Taylor & Francis
Taylor & Francis Group

Insight into a molecular interaction force supporting peptide backbones and its implication to protein loops and folding

Qi-Shi Du[a,b,d]*, Dong Chen[a,b], Neng-Zhong Xie[a], Ri-Bo Huang[a,b]* and Kuo-Chen Chou[c,d]

[a]State Key Laboratory of Non-food Biomass and Enzyme Technology, National Engineering Research Center for Non-food Biorefinery, Guangxi Academy of Sciences, 98 Daling Road, Nanning, Guangxi 530007, China; [b]Nanning Fermentation and Enzyme Engineering Research Center, Life Science and Biotechnology College, Guangxi University, Nanning, Guangxi 530004, China; [c]Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah 21589, Saudi Arabia; [d]Gordon Life Science Institute, 53 South Cottage Road, Belmont, MA 02478, USA

Although not being classified as the most fundamental protein structural elements like α-helices and β-strands, the loop segment may play considerable roles for protein stability, flexibility, and dynamic activity. Meanwhile, the protein loop is also quite elusive; i.e. its interactions with the other parts of protein as well as its own shape-maintaining forces have still remained as a puzzle or at least not quite clear yet. Here, we report a molecular force, the so-called polar hydrogen–π interaction (Hp–π), which may play an important role in supporting the backbones of protein loops. By conducting the potential energy surface scanning calculations on the quasi π-plane of peptide bond unit, we have observed the following intriguing phenomena: (1) when the polar hydrogen atom of a peptide unit is perpendicularly pointing to the π-plane of other peptide bond units, a remarkable Hp–π interaction occurs; (2) the interaction is distance and orientation dependent, acting in a broad space, and belonging to the 'point-to-plane' one. The molecular force reported here may provide useful interaction concepts and insights into better understanding the loop's unique stability and flexibility feature, as well as the driving force of the protein global folding.

Keywords: protein structures; secondary structure; protein loops; Hp–π interaction; hydrogen bond

1. Introduction

More than 60 years ago, the segmental structures of α-helices and β-strands in proteins were first reported and described, rather than by X-ray or NMR experimental observations, but by Pauling and Corey (1951; Pauling, Corey, & Branson, 1951) using the logical deduction based on the concept of hydrogen bond interactions. In justifying their hypothetical structures, Pauling and Corey considered which conformations of peptides were sterically allowed and which kind of hydrogen bond distribution therein is most efficient. Under the light of hydrogen bond concept, Pauling and Corey designed the structures of α-helices and β-strands, in which the hydrogen bonds maintain the structures of these two kinds of regular segments effectively, as shown in Figure 1(A) and (B). In the α-helices, the up cycle and the low cycle of peptide chains are fixed by the hydrogen bonds, and in the β-strands, the two neighboring β-sheets are also fixed by the hydrogen bonds. About three decades later, the handedness of α-helix, β-sheet/barrel formed by β-strands, as well as their packing arrangements in proteins have been in-depth investigated from the energetic point of view by counting various molecular forces such as hydrogen bond

interaction, electronic interaction, and van der Waals interaction, as elaborated in a series of publications during the 80's of last century (see e.g. Chou & Scheraga, 1982; Chou, Nemethy, & Scheraga, 1983a, 1983b, 1984; Chou, Nemethy, Pottle, & Scheraga, 1985; Chou, Nemethy, Rumsey, Tuttle, & Scheraga, 1985, 1986; Chou, Pottle, Nemethy, Ueda, & Scheraga, 1982).

According to the protein structural architecture, the loops are not classified into the regular category. But so far the efforts of using the energetics or molecular forces to analyze the protein secondary structure are successful only for the regular structure elements (Carlacci & Chou, 1990b, 1991; Chou, 1992; Chou, Carlacci, & Maggiora, 1990; Chou, Carlacci, Maggiora, Parodi, & Schulz, 1992; Chou, Maggiora, Nemethy, & Scheraga, 1988; Chou, Nemethy, Pottle, & Scheraga, 1989; Chou, Nemethy, & Scheraga, 1990; Chou et al., 1990; Gerritsen, Chou, Nemethy, & Scheraga, 1985; Scheraga, Chou, & Nemethy, 1982). Actually, in proteins about 47% of all residues belong to this type of irregular structural segments (the loops or coils) (Gordon, Marshall, & Mayot, 1999; Harrison, Pearl, Mott, Thornton, & Orengo, 2002; Joseph, Srinivasan, & de Brevern, 2011; Krissinel &

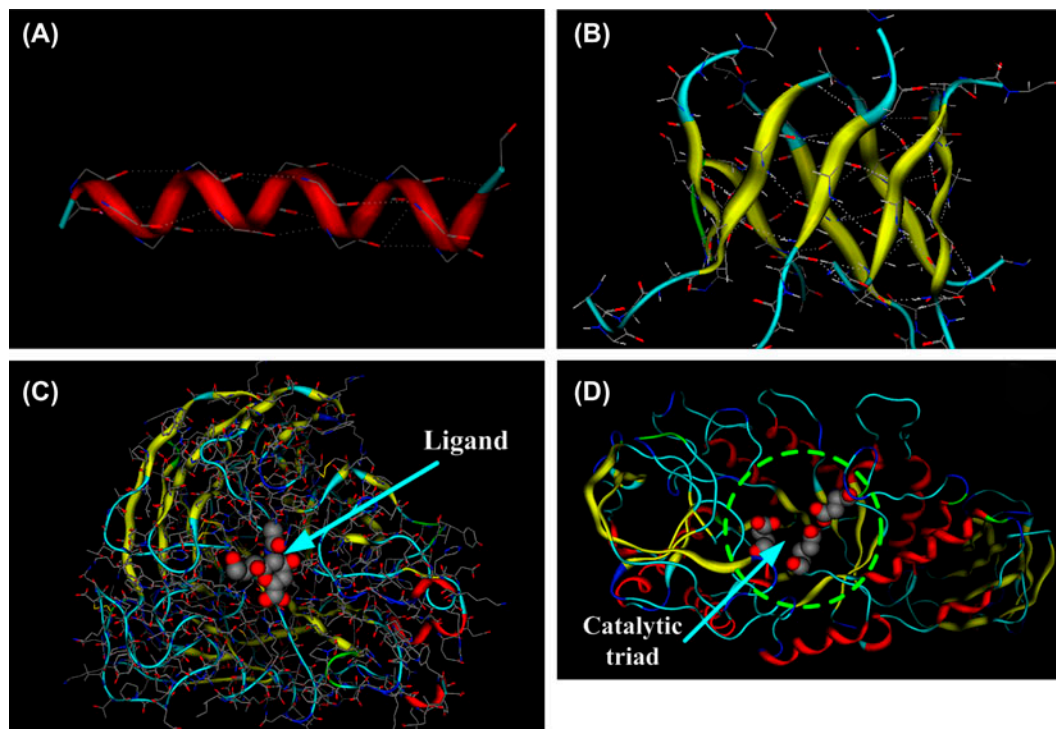Figure 1.   The structures of protein secondary structures α-helices, β-strands, and loops. (A) The hydrogen bonds in α-helix. In α-helix, the π-planes of peptide bond units in the up cycle and the lower cycle are arranged in parallel, and the hydrogen bond donors and acceptors between two cycles are in the best orientation for hydrogen bonds. (B) The hydrogen bonds in TIM $(\beta/\alpha)_8$ barrel of α-amylase. In β-strand, all π-planes of peptide bond units are arranged in parallel. The hydrogen bond donors and acceptors between two neighboring β-strands form stable hydrogen bonds. (C) The ligand binding location in neuraminidase (PDB code 1F8B) of influenza A viruses. All binding sites are in the loops. (D) In α-amylase (pdb code: 1BLI), three key residues (Asp231, Glu261, and Asp328), the so-called the catalytic triad, are located at the loops. In average, the residues in loops are up to 47%.

Henrick, 2004; Lu, 2000). Particularly, when using energetic approach to model many drug-target proteins important for drug development such as bovine somatotropin (Carlacci, Chou, & Maggiora, 1991), one could not avoid dealing with loop structures (Carlacci et al., 1991). Unfortunately, so far we still do not know the interaction force in supporting the loop structures.

In many cases, the loops are the most active parts in enzymes that play a key role in biological functions (Sirois, Sing, & Chou, 2005; Sirois, Touaibia, Chou, & Roy, 2007; Wen, Wang, Yang, & Xie, 2010, 2011). For example, almost all drug binding sites in neuraminidase (NA) of influenza A viruses (PDB code 1F8B) (Palese, Tobita, Ueda, & Compans, 1974; von Itzstein, 2007) are in the loops, as shown in Figure 1(C). In Figure 1(D) the three key residues (Asp231, Glu261, and Asp328) of enzyme α-amylase (PDB code 1BLI) (Machius, 1998), the so-called the catalytic triad, are located in the loops. The structural changes in the 150 loop of NA may cause the drug resistance problem (Du, Wang, & Huang, 2010; Du, Wei, & Huang, 2011; Wang, Du, Huang, & Zhang, 2009). The loops also play a key role for the two most remarkable features in proteins; i.e. the stability and flexibility, such as

in stabilizing the helix bundles (Carlacci & Chou, 1990a; Chou, 1991; Chou & Zheng, 1992; Chou, Maggiora, & Scheraga, 1992; Thompson, Chou, & Zheng, 1995), beta-barrel (Chou & Carlacci, 1991), and ion-channels (Berardi, Shih, Harrison, & Chou, 2011; OuYang et al., 2013; Schnell & Chou, 2008), and allowing the occurrence of low-frequency internal collective motions responsible for many remarkable biological functions of proteins (Chou, Li, Klee, & Bax, 2001; Chou, 1983, 1985, 1987, 1988, 1989; Chou & Chen, 1977; Wang & Chou, 2009, 2010; Wang, Gong, Wei, & Li, 2009). Proteins (or enzymes) are not a rigid object, but an entity that is dynamically quite flexible and structurally relatively stable in order to be able to perform various remarkable biological functions. In contrast, the α-helices, β-strands, and tight turns are relatively rigid parts, although they may move with respect to one another via modulating the conformation of the relevant loop. In other words, the loop can be likened to the 'pivot' or 'hinge,' around which the rigid parts of a protein may move to realize its function without breaking the system owing to the loop's existence.

The knowledge of the interaction forces in protein irregular structures is the necessary theoretical basis for

development of computational methods and tools to predict the protein folding and 3D structures. In the study of protein family evolution, most structural changes happen in the loops (Agarwal, Rajavel, Gopal, & Srinivasan, 2009; Bhaduri, Pugalenthi, & Sowdhamini, 2004; Marti-Renom, Ilyin, & Sali, 2001; Mizuguchi, Deane, Blundell, & Overington, 1998; Sujatha, Balaji, & Srinivasan, 2001). The loops are often the main engineering targets in protein engineering and protein design for improving the bioactivities and physicochemical properties of enzymes (Kuhlman et al., 2003).

