International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 3, Issue 3, 2011

Research Article

COMPUTATIONAL INVESTIGATION OF N-H...Π INTERACTIONS IN THE STRUCTURAL STABILITY OF TRANSMEMBRANE PROTEINS

SOUMYA RANJAN MOHAPATRA, K. RAMANATHAN, V. SHANTHI, SHASHANK SRIVASTAVA AND RAO SETHUMADHAVAN*

School of Bio Sciences and Technology,Vellore Institute of Technology, Vellore 632014, Tamil Nadu, India
Email: rsethumadhavan@vit.ac.in (R. Sethumadhavan)

Received: 09 March 2011, Revised and Accepted: 08 April 2011

ABSTRACT

From the dawn of knowledge about chemical structures in the early 1800's, through the early 1900s revealing the atomic structure, elucidation of the polymer structure in the 1920s, perception of the first protein structure about 53 years ago, to the resolution of the gene structure more recently in the 1970s, structural knowledge has always led the way for major revolutions in the chemical, physical and biological sciences over the past two centuries. The new millenium has brought with it computational proteomics, which promises a high impact revolution in our understanding of the proteome as we know it. The overall stability of the folded structure of peptides and proteins depends on the various interactions that its amino acid residues take part in. Of the various types of stabilizing interactions, in the current study we have analyzed N-H....π interactions in a set of 100 different transmembrane proteins. The present study details the results of N-H...π interactions in relation to other factors like secondary structural elements, conservation score and stabilization centers in transmembrane proteins. The obtained results suggests that the N-H....π interactions contribute significantly to the overall stability of transmembrane proteins.

Keywords: Transmembrane proteins; N-H....π interactions; Secondary structure; Interactions range; Conservation; Stabilizing centers.

INTRODUCTION

A folded protein is stabilized by a number of noncovalent interactions such as hydrophobic interactions, hydrogen bonds, salt bridges, and cation-aromatic interactions [1,2]. Such interactions have been subjected to extensive structural analysis to elucidate the different geometrical criteria required to identify them [3]. In addition, weak interactions also play a major role towards stabilizing a fully processed mature protein structure. Among the several weak intermolecular interactions contributing to the stability of various chemical and biological entities, the N-H...π interaction is one of the most widely known [4-5]. Nonconventional, weak hydrogen bonds are ubiquitous in proteins, protein-ligand complexes, protein-protein complexes, nucleic acids etc. enhancing their overall stability [6]. In particular, transmembrane protein structures contain a large number of weak interactions, which is presumably due to a relatively large fraction of small amino acids such as glycine, alanine and serine, and the interfold distances are thus shorter; facilitating the weak interactions [6]. Positively charged or δ(+) amino groups of lysine, arginine, asparagine, glutamine and histidine are preferentially located within 6Å of the ring centroids of phenylalanine, tyrosine and tryptophan, where they make van der Waals contact with the δ(-) π-electrons and avoid the δ (+) ring edge. This geometric pattern is recognized as N-H...π interaction [4].

These non-covalent interactions involving the π ring system as hydrogen bond acceptor were first described by Wulf et al.[7] through spectroscopic analysis of small molecules and subsequently in peptides by McPhail and Sim [8], but their importance was not immediately appreciated. Much later, the N-H...π interactions in proteins attracted the greater attention, following the observation of the stabilizing effect of such interactions in beta sheets [9], helix termini [10], helices containing proline residues [11], packing of transmembrane helices [12], collagen [13], and DNA [14]. Since then, a greater number of X-H...π interactions in proteins have been found in a single case studies, involved in wide variety of functions such as secondary structure stabilization [15,16], DNA recognition [17], enzymatic action [18], ligand and drug recognition [19].

