NH…π Interactions: Investigations on the Evidence and Consequences in RNA Binding Proteins

Anand Anbarasu, Sudha Anand and Rao Sethumadhavan*

Bioinformatics and Computational Biology Division, School of Biotechnology, Chemical and Bio-medical Engineering, Vellore Institute of Technology, Vellore-632014, India

Abstract: We have analyzed the role of N-H…π interactions in 59 RNA binding protein structures with relation to other factors like amino acid preference, secondary structural elements, solvent accessibility of a particular fold, conservation score, stabilization centers and stabilizing residues. There was an average of 12 N-H…π interactions per protein in the 59 proteins investigated and an average of one significant N-H…π interaction for every 37 amino acid residues in RNA binding proteins. 63% of the N-H…π interactions were computed to be inter-residue interactions, while the remaining 37% of the N-H…π interactions were found to be intra-residue interactions. The contribution of the main chain to side chain N-H…π mediated interactions was higher irrespective of the amino acids involved. Long-range N-H…π interactions appear to be the predominant type of interactions in RNA binding proteins. Arg, Asn, Asp, Gln, Glu, Ile, Leu, Lys and Ser preferred to be in helix. Ala, Cys, Phe, Thr, Trp, Tyr and Val were found to prefer strand conformation. Polar amino acid residues involved in N-H…π interactions were solvent exposed and most of the non polar residues involved in N-H…π interactions were excluded from the solvent. Most of the amino acid residues involved in N-H…π interactions might be conserved in RNA binding residues. The specificity is shown by residues not involved in N-H…π interactions as only a very small percentage of interacting residues are involved in RNA Binding. On the whole, the results of the present investigations on N-H…π interactions might be useful for intra protein-protein interactions and studies on inter-residue and intra-residue interactions.

Keywords: N-H…π interactions, long-range interactions, secondary structure, stabilization centers, conservation score.

1. INTRODUCTION

Although the structure and function of a protein is determined by conventional hydrogen bonds, ionic bonds and hydrophobic interactions a set of weak interactions has also been recognized to play an important role in protein structure and stability [1]. This set includes N-H…π, C-H…π and C-H…O interactions. Though, the over all stabilization energy is much smaller due to reduced electrostatic contribution and exchange repulsion energy, they are still significant simply because of the increased number of weak interactions that occur in proteins [2]. Even though the occurrence of this Non-Canonical Interactions (NCI) was well documented very early in time [3-5], it was not until recently that their importance was completely understood. Several large-scale studies over the last decade have unambiguously revealed the occurrence of these interactions in crystal structures, revealing the importance of such interactions and therefore reviving interest in studying them in greater detail [1, 6-17]. In terms of energetic contribution, theoretical *ab initio* calculations [18-21] have clearly revealed that the energy of these N-H…π interactions is less than the energy of a conventional hydrogen bond (O/N-H…O=C). For instance, N-H…π interactions may contribute up to 3.5 kcal/mol, whereas regular hydrogen bonds may contribute up to 5.5 kcal/mole [8]. Even though the N-H…π interactions is less in energetic terms, it is observed that they occur frequently in proteins and cumulatively these interactions provide significant energy for protein stability. Observations of N-H…π interactions in bovine pancreatic trypsin inhibitor [22] and hemoglobin protein interaction [23] have shown the significance of these interactions in proteins. A greater number of X-H…π interactions in proteins have been involved in a wide variety of functions such as secondary structure stabilization [24], drug recognition [25], DNA recognition [26] and enzyme action [27]. Our group has also reported the influence of cation-π interactions in RNA binding proteins [28] and explored the role of such interactions in glycoproteins, lipid binding proteins and RNA binding proteins [29]. In addition, we have also investigated the role of C-H…π interactions in protein-RNA complexes [30] Though previous studies have investigated the occurrence of NCI in various proteins [2, 6, 8] very few studies have systematically studied the role of N-H…π interactions in relation to other factors like amino acid preference, secondary structural elements, solvent accessibility of a particular fold, conservation score, stabilization centers and stabilizing residues. We have addressed these issues by analyzing the structures of protein-RNA complexes. Probably, ours is the first such report on the analysis of N-H…π interactions in RNA binding proteins.

*Address correspondence to this author at the Bioinformatics and Computational Biology Division, School of Biotechnology, Chemical and Bio-medical Engineering, Vellore Institute of Technology, Vellore-632014, India; Tel: +91 4162202522; Fax: +91 4162243092;
E-mail: rsethumadhavan@vit.ac.in

1874-1991/08 2008 Bentham Open

2. MATERIAL AND METHODS

2.1. Data Set

We have selected 59 RNA binding proteins from Protein Data Bank (PDB) [31] for our investigation. The structures were solved with <2.5A° resolution and the sequence identity among the proteins in the dataset was less than 45%.

The PDB IDS' are as follows:

1A8V, 1ASY, 1ASZ, 1B23, 1B7F, 1C0A, 1CVJ, 1CX0, 1EFU, 1EFW, 1F7U, 1F7V, 1F7Y, 1FFY, 1G59, 1GAX, 1GIK, 1GTS, 1H2D, 1H2T, 1H2V, 1H38, 1H3E, 1H4Q, 1H4S, 1HRO, 1HYS, 1I6U, 1IL2, 1IVS, 1JIU, 1J2B, IJBR, 1JID, 1K8W, 1KNZ, 1KOG, 1KOI, 1LGN, 1M2P, 1MIK, 1MJI, 1MMS, 1MSW, 1N35, 1N38, 1N77, 1N78, 1NB7, 1NH3, 1O0B, 1O0C, 1ONP, 1QF6, 1QLN, 1QU2, 1QU3, 1YRN and 2FMT.

2.2. Computation of N-H…π Interactions

We used a stand-alone program which identifies the N-H…π interactions based on the four different geometric criteria that are defined based on distances and angles of the atoms under consideration [32]. The type of N-H…π interaction is indicated by a two letter code in which the first letter indicates the donor atom and the second the acceptor: M, S and S5 represent the main-chain atom, side-chain atom and side-chain atom in the 5 membered aromatic ring respectively. There are four types of N-H…π interactions namely, main-chain to side-chain (MS), main-chain to 5 member aromatic ring of Trp (MS-5), side-chain to side-chain (SS) and side-chain to the 5 member aromatic ring of Trp (SS-5).