Unlike α-helices and β-strands, in the protein loops, the quasi π-planes of peptide bond units often form the dihedral angles around 90°, which are not the appropriate orientations for the hydrogen bonds. The structures of loops cannot get reasonable explanation using the concept of hydrogen bonds. In the three-dimensional structures of proteins the loops look like soft silk threads. Are there any supporting forces in the loops? What physicochemical factors maintain the structural shapes of loops? In this study, we are to explore this important question based on the concept of polar hydrogen–π interactions (Birchall et al., 2011; Li & Fan, 2010; Mróz, 1995; Pidaparti, Svintradze, Shan, & Yokota, 2009), evaluate the interaction energies and analyze the role of polar hydrogen–π interactions in loops using quantum chemical calculations at the higher level.

2.   Method and theory

Hydrogen–π bonds are a special molecular interaction type, referring to the interactions between hydrogen atoms (attaching to different atomic groups) and the π-electron density of conjugate molecules (or groups). Hydrogen–π interactions can be classified into two types: the non-polar hydrogen–π interactions, in which the hydrogen atoms are attached to the carbon atoms (H–π or CH–π) (Birchall et al., 2011; Pidaparti et al., 2009); and the polar hydrogen–π interactions, in which the hydrogen atoms are attached to the electronegative atoms, such as oxygen, nitrogen, and chlorine (Birchall et al., 2011; Hughes & Waters, 2006). The polar hydrogen–π bonds and non-polar hydrogen–π bonds have different physical nature and properties. In this study, the notation 'Hp-π' is used for the polar hydrogen–π interactions (or bonds), to make it different from the common hydrogen bond (H–b) as well as the disputed non-polar hydrogen–π (H–π or CH–π) interactions. In protein structures, the Hp–π interactions could happen between amino acid side chains, or play roles in the peptide backbone of proteins. The Hp–π interactions between amino acid side chains have been studied in reference (Du, Wang, Du, Chen, & Huang, 2013). This study focuses on the Hp–π interactions in peptide backbones of proteins.

The peptide bond units of protein backbones are π-groups comprised by atom N, C, and O. The peptide bond units possess the interaction elements for both common hydrogen bond interactions and the polar hydrogen–π interactions. Although the hydrogen–π interactions have been studied for many years by many investigators (Birchall et al., 2011; Hughes & Waters, 2006; Pidaparti et al., 2009), the roles of Hp–π interactions in protein backbone structures and in protein folding have not been studied sufficiently. Two simple examples of the Hp–π interaction structures and highest occupied molecular orbital (HOMOs) of $CH_3OH...C_2H_4$ and $CH_3OH...C_6H_6$ are shown in Figure 2(A) and (B), where the polar hydrogen atom of CH_3OH is perpendicularly pointing to the π-plane of the conjugate molecules. In Figure 2, the HOMOs of the two Hp–π interaction systems ($CH_3OH...C_2H_4$ and $CH_3OH...C_6H_6$) are shown

The electron dispersion is an important contribution in the Hp–π interaction energies (Birchall et al., 2011). For better evaluating the electron dispersion and correlation energies involved in the Hp–π interactions, the post Hartree-Fock methods (Cramer & Bickelhaupt, 2003; Sherrill & Schaefer III, 1999) are necessary. In this study,
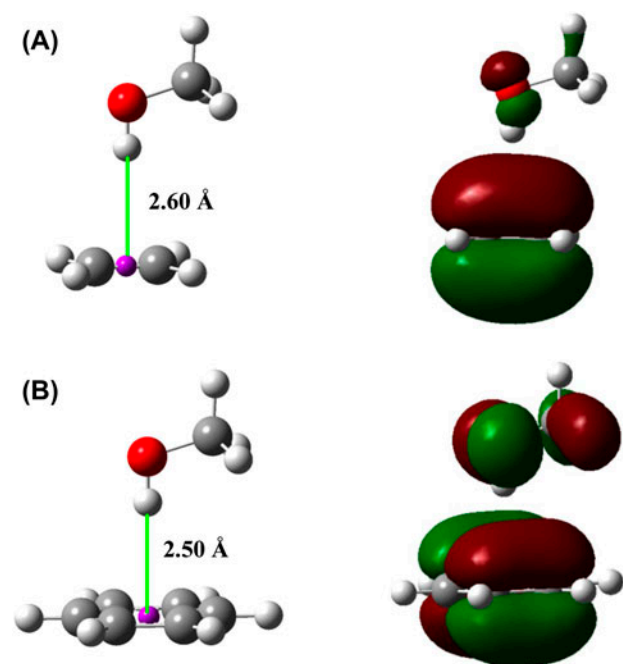

Figure 2.   Typical polar hydrogen–π (Hp–π) interactions. (A) The Hp–π interaction between CH_3OH and C_2H_4 (−12.887 kJ/mol). The polar hydrogen atom of CH_3OH perpendicularly points to the center (pink ball) of double bond of C_2H_4. (B) The Hp–π interaction between CH_3OH and C_6H_6 (−22.731 kJ/mol). The polar hydrogen atom of CH_3OH perpendicularly points to the center of benzene π-plane (pink ball). The polar hydrogen atom is in close touching with the π–MO of aromatic molecules.

the QM methods coupled-cluster singles and doubles (CCSD) and coupled-cluster singles, doubles, and partial triples (CCSD(T)) (Lee & Rice, 1988; Purvis III & Bartlett, 1982; Sherrill & Schaefer III, 1999) are adopted. It essentially takes the basic Hartree–Fock molecular orbital method and constructs multi-electron wave functions using the exponential cluster operator to account for electron correlation and dispersion. This method has been used by some of the most accurate calculations for small-to medium-sized molecules (Agarwal et al., 2009; Cramer & Bickelhaupt, 2003; Scuseria & Schaefer III, 1989; Scuseria, Janssen, & Schaefer, 1988; Sindzingre, Lhuillier, & Fouet, 2001). On the other hand, since the Hp–π bond length (~2.5 Å) is much longer than the common hydrogen bonds (~2.0 Å), large basis sets are necessary, including the polarization functions, diffuse functions, and floating functions. The CCSD(T) is better than the CCSD, because the CCSD uses only single- and double-excited configurations, while the CCSD(T) also includes part triple-excited configurations. The large basis functions make the number of higher excited configurations increasing tremendously. Accordingly, the CPU-time of CCSD(T) calculations is much longer than that of the CCSD calculations. The state-of-the art method CCSD(T) with large basis set is extremely expensive and CPU-time consuming. The test calculation results of a small Hp–π interaction pair $CH_3OH…C_2H_4$ by means of various methods and basis sets are listed in Table 1.

The Hp–π interaction energies, calculated using density functional theory (DFT) method B3LYP, are around 20–30% smaller than that of the two post Hartree-Fock methods (CCSD and CCSD(T)). This is because the common DFT methods fail in evaluating the dispersion energies that is an important contribution in Hp–π interactions. The more advanced method CCSD(T) gives better results than that of the CCSD method. However, the cpu-time of CCSD(T) is much longer than that of the CCSD. In solving this problem, a simple method is to use 'ghost atoms.' The ghost hydrogen atom H-Bq is an empty atom possessing the basis functions of hydrogen atom, but having no nucleus charge and electron (Asturiol, Duran, & Salvador, 2008; Balabin, 2008;

Mayer & Valiron, 1998). As shown in Table 1, the CCSD calculation using 6-311 + G(d,p) basis set plus a hydrogen 'ghost atom' H-Bq yields the result (−11.715 kJ/mol), very close to the result (−11.853 kJ/mol) of the CCSD(T)/cc-pVTZ method. However, the cpu-time taken by CCSD calculation is only 1/8 of the time by CCSD(T) calculation that is from 5.3 to 44.6 h. The CCSD(T) calculation using 6-311 + G(d,p) basis set plus a 'ghost atom' H-Bq obtains Hp–π energy −12.887 kJ/mol, better than that of the CCSD(T)/cc-pVTZ method. However, the cpu-time of CCSD calculation is much shorter than the CCSD(T) calculation. Another advantage of using the ghost atom is reducing the basis set superposition error (BSSE) (Asturiol et al., 2008; Balabin, 2008; Mayer & Valiron, 1998).

In this study, the molecular interaction energy between two molecules A and B are defined as the energy difference between A andB complex and the molecular monomers A and B; i.e.

$$\Delta E_{A-B} = E_{A-B} - (E_A + E_B)$$

Negative interaction energy values represent the attractive interactions, and the positive interaction energy values indicate the repulsive interactions. In this study, the large-scale scanning calculations are performed using CCSD/6-311 + G(d,p) method, and the interaction energies at optimized structures are calculated using the most advanced method CCSD(T)/6-311 + G(d,p)+H-Bq. The 'ghost hydrogen atom' (H–Bq) is attached to the polar hydrogen atom, and the distance to polar hydrogen atom is 0.8 Å. Keep it in mind that in the calculations for Hp–π interaction energies, the same ghost atom is also added to the two molecule monomers. The Hp–π interaction energies in solutions are calculated using the polarizable continuum model (PCM) (Amovilli et al., 1999; Cossi & Barone, 1998; Foresman, Keith, Wiberg, Snoonian, & Frisch, 1996; Miertuš, Scrocco, & Tomasi, 1981). In PCM method, the solvation effect of protein in solution is modeled as a polarizable continuum, rather than individual molecules, makes *ab initio* computation feasible. All calculations were carried out with the

Table 1. Comparison of three methods (DFT, CCSD, and CCSD(T)) and four basis sets (6-311 + G(d,p), TZVP, cc-pVTZ, and 6-311 + G(d,p)+H-Bq) in calculations of polar hydrogen–π (Hp–π) interactions.

Hp–π [a]$CH_3OH…C_2H_4$	B3LYP		CCSD		CCSD(T)	
	Energy (kJ/mol)	cpu-Time (h)	Energy (kJ/mol)	cpu-Time (h)	Energy (kJ/mol)	cpu-Time (h)
6-311G(d,p)	−7.29	0.1	−9.72	4.3	−10.748	14
TZVP	−7.658	0.05	−9.179	3	−10.105	7.6
cc-pVTZ	−8.269	0.2	−10.662	16.5	−11.853	44.6
[b]6-311Gdp+H-Bq	−7.547	0.1	**−11.715**	5.3	**−12.887**	17.9

[a]Structures of molecular monomers are optimized at CCSD/6-311 + G(d,p) level.
[b]A hydrogen 'ghost atom' (H-Bq) is added to the 6-311 + G(d,p) basis functions of polar hydrogen atom. The ghost atom H-Bq is in the middle of Hp–π bond, and the distance to polar hydrogen is 0.8 Å.