Despite the fairly large number of theoretical and experimental investigations of cation-π and C-H...π hydrogen bonds, there are relatively a few studies on systems exhibiting the N-H...π interactions. This is because, the relatively small intermolecular interaction energies of these systems make it very difficult to characterize them experimentally [20,21]. Nevertheless its weak nature makes it one of the most poorly understood interactions. Following the identification of an amino/aromatic hydrogen bond in SH2 domain/peptide binding by Walksman [22], there has been a resurgence of interest in such interactions in proteins.

Theoritical ab initio calculations have also been performed which have shown that the energy of these noncovalent interactions is less than the energy of a conventional hydrogen bond. However, since these interactions can occur more frequently than regular hydrogen bonds, they may well contribute to the protein's stability to the same extent as standard hydrogen bonds [23]. Hence, in this work, an attempt has been made to collect the information concerning N-H...π interactions on the structural stability of transmembrane proteins. In addition, we have systematically studied the role of N-H...π interactions in relation to other factors like amino acid preference, secondary structural elements, interaction range, conservation analysis and stabilization centers.

Transmembrane proteins are, in many respects, easier to investigate computationally than experimentally, due to the uniformity of their structure and interactions i.e. consisting predominantely of nearly parallel helices packed together on one hand and presenting the challenges of solubility on the other; considering this we have undertaken an computational approach in this study to determine the contribution of N-H...π interactions towards the overall stability of transmembrane proteins. The frequency and extent of conservation in the amino acids involved in the presented interactions unambiguously shows that the N-H...π interactions cannot and must not be neglected. We postulate that the incorporation of the entirety of this N-H...π interactions could provide new perspectives and possibly new answers for the structural biologist.

MATERIALS AND METHODS

Data set

A database of membrane proteins was derived from the information about their three-dimensional structures. Membrane proteins are of two kinds, one that spans the cytoplasmic membrane with α-helices (TMH) and the other that consists of β-strands in the outer membranes (TMS). We have selected a set of 100 (50 TMH and 50 TMS) trasmembrane proteins for our analysis. The co-ordinates of the proteins have been taken from the PDB [24]. The PDB codes of the TMH and TMS used for the analysis are shown in Table 1.

N-H....π interactions

N-H...π interactions are calculated using the program available for this purpose namely HBAT [25]. The positions and geometry of donor

and acceptor atom with their default parameters are shown in Fig. 1. The donor group is represented as N-H and the acceptor is the π system. The distances are usually measured from the centroid (M) i.e., centre of the π ring. P1 and P2 are distances from N and H, respectively to M. P3 is the angle between vectors N-H and H-M while P4 is the angle between the NM and MN. Here N is normal to the centre of the π ring. The geometry is adapted from earlier work of Babu [3]. The N-H…π interaction types are represented by a two-letter code in which the first letter indicates the donor atom and the second the acceptor: S and S5 represent the side-chain atom and side-chain atom in the five-membered aromatic ring, respectively. We classified the N-H…π interactions into two types namely, side-chain to side-chain N-H….π interactions (SS-N-H…π) and side-chain to side-chain five member aromatic ring N-H….π interactions (SS5-N-H…π).

Secondary structure analysis

Secondary structure is one of the major criteria to understand the structure and function of proteins. Hence a systematic analysis of each N-H….π interaction forming residue was performed based on their location in different secondary structures of transmembrane proteins. We obtained the information about secondary structures from PROSS program which is available at http://roselab.jhu.edu/utils/runpross.html.

PROSS uses backbone dihedral angles to define secondary structures. Torsion angle space (Ramachandran map) is divided into a Phi/Psi grid with the grid squares referred to as mesostates. Any protein backbone conformation can be approximated by its linear sequence of mesostate identifiers, and regular expressions of mesostate sequences can be used to define α–helices, β-strands, turns and coil conformations of proteins.

Classification by residue-residue contacts

The N-H….π interacting residues coming within a sphere of 8Å was computed as described earlier [26.29]. For a given residue, the comparison of the surrounding residue is analyzed in terms of the location at the sequence level. The residues that are within a distance of two residues are considered to contribute to short-range interactions, whereas those within a distance of ±3 or ±4 residues contribute to medium-range and those more than four residues away contribute to long-range interactions [28]. This classification enables us to evaluate the contribution of short-range, medium-range and long-range contacts in the formation of N-H…π interactions.

