



## COMPUTATIONAL PERSPECTIVE IN THE STRUCTURAL STABILITY OF 'ALL-ALPHA' PROTEINS: THE N-H... $\pi$ INTERACTIONS

V. SHANTHI AND RAO SETHUMADHAVAN\*

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

Received: 18 Dec 2010, Revised and Accepted: 20 Jan 2011

### ABSTRACT

Interactions between NH moiety and aromatic side chains of amino acid residues, the so-called N-H...  $\pi$  interaction, are supposed to contribute to the overall stability of the folded structure of peptides and proteins. The present study details the results of N-H...  $\pi$  interactions in relation to other factors like secondary structural elements, solvent accessibility, conservation score and stabilization centers in 'all-alpha' proteins. We observed 160 N-H... $\pi$  interactions in a data set of 75 'all-alpha' proteins. Side-chain to side-chain N-H... $\pi$  interactions are the predominant type of interactions in the data set. Among the various types of folds of 'all-alpha' proteins considered for the analysis, proteins belonging to alpha-alpha superhelix fold have the highest percentage of N-H... $\pi$  interactions. The secondary structure preference, solvent accessibility and stabilization centers of N-H... $\pi$  interacting residues were estimated. These interactions are mainly formed by long range contacts. More than 50% of the N-H... $\pi$  interacting residues are highly conserved. It is likely that the N-H... $\pi$  interactions contribute significantly to the overall stability of 'all-alpha' proteins.

**Keywords:** N-H... $\pi$  interactions, Secondary structure, Interactions range, Conservation, Stabilizing centers.

### INTRODUCTION

With the understanding of protein structure in 1957 several groups have carried out data base analysis to establish the stability of proteins. A folded protein is stabilized by a number of noncovalent interactions such as hydrophobic interactions, hydrogen bonds, salt bridges, and cation-aromatic interactions<sup>1,2</sup>. In addition, there has been a tremendous interest to obtain information on the nature of weak interactions. Among the several weak intermolecular interactions pervading chemistry and biology, the N-H... $\pi$  interaction is one of the most widely known<sup>3,4</sup>. Positively charged or  $\delta(+)$  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  $\delta(-)$   $\pi$ -electrons and avoid the  $\delta(+)$  ring edge. This geometric pattern is recognized as N-H... $\pi$  interaction<sup>3</sup>.

Probably the first example of X-H... $\pi$  interactions in peptide crystal structure was reported by McPhail and Sim<sup>5</sup>, but it had a little impact on structural science at that time. Much later, the N-H... $\pi$  interactions in proteins attracted the greater attention, following the observation of such interactions in bovine pancreatic trypsin inhibitor (BPTI)<sup>6</sup>, a structure refined at 1.2Å resolution and in hemoglobin protein-ligand interactions<sup>7</sup>. Since then, a greater number of X-H... $\pi$  interactions in proteins have been found in a single case studies, involved in wide variety of functions such as secondary structure stabilization<sup>8,9</sup>, DNA recognition<sup>10</sup>, enzymatic action<sup>11</sup> and drug recognition<sup>12</sup>. The large surface makes  $\pi$ -acceptors a "target that is easy to hit"<sup>13</sup>. Indeed these interactions are widely prevalent in a number of protein crystal structures<sup>3,14,15</sup>. Indeed, the intermolecular interactions involving aromatic rings ( $\pi$ - $\pi$ <sup>16</sup>, cation- $\pi$ <sup>17</sup>, X-H... $\pi$ <sup>18,19</sup> etc.) are important in modulating the high-grade structures and functions of many proteins<sup>20</sup>. Recently, we have published the role of C-H... $\pi$  interactions on the structural stability of single chain 'all alpha' proteins<sup>21</sup>.

Despite the fairly large number of theoretical and experimental investigations of cation- $\pi$  and C-H... $\pi$  hydrogen bonds, there are relatively a few studies on systems exhibiting the N-H... $\pi$  interactions. This is because, the relatively small intermolecular interaction energies of these systems make it very difficult to characterize them experimentally<sup>22,23</sup>. 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<sup>24</sup>, there has been a resurgence of interest in such interactions in proteins. We believe