Sugon-5000A computer and TH-1A super computer (www.nscc-tj.gov.cn) using Gaussian 09 and Gauss View 5 software packages (Frisch et al., 2010).

3. Results

All the results obtained are reported and summarized using tables and figures, followed up by brief comparisons and illustrations.

3.1. *Hp–π interactions in protein backbones*

The peptide bond unit in protein backbones can be represented by the *n*-methyl acetamide (NMA), in which the atom N, C, and O compose a π-plane, with the elements for both common hydrogen bond interactions and the polar hydrogen–π interactions, as shown in Figure 3(A). The polar hydrogen atom, attached to the nitrogen atom, is the donor of both common hydrogen bonds and the polar hydrogen–π bonds. The carbonyl oxygen atom is the acceptor of common hydrogen bonds, and the N–C–O triangle π-plane is the acceptor of the polar hydrogen–π interactions. In the protein backbone structures, the interaction space for hydrogen bond interaction is in the front of oxygen atom on the π-plane, and the Hp–π interaction space is on the perpendicular direction of the N–C–O triangle π-plane, as shown in Figure 3(B). The interaction space of Hp–π bond is much larger than that of hydrogen bond interactions. The Hp–π interaction structure, between two NMA molecules is shown in Figure 3(C). In the Hp–π interaction structure the polar hydrogen atom of NMA-1 is perpendicularly pointing to the π-plane of NMA-2. The HOMO of Hp–π interaction is illustrated in Figure 3(D), in which the polar hydrogen atom of NMA-1 is in closely touching with the π-MO of peptide bond unit. The main physical contributors in the Hp–π interactions are the electrostatic interaction and the molecular orbital coordination between the polar hydrogen atom and the π-MO of conjugate molecules (Du, Long, Meng, & Huang, 2012). Shown in Figure 3(E) and (F) is a comparison between the two NMA molecules in the common hydrogen bond interaction structure.

The interaction energies of some typical Hp–π interactions are listed in Table 2. Comparing the Hp–π interaction energies of NMA–C_2H_4 (−12.887 kJ/mol) with NMA–$C_6H_5CH_3$ (−22.731 kJ/mol), we can see that the Hp–π energies increase with the size of conjugate (or aromatic) molecules. The Hp–π interactions are the point to π-plane interactions, which could happen at all π-plane. In the NMA–NMA Hp–π interaction, the interaction energies at three atoms N, C, O, and at the mass center of π-plane are calculated individually. For facilitating comparison, in Table 2 we also list the interactions energies of the cation–π interaction $CH_3NH_3^+$–$C_6H_5CH_3$ (−50.86 kJ/mol)

(Du, Long, et al., 2012; Du, Meng, Liao, & Huang, 2012), the hydrogen bond energies of water–water (−21.26 kJ/mol) and NMA–NMA, (−24.39 kJ/mol). In the hydrogen bond interaction of NMA–NMA the two π-planes share the same plane, and the polar hydrogen atom of NMA-1 points to the carbonyl oxygen atom of the NMA-2. The cation–πn–π interaction energy (−50.86 kJ/mol) of $CH_3NH_3^+$–$C_6H_5CH_3$ is much larger than the energies of three Hp–π interactions because of the large positive charge of proton than the charge of polar hydrogen atom. The NMA–NMA Hp–π interaction energies at N (−17.33 kJ/mol), C (−16.46 kJ/mol), and the mass center (−16.00 kJ/mol) are around 2/3 of the NMA–NMA hydrogen bond energy (−24.39 kJ/mol), and the Hp–π energy at O (−24.25 kJ/mol) is very close to the hydrogen bond energy of NMA–NMA. The Hp–π interaction energies are comparable to the common hydrogen bonds, therefore are one of the main contributors that maintains the structures of protein backbones. In Table 2 the Hp–π interaction energies of NMA–NMA in water solution are calculated using the PCM (Amovilli et al., 1999; Cossi & Barone, 1998; Foresman et al., 1996; Miertuš et al., 1981), which are around half of the Hp–π interaction energies in vacuum. Like common hydrogen bond interactions, the Hp–π interactions are less affected by the solvation effect and play significant roles in the aqueous solutions of proteins.

3.2. *Potential energy surface of NMA–NMA Hp–π interaction*

As illustrated in above section, the Hp–π interactions are the point to π-plane interactions, which could occur at all π-plane. It is helpful to scan the potential energy surface of NMA–NMA Hp–π interactions at the π-plane. The scanning calculations are performed in a 4×8 Å^2 area at the π-plane with a step of 0.1 Å increment. A total of 3321 points were computed, which was a computation intensive and cpu-time consuming procedure. In the scanning calculations, the distance between the polar hydrogen atom and the π-plane was kept at 2.5 Å that was the typical Hp–π interaction distance. The contour map of potential energy surface of NMA–NMA Hp–π interactions is shown in Figure 4.

In general, the potential energy surface of NMA–NMA Hp–π interaction is like a curve sliding board, where a potential energy groove is along the chemical bond direction N→C→O, and bounded by potential energy barriers. In the potential energy surface, there are three narrow potential energy wells at the positions of N (−17 kJ/mol), C (−16 kJ/mol), and O (−24 kJ/mol), and a broad potential energy pool at the front of O atom. The Hp–π interaction energies of NMA–NMA are listed in Table 2. In the process of calculation, the polar hydrogen was kept perpendicular direction to the π-plane. If
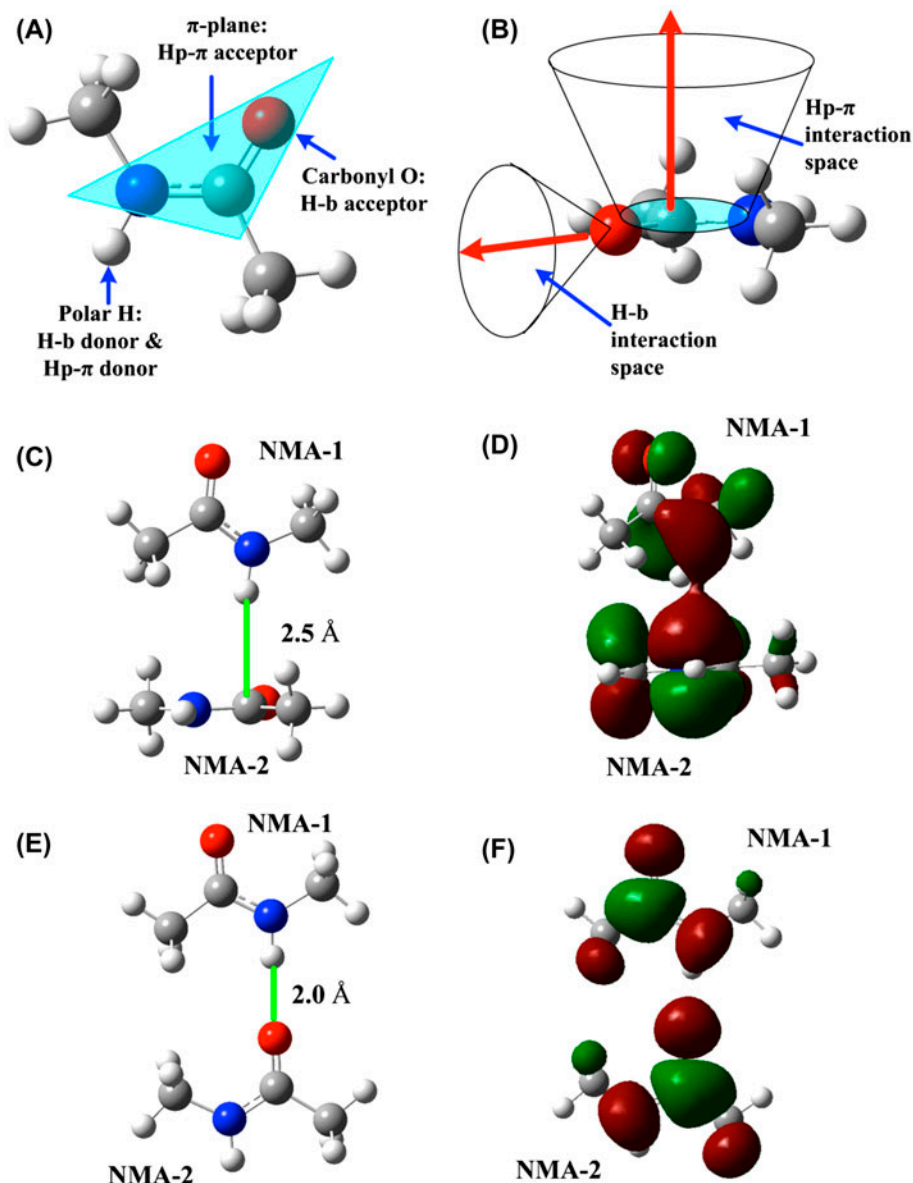
Figure 3.   The structure of NMA is the model of peptide bond unit in protein backbones. (A) In NMA, the atom N, C, and O comprise a π-plane, possessing the interaction elements for both common hydrogen bond interactions and for polar hydrogen–π interactions. The polar hydrogen atom, attaching on the nitrogen atom, is the donor for both common hydrogen bonds and polar hydrogen–π bonds. The carbonyl oxygen atom is the acceptor of common hydrogen bonds, and the N–C–O triangle π-plane is the acceptor of the polar hydrogen–π interactions. (B) The interaction spaces of common hydrogen bonds and of the Hp–π interactions in protein backbones. The common hydrogen bond interaction space is in the front of oxygen atom on the π-plane, and the Hp–π interaction space is on the top of the N–C–O triangle π-plane. (C) The polar hydrogen–π interaction (Hp–π) structure between two NMA molecules. The polar hydrogen atom of NMA-1 perpendicularly points to the π-planes of the NMA-2. (D) The HOMO of Hp–π interaction system. The polar hydrogen atoms are in closely touching with the NMA π–MO. (E) The common hydrogen bond (H–b) interaction structure between two NMA molecules. (F) The HOMO of H-b interaction between two NMA molecules.

the hydrogen atom changed its direction gradually, and eventually went to the π-plane of NMA-2, the NMA–NMA Hp–π interaction would change to NMA–NMA hydrogen bond interaction. In this case, the fourth broad energy well might spread even deeper and broader.

3.3. Potential energy surface scanning for torsion angles Φ and Ψ

Shown in Figure 4 is the potential energy scanning of NMA–NMA Hp–π interactions, where the two NMAs are independent molecules, which cannot be directly used to

Table 2. The interaction energies[a] and geometry parameters of some typical Hp–π interactions, cation–π interactions, and hydrogen bond interactions.