Evolutionary Pattern Analysis

We have used the conservation score as a parameter to study the evolutionary relationship for the interacting residues. The conservation score of N-H…π interacting amino acid residues in each protein using the ConSurf program [30]. This program computes the conservation based on the comparison of the sequence of a PDB chain with the proteins deposited in Swiss-Prot [31] and finds the ones homologous to the PDB sequence. The number of PSI-BLAST iterations and the *E*-value cutoff used in all similarity searches were 1 and 0.001, respectively. All the sequences that are evolutionarily related with each one of the proteins in the data set were used in the subsequent multiple alignments. Based on these protein sequence alignments, the number of highly conserved regions was found out among the donor and acceptor residues separately.

Stabilizing centers

Stabilization centers are clusters of residues that are involved in medium or long-range interactions [32]. Residue clusters are identified in protein contact maps where an accumulation of long range interactions is observed. The residues in these cores are called stabilization center (SC) residues, referring to their suspected role in 3D structure stabilization, and are identified as follows. The sequence environment of each residue pair involved in a long range interaction is analyzed. For each such residue pair we locate two additional pairs, one in the N-terminal flanking tetra peptide and one in the C-terminal tetra peptide of the original interacting residue pair making the most long range interactions with each other. If the number of interactions of these two triplets, the central interacting residues plus the two additional ones, one on each flanking side is equal to or greater than seven of the possible nine contacts, then the two central residues are accepted as members of an SC. The stabilization centers for the N-H…π interacting amino acid residues were computed using the SCide program [33] for computing the stabilization centers. If a residue is involved in a stabilization center, its SC value becomes 1; otherwise 0.

RESULTS

N-H…π interactions

There was a total of 175 N-H…π interactions in the set of 50 transmembrane helices and a total of 377 N-H…π interactions in the set of 50 transmembrane strands. he most prominent representatives are the interactions between aromatic N–H donor groups and aromatic π acceptors (ie, SS-N–H…π interactions). Though N-H….π interaction has been reported with His acting as an acceptor [34], the frequency of occurrence of such bonds is low owing to the unsuitability of imidazole ring in this role when charged (His may accept such an interaction only in neutral form). In order to identify the percentage contribution by an amino acid to the stability, the ratio between the numbers of interactions involving a particular amino acid to the total number of interactions involving all the amino acids was calculated, and was denoted as S.

$$S = \frac{\text{Interactions involving a particular amino acid}}{\text{Total number of interactions}} \times 100$$

The values of S obtained for all the amino acids in both transmembrane helices and strands were plotted in Fig. 2. The percentage ratio calculated shows that Arg make the maximum contribution to N-H…π interactions in both the datasets (81 interactions in a total of 175 interactions in helices and interactions in a total of 377 interactions in strands). It might be due to the fact that the side chain of arginine is larger and less well water-solvated than that of other amino acid residues, it likely benefits from better van der Waals interactions with the aromatic ring. In addition, as suggested by Thornton and colleagues [35], the side chain of Arg may still donate several hydrogen bonds while simultaneously binding to an aromatic ring (if it is stacked). Amongst the aromatic residues, Tyr is the most common amino acid involved in such interactions (53 interactions in a total of 175 interactions in helices and 182 interactions in a total of 377 interactions in strands). Hence, Arg and Tyr residues may be quite important for the stability of transmembrane proteins.

Secondary structure preferences

The propensity of the amino acid residues to favor a particular conformation has been well documented. Such conformational preference is not only dependent on the amino acid but also dependent on the local amino acid sequence. We analyzed the secondary structure preference of each residue type i.e. donor and acceptor residues, which participated in the N-H…π interactions. The secondary structure preference of each of the residue types involved in all the N-H….π interactions for both the datasets were obtained using PROSS program. It is interesting to note that as expected, donor and acceptor residues preferred to be in helix in the first dataset containing TMH and in the second dataset containing TMS, both residue types preferred to be in β-sheet configuration.

