different parameters calculated separately in metallo- and nonmetalloproteins. Average DPX is the average side chain residue depth (dpxr) calculated using the DPX server (*69*). ∑DPX is the sum of dpxr for all proteins for a particular residue. RSA is the relative SASA of the side chain residues calculated using the "naccess" program (*70*). As expected, for both metallo- and nonmetalloprotein, the properties mentioned above follow the hydropathy scale in general (not shown here). However, when compared between metallo- and nonmetalloproteins, we observe that there is an increase of dpxr and ∑DPX for His in metalloproteins. There is a minor decrease in RSA for His in metalloproteins. Therefore, we can analyze from these results that although His is not the residue which primarily contributes to the overall stability of a protein, it has certainly an affinity toward a metal protein.

One more interesting observation is the plot which compares the average number of residues in metallo- and nonmetalloproteins. It can be observed here also that His is not the most or least abundant protein. However, there is a definite increase in the average number of residues of His in metalloproteins. The increase in average number together with the increase in average atom depth (dpxr) reflects a definite increase of His in the total atom depths for all proteins taken into consideration under metallo- and nonmetalloprotein here (∑DPX). Since the total atom depth correlates well with protein stability, the above analysis can throw some light on the apparent increase in the effective interaction among His residues in metalloproteins.

Conclusion. Let us first summarize the main results of this work. With the help of a orientation-dependent potential of mean force obtained by a statistical mechanical analysis of protein structures from the Protein Data Bank, we have observed that the presence of metal ions can induce a strong influence on the effective interaction potential among the amino acid residues. These results cannot be guessed from the hydropathy scale alone. In defiance of the hydropathy scale often used in the minimalistic models of protein folding,

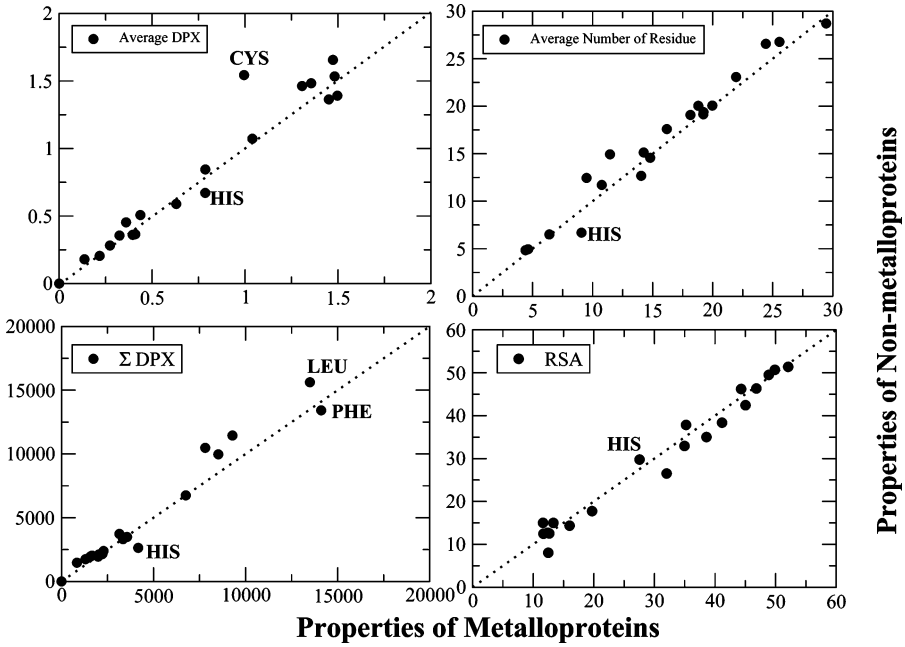

FIGURE 17: Four different properties (shown as the legend) of the amino acid side chain residues are compared among the sets of metallo- and nonmetalloproteins. The results corresponding to His are marked. The diagonal lines are just to guide the eye.

the effective interaction between some hydrophilic groups (like arginine and tryptophan) can be attractive at short separation! The case of arginine is particularly revealing because it is strongly hydrophilic, but the effective potential is attractive at short separation.

The calculations presented here allow for quantification of the effective interaction. In the case of histidine−histidine interaction, the maximum effective attractive interaction increases from $-16k_BT$ in nonmetalloprotein to $-48k_BT$ in metalloproteins. Such strong local interaction is expected to play a critical role in the stability and folding of proteins. Increase in the total atom depth and the average number suggest that this enhancement may be due to the apparent increase in the number of His residues in the core of the metalloproteins. Negative influences are also present. This may be due to the less preference for a particular pair in the presence of metals in a protein.