Interaction	Molecule pair	Energy (kJ/mol)	Distance (Å)
	NMA–C_2H_4	−12.887	2.535
	NMA–C_6H_6	−22.731	2.405
Hp–π	[b]NMA–NMA	in vacuum	
in vacuum	NMA–N	−17.329	2.310
	NMA–C	−16.464	2.514
	NMA–O	−24.252	2.368
	NMA-Center	−16.001	2.550
	[b]NMA–NMA	in water	
Hp–π	NMA–N	−12.207	2.654
in water	NMA–C	−8.276	3.110
	NMA–O	−15.759	2.319
Cation–π	$CH_3NH_3^+$–C_6H_6	−50.860	3.140
H-bond	[c]NMA–NMA	−24.391	2.019
	H_2O–H_2O	−21.258	1.997

[a]Hp–π interaction energies are calculated at CCSD(T)/6-311 + G(d,p)+H-Bq level.
[b]In the Hp–π interaction the polar hydrogen atom of NMA-1 perpendicularly points to the N, C, and O of NMA-2, respectively.
[c]In the common hydrogen bond between two NMAs, the two π-planes are in parallel.

illustrate the backbones of proteins, because the peptide bond units in protein backbones are connected by the single bonds of the C_α atoms. Between the two connected peptide bond units (denoted by p1 and p2), there are many possible combinations of the two torsion angles Φ (rotation along the N–C_α bond) and Ψ (rotation along the C_α–C bond) (Lovell et al., 2003; Perczel, Farkas, Jákli, Topol, & Csizmadia, 2003; Wi & Spano, 2011). The

connected peptide bond units and the definition of torsion angles Φ and Ψ are shown in Figure 5(C).

A potential energy surface scanning for the torsion angles Φ and Ψ is very helpful for exploring the full aspects of the Hp–π interactions in protein backbones. Shown in Figure 5(A) is the Φ–Ψ contour map of potential energy surface between two connected peptide bond units from −180° to 180° for both Φ and Ψ in 5° increment. The potential energy barriers, wells, and platforms in the Φ–Ψ contour map are summarized and illustrated in Table 3. The Φ–Ψ contour map is separated by three high-potential energy barriers from north to south in the middle of the map. The highest mountain peak in the center of the map is the conformation ($\Phi = 0°$, $\Psi = 0°$), in which the carbonyl oxygen atom of p1 is in conflict with the polar hydrogen atom of p2, as shown in Table 3. The other two energy barriers (at the top and bottom), actually, are the same conformation ($\Phi = 0°$, $\Psi = \pm 180°$), in which the two carbonyl oxygen atoms of p1 and p2 are very close. The two smaller energy barriers in the middle left and right are in the same conformation ($\Phi = \pm 180°$, $\Psi = 0°$), in which the two polar hydrogen atoms of p1 and p2 are very close to each other (see Table 3).

The Φ-Ψ conformations of β-strand are concentrated in the up-left corner ($\Phi = -180°$ and $\Psi = 180°$, deep blue). Actually, the other three corners are the same conformation ($\Phi = \pm 180°$ and $\Psi = \pm 180°$). In the area from ($\Phi = -180°$ and $\Psi = 180°$) to ($\Phi = -135°$ and $\Psi = 135°$) is the conformations for β-strands. The most stable β-strand conformation is ($\Phi = -165°$ and $\Psi = 165°$, $\Delta E = -46.03$ kJ/mol).
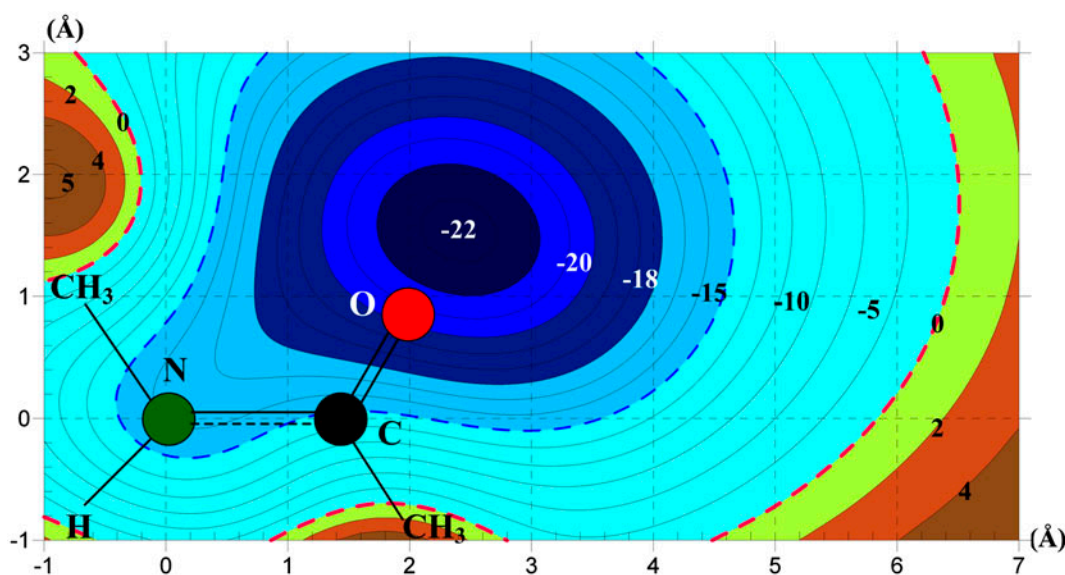

Figure 4. The potential energy surface of NMA–NMA Hp–π interactions. The NMA–NMA Hp–π interaction potential surface is like sliding board, where a potential energy groove is in the N–C–O triangle π-plane. The potential energies at three atoms N, C, and O are −17.329, 16.464, and −24.252 kJ/mol, respectively. In the calculations the polar hydrogen atom of NMA-1 keeps perpendicular direction to the π-plane of NMA-2.
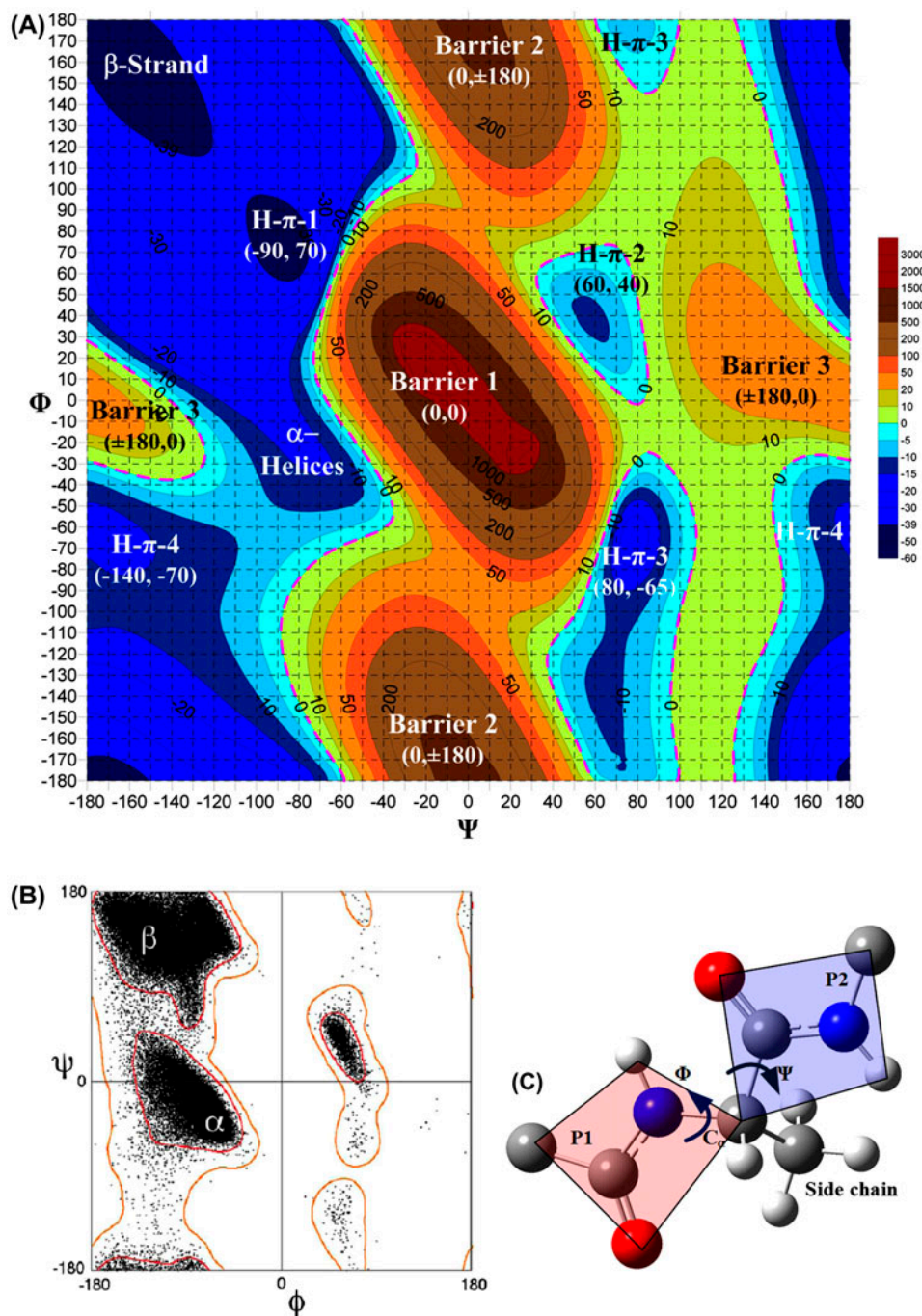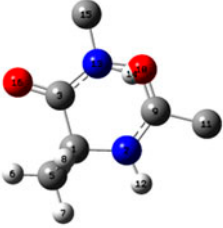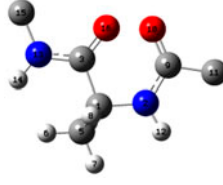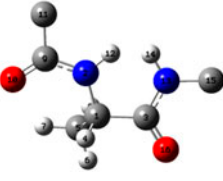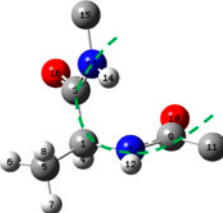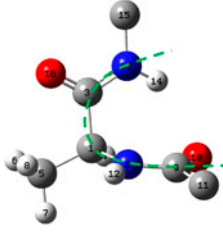
Figure 5.   Four loop structural conformations and corresponding Hp–π interaction types (H–π–1, H–π–2, H–π–3, and H–π–4) in dipeptide Φ–Ψ potential map. (A) Four Hp–π conformations distribute in the four quadrants of the potential map. The H–π–1 conformation is in the up-left quadrant, centered in th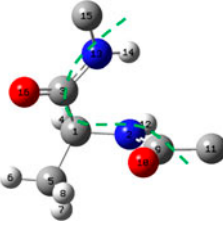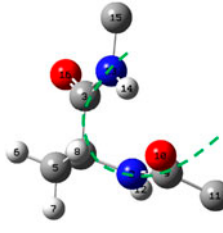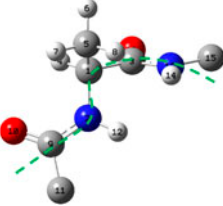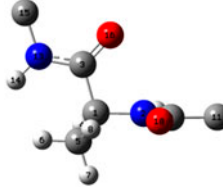e position ($\Phi = -90°$, $\Psi = 70°$). The H–π–2 conformation is in the up-right quadrant, centered in the position ($\Phi = 60°$ and $\Psi = 40°$), bounded by the -10 kJ/mol contour line. The center of H–π–3 conformation is in the lower-right quadrant, located in the ($\Phi = 80°$ and $\Psi = -65°$) area. The fourth Hp–π conformation (H–π–4) is in the lower-left quadrant, centered in the area ($\Phi = -140°$ and $\Psi = -70°$). (B) The Ramachandran plot (http://en.wikipedia.org/wiki/Ramachandran) was generated using density-dependent smoothing for 81,234 non-Gly, non-Pro, and non prePro residues with $B < 30$ from 500 high-resolution proteins (45). The Φ–Ψ contour map of theoretical potential energy surface is basically consistent with the Φ-Ψ torsion conformations in Ramachandran plots. (C) The definition of two torsion angles Φ and Ψ. The two connected peptide bond units are denoted by P1 and P2, respectively.