Sequential separation

The contribution of N-H….π interactions in transmembrane proteins could define either the local or the global stability of the proteins. Therefore, the need to evaluate the contribution of inter-residual N-H…π interactions arises. The sequential distance between residues that contributed to N-H…π interactions were calculated and results were depicted in Fig. 3. It reveals that 24.03%, 30.23% and 45.74% of the N-H…π interactions were found to be long-range, medium-range and short-range interactions respectively in transmembrane helices (TMH) while the individual contibution of the aforesaid three types of interaction ranges in transmembrane strands (TMS) runs to be 42.7%, 5.3% and 52.0% respectively.

Relationship between Conservation Score and N-H....π interaction

We used the ConSurf program to compute the conservation score of amino acid residues involved in N-H...π interactions in the transmembrane protein datasets, and the results were shown in Fig. 4. 38% of the donor amino acid residues (20.83% from TMH and 17.79% from TMS) and 31% of the acceptor residues (16.66% from TMH and 14.4% from TMS) were highly conserved regions of the protein sequences.

Stabilizing centers

We used the SCide program for computing the stabilization centers involved in N-H...π interactions in the transmembrane proteins data set, and the results were shown in Fig. 5. In TMS, 28.9% of the amino acid residues that contributed donor atoms in N-H...π interactions had one or more stabilization centers and 11.4% of TMH donor residues had one or more stabilization centers in addition to their contribution to N-H...π interactions. In case of amino acid residues that contributed acceptor atoms in N-H...π interactions, 11.4% of the TMH amino acid residues had one or more stabilization centers, while 13.7% of the acceptor amino acid residues from TMS had a high stabilizing effect on the protein structure through their participation in stabilization centers.

DISCUSSION

We have investigated the influence of N-H...π interactions on the structural stability of transmembrane proteins. There was a total of 552 N-H...π interactions in the set of transmembrane proteins. We find that Arg residue plays an important role in forming such interactions. The most prominent representatives are the interactions between aromatic N–H donor groups and aromatic π acceptors (ie, SS-N–H...π interactions). Though N-H....π interaction has been reported with His acting as an acceptor [41], the frequency of occurrence of such bonds is low owing to the unsuitability of imidazole ring in this role when charged. In the first dataset, TMH, most of the residues involved in N-H...π interactions prefer the secondary structure of alpha helical segments, thus justifying their transmembrane protein type i.e. α-helices. This indicates that either direct neighbors along the sequence or close neighbors in helices or sometimes coils preferably display this kind of interaction. Thus, the transmembrane α-helix proteins are, therefore, confronted with a very large number of helices in their three dimensional arrangements. In the second dataset containing the transmembrane β-strands, most of the involved residues in N-H...π interactions as expected prefer the secondary structure of β-sheets. The N-H...π interactions are formed mainly by long range contacts as evident from the inter-residual distance calculations in these interactions. From the conservation score of each amino acid residues, we were able to infer that more than 30% of the interacting residues might be highly conserved in transmembrane proteins. The conservation of amino acid residues with π-systems in some cases may be linked to their involvement in N-H...π interactions and to the stability or the function of the protein. Furthermore, significant percentage of both donor and acceptor residues in N–H...π interactions had one or more stabilizing centers in transmembrane proteins.

This kind of statistical reports unequivocally shows that the weaker interactions cannot and must not be neglected. This interaction, which is about half as strong as a normal hydrogen bond, contributes approximately 3 kcal/mol of stabilizing energy and is expected to play a significant role in molecular associations. Albeit weak, but cumulatively can make a quantitatively greater energetic contribution to folding and stability. All these show that N-H...π interactions are typically an integral part of hydrogen bonding in proteins. The consideration of these important interactions might enhance the usefulness of protein stabilities, interaction energies and folding energies calculations in general and further our understanding of protein structures and their functions.