2.3. Sequential Separation

The N-H…π interacting residues coming within a sphere of 8A° was computed as described earlier [33-35]. For a given residue, the comparison of the surrounding residue is analyzed in terms of the location at the sequence level. The contribution from <± 3 residues are treated as short-range contacts, ± 3 or ± 4 residues as medium-range contacts and >± 4 residues are treated as long range contacts. 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.

2.4. Secondary Structure and Solvent Accessibility

Secondary structure and solvent accessibility are the two major intermediate steps 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 RNA binding proteins and their solvent accessibility. We obtained the information about secondary structures from PDB and identified solvent accessibility of the proteins using the program ASA-View [36]. Solvent accessibility was 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 [37, 38].

2.5. Conservation Score

We computed the conservation score of N-H…π interacting amino acid residues in each protein using the ConSurf server [39]. This server computes the conservation based on the comparison of the sequence of a PDB chain with the proteins deposited in Swiss-Prot [40] and finds the ones that are 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 evolutionary 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 residues with a score of 9 are considered highly conserved.

2.6. Stabilization Centers

Stabilization centers are clusters of residues that are involved in medium or long range interactions [41]. Residues can be considered part of stabilization centers if they are involved in medium or long range interactions and if two supporting residues can be selected from both of their flanking tetra peptides, which together with the central residues form at least seven out of the nine possible contacts. The stabilization centers for the N-H…π interacting amino acid residues were computed using the SCide server [42] in RNA binding proteins.

2.7. Binding Residues

Binding residues were identified using the available information in amino acid – nucleotide interaction database [43]. An amino acid and a nucleotide base is in contact with each other if any one of the atoms in the amino acid is within the distance of 3.5 Å to any atom in the base. Accordingly, the amino acid is considered to be a binding site residue and it has the affinity to interact with RNA. The analysis of the nature of N-H…π interaction forming residues with respect to binding affinity has been computed.

3. RESULTS AND DISCUSSION

3.1. N-H…π Interactions

There was an average of 12 N-H…π interactions per protein, in the 59 proteins investigated and an average of one significant N-H…π interaction for every 37 amino acid residues in RNA binding proteins. The number of N-H…π interactions were more when compared to cation-π interactions [28, 29] and less than the number of C-H…π interactions [30] in the same data set. The results of the four different N-H…π interactions are depicted in Fig. (**1**) and it was found that, 47% of the interactions were MS interactions, 17% of the interactions were MS5 interactions, 27% of the interactions were SS interactions and the remaining 9% interactions were SS5 interactions. The major contribution to N-H…π interactions was mainly from MS interactions. The N-H…π interactions in RNA binding protein PDB ID 1ASZ: Aspartyl tRNA synthetase G76 to W156 (main chain – side chain) and R453 to F448 (side chain – side chain) is depicted in Fig. (**2**). We classified the

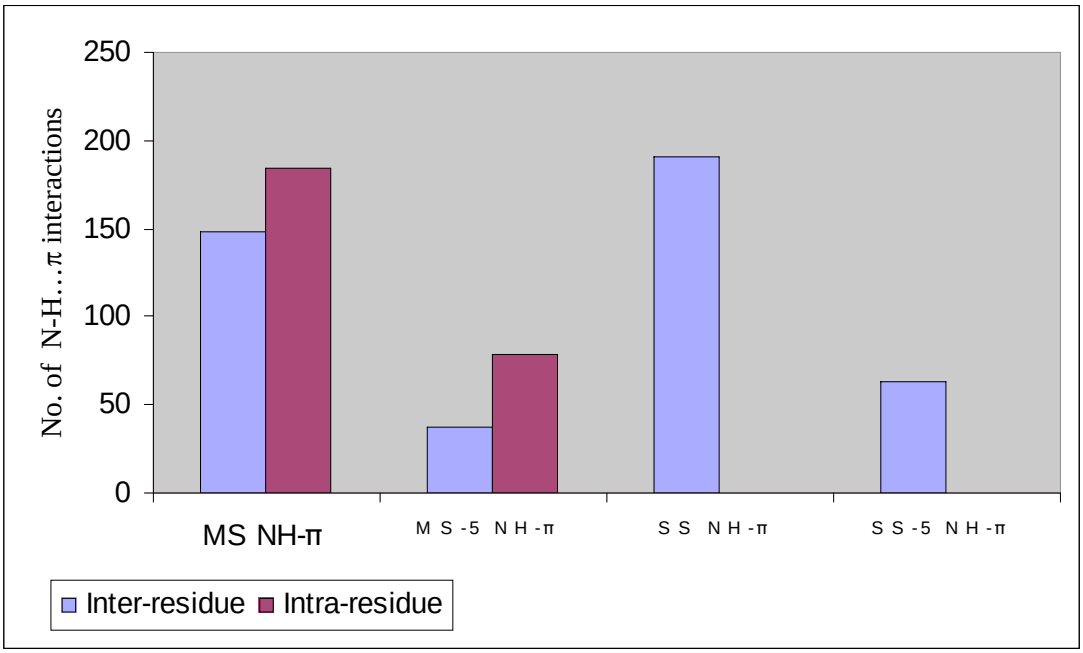

Fig. (1). N-H…π interactions in RNA Binding Proteins.

N-H…π interactions into inter-residue and intra-residue interactions based on the acceptor and donor atoms involved in the interaction. If the donor and acceptor residues involved in N-H…π interactions were from different amino-acid residues, then it is referred to as inter-residue N-H…π interactions and if the donor and acceptor residues involved in N-H…π interactions were from the same amino-acid residue then it is referred to as intra-residue N-H…π interactions. The results of the inter-residue and intra-residue interactions in the RNA binding proteins investigated in the present study are depicted in Fig. (**1**). 63% of the N-H…π interactions were computed to be inter-residue interactions, while the remaining 37% of the N-H…π interactions were found to be intra-residue interactions. Among the inter residue N-H…π interactions, 34% of the interactions were MS interactions, 9% of the interactions were MS5 interactions, 43% of the interactions were SS interactions and 14% of the interactions were found to be SS5 interactions. Among the intra residue N-H…π interactions, 70% of the interactions were found to be MS interactions and 30% of the interactions were found to be MS5 interactions. We consider this to be an important observation in RNA binding proteins, in the sense that, N-H…π interactions may play an important role in determining the structural stability of RNA binding proteins. Hence, these results will be a good starting point for studies on intra protein-protein interactions and for investigations on inter-residue and intra-residue interactions with protein molecules.