Table 3. The structures and energies of some key interaction conformations in potential energy contour map of Φ-Ψ torsion angle scanning.

[a]Energy barrier 1 Φ = 0°, Ψ = 0° ΔE = ∞ Collision between H and O	Energy barrier 2 Φ = 0°, Ψ = ±180° ΔE = 814.6521 Repulsion between two O atoms
Energy barrier 3 Φ = ±180°, Ψ = 0° ΔE = 38.2561 Repulsion between two H atoms	β-Sheet Φ = −165°, Ψ = 165° 2.14 Å ΔE = −46.0318 Most stable inner H-b
[b]α-Helices Φ = −65°, Ψ = −40° ΔE = −16.0318 Overlapped by H–π–1	H–π–1 Φ = −90°, Ψ = 70° R_{H-O} = 2.38 Å ΔE = −43.4913 H–π bond Loop
H–π–2 Φ = 60°, Ψ = 40° R_{H-N} = 2.31 Å ΔE = −12.9928 H–π bond Loop	H–π–3 Φ = 80°, Ψ = −65° R_{H-O} = 2.03 Å ΔE = −23.9234 H–π bond Loop
H–π–4 Φ = −140°, Ψ = −70° R_{H-C} = 2.51 Å ΔE = −14.6219 H–π bond Loop	[c]Reference state Φ = 90°, Ψ = 180° ΔE = 0.0000 Unfavorable for H–π interactions

[a]Energies in kJ/mol.
[b]The structural conformations of α-Helices and Hp–π bonds (H–π–1) partially overlap each other.
[c]The interaction energies (ΔE) of different interaction conformations are the energy differences between the energies of interaction conformations and the energy of a reference state. In this study, an unfavorable configuration (Φ = 90° and Ψ = 180°) for the Hp–π interactions is selected as the reference state.

The conformation of α-helices is in the region from (Φ = −90°, Ψ = −15°) to (Φ = −35°, Ψ = −70°), centered at (Φ = −65°, Ψ = −40°), keeping the appropriate 105° curve for the shape of helices. Actually, the conformations of α-helices are overlapped by the Hp–π conformations of loops. We will discuss this problem in more detail below.

In the Φ–Ψ contour map, there are four Hp–π conformations, located in the four quadrants of the map. In

Figure 5 and Table 3 the four Hp–π conformations are denoted by H–π–1, H–π–2, H–π–3, and H–π–4, respectively. The H–π–1 conformation is in the up-left quadrant, centered in the position (Φ = −90°, Ψ = 70°, deep blue), in which the polar hydrogen atom of p2 is perpendicular to the π-plane of p1, and the distance between polar hydrogen atom and oxygen atom is 2.38 Å, a stable Hp–π bond forming (ΔE = −43.49 kJ/mol). The H–π–1 is one of the most stable conformations in the

potential energy map, extending to a broad area shown in blue color (−20 kJ/mol) in Figure 5, covering the conformation of α-helices (Φ = −65° and Ψ = −40°) and β-strands (Φ = −165° and Ψ = 165°).

The H–π–2 configuration is in the up-right quadrant, centered in the position (Φ = 60° and Ψ = 40°), bounded by the −10 kJ/mol contour line. In the H–π–2 conformation, the polar hydrogen atom of p2 perpendicularly points to the nitrogen atom of p1, forming a Hp–π bond (ΔE = −12.993 kJ/mol, R_{H-N} = 2.31 Å). The center of H–π–3 configuration is in the lower-right quadrant, located in the (Φ = 80° and Ψ = −65°) area. In the H–π–3 configuration, the polar hydrogen atom of p2 perpendicularly points to the oxygen atom of p1, forming a stable Hp–π bond (ΔE = −23.923 kJ/mol, R_{H-O} = 2.03 Å).

The fourth Hp–π configuration (H–π–4) is in the lower-left quadrant, centered in the area (Φ = −140° and Ψ = −70°). Unlike other three Hp–π configurations, in the H–π–4 conformation, the Hp–π bond is formed between the polar hydrogen atom of p1 and the carbonyl carbon atom of p2. In the H–π–4 configuration, the Hp–π bond is in a four-atom ring. In Table 3, the curve shapes of the four Hp–π configurations are indicated by green dashed lines.

In the Φ–Ψ scanning calculations, a methyl (–CH$_3$) is attached to the C$_\alpha$, which is the side chain of alanine. If other side chains are used, we may get different Φ–Ψ contour maps. The interaction energies of different Hp–π conformations are the energy differences between the energies of Hp–π conformations and the energy of a reference state. In this study, we select an unfavorable configuration (Φ = 90° and Ψ = 180°) to the Hp–π interactions as the reference state, in which the polar hydrogen atom of p2 is in the opposite direction of the p1 π-plane. The structure of reference state is listed in the last cell of the Table 3.

3.4. Comparison between Φ–Ψ potential map and Ramachandran plot

The Ramachandran plot (Ramachandran, Ramakrishnan, & Sasisekharan, 1963) was first developed in 1963 by Ramachandran et al. that shows the backbone dihedral angles Ψ against Φ of amino acid residues in protein structure. The original Ramachandran plot was generated using hard-sphere calculations. Lately, such Φ–Ψ maps of dipeptide conformations were drawn based on the high-resolution protein structures (Lovell et al., 2003; Ting et al., 2010). The Ramachandran plot (http://en.wikipedia. org/wiki/Ramachandran) in Figure 5(B) was generated using density-dependent smoothing for 81,234 non-Gly, non-Pro, and non prePro residues with B < 30 from 500 high-resolution proteins (Lovell et al., 2003). The Φ–Ψ contour map of theoretical potential energy surface (Figure 5(A)) is basically consistent with the Φ–Ψ torsion

conformations in the Ramachandran plot (Figure 5(B)). In the Ramachandran plot, the population density of α-helix and β-strand conformations are indicated; however, there are no population distributions of loops.

The population of loops in Ramachandran plot should be larger than the population of α-helices and β-strands, because 47% residues are in the loops. Actually, the population of loops spread in broad area where the Hp–π conformations (H–π–1, H–π–2, H–π–3, and H–π–4) are located, and overlap part population of α-helices and β-strands. In the Ramachandran plot, the H–π–1 loops cover partial area of the β-strands in the up-left quadrant, and also overlap the area of α-helices between the low-left quadrant and up-left quadrant. The conformations of H–π–2 loops occupy the area in the up-right quadrant, and the conformations of H–π–3 loops are located at the area in the low-right quadrant, where is the population of the L-α-helices (Lovell et al., 2003). The conformations of H–π–4 loops spread in large area at the low-left quadrant of the Ramachandran plot, where the sparse population corresponds with the small interaction energy of H–π–4 interactions.

The Φ–Ψ population density in the Ramachandran plot represents the total folding free energy. However, the Φ–Ψ population density in the current theoretical Φ–Ψ potential map only represents the folding energy without taking into account the energies from other effects, such as entropy and solvation effects. A remarkable difference between the Ramachandran plot and theoretical Φ–Ψ potential map is for the α-helices. In the location of α-helices in the theoretical Φ–Ψ potential map, the potential well of α-helices is shallower than that of the β-strands. A possible reason is that more contribution from the solvent effect might be involved in forming α-helices than β-strands.

3.5. The Hp–π interactions in protein loops

In this section, the roles of Hp–π bonds in a real protein loop are analyzed. The loop in Figure 6(A) is taken from the structure of α-amylase (PDB code 1BLI) (Wang, Du, et al., 2009), containing 19 residues from Val324 to Phe343, in which the Asp328 is a key residue of the catalytic triad. Among the 19 residues, there are 15 Hp–π bonds, indicated by thin green lines. The Gln330 and Gly332 are fixed by the special peptide bonds of proline (Pro331). The residues Val339 and Gln340 are located at the straight segment of the loop, not supported by the Hp–π interactions. The Val339 and Thr329 are supported hydrogen bonds between side chain of Asp325 and peptide backbone, indicated by thick yellow lines. All other 14 residues are fixed by at least one Hp–π bond, indicated by thin green lines.