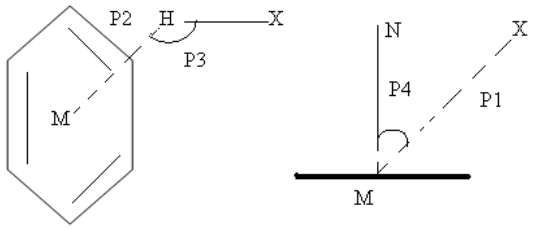

Fig. 1: Parameters for X-H.... π interaction (X=N): P1 ≤ 5.00 Å; P2 ≤ 4.50 Å; P3 ≥ 120°; P4≤30°

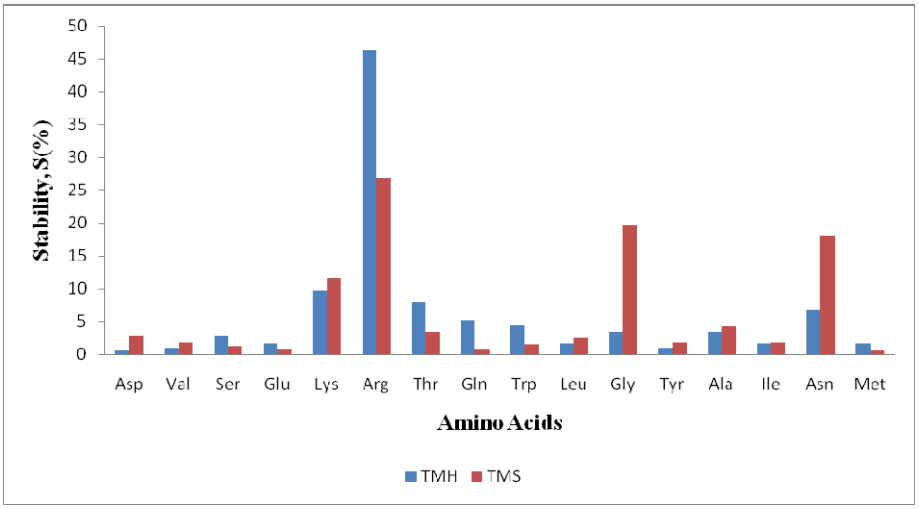

Fig. 2: Amino acids contribution to the stability of transmembrane **proteins**

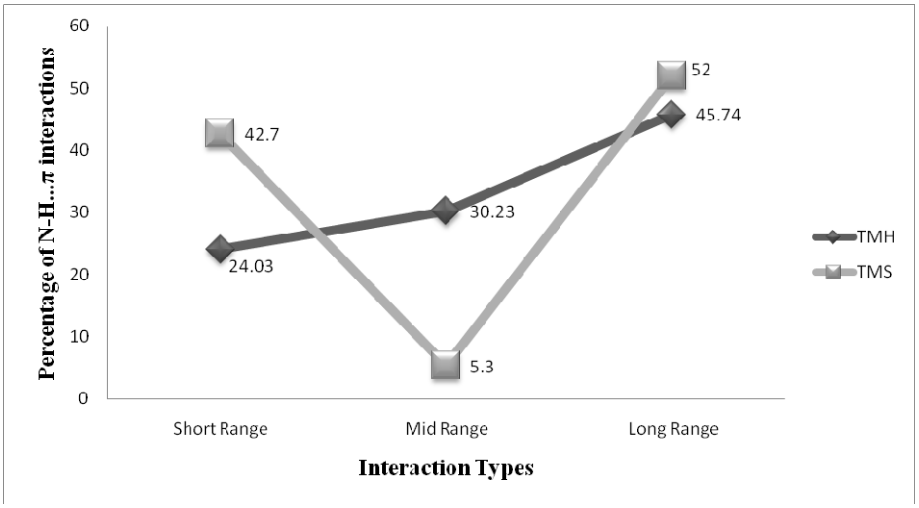

Fig. 3: N-H...π interactions range in transmembrane **proteins**

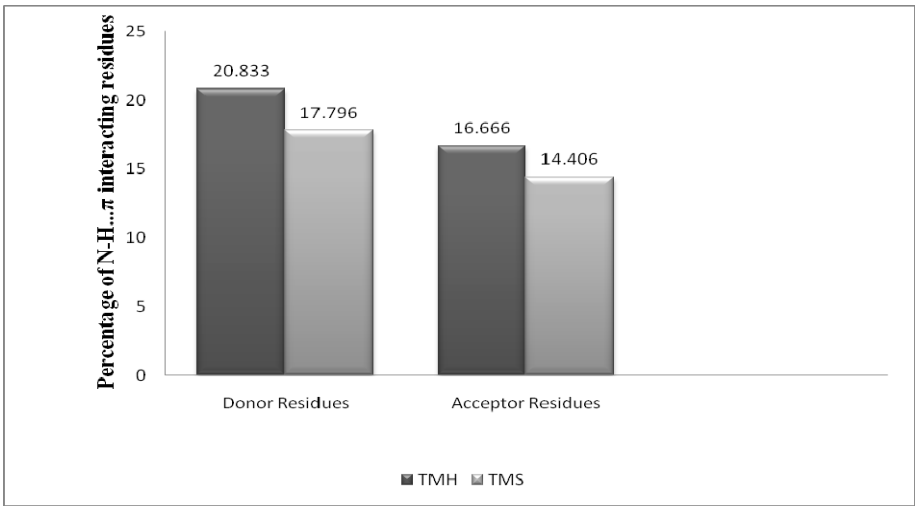

Fig. 4: Conservation score for interacting residues

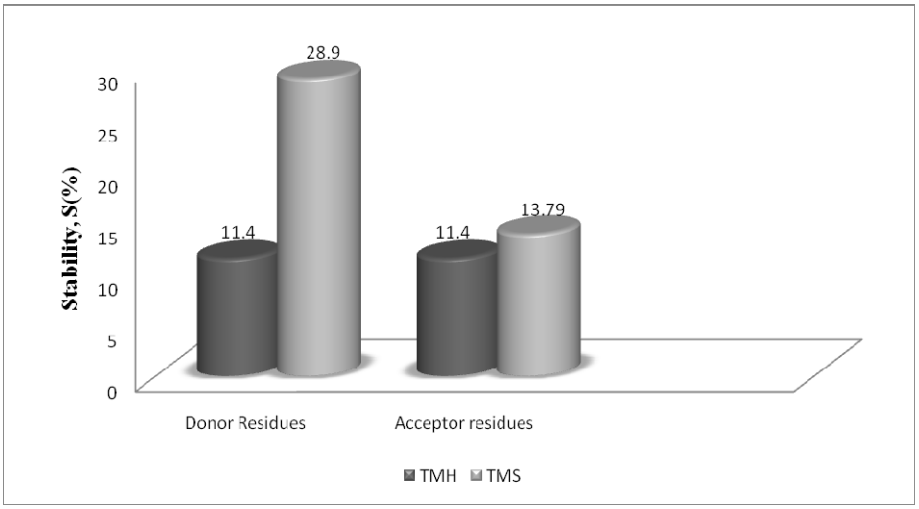

Fig. 5: Stabilization centers in transmembrane proteins

Table 1: Data Set of Transmembrane Proteins

Transmembrane Proteins			
Transmembrane Helices (TMH)		**Transmembrane Strands (TMS)**	
1A91	1QKP	1BH3	1BT9
1B9U	1XP5	1FIL	1FW2
1C3W	2EAT	1HXT	1ILD
1FBK	2RDD	1KMP	1MM4
1KY0	3GWV	1ORM	1P4T
1ORQ	1ATA	1QFF	1QJ8
1VJM	1BZK	1UYN	1YC9
3MRA	1E0P	2HDF	2IAH
2NOP	1IIJ	2JQY	2K4T
2Z5X	1MGY	2ZFG	3CSP
1A11	1SKH	1BXW	1FCP
1B11	1XRD	1G90	1GFM
1C8R	2KAM	1IMO	1K24
1FJP	2UUH	1MPF	1NQE
1L0M	3HGC	1PHO	1PRN
1PXS	1B0K	1QKC	1THQ
1VRY	1C0V	2FGQ	2GE4
2AU1	1F42	2JK4	2JLN
2OA0	1KG9	2OMF	2POR
2ZUP	1NEN	2EMN	3FHH
1APA	1SOR	1FEP	1TLW
1BTQ	1YMG	1H6S	2GUF
1C8S	2NEO	1KMO	2JMM
1HZK	2W2E	1OPF	2WJQ
1LIJ	3JYC	1QD5	3FYX

ACKNOWLEDGMENT

The authors thank the management of Vellore Institute of Technology, for providing the facilities to carry out this work.

REFERENCES

1. Gallivan JP, Dougherty DA. Cation-π Interactions in Structural Biology. Proceedings of the National Academy of Sciences USA 1999; 96:9459–9464.
2. Dill KA. Dominant forces in protein folding. Biochemistry 1990; 29(31):7133–7155.
3. Babu MM. NCI: A program to identify non-canonical interactions in protein structures. Nucleic Acids Research 2003; 31:3345–3348.