that this interest leads to a more complete understanding of the forces that govern the protein structure, stability and folding. Hence, in this work, an attempt has been made to collect the information concerning N-H... $\pi$  interactions on the structural stability of 'all-alpha' proteins. In addition, we have systematically studied the role of N-H... $\pi$  interactions in relation to other factors like amino acid preference, secondary structural elements, solvent accessibility, conservation score and stabilization centers. For all this study, we have chosen only single chain alpha proteins. These represent relatively simpler systems in which all the weaker interactions can be studied in the absence of the effects of a complex quaternary structure and the occurrence of redundancy in the data set. Also the statistics become harder, if the pattern we detect persists in other classes of proteins. We also believe that this N-H... $\pi$  interactions expected not to play any role in inter-subunit interactions. It is noteworthy to mention here that the percentage of N-H... $\pi$  interactions was higher than that of cation- $\pi$  interactions in the same set of proteins studied by Remila L. Martis<sup>25</sup>. The frequency and extent of conservation presented unequivocally show that the N-H... $\pi$  interactions cannot and must not be neglected. We postulate that the incorporation of the entirety of this N-H... $\pi$  interactions could provide new perspectives and possibly new answers for the structural biologist. Hence without ambiguity, we can confirm that N-H... $\pi$  interactions play an important role on the structural stability of 'all-alpha' proteins.

### MATERIALS AND METHODS

#### Data set

We have selected a set of 75 non-redundant 'all-alpha' proteins, with sequence identity less than 25%, using the sequence analysis package EMBOSS. EMBOSS is an EBI tool that can be used for pair wise alignment of protein sequences. The co-ordinates of the proteins have been taken from the PDB<sup>26</sup>. The PDB codes of the proteins used for the analysis are shown in Table 1. The 'all-alpha' proteins have been selected from the following five folds as classified by the SCOP<sup>27</sup>: (1) cytochrome c (a.3), (2) DNA/RNA binding 3-helical bundle (a.4), (3) four helical up and down bundle (a.24), (4) fold:EF hand like (a.39) and (5) alpha-alpha super helix (a.118).

#### N-H... $\pi$ interactions

N-H... $\pi$  interactions are calculated using the program available for this purpose namely HBAAT<sup>28</sup>. 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  $\pi$  system. The distances are usually measured from the centroid (M) i.e. centre of the  $\pi$  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  $\pi$  ring. The geometry is adapted from earlier work of Babu<sup>29</sup>. The N-H... $\pi$  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... $\pi$  interactions into two types namely, side-chain to side-chain N-H... $\pi$  interactions (SS-N-H... $\pi$ ) and side-chain to side-chain five member aromatic ring N-H... $\pi$  interactions (SS5-N-H... $\pi$ ).

### Secondary structure and solvent accessibility analysis

Secondary structure and solvent accessibility are the two major criteria to understand the structure and function of proteins. Hence a systematic analysis of each N-H... $\pi$  interaction forming residue was performed based on their location in different secondary structures of 'all-alpha' proteins and their solvent accessibility. We obtained the information about secondary structures from PROSS program which is available at <http://roselab.jhu.edu/utis/runpross.html>. Solvent accessibility of the proteins is identified using the program ASA-View<sup>30</sup>. Solvent accessibility is divided into three classes, buried, partially buried and exposed indicating, respectively, the least, moderate and high accessibility of the amino acid residues to the solvent<sup>31,32</sup>.

### Classification by residue-residue contacts

The N-H... $\pi$  interacting residues coming within a sphere of 8Å was computed as described earlier<sup>33-36</sup>. 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  $\pm 3$  or  $\pm 4$  residues contribute to medium-range and those more than four residues away contribute to long-range interactions<sup>35</sup>. This classification enables us to evaluate the contribution of short-range, medium-range and long-range contacts in the formation of N-H... $\pi$  interactions.

### Conservation score

We computed the conservation score of N-H... $\pi$  interacting amino acid residues in each protein using the ConSurf program<sup>37</sup>. This program computes the conservation based on the comparison of the sequence of a PDB chain with the proteins deposited in Swiss-Prot<sup>38</sup> 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 residues are classified into nine categories from highly variable to highly conserved. Residues with a score of 1 are considered highly variable and those with a score of 9 are considered highly conserved.

### Stabilizing centers

Stabilization centers are clusters of residues that are involved in medium or long-range interactions<sup>39</sup>. 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... $\pi$  interacting amino acid residues were computed using the SCide program<sup>40</sup> for computing the stabilization centers.