This study also draws our attention to the necessity to reconsider the construction of statistical potentials. We may not use a data set of protein without any classification except for their nonredundancy in the sequence. In a particular nonredundant set, it is important to consider the number of metallo- and nonmetalloproteins separately while calculating the *Z*-score or the total potential of a protein because almost one-third of the known proteins are metalloproteins (*31*). Moreover, the number of residues for a particular set of proteins can create an unbiased distribution. The specific pair interactions observed among the amino acid residues imply that although the standard notion of hydrophobicity is applicable to understand the interaction of residues with water, the protein environment creates a different situation for interaction phenomena among the residues.

In conclusion, we have used a new orientation-dependent potential of mean force to examine effective interaction between amino acid residues in proteins. We have observed several interesting new features which point toward the inadequacy of the hydropathy scale in describing the interaction between amino acid residues in the protein environment. We found that the effective interaction is different for histidine and tryptophan in metallo- and nonmetalloproteins. We have provided a molecular explanation for the unusual results presented in this work.

ACKNOWLEDGMENT

We thank P. Bhimalapuram, P. Balaram, P. K. Das, P. Chakrabarti, M. Vijayan, João Malhado, and Suyong Re for helpful discussions.

REFERENCES

1. (a) Richards, F. M., Eisenberg, D. S., Kuriyan, J., and Daggert, V., Eds. (2003) *Advances in Protein Chemistry*, Vol. 66, Publisher, Address; (b) Matthews, C. R., Ed. (2000) *Protein Folding Mechanisms*, Vol. 53, Publisher, Address.
2. Duan, Y., and Kollman, P. A. (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution, *Science 282*, 740−744.
3. Snow, C. D., Nguyen, H., Pande, V. S., and Gruebele, M. (2002) Absolute comparison of simulated and experimental protein-folding dynamics, *Nature 420*, 102−106.
4. Levitt, M., and Warshel, A. (1975) Computer simulation of protein folding, *Nature 253*, 694−698.
5. Levitt, M. (1976) A simplified representation of protein conformations for rapid simulation of protein folding, *J. Mol. Biol. 104*, 59−107.
6. Dill, K. A., Alonso, D. O. V., and Hutchinson, K. (1989) Thermal stabilities of globular proteins, *Biochemistry 28*, 5439−5449.
7. Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry 29*, 7133−7155.
8. Bryngelson, J. D., and Wolynes, P. G. (1987) Spin glasses and the statistical mechanics of protein folding, *Proc. Natl. Acad. Sci. U.S.A. 84*, 7524−7528.
9. Honeycutt, J. D., and Thirumalai, D. (1990) Metastability of the folded states of globular proteins, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3526−3529.
10. Sali, A., Shakhnovich, E., and Karplus, M. (1994) How does a protein fold?, *Nature 369*, 248−251.
11. Skolnick, J., and Kolinski, A. (1990) Simulations of the folding of a globular protein, *Science 250*, 1121−1125.
12. Tanford, C. (1973) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, John Wiley & Sons, New York.
13. Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation, *Adv. Protein Chem. 14*, 1−63.
14. Tanford, C. (1978) The hydrophobic effect and the organization of living matter, *Science 200*, 1012−1018.
15. Thomas, P., and Dill, K. (1996) Statistical potentials extracted from protein structures: How accurate are they?, *J. Mol. Biol. 257*, 457−469.
16. Levy, Y., and Onuchic, J. N. (2004) Water and proteins: A love-hate relationship, *Proc. Natl. Acad. Sci. U.S.A. 101*, 3325−3326.
17. Fersht, A. R. (1995) Optimization of rates of protein folding: The nucleationcondensation mechanism and its implications, *Proc. Natl. Acad. Sci. U.S.A. 92*, 10869−10873.
18. Stryer, L. (1995) *Biochemistry*, 4th ed., W. H. Freeman, New York.
19. Nelson, D. L., and Cox, M. M. (2005) *Lehninger Principles of Biochemistry*, 4th ed., W. H. Freeman, New York.
20. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol. 157*, 105−132.
21. Srinivas, G., and Bagchi, B. (2002) Foldability and the funnel of HP-36 protein sequence: use of hydropathy scale in protein folding, *J. Chem. Phys. 116*, 8579−8587.
22. Mukherjee, A., and Bagchi, B. (2003) Correlation between rate of folding, energy landscape, and topology in the folding of a model protein HP-36, *J. Chem. Phys. 118*, 4733−4747.
23. Juretic, D., Zoranic, L., and Zucic, D. (2002) Basic charge clusters and predictions of membrane protein topology, *J. Chem. Inf. Comput. Sci. 42*, 620−632.
24. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., and Zehfus, M. H. (1985) *Science 229*, 834−838.
25. Casari, G., and Sippl, M. J. (1992) Structure-derived hydrophobic potential, *J. Mol. Biol. 224*, 725−732.
26. Chandler, D. (2005) Interfaces and the driving force of hydrophobic assembly, *Nature 437*, 640−647.
27. Lum, K., Chandler, D., and Weeks, J. D. (1999) Hydrophobicity at small and large length scales, *J. Phys. Chem. B 103*, 4570−4577.
28. DeGardo, W. F., Summa, C. M., Pavone, V., Nastri, F., and Lombardi, A. (1999) De novo design and structural characterization of proteins and metalloproteins, *Annu. Rev. Biochem 68*, 779−819.
29. Xing, G., and DeRose, V. J. (2001) Designing metal peptide models for protein structure and function, *Curr. Opin. Chem. Biol. 5*, 196−200.
30. Lu, Y., Berry, S. M., and Pfister, T. D. (2001) Engineering novel metalloproteins: Design of metal-binding sites into native protein scaffolds, *Chem. Rev. 101*, 3047−3080.
31. Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A., and Pique, M. E. (2002) MDB: the Metalloprotein Database and Browser at The Scripps Research Institute, *Nucleic Acids Res. 30*, 379−382.
32. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) Three-dimensional solution structure of a single zinc finger DNA-binding domain, *Science 245*, 635−637.
33. Tanaka, S., and Scheraga, H. (1976) Medium- and long-range interaction parameters between amino acids for predicting threedimensional structures of proteins, *Macromolecules 9*, 945−950.
34. Miyazawa, S., and Jernigan, R. L. (1985) Estimation of effective interresidue contact energies from protein crystal structures: quasichemical approximation, *Macromolecules 18*, 534−552.
35. Sippl, M. J. (1990) Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins, *J. Mol. Biol. 213*, 859−883.