4. Burley SK, Petsko GA. Amino-aromatic interactions in proteins. FEBS Letters 1986; 203:139-143.
5. Burley SK, Petsko GA. Weakly polar interactions in proteins. Advances in Protein Chemistry 1988; 39:125-189.
6. Park H, Yoon J, Seok C. Strength of C-H...O==C hydrogen bonds in transmembrane proteins J Phys Chem 2008; 112:1041-1048.
7. Wulf OR, Liddel U, Hendricks SB. The effect of ortho substitution on the absorption of the OH group of phenol in the infrared. J Amer Chem Socie 1936; 56:2287-2293.
8. McPhail AT, Sim GA. Hydroxyl-benzene hydrogen bonding. An X-ray study. Chemical Communications 1965; 00,124-125.
9. Fabiola GF, Krishnaswamy S, Nagarajan V, Pattabhi V. C-H...O hydrogen bonds in beta sheets. Acta Crystallog Sect 1997; 53:316-320.
10. Babu MM, Singh SK, Balaram P. A C-H...O hydrogen bond stabilized polypeptide chain reversal motif at the C terminus of helices in proteins. J Mol Biol 2002; 322:871-880.
11. Chakrabarti P, Chakrabarti S. The C-H....O hydrogen bond involving proline residues in alpha-helices. J Mol Biol 1998; 284:867-873.
12. Senes A, Ubarretxena-Belandia I, Engelman DM. The C-H...o hydrogen bond: A determinant of stability and specificity in the transmembrane helix interactions. Proc Natl Acad Sci USA 2001; 98: 9056-9061.
13. Bella j, Berman MH. Crystallographic evidence for C-H...O hydrogen bonds in a collagen triple helix. J Mol Biol 1996; 264:734-742.
14. Ghosh A, Bansal M. C-H...o hydrogen bonds in the minor groove of A-tracts in DNA double helices. J Mol Biol 1999; 294:1149-1158.
15. Armstrong KM, Fairman R, Baldwin RL. The (i, i + 4) Phe-His interaction studied in an alanine-based alpha-helix. Journal of Molecular Biology 1993; 230:284-291
16. Burley SK, Petsko GA. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science 1985; 229:23–28.