## RESULTS

### N-H... $\pi$ interactions

There was a total of 160 N-H... $\pi$  interactions in the set of 75 'all-alpha' proteins. The number of N-H... $\pi$  interactions was higher than that of cation- $\pi$  interactions<sup>25</sup> and less than the number of C-H... $\pi$  interactions<sup>21</sup> in the same set of protein studied. Of the different types of N-H... $\pi$  interactions, majority 84% of the N-H... $\pi$  interactions were SS-N-H... $\pi$  interactions and remaining 16% of the interactions were from SS5-N-H... $\pi$  interactions. Though N-H... $\pi$  interaction has been reported with His acting as an acceptor<sup>41</sup>, 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). As a representative picture, the SS-N-H... $\pi$  interactions in 'all-alpha' protein PDB ID 1CYJ (between Lys 33 and Phe 44) is shown in Fig. 2. 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 were plotted in Fig. 3. The percentage ratio calculated shows that Arg make the maximum contribution to this N-H... $\pi$  interaction (52 interactions in a total of 160 interactions). 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<sup>42</sup>, 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, Phe is the most common amino acid involved in such interactions (56 interactions in a total of 160 interaction). It might be due to the highest occurrence of Phe among the aromatic amino acids. Hence, Arg and Phe residues may be quite important for the stability of 'all-alpha' proteins. Another important issue is related to the interatomic distances. The analysis of the distribution of interatomic N... $\pi$  and H... $\pi$  distances were shown in Fig. 4. The pattern of the N... $\pi$  distances shows that, majority 31% of the interactions were found in between the range of 3.76 Å to 4.00 Å. The pattern of the H... $\pi$  distances shows that, the majority 24% of the interactions were found in between the range of 3.76 Å to 4.00 Å. We have also analyzed the percentage of N-H... $\pi$  interactions in various folding types of 'all-alpha' proteins. This result was shown in Fig. 5. Among the various folding types, it is observed that a.118 fold has the maximum number of proteins showing various N-H... $\pi$  interactions. This result is consistent with the earlier ones on cation- $\pi$  interaction studied in the same set of proteins studied by Martis<sup>25</sup>.

### 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 amino acid, which participated in the two types of N-H... $\pi$  interactions namely, SS-N-H... $\pi$  and SS5-N-H... $\pi$  interactions. The secondary structure preference of each of the amino acids involved in all the above said types of N-H... $\pi$  interactions were obtained using PROSS program and the results were depicted in Table 2 and Table 3. It is interesting to note that donor residues such as Gly, Ala, Asp, Glu, Phe and Trp preferred to be in coil and Val, Leu, Ile, Ser, Thr, Cys, Asn, Gln, Arg, Lys and Tyr preferred to be in helix. The acceptor residues such as, Phe, Tyr, Trp and His were preferred to be in helix.

### Solvent accessibility

We have estimated the solvent accessibility of all the amino acid residues that were involved in N-H... $\pi$  interactions using the program ASA-View<sup>30</sup>. We found that of the different amino acids that were involved in N-H... $\pi$  interactions, Gly, Ala, Val, Leu, Cys,

Asn, Gln, Phe, Tyr, Trp and His were in the buried regions. Ser, Thr, Arg, Lys, Ile, Asp and Glu residues involved in N-H... $\pi$  interactions were in the partially buried regions. We found that most of the polar amino acid residues involved in N-H... $\pi$  interactions were solvent exposed and most of the aromatic residues involved in N-H... $\pi$  interactions were excluded from the solvent (Table 4 and Table 5).

### Sequential separation

The contribution of N-H... $\pi$  interactions in 'all-alpha' proteins could define either the local or the global stability of the proteins. Therefore, there is a need to evaluate the contribution of inter-residual N-H... $\pi$  interactions. The sequential distance between residues that contributed to N-H... $\pi$  interactions were calculated and results were depicted in Fig. 6. It reveals that 67%, 20% and 13% of inter-residue N-H... $\pi$  interactions were found to be long-range, medium-range and short-range interactions respectively.

### Conservation score

We used the ConSurf program to compute the conservation score of amino acid residues involved in N-H... $\pi$  interactions in 'all-alpha'

proteins, and the results were shown in Fig. 7. 27% of the amino acid residues that contributed donor atoms in N-H... $\pi$  interactions had the highest conservation score of 9, while 40% of the amino acid residues had a conservation score, in the range of 6-8. Thus, 67% of the donor amino acid residues had a higher conservation score. In case of amino acid residues that contributed acceptor atoms in N-H... $\pi$  interactions, 23% of the acceptor amino acid residues had the highest conservation score of 9, while 43% of the amino acid residues had a conservation score, in the range of 6-8. Thus, 66% of the acceptor amino acid residues had a higher conservation score.