36. Hendlich, M., Lackner, P., Weitckus, S., Floechner, H., Froschauer, R., Gottsbacfher, K., Casari, G., and Sippl, M. J. (1990) Identification of native protein folds amongst a large number of incorrect models: the calculation of low energy conformations from potentials of mean force, *J. Mol. Biol. 216*, 167−180.
37. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) A new approach to protein fold recognition., *Nature 358*, 86−89.
38. Godzik, A., and Skolnick, J. (1992) Sequence-structure matching in globular proteins: Application to supersecondary and tertiary structure determination, *Proc. Natl. Acad. Sci. U.S.A. 89*, 12098−12110.
39. Nishikawa, K., and Matsuo, Y. (1993) Development of pseudoenergy potentials for assessing protein 3D-1D compatibility and detecting weak homologies, *Protein Eng. 6*, 811−820.
40. Kocher, J. P., Rooman, M. J., and Wodak, S. J. (1994) Factors influencing the ability of knowledge-based potentials to identify native sequence-structure matches, *J. Mol. Biol. 235*, 1598−1613.
41. Masunov, A., and Lazaridis, T. (2003) Potential of mean force between ionizable amino acid side chains in water, *J. Am. Chem. Soc. 125*, 1722−1730.
42. Jiang, L., Gao, Y., Mao, F., Liu, Z., and Lai, L. (2002) Potential of mean force for protein−protein interaction studies, *Proteins 46*, 190−196.
43. Liwo, A., Oldziej, S., Pincus, M., Wawak, R., Rackovsky, S., and Scheraga, H. (1997) A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data, *J. Comput. Chem. 18*, 849−873.
44. Melo, F., and Feytsman, E. (1997) Novel knowledge-based mean force potential at atomic level, *J. Mol. Biol. 267*, 207−222.
45. Melo, F., and Feytsman, E. (1998) Assessing protein structures with a non-local atomic interaction energy, *J. Mol. Biol. 277*, 1141−1152.
46. Gay, J. G., and Berne, B. J. (1981) Modification of the overlap potential to mimic a linear site-site potential, *J. Chem. Phys. 74*, 3316−3319.
47. Buchete, N.-V., Straub, J. E., and Thirumalai, D. (2003) Anisotropic coarsegrained statistical potentials improve the ability to identify native-like protein structures, *J. Chem. Phys. 118*, 7658−7671.
48. Buchete, N.-V., Straub, J. E., and Thirumalai, D. (2004) Orientational potentials extracted from protein structures improve native fold recognition, *Protein Sci. 13*, 862−874.
49. Mukherjee, A., Bhimalapuram, P., and Bagchi, B. (2005) Orientation-dependent potential of mean force for protein folding, *J. Chem. Phys. 123*, 14901−14911.
50. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
51. Eyrich, V. A., Friesner, R. A., Standley, D. M. (2002) Ab initio protein structure prediction using a size-dependent tertiary folding potential, *Adv. Chem. Phys. 120*, 223−263.
52. Scheiner, S., Kar, T., and Pattanayak, J. (2002) Comparison of various types of hydrogen bonds involving aromatic amino acids, *J. Am. Chem. Soc. 124*, 13257−13264.