17. Parkinson G, Gunasekara A, Vijtechovsky J, Zhang X, Kunkel TA, Berman H, Ebright RH. Aromatic hydrogen bond in sequence specific protein NA recognition. Nature Structural and Molecular Biology 1996; 3:837-841.
18. Liu S, Ji X, Gilliland GL, Stevens WJ, Armstrong RN. Second sphere electrostatic effects in the active site of glutathione S-transferase. Observation of an on-face hydrogen bond between the side-chains of threonine 13 and the π cloud of zyrosine 6 and its influence on catalysis. Journal of the American Chemical Society 1993; 115:7910-7911.
19. Kryger G, Silman I, Sussman JL. Structure of acetylcholinesterase complexed with E2020 (Aricept®): implications for the designs of new anti-Alzheimer drugs. Structure 1999; 7:297-307.
20. Kim KS, Tarakeshwar P, Lee JY. Molecular clusters of π-systems: theoretical studies of structures, spectra, and origin of interaction energies. Chemical Reviews 2000; 100:4145-4185.
21. Meyer EA, Castellano RK, Diederich F. Interactions with Aromatic Rings in Chemical and Biological Recognition. Angewandte Chemie International Edition in English 2003; 42:1210–1250.
22. Waksman G, Kominos D, Robertson S, Pant N, Baltimore D, Birge R, Cowburn D, Hanafusa H, Mayer B, Overduin M, Resh M, Rios C, Silverman L, Kuriyan J. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. Nature 1992; 358:646-653.
23. Jiang L, Lai L. C-H...O hydrogen bonds at protein –protein interfaces. J Biol Chem 2002; 277:37732-37740.
24. Berman HM, Westbrook JZ, Feng G, Gillilandm TN, Bhat H, Weissig IN. The Protein Data Bank. Nucleic Acids Research 2000; 28:235–242.
25. Tiwari A, Panigrahi SK. HBAT: A complete package for analysing strong and weak hydrogen bonds in macromolecular crystal structures. In Silico Biology 2007; 7:0057.