### Stabilizing centers

We used the SCide program for computing the stabilization centers in the 'all-alpha' proteins data set. We found that 32% of the amino acid residues that contribute donor atoms to N-H... $\pi$  interactions had one or more stabilization centers in addition to their contribution to N-H... $\pi$  interactions and similarly 38% of the amino acid residues that contribute acceptor atoms to N-H... $\pi$  interactions had one or more stabilization centers in addition to their contribution to N-H... $\pi$  interactions.

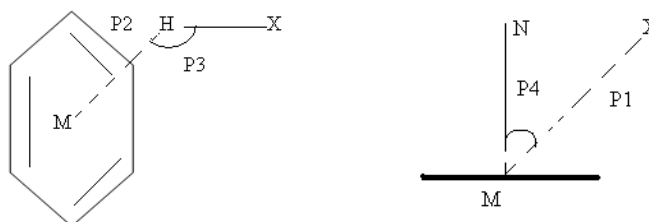


Fig. 1: Parameters for X-H... $\pi$  interaction (X=N): P1  $\leq$  5.00 Å; P2  $\leq$  4.50 Å; P3  $\geq$  120°; P4  $\leq$  30°.

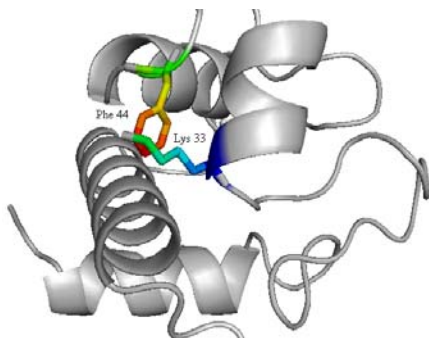


Fig. 2a: Pymol view of SS-N-H... $\pi$  interacting pairs in PDB ID 1CCR

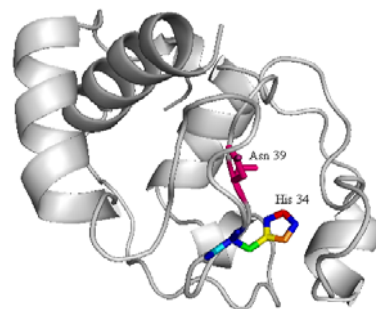


Fig. 2b: Pymol view of SS5-N-H... $\pi$  interacting pairs in PDB ID 1CYJ

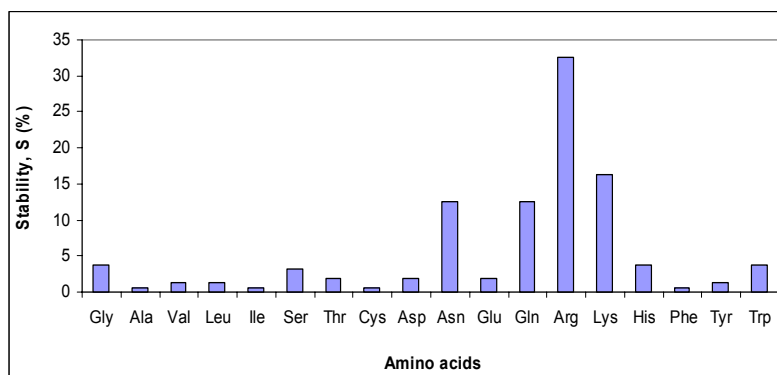


Fig. 3: Amino acids contribution to the stability of 'all-alpha' proteins

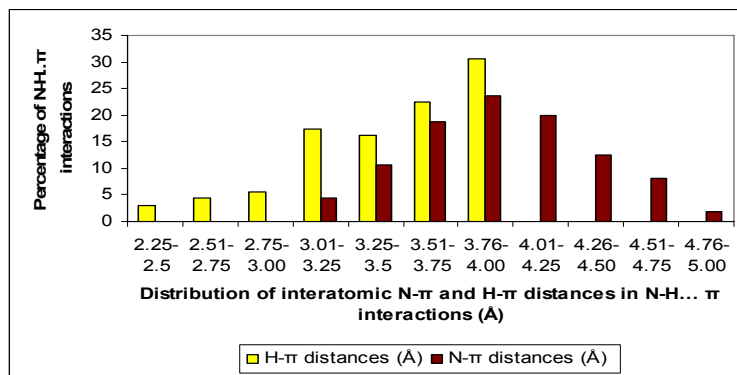


Fig. 4: Observed distribution of N-H...π interactions as a function of interatomic (N-π and H-π) distances in the 'all-alpha' proteins