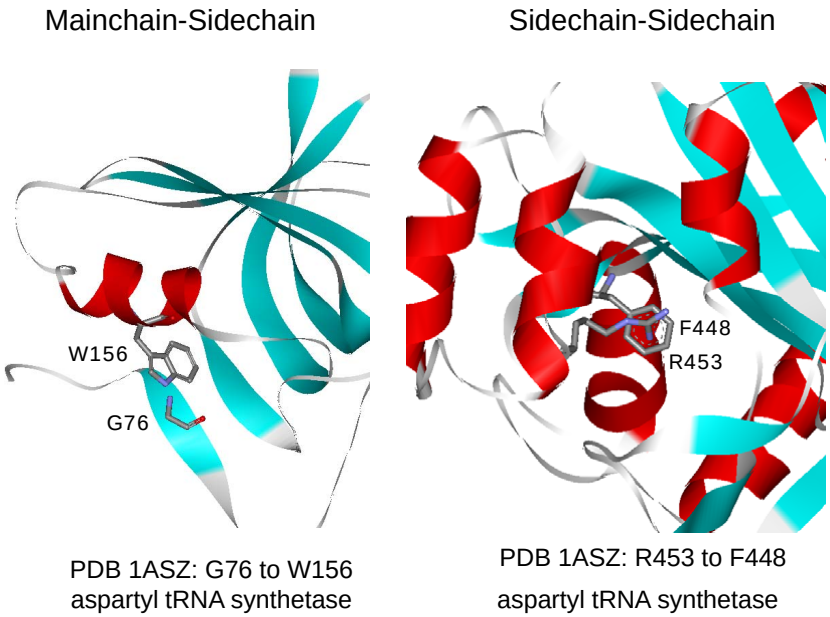

Fig. (2). N-H… π interactions in Aspartyl tRNA synthetase.

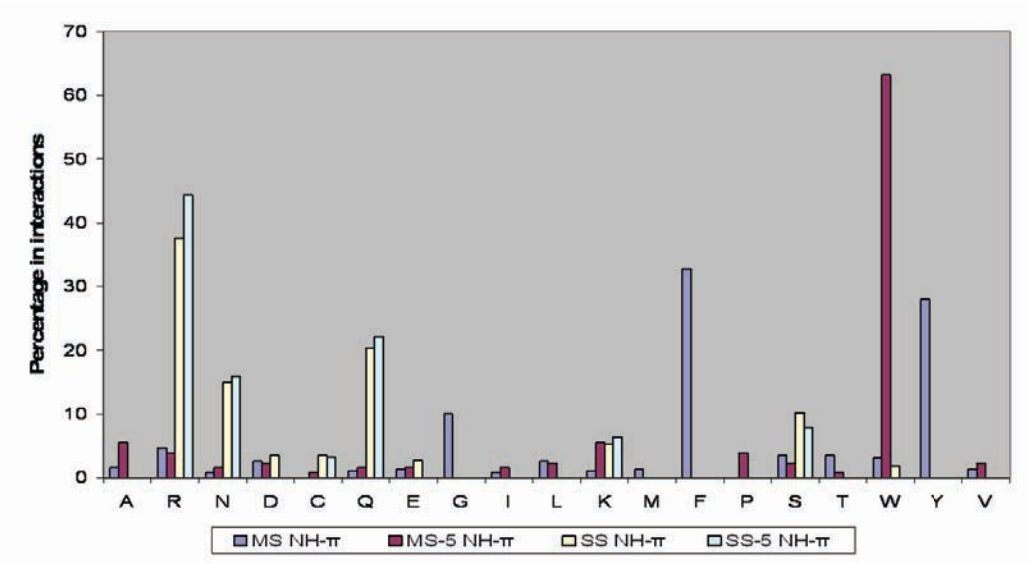

Fig. (3). Donor residues in N-H…π interactions.

The contribution of donor residues in the four different types of N-H…π interactions is depicted in Fig. (**3**). The contribution of donor residues in MS interactions was mainly from Phe, Tyr and Gly residues. The contribution of acceptor residues in the four different types of N-H…π interactions is depicted in Fig. (**4**). The acceptor atom contribution for MS interactions was mainly from the π residues namely Phe, and Tyr. The contribution of acceptor residues for MS interactions from Trp was very minimal. The contribution of donor and as well as acceptor residues in MS5 interactions was mainly from Trp. In the case of SS interactions the donor atom contribution was mainly from Arg, Gln, Asn and Ser. The acceptor atom contribution in SS interactions was similar to MS interactions and it was mainly from Phe, and Tyr. In SS5 interactions the donor atom contribution was similar to SS interactions and the donor atom contribution was mainly from Arg, Gln and Asn. Phe and Tyr did not contribute acceptor residues in MS5 interactions and SS5 interactions.

All the naturally occurring amino-acids except for His had donor residues that were involved in N-H…π interactions. The contribution of MS interactions was higher in N-H…π interactions irrespective of the amino acids involved. In terms of total donor atom contribution to N-H…π interactions, the highest contribution of donor residues was mainly from Arg, Phe, Tyr, Trp, Gln and Asn residues. In the case of acceptor atom contribution to N-H…π interactions, Trp was involved in all the four types of N-H…π interactions, whereas Phe and Tyr were involved only in MS and SS interactions. It was interesting to note that, Trp contributed both donor and acceptor residues in all the four types of N-H…π interactions. From this we infer that, the contribution of Trp to N-H…π interactions might be significant in RNA binding proteins.