26. Gromiha MM, Selvaraj S. Influence of Medium and Long Range Interactions in Different Structural Classes of Globular Proteins. Journal of Biological Physics 1997; 23:151–162.
27. Gromiha MM, Santhosh C, Ahmed S. Structural Analysis of Cation-π Interactions in DNA Binding Proteins. International Journal of Biological Macromolecules 2004; 34:203–211.
28. Gromiha MM, Selvaraj S. Inter-residue interactions in protein folding and stability. Progress in Biophysics and Molecular Biology 2004; 86:235–277.
29. Selvaraj S, Gromiha MM. Role of Hydrophobic Clusters and Long-range Contact Networks in the Folding of (alpha/beta)8 Barrel Proteins. Biophysical Journal 2003; 84:1919–1925.
30. Glaser F, Pupko T, Paz I, Bell RE, Bechor D, Martz E, et al. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 2003; 19:163–164.
31. Boeckman B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M. The swiss prot protein knoweldge base and its supplement TrEMBL in 2003. Nucleic Acids Research 2003; 31:365–370.
32. Dosztanyi Z, Fiser A, Simon I. Stabilization centers in proteins: identification, characterization and predictions. Journal of Molecular Biology 1997; 272:597-612.
33. Dosztanyi Z, Magyar C, Tusnady GE, Simon I. Scide: Indentification of stabilization centers in proteins. Bioinformatics 2003; 19:899-900.
34. Vasquez GB, Ji X, Fronticelli C, Gilliland GL. Human Carboxyhemoglobin at 2.2 Å Resolution: Structure and Solvent Comparisons of R-State, R2-State and T-State Hemoglobins. Acta Crystallographica. Section D, Biological Crystallography 1998; 54: 355–366.
35. Mitchell JBO, Nandi CL, McDonald IK, Thornton JM, Price SL. Amino / Aromatic Interactions in Proteins: Is the Evidence Stacked Against Hydrogen Bonding? Journal of Molecular Biology 1994; 239:315–331.