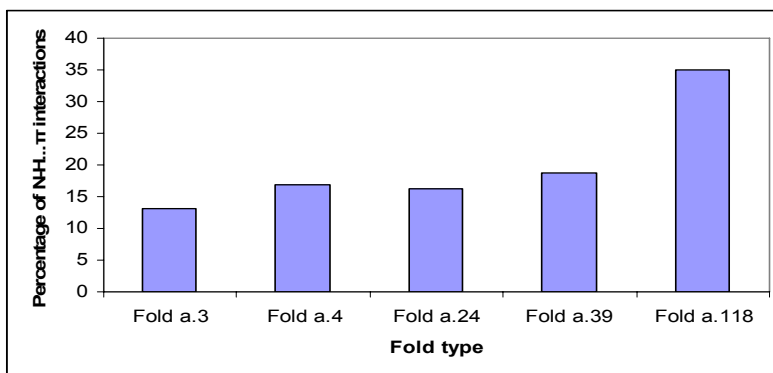


Fig. 5: Percentage of N-H...π interactions in various folding type of 'all-alpha' proteins

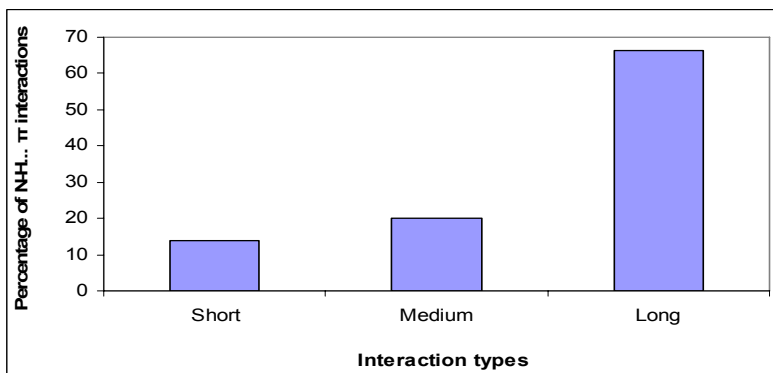


Fig. 6: N-H...π interactions range in 'all-alpha' proteins

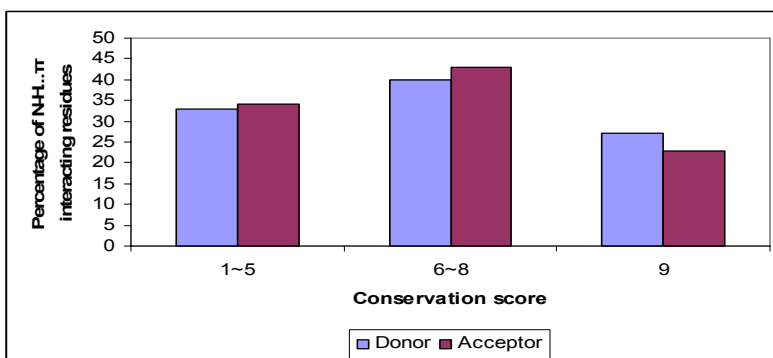


Fig. 7: Conservation score for N-H...π interacting residues in 'all-alpha' proteins

Table 1: List of PDB codes of 'all-alpha' proteins considered for analysis of N-H... $\pi$  interactions

| a.3 Fold | a.4 Fold | a.24 Fold | a.39 Fold | a.118 Fold |
|----------|----------|-----------|-----------|------------|
| 1A56     | 1AOY     | 1A7D      | 1CDP      | 1B89       |
| 1C52     | 1C20     | 1CGN      | 1E14      | 1EYH       |
| 1C2N     | 1D5V     | 1CPQ      | 1IG5      | 1HF8       |
| 1CC5     | 1D8J     | 1DOV      | 1IJ5      | 1HO8       |
| 1CCH     | 1G2H     | 1G5Z      | 1K9P      | 1HU3       |
| 1CCR     | 1GVD     | 1GS9      | 1MHO      | 1HZ4       |
| 1CRY     | 1GXQ     | 1KTM      | 1Q80      | 1IB2       |
| 1CYJ     | 1IG6     | 1LPE      | 1RK9      | 1KLX       |
| 1E8E     | 1JGS     | 1NFN      | 1RRO      | 1LRV       |
| 1F1F     | 1LEA     | 1NZE      | 1SRA      | 1M8Z       |
| 1GDV     | 1LFB     | 1O3U      | 1TOP      | 1OYZ       |
| 1GKS     | 1MIJ     | 1SR2      | 2SAS      | 1PBV       |
| 1LS9     | 1P4W     | 1TQG      | 3PAT      | 1PAQ       |
| 1YCC     | 2EZI     | 2A0B      | 5PAL      | 1TE4       |
| 451C     | 2HTS     | 2MHR      | 5TNC      | 2BCT       |