3.2. Sequential Separation

The contribution of N-H…π interactions in RNA binding 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…π interactions. The contribution from <± 3 residues are treated as short-range contacts, ± 3 or ± 4 residues as medium-range contacts and

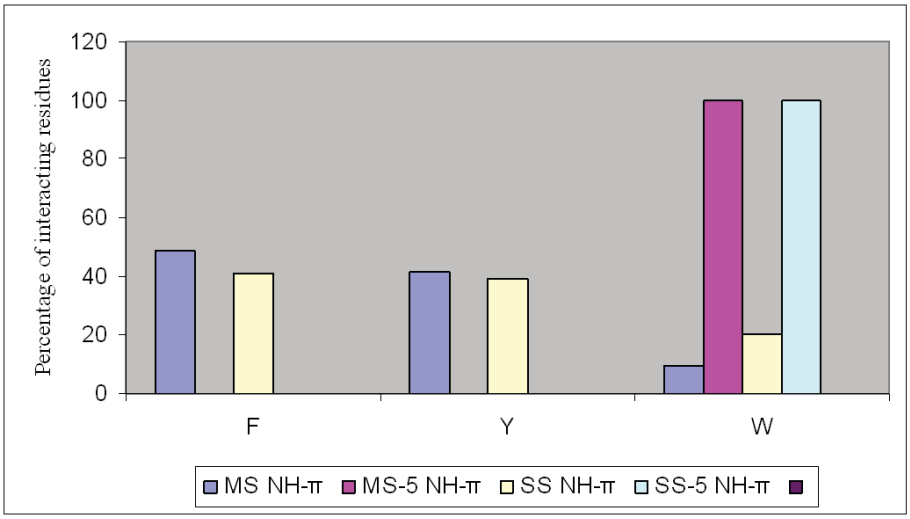

Fig. (4). Acceptor residues in N-H…π interactions.

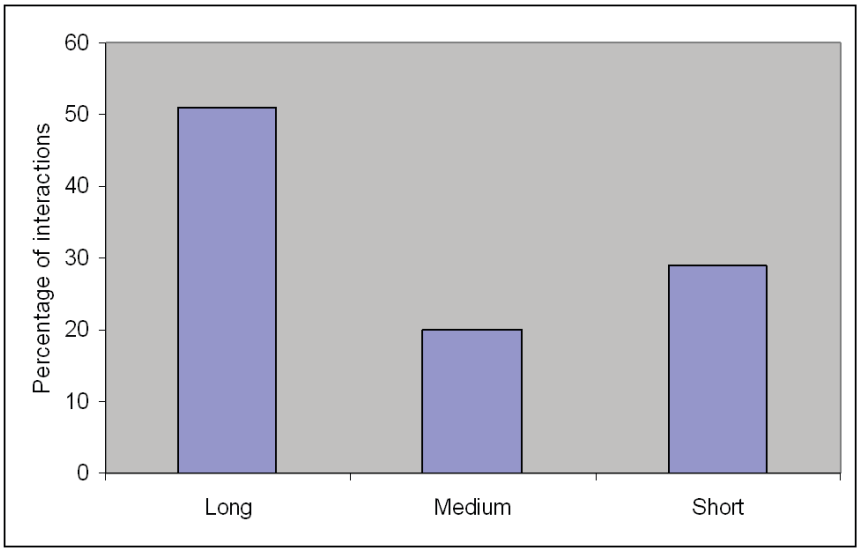

Fig. (5). Sequential separation in N-H…π interacting residues.

>± 4 residues are treated as long range contacts [33-34]. 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. The sequential distance between residues that contributed donor and acceptor residues to inter-residue N-H…π interactions was calculated and results are depicted in Fig. (**5**). 51%, 20% and 29% of inter-residue N-H…π interactions were found to be long-range, medium-range and short-range interactions. Long-range N-H…π interactions are the predominant type of interactions in RNA binding proteins. These observations were consistent with the results obtained with cation-π [28, 29] and C-H…π interactions [30] in the same data set. The contribution of the medium-range and short-range interactions is comparatively less in all the four types of N-H…π interactions studied. Hence the folding of the protein molecule might be mediated by N-H…π interacting amino acid residues that are located far apart at the sequence level and thus these interactions contribute significantly to the global conformational stability of RNA binding proteins.

3.3. Secondary Structure Preference

The occurrence of weak interactions has been observed at the terminus of the secondary structural units, in particular α-helix and β-sheets [13-14]. These interactions play a definitive role in stabilizing these structures of proteins. The propensity of the amino acid residues to favor a particular conformation has been well documented [44]. Such conformational preference is not only dependent on the amino acid alone but is also dependent on the local amino acid sequence. We analyzed the secondary structure preference of each amino acid which participates in all the four different types of N-H…π interactions. The secondary structure preference of each of the amino acids involved in all the above said types of N-H…π interactions were obtained from PDB and the results are depicted in Fig. (**6**). We found that, Arg, Asn, Asp, Gln, Glu, Ile, Leu, Lys and Ser preferred to be in helix. Ala, Cys, Phe, Thr, Trp, Tyr and Val were found to prefer strand conformation. Gly, Met and Pro preferred to be in turns. The secondary structure preferences for most of

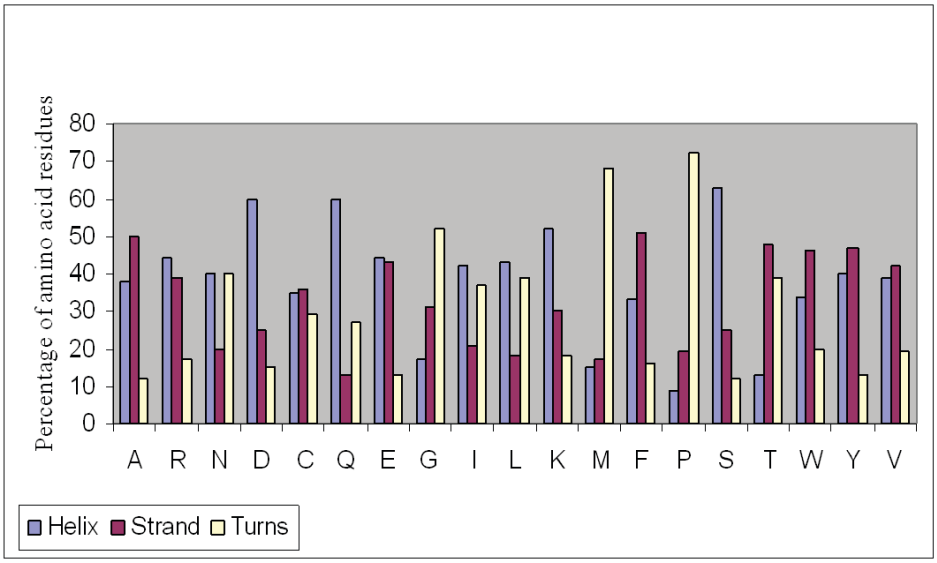

Fig. (6). N-H…π interacting residues in secondary structure.