Figure 6(B) is the portrait of Hp–π bond between two connected residues Trp342 and Phe343. For a stable
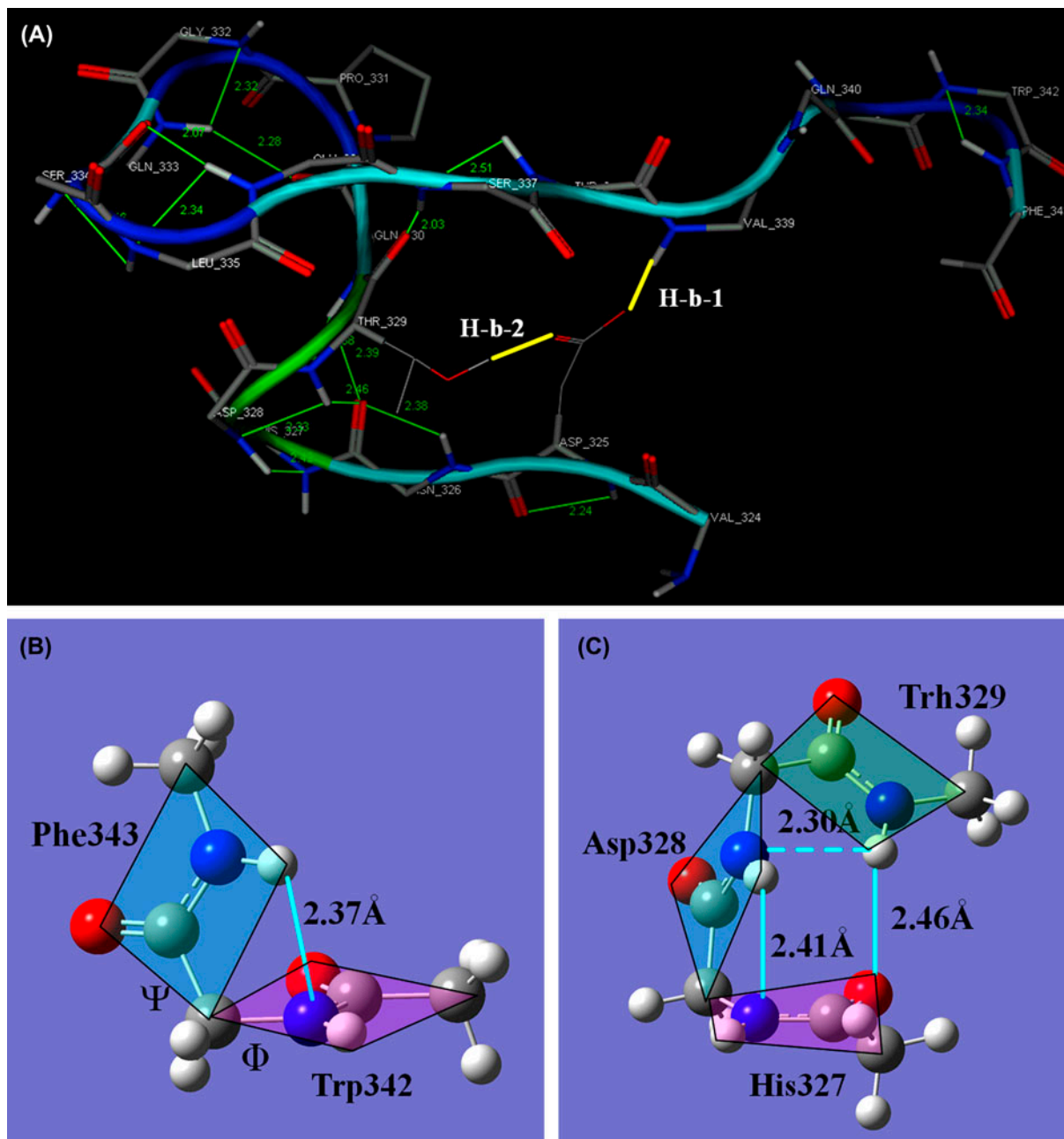
Figure 6.   The structure of a loop (Asp325–Phe343) in α-amylase (pdb code: 1BLI). (A) The Hp–π bonds in the loop (Asp325–Phe343) of α-amylase 1BLI. Among the 19 residues, there are 15 Hp–π bonds, indicated by thin green lines. The Gln330 and Gly332 are fixed by the special peptide bonds of Pro331. The residues Thr338 and Val339 are located at the straight segment. No Hp–π bond supports these two residues. The Val339 and Thr329 are supported hydrogen bonds between side chain of Asp325 and peptide backbone, indicated by thick yellow lines. All other 14 residues are fixed by at least one Hp–π bond. (B) The Hp–π bond between two connected residues Trp342 and Phe343 ($\Phi = -60°$ and $\Psi = -31°$). (C) The Hp–π bonds among three continuous residues (His327, Asp328, and Thr329). Except the two Hp–π bonds between peptide bond unit 1 and unit 2, and between unit 2 and unit 3, the peptide bond unit 3 takes a very good orientation forming a stable Hp–π bond with the peptide bond unit 1.

Hp–π bond between two connected peptide bond units, the two torsion angles Φ and Ψ have to fall in a suitable range. In Figure 6(B) the two torsion angles are $\Phi = -60°$ and $\Psi = -31°$. Figure 6(C) shows the Hp–π

bonds in three continuous residues (His327, Asp328, and Thr329). The interaction geometry of the first two peptide bond units ($\Phi = -68°$ and $\Psi = -16°$) is the same as that in Figure 6(B), and the interaction geometry

between the peptide bond unit-2 and unit-3 ($\Phi = -83°$ and $\Psi = -32°$) is just the same as the geometry between unit-1 and unit-2. In addition, the peptide bond unit-3 takes a very good orientation forming a stable Hp–π bond with the peptide bond unit-1, as shown in Figure 6(C). In this loop, a similar three-residue Hp–π structure is found between the continuous residues Ser334, Leu335, and Glu336.

The Hp–π bonds are concentrated in the curve parts of the loop. Usually, the Hp–π bonds act in short and continuously connected peptide bond units no more than 4 residues. In long loops, especially in the straight segment, the interactions between side chains and peptide backbones provide the supporting forces. In Figure 6(A) the side chain of Asp325 forms a stable hydrogen bond with the nitrogen atom in peptide backbone of Val 339, which is in the straight segment of the loop. Also a hydrogen bond is formed between the side chains of Asp325 and Thr329, indicated by thick yellow lines.

4. Discussion

In a recent paper by Kumar and Das (2012), the Hp–π interactions were theoretically and experimentally studied using resonant two photon ionization (R2PI), IR–UV, and UV–UV double resonance spectroscopic techniques, and high-level QM calculations. According to their experiments, the N–H…π hydrogen bonds and slanted T-shaped structures were observed in molecular dimer. The experimental observations could be the evidence of Hp–π bonds in molecular interactions. The interaction energies of polar hydrogen–π interactions (Hp–π) are much stronger than that of the non polar hydrogen–π interactions (H–π) (Du et al., 2010; Guan, Zhou, & Yan, 2012; Mróz, 1995). However, the roles of Hp–π interactions in protein structures have not been studied sufficiently yet.

The theoretical Φ–Ψ potential surface (Figure 5(A)) is consistent with the statistical Ramachandran Φ–Ψ plot (Lovell et al., 2003; Perczel et al., 2003; Ramachandran et al., 1963; Ting et al., 2010) to a certain extent. However, there are some subtle differences between the two plots. In Ramachandran plots, the Φ-Ψ conformation population density corresponds to the folding free energy ΔG, the higher population density, the lower folding free energy. On the other hand, the Φ–Ψ potential surfaces are the distribution of interaction energies ΔE between the two connected amino acids. The folding energies ΔE are the main contribution to the statistical folding free energies ΔG. However, there are other factors in the folding free energies, such as entropy contribution and the solvation effect. The Φ-Ψ potential surfaces cannot be exactly the same as the Ramachandran plots. The differences between two type plots may be caused by the entropy contribution and the solvation effect. The other

reason is that the Ramachandran plot is the statistical results for general amino acid types (no Gly and Pro included), and the theoretical Φ–Ψ potential surface (Figure 5) is calculated from a specific amino acid alanine.

The unique properties of Hp–π interactions in protein loops, explored in this study may have many potential applications. A potential application of Hp–π interactions is in the construction of physical effective energy function (PEEF) (Balabin, 2008; Lazaridis & Karplus, 2000) to evaluate the free energies involved in protein structures and protein folding. PEEF has been used to refine X-ray crystal structures and solution NMR structures. However, these type methods generally cannot provide accurate protein folding free energies in protein design and engineering (Balabin, 2008; Lazaridis & Karplus, 2000). One reason might be the missing of Hp–π interactions in loops. The Hp–π interactions certainly can help to improve the protein structure-free energy calculations. The other potential applications of Hp–π interactions include the construction of protein blocks (PB) (Joseph et al., 2011; Tyagi, Bornot, Offmann, & de Brevern, 2009). Actually, the protein loops are not completely 'irregular.' Brevern's group (Joseph et al., 2011; Tyagi et al., 2009) suggested that the protein structures can be classified into 16 basic construction blocks. Among them 14 PB are for loops. As shown in Figure 5(A), there are only limited Φ-Ψ conformation types (including α-helices, β-strands, and four loop types: H–π–1, H–π–2, H–π–3, and H–π–4). The basic loop types and the Hp–π interactions may provide a theoretical basis for deducing the folding free energies in the PB.

It is instructive to point out that, besides α-helix, β-strand and loop, the tight turn is also an important element in proteins. A tight turn in protein is defined as a site where a polypeptide chain reverses its overall direction and that the amino acid residues directly involved in forming the turn are no more than six (Chou, 2000; Richardson, 1981). According to the different manners of intra hydrogen bond interaction along a protein chain and the number of amino acid residue involved (Chou, 2000; Kabsch & Sander, 1983), tight turn can be classified into the following five types: (1) δ or 2-residue turn, (2) γ or 3-residue turn, (3) β or 4-residue turn, (4) α or 5-residue turn, and (5) π or 6-residue turn. See Equation (2) given in a comprehensive review (Chou, 2000) for a rigorous formulation in this regard. Therefore, protein loops are completely different with the tight turns as reflected by the following three facts: (1) the number of amino acid residues involved in forming a loop is much larger than that in forming a tight turn; (2) the conformation of a loop is much more irregular than that of a tight turn; and (3) the flexibility of a loop is much higher than that of a tight turn.

In the last two decades, many efforts have been made to predict tight turns, such as alpha-turn (Cai & Chou,

1999; Cai, Feng, & Li, 2003; Chou, 1997b; Pavone, Gaeta, Lombardi, Nastri, & Maglio, 1996), beta-turn, (Cai, Yu, & Chou, 1998; Chou, 1997a; Chou & Blinn, 1997; Shi, Hu, Li, & Liu, 2011; Zhang & Chou, 1997) and gamma-turn (Jahandideh, Hoseini, Jahandideh, Hoseini, & Disfani, 2009; Jahandideh, Sarvestani, Abdolmaleki, Jahandideh, & Barfeie, 2007). Recently, a unified model (Song et al., 2012) was proposed for simultaneously predicting all the five types of tight turns. However, all the aforementioned prediction methods in identifying tight turns and their types were developed by machine-learning approaches and statistical principle rather than energetics. Thus, a natural question might be raised as asking: Can the new molecular force also provide us more insights into the tight turns? This is indeed a very intriguing question. Since tight turns are completely different from loops, however, to address this problem, we need a large amount of calculations and a lot of in-depth analyses. We will investigate this in our future research.

5. Conclusion

Summarizing in short words, in protein backbones, when the π-planes of peptide bond units are in parallel orientation, they may form the common hydrogen bonds that are the supporting force in α-helices and β-strands, and when the π-planes are in perpendicular direction, they may form the Hp–π bonds that are the force maintaining the shape of loops. The important role of Hp–π interactions in protein folding is supported by the Ramachandran Φ–Ψ plots. The Hp–π interactions may provide useful theoretical concept and are helpful for development of computational methods evaluating the folding free energies of loops in protein structures. The basic PB and the structural limitations in protein folding may be derived and illustrated using the concept of Hp–π interactions.