Table 2: Frequency of occurrence of N-H... $\pi$  interactions forming donor residue in different secondary structures

| Amino acids | Coil | $\beta$ Strand | Polyproline II | $\beta$ Turn | $\alpha$ Helix |
|-------------|------|----------------|----------------|--------------|----------------|
| Gly         | 83.3 | -              | -              | 16.7         | -              |
| Ala         | 100  | -              | -              | -            | -              |
| Val         | -    | -              | -              | -            | 100            |
| Leu         | -    | -              | -              | -            | 100            |
| Ile         | -    | -              | -              | -            | 100            |
| Ser         | 25   | -              | -              | -            | 75             |
| Thr         | -    | -              | -              | -            | 100            |
| Cys         | -    | -              | -              | -            | 100            |
| Asp         | 66.6 | -              | -              | -            | 33.3           |
| Asn         | 5    | -              | 21             | 31.6         | 42.2           |
| Glu         | 66.6 | -              | -              | -            | 33.3           |
| Gln         | -    | -              | -              | -            | 100            |
| Arg         | 22.2 | -              | 2.2            | 15.5         | 60             |
| Lys         | 16.6 | 8.3            | -              | 33.3         | 41.6           |
| Phe         | 100  | -              | -              | -            | -              |
| Tyr         | -    | -              | -              | -            | 100            |
| Trp         | 75   | -              | -              | -            | 25             |

Table 3: Frequency of occurrence of N-H... $\pi$  interactions forming acceptor residue in different secondary structures

| Amino acids | Coil  | $\beta$ Strand | Polyproline II | $\beta$ Turn | $\alpha$ Helix |
|-------------|-------|----------------|----------------|--------------|----------------|
| Phe         | 20.75 | 5.6            | 3.7            | 18.8         | 50.9           |
| Tyr         | -     | 7.4            | -              | 11.1         | 81.5           |
| Trp         | 16.2  | -              | 6.9            | 2.3          | 74.4           |
| His         | 22.7  | -              | 22.7           | 9            | 45.45          |

Table 4: Solvent accessibility of N-H... $\pi$  interactions forming donor residues in 'all-alpha' proteins

| Amino acids | Buried | Partially buried | Exposed |
|-------------|--------|------------------|---------|
| Gly         | 83.3   | 16.6             | -       |
| Ala         | 100    | -                | -       |
| Val         | 100    | -                | -       |
| Leu         | 100    | -                | -       |
| Ile         | -      | 100              | -       |
| Ser         | -      | 75               | 25      |
| Thr         | -      | 75               | 25      |
| Cys         | 100    | -                | -       |
| Asp         | -      | 66.6             | 32.3    |
| Asn         | 57.8   | 36.8             | -       |
| Glu         | -      | 66.6             | 33.3    |
| Gln         | 61.1   | 22.2             | 16.6    |
| Arg         | 40     | 46.6             | 13.3    |
| Lys         | 44     | 48               | 8       |
| Phe         | 100    | -                | -       |
| Tyr         | 100    | -                | -       |
| Trp         | 100    | -                | -       |

Table 5: Solvent accessibility of N-H... $\pi$  interactions forming acceptor residues in 'all-alpha' proteins

| Amino acids | Buried | Partially buried | Exposed |
|-------------|--------|------------------|---------|
| Phe         | 71.7   | 24.5             | 3.7     |
| Tyr         | 61.3   | 35.4             | 3.2     |
| Trp         | 64.1   | 25.6             | 10.2    |
| His         | 40.9   | 54.5             | 4.5     |