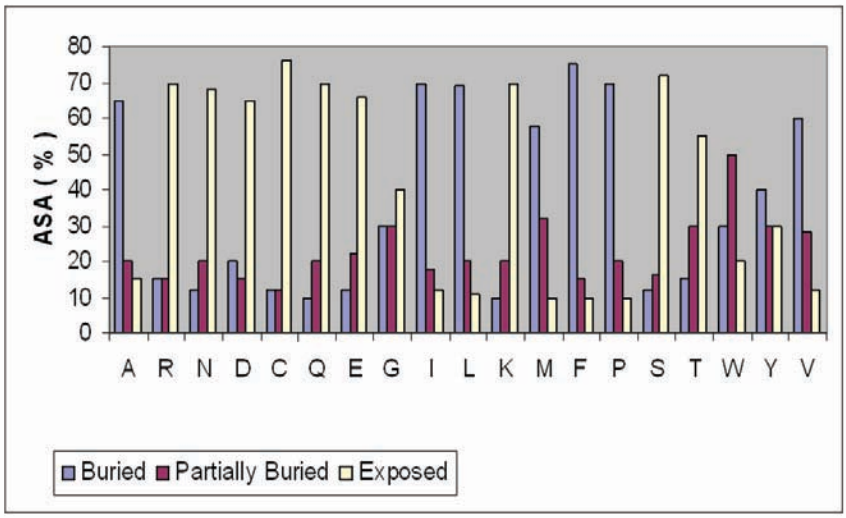

Fig. (7). Solvent accessibility range in N-H…π interacting residues.

the N-H…π interacting residues were consistent with the information available in literature [44].

3.4. Solvent Accessibility Calculations

We have estimated the solvent accessibility of all the amino acid residues that were involved in N-H…π interactions with the aid of ASA-View [36]. The relation between the amino acid residues in N-H…π interactions and solvent accessibility are shown in Fig. (**7**). The solvent accessibility of amino acid residues has been categorized as buried, partially buried and exposed [38-39]. We found that, of the different amino acids that were involved in N-H…π interactions, Arg, Asn, Asp, Cys, Gln, Glu, Gly, Lys, Ser and Thr were in the exposed regions. Ala, Ile, Leu, Met, Phe, Pro, Tyr and Val residues involved in N-H…π interactions were in the buried regions while Trp preferred to be in partially buried region. We found that most of the polar amino acid residues involved in N-H…π interactions were solvent exposed and most of the non polar residues involved in N-H…π interactions were excluded from the solvent. Since the percentage of polar residues involved in N-H…π interactions is much higher than the non-polar residues, the polar residues might contribute significantly to the stability of the RNA binding proteins as the contribution of the global energy is much greater in solvation.

3.5. Conservation Score

We used the ConSurf server [39] to compute the conservation score of amino acid residues involved in N-H…π interactions in RNA binding proteins, and the results are shown in Fig. (**8**). The server computes the conservation based on the comparison of the sequence of a PDB chain with the proteins deposited in Swiss-Prot [40] and finds the ones that are 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 evolutionary 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 residues with a score of 9 are considered highly conserved. Conservation score of ≥ 6 is the cutoff value used to identify the stabilizing residues. 25% of the amino acid residues that contributed donor residues in N-H…π interactions had the highest conservation score of 9,

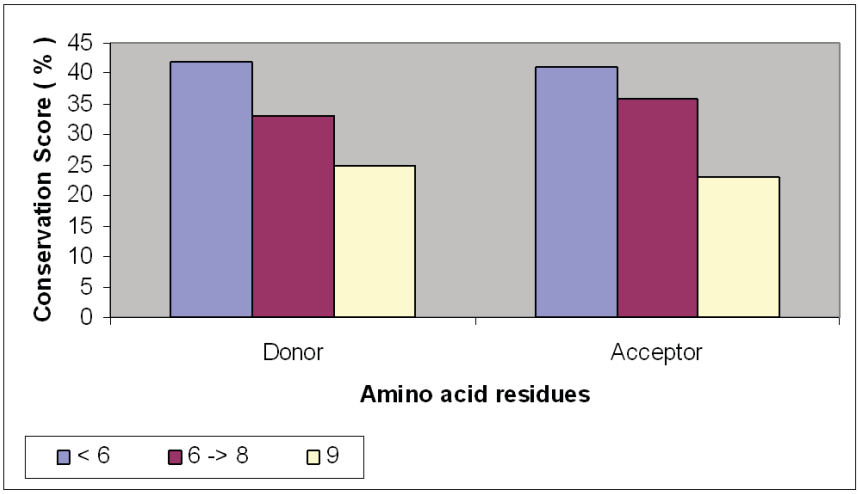

Fig. (8). Conservation score in N-H…π interacting residues.

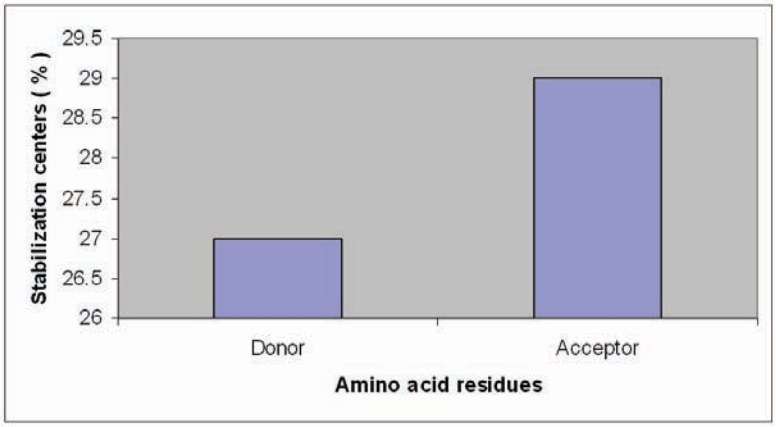

Fig. (9). Stabilization centers in N-H…π interacting residues.

while 33% of the amino acid residues had a conservation score, in the range of 6 to 8. Thus, 58% of the donor amino acid residues had a higher conservation score. In the case of amino acid residues, that contributed acceptor residues in N-H…π interactions, 23% of the acceptor amino acid residues had the highest conservation score of 9, while 36% of the amino acid residues had a conservation score, in the range of 6 to 8. Thus, 59% of the acceptor amino acid residues had a higher conservation score. From this we were able to infer that, most of the amino acid residues involved in N-H…π interactions might be conserved in RNA binding proteins.