Some useful concluded points can be drawn as follows. (1) Peptide bond units possess the interaction donors and acceptors of both hydrogen bond interactions and Hp–π bond interactions. The most unfavorable direction to the H–b interactions (perpendicular to π-plane) is just the most favorable direction to the Hp–π interactions. (2) In proteins, the energies of Hp–π interactions are within the range -16 to -24 kJ/mol, the bond lengths are around 2.5 Å (from polar hydrogen to π-plane). The Hp–π interaction space is larger than that of common hydrogen bond interactions. (3) The H–b interactions are the supporting force in α-helices and β-strands, and the Hp–π interactions are the force maintaining the structural shapes of protein loops. (4) The Hp–π bonds are point to π-plane interactions, acting on a larger area and possessing broader energy range. Consequently, the loops have more structural conformations,

and more flexible structures than that of the α-helices and β-strands. (5) In water solution, the interaction energies of Hp–π interactions are only little smaller than in vacuum. Similar to the common hydrogen bond interactions, the Hp–π interactions are less affected by the solvent dielectric constants.

Competing interests

We declare that there are no any competing interests.

References

Agarwal, G., Rajavel, M., Gopal, B., & Srinivasan, N. (2009). Structure-based phylogeny as a diagnostic for functional characterization of proteins with a cupin fold. *PLoS One, 4*, e5736.

Amovilli, C., Barone, V., Cammi, R., Cancès, E., Cossi, M., Mennucci, B., … Tomasi, J. (1999). Recent advances in the description of solvent effects with the polarizable continuum model. *Advances in Quantum Chemistry, 32*, 227–261.

Asturiol, D., Duran, M., & Salvador, P. (2008). Intramolecular basis set superposition error effects on the planarity of benzene and other aromatic molecules: A solution to the problem. *The Journal of Chemical Physics, 128*, 144108.

Balabin, R. M. (2008). Enthalpy difference between conformations of normal alkanes: Intramolecular basis set superposition error (BSSE) in the case of n-butane and n-hexane. *The Journal of Chemical Physics, 129*, 164101.

Berardi, M. J., Shih, W. M., Harrison, S. C., & Chou, J. J. (2011). Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature, 476*, 109–113.

Bhaduri, A., Pugalenthi, G., & Sowdhamini, R. (2004). PASS2: An automated database of protein alignments organised as structural superfamilies. *BMC Bioinformatics, 5*, 35.

Birchall, L. S., Roy, S., Jayawarna, V., Hughes, M., Irvine, E., Okorogheye, G. T., … Edwards, A. A. (2011). Exploiting CH-π interactions in supramolecular hydrogels of aromatic carbohydrate amphiphiles. *Chemical Science, 2*, 1349–1355.

Cai, Y. D., & Chou, K. C. (1999). Artificial neural network model for predicting α-turn types. *Analytical Biochemistry, 268*, 407–409.

Cai, Y. D., Feng, K. Y., & Li, Y. X. (2003). Support vector machine for predicting α-turn types. *Peptides, 24*, 629–630.

Cai, Y. D., Yu, H., & Chou, K. C. (1998). Prediction of beta-turns. *Journal of Protein Chemistry, 17*, 363–376.

Carlacci, L., & Chou, K. C. (1990a). Electrostatic interactions between loops and α-helices in four-helix bundle proteins. *Protein Engineering, 4*, 225–227.

Carlacci, L., & Chou, K. C. (1990b). Monte Carlo method applied in the search for low energy conformations of ßαßαß structures. *Biopolymers, 30*, 135–150.

Carlacci, L., & Chou, K. C. (1991). New development on energetic approach to the packing in proteins. *Journal of Computational Chemistry, 12*, 410–415.

Carlacci, L., Chou, K. C., & Maggiora, G. M. (1991). A heuristic approach to predicting the tertiary structure of bovine somatotropin. *Biochemistry, 30*, 4389–4398.

Chou, K. C. (1983). Identification of low-frequency modes in protein molecules. *Biochemical Journal, 215*, 465–469.

Chou, K. C. (1985). Low-frequency motions in protein molecules. Beta-sheet and beta-barrel. *Biophysical Journal, 48*, 289–297.

Chou, K. C. (1987). The biological functions of low-frequency vibrations (phonons). VI. A possible dynamic mechanism of allosteric transition in antibody molecules. *Biopolymers, 26*, 285–295.

Chou, K. C. (1988). Low-frequency collective motion in biomacromolecules and its biological functions. *Biophysical Chemistry, 30*, 3–48.

Chou, K. C. (1989). Low-frequency resonance and cooperativity of hemoglobin. *Trends in Biochemical Sciences, 14*, 212–213.

Chou, K. C. (1991). Reply from K.-C.Chou: The role of loops in stabilizing bundle motif protein structures. *Protein Engineering, 4*, 849–850.

Chou, K. C. (1992). Energy-optimized structure of antifreeze protein and its binding mechanism. *Journal of Molecular Biology, 223*, 509–517.

Chou, K. C. (1997a). Prediction of beta-turns in proteins. *Journal of Peptide Research, 49*, 120–144.

Chou, K. C. (1997b). Prediction and classification of α-turn types. *Biopolymers, 42*, 837–853.

Chou, K. C. (2000). Prediction of tight turns and their types in proteins. *Analytical Biochemistry, 286*, 1–16.

Chou, K. C., & Blinn, J. R. (1997). Classification and prediction of beta-turn types. *Journal of Protein Chemistry, 16*, 575–595.

Chou, K. C., & Carlacci, L. (1991). Energetic approach to the folding of α/β barrels. *Proteins: Structure, Function, and Genetics, 9*, 280–295.

Chou, K. C., Carlacci, L., & Maggiora, G. M. (1990). Conformational and geometrical properties of idealized β-barrels in proteins. *Journal of Molecular Biology, 213*, 315–326.

Chou, K. C., Carlacci, L., Maggiora, G. M., Parodi, L. A., & Schulz, M. W. (1992). An energy-based approach to packing the 7-helix bundle of bacteriorhodopsin. *Protein Science, 1*, 810–827.

Chou, K. C., & Chen, N. Y. (1977). The biological functions of low-frequency phonons. *Scientia Sinica, 20*, 447–457.

Chou, K. C., Heckel, A., Némethy, G., Rumsey, S., Carlacci, L., & Scheraga, H. A. (1990). Energetics of the structure and chain tilting of antiparallel β-barrels in proteins. *Proteins: Structure, Function, and Genetics, 8*, 14–22.

Chou, J. J., Li, S., Klee, C. B., & Bax, A. (2001). Solution structure of Ca2+-calmodulin reveals flexible hand-like properties of its domains. *Nature Structural Biology, 8*, 990–997.

Chou, K. C., Maggiora, G. M., Nemethy, G., & Scheraga, H. A. (1988). Energetics of the structure of the four-alpha-helix bundle in proteins. *Proceedings of the National Academy of Sciences, 85*, 4295–4299.

Chou, K. C., Maggiora, G. M., & Scheraga, H. A. (1992). Role of loop-helix interactions in stabilizing four-helix bundle proteins. *Proceedings of the National Academy of Sciences, 89*, 7315–7319.

Chou, K. C., Nemethy, G., Pottle, M. S., & Scheraga, H. A. (1985). Folding of the twisted .beta.-sheet in bovine pancreatic trypsin inhibitor. *Biochemistry, 24*, 7948–7953.

Chou, K. C., Nemethy, G., Pottle, M., & Scheraga, H. A. (1989). Energy of stabilization of the right-handed βαβ crossover in proteins. *Journal of Molecular Biology, 205*, 241–249.

Chou, K. C., Nemethy, G., Rumsey, S., Tuttle, R. W., & Scheraga, H. A. (1985). Interactions between an α-helix and a β-sheet. *Journal of Molecular Biology, 186*, 591–609.

Chou, K. C., Nemethy, G., Rumsey, S., Tuttle, R. W., & Scheraga, H. A. (1986). Interactions between two β-sheets energetics of β/β packing in proteins. *Journal of Molecular Biology, 188*, 641–649.

Chou, K. C., Nemethy, G., & Scheraga, H. A. (1983a). Effect of amino acid composition on the twist and the relative stability of parallel and antiparallel .beta.-sheets. *Biochemistry, 22*, 6213–6221.

Chou, K. C., Nemethy, G., & Scheraga, H. A. (1983b). Role of interchain interactions in the stabilization of the right-handed twist of β-sheets. *Journal of Molecular Biology, 168*, 389–407.

Chou, K. C., Nemethy, G., & Scheraga, H. A. (1984). Energetic approach to the packing of .alpha.-helixes. 2. General treatment of nonequivalent and nonregular helixes. *Journal of the American Chemical Society, 106*, 3161–3170.

Chou, K. C., Nemethy, G., & Scheraga, H. A. (1990). Energetics of interactions of regular structural elements in proteins. *Accounts of Chemical Research, 23*, 134–141.

Chou, K. C., Pottle, M., Nemethy, G., Ueda, Y., & Scheraga, H. A. (1982). Structure of β-sheets. *Journal of Molecular Biology, 162*, 89–112.

Chou, K. C., & Scheraga, H. A. (1982). Origin of the right-handed twist of beta-sheets of poly(LVal) chains. *Proceedings of the National Academy of Sciences, 79*, 7047–7051.

Chou, K. C., & Zheng, C. (1992). Strong electrostatic loop-helix interactions in bundle motif protein structures. *Biophysical Journal, 63*, 682–688.

Cossi, M., & Barone, V. (1998). Analytical second derivatives of the free energy in solution by polarizable continuum models. *The Journal of Chemical Physics, 109*, 6246–6254.

Cramer, C. J., & Bickelhaupt, F. (2003). Essentials of computational chemistry. *Angewandte Chemie-International Edition in English, 42*, 381–381.

Du, Q. S., Long, S. Y., Meng, J. Z., & Huang, R. B. (2012b). Empirical formulation and parameterization of cation–π interactions for protein modeling. *Journal of Computational Chemistry, 33*, 153–162.

Du, Q.-S., Meng, J.-Z., Liao, S.-M., & Huang, R.-B. (2012a). Energies and physicochemical properties of cation–π interactions in biological structures. *Journal of Molecular Graphics and Modelling, 34*, 38–45.

Du, Q. S., Wang, Q. Y., Du, L. Q., Chen, D., & Huang, R. B. (2013). Theoretical study on the polar hydrogen-π (Hp-π) interactions between protein side chains. *Chemistry Central Journal, 7*, 92.

Du, Q.-S., Wang, S.-Q., & Huang, R.-B. (2010). Computational 3D structures of drug-targeting proteins in the 2009-H1N1 influenza A virus. *Chemical Physics Letters, 485*, 191–195.

Du, Q.-S., Wei, H., & Huang, R.-B. (2011). Progress in structure-based drug design against influenza A virus. *Expert Opinion on Drug Discovery, 6*, 619–631.