## DISCUSSION

We have investigated the influence of N-H... $\pi$  interactions on the structural stability of 'all-alpha' proteins. We find that 68% of the considered 'all-alpha' proteins exhibit N-H... $\pi$  interactions and 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  $\pi$  acceptors (ie, SS-N-H... $\pi$  interactions). Though N-H... $\pi$  interaction has been reported with His acting as an acceptor<sup>41</sup>, the frequency of occurrence of such bonds is low owing to the unsuitability of imidazole ring in this role when charged. There was a total of 160 N-H... $\pi$  interactions observed in the set of 75 'all-alpha' proteins. The geometric parameters calculated for these interactions suggest that N-H... $\pi$  interactions could be classified as weak H bonds and occur mainly in the distances greater than 3.76 Å and 3.50 Å from the N and H atoms, respectively. Most of the residues involved in N-H... $\pi$  interactions prefer the secondary structure of alpha helical segments. This indicates that either direct neighbors along the sequence or close neighbors in helices or sometimes coils preferably display this kind of interaction. This result is consistent with previous computational predictions regarding secondary structure preference of C-H... $\pi$  interactions in the same set of proteins reported by our group earlier<sup>21</sup>. Thus, the 'all-alpha' proteins are, therefore, confronted with a very large number of helices in their three dimensional arrangements. The solvent accessibility analysis is quite reasonable in the sense that the aromatic residues are in principle, non polar residues, and tend to be buried. Since Arg and Lys are polar in nature they tend to be partially exposed to the solvent surface. Precisely the same trends were observed in the C-H... $\pi$  interactions<sup>21,43</sup>. Perhaps N-H... $\pi$  interactions involving residues either as donor or as acceptor groups are found mostly in the interior of the protein and tend to be buried in nature. These might be some of the reasons for their nature of solvent accessibility. These interactions are formed mainly by long range contacts. From the conservation score of each amino acid residues, we were able to infer that more than 50% of the interacting residues might be conserved in 'all-alpha' proteins. The conservation of amino acid residues with  $\pi$ -systems may in some cases be linked to their involvement in N-H... $\pi$  interactions and to the stability or the function of the protein. Furthermore, 32 percentages of donor residues and 38 percentages of acceptor residues in N-H... $\pi$  interactions act as stabilizing centers in 'all-alpha' proteins. Among the various types of folds of 'all-alpha' proteins considered for the analysis, proteins belonging to a118 fold have the highest percentage of N-H... $\pi$  interactions. The numbers presented 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... $\pi$  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.

## ACKNOWLEDGMENT

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

## REFERENCES

- Gallivan JP, Dougherty DA. Cation- $\pi$  Interactions in Structural Biology. Proceedings of the National Academy of Sciences USA 1999; 96:9459-9464.
- Dill KA. Dominant forces in protein folding. Biochemistry 1990; 29(31):7133-7155.