3.6. Stabilization Centers

Stabilization centers are clusters of residues that are involved in medium or long range interactions [41]. Residues can be considered part of stabilization centers if they are involved in medium or long range interactions and if two supporting residues can be selected from both of their flanking tetra peptides, which together with the central residues form at least seven out of the nine possible contacts. We used the SCide [42] server for computing the stabilization centers in the RNA binding proteins and the results are depicted in Fig. (**9**). We found that 27% of the amino acid residues that contribute donor residues to N-H…π interactions had one or more stabilization centers in addition to their contribution to N-H…π interactions and similarly 29% of the amino acid residues that contribute acceptor residues to N-H…π interactions had one or more stabilization centers in addition to their contribution to N-H…π interactions. From this we infer that, these residues might contribute additional stability to the RNA binding proteins in addition to their participation in N-H…π interactions.

3.7. Binding Residues

The binding sites were available for 45 protein-RNA complexes studied and the results for N-H…π forming residues are shown in Table **1**. We observed binding at the interface which was also involved in the N-H…π interactions. Only 2% of the RNA binding amino acid residues were involved in N-H…π interactions. Hence the specificity is shown by residues that are not involved in the

Table 1. Binding Residues in RNA Binding Proteins

PDB ID	Binding Residues with RNA Bases
1ASY-A	Gln 120, Gln 121, Lys 142, Lys 155, Pro 178, Lys 180, Ser 181, Glu 188, Glu 202, Asn 227, Thr 230, Asn 328, Ser 329, Asn 330, Thr 331, His 334, Met 335, Thr 424, Glu 478, Ser 481, Arg 531.
1ASZ-A	Arg 119, Thr 124, Gln 138, Pro 178, Lys 180, Ser 181, Asn 227, Thr 230, Ser 280, Glu 281, Ser 284, Ser 301, Glu 327, Asn 328, Ser 329, Asn 330, Thr 424, Lys 553.
1B23-P	Tyr 88, Asn 91, Trp 200 Thr 232, Arg 330, His 331, Thr 332, Gly 337, Arg 339, Thr 350, Lys 376, Glu 390, Gly 391.
1B7F-A	**Gln 134**, Arg 155, **Arg 158, Tyr 164**, Ser 165, Arg 195, Lys 197, Arg 202, Asn 212, Tyr 214, Asn 217, Gln 239, Asn 241, Arg 252, Arg 287, Ala 289.
1C0A-A	Arg 28, Asp 29, Ser 32, Gln 46, Arg 64, Asn 82, Asn 84, Leu 108, Asp 111, Asn 113, His 114, Asn 116, Thr 117, Glu 119, Thr 169, Pro 170, Ser 193, Glu 219, Asp 220, Arg 222, Ala 223, Arg 225, Arg 549, Thr 557, Thr 558.
1CVJ-A	Ser 12.
1CX0-A	Asn 15, Asn 16, Glu 19, Leu 49, Lys 50, Arg 52, Lys 80, Arg 83, Gln 85, Tyr 86, Lys 88, Asp 90, Ser 91, Asp 92.
1EFW-A	Arg 29, Asp 30, Leu 31, Gly 33, Gln 47, **Arg 64**, Arg 78, Glu 80, Asn 82, Glu 125, Lys 552.
1F7U-A	Glu 59, Lys 102, Phe 113, Asn 153, Lys 156, **Tyr 224**, Glu 294, Lys 319, Lys 340, Ser 341, Tyr 347, Arg 350, Gln 374, Gln 406, Lys 439, Lys 466, Lys 469, Gly 483, Asp 484, Tyr 488, Tyr 491, Arg 495, Ser 498, His 559, Ser 563, Tyr 565, Asp 566, Leu 568, Val 570, Ala 571, Met 607.
1F7V-A	Lys 439, Lys 466, Asn 469, Gly 483, Asp 484, Tyr 488, Tyr 491, Ser 494, Arg 495, Ser 498, His 559, Ser 563, Tyr 565, Leu 568, Val 570, Ala 571, Met 607.

(Table 1) contd…..