53. Burley, S. K., and Petsko, G. A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization, *Science, 229*, 23−28.
54. Gould, R. O., Gray, A. M., Taylor, P., and Walkinshaw, M. D. (1985) Crystal environments and geometries of leucine, isoleucine, valine and phenylalanine provide estimates of minimum nonbonded contact and preferred van der Waals interaction distances, *J. Am. Chem. Soc. 107*, 5921−5927.
55. Burley, S. K., and Petsko, G. A. (1985) Weakly polar interactions in proteins, *Adv. Protein Chem. 39*, 125−189.
56. Cacelli, I., Cinacchi, G., Prampolini, G., and Tani, A. (2004) Modeling benzene with single-site potentials from ab initio calculations: A step toward hybrid models of complex molecules, *J. Chem. Phys. 120*, 3648−3656.
57. Chothia, C. (1976) The nature of the accessible and buried surfaces in proteins, *J. Mol. Biol. 105*, 1−14.
58. Chang, Y., Zajicek, J., and Castellino, F. J. (1997) Role of tryptophan-63 of the kringle 2 domain of tissue-type plas minogen activator in its thermal stability, folding, and ligand binding properties, *Biochemistry 36*, 7652−7663.
59. Samanta, U., Pal, D., and Chakrabarti, P. (2000) Environment of Tryptophan side chains in proteins, *Proteins 38*, 288−300.
60. Baker, N. A. (2004) Poisson−Boltzmann methods for biomolecular electrostatics, *Methods Enzymol. 383*, 94−118.
61. Bhattacharyya, R., Saha, R. P., Samanta, U., and Chakrabarti, P. (2003) Geometry of interaction of the histidine ring with other planar and basic residues, *J. Proteome Res. 2*, 255−263.
62. Sundberg, R. J., and Martin, R. B. (1974) Interactions of histidine and other imidazole derivatives with transition metal ions in chemical and biological systems, *Chem. Rev. 74*, 471−517.
63. Chakrabarti, P. (1990) Geometry of interaction of metal ions with histidine residues in protein structures, *Protein Eng. 4*, 57−63.
64. Reynolds, W. F., Peat, I. R., Freedman, M. R., and Lyerla, J. J. R. (1973) Determination of the tautomeric form of the imidazole ring of L-histidine in basic solution by carbon-13 magnetic resonance spectroscopy, *J. Am. Chem. Soc. 95*, 328−331.
65. Humphrey, W., Dalke, A., and Schulten, K. A. (1996) VMD−Visual Molecular Dynamics, *J. Mol. Graphics 14*, 33−38.
66. Lee, B., and Richards, F. M. (1971) The interpretation of protein structure: Estimation of static accessibility, *J. Mol. Biol. 55*, 379−400.
67. Chakravarty, S., and Varadarajan, R. (1999) Residue depth: a novel parameter for the analysis of protein structure and stability, *Structure 7*, 723−732.
68. Pintar, A., Carugo, O., and Pongor, S. (2003) Atom depth as a descriptor of the protein interior, *Biophys. J. 84*, 2553−2561.
69. Pintar, A., Carugo, O., and Pongor, S. (2003) DPX: for the analysis of the protein core, *Bioinformatics 19*, 313−314.
70. Hubbard, S. J., and Thornton, J. M. (1993) Naccess Version 2.1.1, Department of Biochemistry and Molecular Biology, University College, London.

BI0522899