- Burley SK, Petsko GA. Amino-aromatic interactions in proteins. FEBS Letters 1986; 203:139-143.
- Burley SK, Petsko GA. Weakly polar interactions in proteins. Advances in Protein Chemistry 1988; 39:125-189.
- McPhail AT, Sim GA. Hydroxyl-benzene hydrogen bonding. An X-ray study. Chemical Communications 1965; 00,124-125.
- Wlodower A, Walter J, Huber R, Sjolín L. Structure of bovine pancreatic trypsin inhibitor: Results of joint neutron and X-ray refinement of crystal form II. Journal of Molecular Biology 1984; 180:301-329.
- Perutz MF, Fermi G, Abraham DJ, Poyart C, Bursaux E. Hemoglobin as a receptor of drugs and peptides: x-ray studies of the stereochemistry of binding. Journal of the American Chemical Society 1986; 108:1064-1078.
- 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.
- Burley SK, Petsko GA. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science 1985; 229:23-28.
- Parkinson G, Gunasekara A, Vjtechovsky J, Zhang X, Kunkel TA, Berman H, Ebright RH. Aromatic hydrogen bond in sequence specific protein DNA recognition. Nature Structural and Molecular Biology 1996; 3:837-841.
- 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  $\pi$  cloud of tyrosine 6 and its influence on catalysis. Journal of the American Chemical Society 1993; 115:7910-7911.
- 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.
- Suzuki S, Green PG, Bumgarner RE, Dasgupta S, Goddard WA, Blake GA. Benzene forms hydrogen bonds with water. Science 1992; 257:924-945.
- Levitt M, Perutz MF. Aromatic rings act as hydrogen bond acceptors. Journal of Molecular Biology 1988; 201:751-754.
- Steiner T, Koellner G. Hydrogen bonds with  $\pi$ -acceptors in proteins: frequency and role in stabilizing local 3D structures. Journal of Molecular Biology 2001; 305:535-557.
- Janiak C. A critical account on  $\pi$ - $\pi$  stacking in metal complexes with aromatic nitrogen-containing ligands. Journal of the Chemical Society-Dalton Transactions 2000; 21:3885-3896.
- Ma JC, Dougherty DA. The cation- $\pi$  interaction. Chemical Reviews 1997; 97: 303- 1324.
- Nishio M, Hirota M, Umezawa Y. (Eds.). The C-H... $\pi$  Interaction. New York: Wiley; 1998.
- Umezawa Y, Tsuboyama S, Takahashi H, Uzawa J, Nishio M. C-H... $\pi$  Interaction in the Conformation of Organic Compounds. A Database Study Tetrahedron 1999; 55:10047-10056.
- Chakrabarti P, Bhattacharyya R. Geometry of nonbonded interactions involving planar groups in proteins. Progress in Biophysics and Molecular Biology 2007; 95:83-137.
- Shanthi V, Ramanathan K, Rao Sethumadhavan. Exploring the role of C-H... $\pi$  interactions on the structural stability of single chain 'all-alpha' proteins. Applied Biochemistry and Biotechnology 2010; 160:1473-1483.
- Kim KS, Tarakeshwar P, Lee JY. Molecular clusters of  $\pi$ -systems: theoretical studies of structures, spectra, and origin of interaction energies. Chemical Reviews 2000; 100:4145-4185.
- 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.
- 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.
- Martis RL, Singh SK, Michael Gromiha M, Santhosh C. Role of cation- $\pi$  interactions in single chain 'all-alpha' proteins. Journal of Theoretical Biology 2008; 250:655- 662.

26. Berman HM, Westbrook JZ, Feng G, Gillilandm TN, Bhat H, Weissig IN. The Protein Data Bank. *Nucleic Acids Research* 2000; 28:235–242.
27. Conte LL, Ailey B, Hubbard TJP, Brenner SE, Murzin AG, Chothia C. SCOP: a structural classification of proteins database. *Nucleic Acids Research* 2000; 28:257–259.
28. 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.
29. Babu MM. NCI: A program to identify non-canonical interactions in protein structures. *Nucleic Acids Research* 2003; 31:3345–3348.
30. Shandar A, Gromiha MM, Fawareh H, Sarai A. ASAView: Database and tool for solvent accessibility representation in proteins. *BMC Bioinformatics* 2004; 5:51.
31. Gilis D, Rooman M. Stability changes upon mutation of solvent-accessible residues in proteins evaluated by database-derived potentials. *Journal of Molecular Biology* 1996; 257:1112–1126.
32. Gilis D, Rooman M. Predicting protein stability changes upon mutation using database-derived potentials: solvent accessibility determines the importance of local versus non-local interactions along the sequence. *Journal of Molecular Biology* 1997; 272:276–290.
33. 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.
34. Gromiha MM, Santhosh C, Ahmed S. Structural Analysis of Cation- $\pi$  Interactions in DNA Binding Proteins. *International Journal of Biological Macromolecules* 2004; 34:203–211.
35. Gromiha MM, Selvaraj S. Inter-residue interactions in protein folding and stability. *Progress in Biophysics and Molecular Biology* 2004; 86:235–277.
36. Selvaraj S, Gromiha MM. Role of Hydrophobic Clusters and Long-range Contact Networks in the Folding of (alpha/beta)<sub>8</sub> Barrel Proteins. *Biophysical Journal* 2003; 84:1919–1925.
37. 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.
38. 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 knowledge base and its supplement TrEMBL in 2003. *Nucleic Acids Research* 2003; 31:365–370.
39. Dosztanyi Z, Fiser A, Simon I. Stabilization centers in proteins: identification, characterization and predictions. *Journal of Molecular Biology* 1997; 272:597-612.
40. Dosztanyi Z, Magyar C, Tusnady GE, Simon I. Scide: Identification of stabilization centers in proteins. *Bioinformatics* 2003; 19:899-900.
41. 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.
42. 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.
43. Brandl M, Weiss MS, Jabs A, Sühnel J, Hilgenfeld R. C-H... $\pi$ -Interactions in Proteins. *Journal of Molecular Biology* 2001; 307:357–377.