Foresman, J. B., Keith, T. A., Wiberg, K. B., Snoonian, J., & Frisch, M. J. (1996). Solvent Effects. 5. Influence of cavity shape, truncation of electrostatics, and electron correlation on ab initio reaction field calculations. *The Journal of Physical Chemistry, 100*, 16098–16104.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., … Fox, D. J. (2010). *Gaussian*. Wallingford: Gaussian.

Gerritsen, M., Chou, K. C., Nemethy, G., & Scheraga, H. A. (1985). Energetics of multihelix interactions in protein folding: Application to myoglobin. *Biopolymers, 24*, 1271–1291.

Gordon, D. B., Marshall, S. A., & Mayot, S. L. (1999). Energy functions for protein design. *Current Opinion in Structural Biology, 9*, 509–513.

Guan, P. P., Zhou, W., & Yan, H. (2012). The relationship between geometric patterns of hydrogen bonds and periodic dinucleotides in nucleosome structures. *Journal of Theoretical Biology, 313*, 136–141.

Harrison, A., Pearl, F., Mott, R., Thornton, J., & Orengo, C. (2002). Quantifying the similarities within fold space. *Journal of Molecular Biology, 323*, 909–926.

Hughes, R. M., & Waters, M. L. (2006). Effects of lysine acetylation in a β-hairpin peptide: comparison of an amide-π and a cation-π interaction. *Journal of the American Chemical Society, 128*, 13586–13591.

Jahandideh, S., Hoseini, S., Jahandideh, M., Hoseini, A., & Disfani, F. M. (2009). γ-turn types prediction in proteins using the two-stage hybrid neural discriminant model. *Journal of Theoretical Biology, 259*, 517–522.

Jahandideh, S., Sarvestani, A. S., Abdolmaleki, P., Jahandideh, M., & Barfeie, M. (2007). γ-Turn types prediction in proteins using the support vector machines. *Journal of Theoretical Biology, 249*, 785–790.

Joseph, A. P., Srinivasan, N., & de Brevern, A. G. (2011). Improvement of protein structure comparison using a structural alphabet. *Biochimie, 93*, 1434–1445.

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers, 22*, 2577–2637.

Krissinel, E., & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica, Section D: Biological Crystallography, 60*, 2256–2268.

Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., & Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science, 302*, 1364–1368.

Kumar, S., & Das, A. (2012). Effect of acceptor heteroatoms on π-hydrogen bonding interactions: A study of indole···thiophene heterodimer in a supersonic jet. *The Journal of Chemical Physics, 137*, 094309.

Lazaridis, T., & Karplus, M. (2000). Effective energy functions for protein structure prediction. *Current Opinion in Structural Biology, 10*, 139–145.

Lee, T. J., & Rice, J. E. (1988). An efficient closed-shell singles and doubles coupled-cluster method. *Chemical Physics Letters, 150*, 406–415.

Li, X.-Q., & Fan, P. (2010). A duplex DNA model with regular inter-base-pair hydrogen bonds. *Journal of Theoretical Biology, 266*, 374–379.

Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., … Richardson, D. C. (2003). Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics, 50*, 437–450.

Lu, G. (2000). TOP: A new method for protein structure comparisons and similarity searches. *Journal of Applied Crystallography, 33*, 176–183.

Machius, M., Declerck, N., Huber, R., & Wiegand, G. (1998). Activation of Bacillus licheniformis α-amylase through a disorder→order transition of the substrate-binding site mediated by a calcium–sodium–calcium metal triad. *Structure, 6*, 281–292.

Marti-Renom, M. A., Ilyin, V. A., & Sali, A. (2001). DBAli: A database of protein structure alignments. *Bioinformatics, 17*, 746–747.

Mayer, I., & Valiron, P. (1998). Second order Møller–Plesset perturbation theory without basis set superposition error. *The Journal of Chemical Physics, 109*, 3360–3373.

Miertuš, S., Scrocco, E., & Tomasi, J. (1981). Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chemical Physics, 55*, 117–129.

Mizuguchi, K., Deane, C. M., Blundell, T. L., & Overington, J. P. (1998). HOMSTRAD: A database of protein structure alignments for homologous families. *Protein Science, 7*, 2469–2471.

Mróz, I. (1995). Theoretical consideration of the relation between hydrogen bond types and possible secondary structures of a polypeptide chain. *Journal of Theoretical Biology, 175*, 245–255.

OuYang, B., Xie, S., Berardi, M. J., Zhao, X. M., Dev, J., Yu, W., Sun, B., & Chou, J. J. (2013). Unusual architecture of the p7 channel from hepatitis C virus. *Nature, 498*, 521–525.

Palese, P., Tobita, K., Ueda, M., & Compans, R. W. (1974). Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology, 61*, 397–410.

Pauling, L., & Corey, R. B. (1951). The pleated sheet, a new layer configuration of polypeptide chains. *Proceedings of the National Academy of Sciences, 37*, 251–256.

Pauling, L., Corey, R. B., & Branson, H. R. (1951). The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences, 37*, 205–211.

Pavone, V., Gaeta, G., Lombardi, A., Nastri, F., & Maglio, O. (1996). Discovering protein secondary structures: Classification and description of isolated α-turns. *Biopolymers, 38*, 705–721.

Perczel, A., Farkas, Ö., Jákli, I., Topol, I. A., & Csizmadia, I. G. (2003). Peptide models. XXXIII. Extrapolation of low-level Hartree-Fock data of peptide conformation to large basis set SCF, MP2, DFT, and CCSD(T) results. The Ramachandran surface of alanine dipeptide computed at various levels of theory. *Journal of Computational Chemistry, 24*, 1026–1042.

Pidaparti, R. M., Svintradze, D. V., Shan, Y., & Yokota, H. (2009). Optimization of hydrogen bonds for combined DNA/collagen complex. *Journal of Theoretical Biology, 256*, 149–156.

Purvis, G. D., III, & Bartlett, R. J. (1982). A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *The Journal of Chemical Physics, 76*, 1910–1918.

Ramachandran, G., Ramakrishnan, C. t., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology, 7*, 95–99.

Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Advances of Protein Chemistry, 34*, 167–339.

Scheraga, H. A., Chou, K. C., & Nemethy, G. (1982). Interactions between the fundamental structures of polypeptide chains. In R. Srinivasan & R. H. Sarma (Eds.), *Conformation in Biology* (pp. 1–10). Gilderland, NY: Adenine Press.

Schnell, J. R., & Chou, J. J. (2008). Structure and mechanism of the M2 proton channel of influenza A virus. *Nature, 451*, 591–595.

Scuseria, G. E., Janssen, C. L., & Schaefer, H. F. (1988). An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations. *The Journal of Chemical Physics, 89*, 7382–7387.

Scuseria, G. E., & Schaefer, H. F., III (1989). Is coupled cluster singles and doubles (CCSD) more computationally intensive than quadratic configuration interaction (QCISD)? *The Journal of Chemical Physics, 90*, 3700–3703.

Sherrill, C. D., & Schaefer, H. F., III. (1999). The configuration interaction method: Advances in highly correlated approaches. *Advances in Quantum Chemistry, 34*, 143–269.

Shi, X., Hu, X., Li, S., & Liu, X. (2011). Prediction of β-turn types in protein by using composite vector. *Journal of Theoretical Biology, 286*, 24–30.

Sindzingre, P., Lhuillier, C., & Fouet, J. (2001). *Proceedings of the 11th International Conference on Recent Progress in Many-Body Theories*. Manchester, UK.

Sirois, S., Sing, T., & Chou, K. C. (2005). HIV-1 gp120 V3 loop for structure-based drug design. *Current Protein and Peptide Science, 6*, 413–422.

Sirois, S., Touaibia, M., Chou, K. C., & Roy, R. (2007). Glycosylation of HIV-1 gp120 V3 loop: Towards the rational design of a synthetic carbohydrate vaccine. *Current Medicinal Chemistry, 14*, 3232–3242.

Song, Q., Li, T., Cong, P., Sun, J., Li, D., & Tang, S. (2012). Predicting Turns in Proteins with a Unified Model. *PLoS One, 7*, e48389.

Sujatha, S., Balaji, S., & Srinivasan, N. (2001). PALI: a database of alignments and phylogeny of homologous protein structures. *Bioinformatics, 17*, 375–376.

Thompson, T. B., Chou, K. C., & Zheng, C. (1995). Analysis of the loop-helix interaction in bundle motif protein structures *Journal of Protein Chemistry, 14*, 559–566.

Ting, D., Wang, G., Shapovalov, M., Mitra, R., Jordan, M. I., & Dunbrack, R. L., Jr (2010). Neighbor-dependent ramachandran probability distributions of amino acids developed from a hierarchical dirichlet process model. *PLoS Computational Biology, 6*, e1000763.

Tyagi, M., Bornot, A., Offmann, B., & de Brevern, A. G. (2009). Protein short loop prediction in terms of a structural alphabet. *Computational Biology and Chemistry, 33*, 329–333.

von Itzstein, M. (2007). The war against influenza: discovery and development of sialidase inhibitors. *Nature Reviews Drug Discovery, 6*, 967–974.

Wang, J. F., & Chou, K. C. (2009). Insight into the molecular switch mechanism of human Rab5a from molecular dynamics simulations. *Biochemical and Biophysical Research Communications, 390*, 608–612.

Wang, J. F., & Chou, K. C. (2010). Insights from studying the mutation-induced allostery in the M2 proton channel by molecular dynamics. *Protein Engineering, Design & Selection, 23*, 663–666.

Wang, S.-Q., Du, Q.-S., Huang, R.-B., & Zhang, D.-W. (2009). Insights from investigating the interaction of oseltamivir (Tamiflu) with neuraminidase of the 2009 H1N1 swine flu virus. *Biochemical and Biophysical Research Communications, 386*, 432–436.

Wang, J. F., Gong, K., Wei, D. Q., & Li, Y. X. (2009). Molecular dynamics studies on the interactions of PTP1B with inhibitors: from the first phosphate-binding site to the second one. *Protein Engineering, Design & Selection, 22*, 349–355.

Wen, Z. Z., Wang, Y. H., Yang, B., & Xie, M. Q. (2010). Roles of L5-7 loop in the structure and chaperone function of SsHSP14.1. *Protein and Peptide Letters, 18*, 275–281.

Wen, Z. Z., Wang, Y. H., Yang, B., & Xie, M. Q. (2011). Roles of L5-7 loop in the structure and chaperone function of SsHSP14.1. *Protein and Peptide Letters, 18*, 275–281.

Wi, S., & Spano, J. (2011). Site-specific ϕ- and ψ-torsion angle determination in a uniformly/extensively 13C- and 15N-labeled peptide. *Journal of Magnetic Resonance, 212*, 431–439.

Zhang, C. T., & Chou, K. C. (1997). Prediction of β-turns in proteins by 1-4 and 2-3 correlation model. *Biopolymers, 41*, 673–702.