PDB ID	Binding Residues with RNA Bases
1F7Y- A	Lys 7, Arg 16, Thr 21, Gly 22, Thr 24, Arg 34, His 41, Asp 48, His 50, Ser 51, Arg 64, Tyr 68, Arg 71.
1FFY-A	Arg 440, Asn 443, Ser 556, Arg 560, Glu 593, Lys 595, Ser 624, Asp 626, Asp 630, Arg 632, Glu 636, Gln 640, Lys 647, Asn 650, Arg 653, **Gln 709**, Asn 710, Ser 717, Lys 725, Tyr 729, Arg 805, Lys 823, Arg 888.
1G59-A	Arg 47, Glu 107, Arg 163, Glu 172, Ser 181, Glu 207, Lys 243, Leu 272, Glu 282, Ser 299, Val 304, Pro 357, Arg 358, Arg 417, Gln 432, Arg 435, Leu 442, Thr 444, Leu 447.
1GAX-A	Thr 214, Glu 261, Tyr 337, Thr 560, Arg 566, Asp 568, Arg 570, Arg 576, Asn 580, Lys 581, Asn 584, Arg 587, Trp 642, Glu 651, Arg 818, Pro 826, Lys 831, Arg 843, Asn 847.
1GTS-A	Gln 13, Asp 66, Thr 68, Asn 69, Leu 124, Arg 133, Glu 168, Arg 192, Lys 194, Tyr 211, Gln 234, Glu 235, Ile 313, Thr 316, Lys 317, Gln 318, Asp 319, Thr 321, Asn 336, Arg 341, Asn 370, Gln 399, Lys 401, Arg 402, Arg 412, Asn 413, Gln 517, Arg 520, Arg 545, Thr 547.
1H2D-A	Gly 126, **Arg 134**.
1H3E-A	Asp 259, Trp 370, Arg 373, Ser 383, Asn 384, Ala 385, Arg 389, Asn 393, Gln 409, Arg 420.
1H4Q-A	Ser 126, Trp 127, Arg 128, Gln 245, Asp 246.
1H4S-A	Arg 125, Ser 126, Trp 127, Arg 128.
1HR0-B	Gly 99, Asn 104, Lys 179.
1I6U-A	Lys 32, Arg 36, Arg 78, Lys 82, Lys 83, Ser 105, Thr 107, Gly 122.
1IL2-A	Arg 28, Asp 29, Ser 32, Gln 46, Asn 82, Asn 83, Glu 93, Ile 108, Asp 111, Arg 225, Met 447, Glu 482, Arg 549, Thr 557, Thr 558.
1IVS-A	Thr 214, Val 215, Glu 261, Tyr 337, Thr 560, Arg 566, Asp 568, Arg 570, Arg 576, Lys 581, Asn 584, Arg 587, Trp 642, Cys 646, Lys 654, Arg 818, Arg 843.
1J2B-A	Ala 418, Gln 419, Glu 421, Asp 425, Lys 430, Thr 466, Arg 470, Lys 473, Arg 478, Thr 481, Asp 485, Thr 490, Glu 515, Asp 525, Lys 529, Arg 578.
1JBR-A	Arg 65, Asp 143.
1JID-A	Arg 14, Phe 15, Cys 17, **Tyr 19, Tyr 22**, Ile 29, Arg 33, Arg 70, Arg 101.
1K8W-A	Ser 24, His 43, Ala 46, Thr 63, Gln 67, Ala 128, Lys 130, Tyr 131, Gln 132, Gly 133, Lys 135, Arg 141, Arg 151, Lys 176, Tyr 179.
1KNZ-A	Asn 62, Thr 69, Ser 80, Arg 83, Asn 84, **Trp 87**, Asp 100, Lys 132, Ser 133, Ser 134, Ser 135.
1KOG-A	Lys 249, Thr 548, Asn 575, Lys 577, Arg 609.
1LNG-A	Met 1, Ile 2, Trp 4, Tyr 7, Ser 13, Arg 14, Arg 15, Arg 18, Lys 19, Glu 22, Lys 51, Arg 52, Pro 54, Arg 55, His 57, Asp 67, Tyr 68, Lys 69, Asn 71, Lys 72, Lys 77.
1MJI-A	Lys 36, Thr 71, Asp 126, **Asn 132**, Arg 153.
1MMS-A	Lys 10, Gln 30, Ala 75, Ser 76, Lys 80, Lys 87, Gly 88, Ser 89, Ser 90, Pro 92, Lys 93, Arg 94, Lys 112, Asn 117, Asn 119, Ile 127, Gly 130, Thr 131, Lys 133, Ser 134.
1N35-A	Ser 561, Arg 562, Asn 827.
1N38-A	Gly 560, Ser 561, Lys 566.
1N77-A	Thr 43, Asp 44, Arg 47, Glu 107, Arg 147, Glu 172, Lys 180, Ser 181, Tyr 187, Glu 207, Lys 243, Leu 272, Glu 282, Arg 297, Val 304, Pro 357, Arg 358, Arg 417, Arg 435, Leu 442, Thr 444, Leu 447.
1N78-A	Ser 9, Thr 43, Asp 44, Arg 47, Glu 107, Tyr 168, Lys 180, Ser 181, Glu 207, Lys 241, Leu 272, Glu 282, Arg 297, Val 304, Lys 309, Pro 357, Arg 358, Arg 417, Arg 435.
1NB7-A	Cys 14, Lys 141, Arg 158, Gln 446, Cys 451, Ser 556, Asp 559.
1NH3-A	Arg 105, Ser 118, Asn 159, Arg 196, Thr 215.
1O0B-A	Gln 996.
1O0C-A	Glu 990.
1QF6-A	Gly 203, Tyr 205, Met 214, Tyr 219, Arg 245, His 309, Ala 316, Gly 368, Arg 375, Tyr 462, Arg 476, Gln 484, Asp 549, Asn 575, Lys 599, Glu 600, Arg 609.
1QU2-A	Arg 440, Asn 443, Ser 556, Arg 560, Gly 593, Lys 695, Ser 624, Asp 626, Arg 632, Glu 636, Gln 640, Lys 647, Asn 650, Arg 653, **Gln 709**, Asn 710, Ser 717, Lys 725, Tyr 729, Arg 805, Lys 823, Arg 888.
1QU3-A	Phe 14, Arg 440, Ser 556, Arg 560, Gly 593, Arg 632, Asn 650, Arg 653, Asn 809.
1YRN-A	Ile 77, Arg 115, Asn 120, Arg 122, Arg 124.
2FMT-A	Thr 11, Pro 39, Arg 42, Gly 43, Lys 44, Gly 91, Leu 207, Lys 209, Lys 246, Gly 290, Lys 291, Asp 298, Asn 301, Ser 302, Arg 304.

Bolded residues are amino acid residues involved in N-H…π interactions.

N-H…π interactions. Thus, the N-H…π interaction forming residues play an important role in the stability of the RNA binding proteins, while the interaction with RNA is determined by residues that are not involved in N-H…π interactions.

4. CONCLUSION

In summary, the present work on the Non-canonical N-H…π interactions in RNA binding proteins indicates that there was an average of 12 N-H…π interactions per protein in the 59 proteins investigated and an average of one significant N-H…π interaction for every 37 amino acid residues in RNA binding proteins. We observed that 63% of the N-H…π interactions were inter-residue interactions, while the remaining 37% of the N-H…π interactions were found to be intra-residue interactions. We found that 47% of the interactions were MS interactions, 17% of the interactions were MS5 interactions, 27% of the interactions were SS interactions and the remaining 9% interactions were SS5 interactions. We conclude that the contribution of MS mediated interactions might be significant in N-H…π interactions irrespective of the amino acids involved. The sequential distance calculation revealed that long-range N-H…π interactions are the predominant type of interactions in RNA binding proteins. The secondary structure analysis of the amino acid residues involved in N-H…π interactions showed that Arg, Asn, Asp, Gln, Glu, Ile, Leu, Lys and Ser preferred to be in helix, while Ala, Cys, Phe, Thr, Trp, Tyr and Val preferred to be in strand conformation. The solvent accessibility results indicated that the polar amino acid residues involved in N-H…π interactions were solvent exposed and the non polar residues involved in N-H…π interactions were excluded from the solvent. Based on the conservation score we infer that most of the amino acid residues involved in N-H…π interactions might be conserved in RNA binding residues. Only 2% of the interacting residues are involved in RNA binding and hence the specificity is shown by residues that are not involved in N-H…π interactions. On the whole, these results will be a good starting point for studies on intra protein-protein interactions and for investigations on inter-residue and intra-residue interactions with protein molecules.

ACKNOWLEDGEMENT

The authors thank the management of the VIT University for the support and encouragement to carry out this work. The authors would also like to gratefully acknowledge Dr. Madan M Babu, Group Leader, MRC Laboratory, Cambridge, UK for helping with the NCI server.

REFERENCES

[1] Desiraju GR, Steiner T. The weak hydrogen bond in structural chemistry and Biology. Oxford University Press: Oxford, 1999.

[2] Brandl M, Weiss MS, Jabs A, Suhnel J, Hilgenfeld R. C-H…π interactions in proteins. J Mol Biol 2001; 307: 357-77.

[3] Wulf OR, Liddel U, Hendricks SB. The effect of ortho substitution on the absorption of OH group of phenol in the infrared. J Am Chem Soc 1936; 58: 2287-93.

[4] McPhail AT, Sim GA. Hydroxyl-benzene hydrogen bonding. An X-ray study. Chem Commun 1965; 124-5.

[5] Sutor DJ. The C-H…O hydrogen bonds in crystals. Nature 1962; 195: 68-9.

[6] Derewenda ZS, Lee L, Derewenda U. The occurrence of C-H…O hydrogen bonds in proteins. J Mol Biol 1995; 252: 248-62.

[7] Bella J, Humphries MJ. C_α-H…O=C hydrogen bonds contribute to the specificity of RGD cell-adhesion interactions. BMC Struct Biol 2005; 5: 4.

[8] Steiner T, Koellner G. Hydrogen bonds with pi-acceptors in proteins: frequencies and role in stabilizing local 3D structures. J Mol Biol 2001; 305: 535-57.

[9] Levitt M, Perutz MF. Aromatic rings act as hydrogen bond acceptors. J Mol Biol 1998; 201: 751-4.

[10] Burley SK, Petsko GA. Aromatic-aromatic interaction – A mechanism of protein-structure stabilization. Science 1985; 229: 23-8.

[11] Toth G, Watts CR, Murphy RF, Lovas S. Significance of aromatic-backbone amide interactions in protein structure. Proteins 2001; 43: 373-81.

[12] Taylor R, Olga K. Crystallographic evidence for the existence of C-H…O, C-H…N, and C-H…Cl hydrogen bonds. J Am Chem Soc 1982; 104: 5063-70.

[13] Fabiola GF, Krishnaswamy, Nagarajan V, Pattabi V. C-H…O hydrogen bonds in beta sheets. Acta Crystallog Sect D 1997; 53: 316-20.

[14] Babu MM, Singh KS, Balaram P. C-H…O hydrogen bond stabilized polypeptide chain reversal motif at the C terminus of helices in proteins. J Mol Biol 2002; 322: 871-80.

[15] Jiang L, Lai L. C-H…O hydrogen bonds at protein-protein interfaces. J Biol Chem 2002; 277: 37732-40.

[16] Chakrabarti P, Chakrabarti S. C-H…O hydrogen bond involving praline residues in alpha helices. J Mol Biol 1998; 284: 867-73.

[17] Singh SK, Babu MM, Balaram P. Registering alpha-helices and beta-strands using backbone C-H…O interactions. Proteins 2003; 51: 167-71.

[18] Chakrabarti P, Bhattacharyya R. Geometry of non-bonded interactions involving planar groups in proteins. Prog Biophys Mol Biol 2007; 95: 83-137.

[19] Scheiner S, Kar T, Gu Y. Strength of the C-H…O hydrogen bond of amino acid residues. J Biol Chem 2001; 276: 9832-37.

[20] Vargas R, Garza J, Dixon DA, Hay BP. How strong is the C-H…O=C hydrogen bond? J Am Chem Soc 2000; 122: 4750-5.

[21] Duan G, Smith VH Jr, Weaver DF. An *ab initio* and data mining study on aromatic-amide interactions. Chem Phys Lett 1999; 310: 323-32.

[22] Wlodawer A, Waler J, Huber R, Sjolin L. Structure of bovine pancreatic trypsin inhibitor results of joint neutron and X-ray refinement of crystal form II. J Mol Biol 1984; 180: 301-29.

[23] Perutz MF. The role of aromatic rings as hydrogen-bond acceptors in molecular recognition. Phil Trans R Soc Ser A 1993; 345: 105-12.

[24] Armstrong KM, Fairman R, Baldwin RL. The (i, i+4) Phe-His interaction studied in an alanine-based α-helix. J Mol Biol 1993; 230: 284-91.

[25] Kryger G, Silman I, Sussman JL. Structure of acetyl choline estrase complexed with E2020 (Aricept): implication for the design of new-Alzheimer drug. Structure 1999; 7: 297-307.

[26] Parkinson G, Gunasekera A, Vijtechovsky J, *et al*. Aromatic hydrogen bond in sequence specific protein DNA recognition. Nat Struct Biol 1996; 3: 837-41.

[27] 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 tyrosine 6 and its influence in catalysis. J Am Chem Soc 1993; 115: 7910-11.

[28] Anbarasu A, Anand S, Mathew L, Sethumadhavan R. Influence of cation- π interactions on RNA-binding proteins. Int J Biol Macromol 2007; 40: 479-83.

[29] Anbarasu A, Sethumadhavan R. Invesigations on the role of Cation- π interactions in glycoproteins lipid binding proteins and RNA-Binding Proteins. J Theor Biol 2007; 247: 346-53.

[30] Anbarasu A, Anand S, Babu MM, Sethumadhavan R. Invesigations on C-H…π interactions in RNA-binding proteins. Int J Biol Macromol 2007; 41: 251-9.

[31] Berman HM, Westbrook JZ, Feng Z, *et al.* The protein data bank. Nucleic Acids Res 2000; 28: 235-42.

[32] Babu MM. NCI: A server to identify non-canonical interactions in protein structures. Nucleic Acids Res 2003; 31: 3345-8.

[33] Gromiha MM, Selvaraj S. Influence of medium and long range interaction in different structural classes of globular proteins. J Biol Phys 1997; 23: 151-62.
[34] Gromiha MM, Selvaraj S. Inter-residue interactions in protein folding and stability. Prog Biophys Mol Biol 2004; 86: 235-77.
[35] Selvaraj S, Gromiha MM. Role of hydrophobic clusters and long-range contact networks in the folding of (alpha/beta) 8 barrel proteins. Biophys J 2003; 84: 1919-25.
[36] Shandar A, Gromiha MM, Fawareh H, Sarai A. ASAView: solvent accessibility graphics for proteins. BMC Bioinformatics 2004; 5: 51.
[37] Gilis D, Rooman M. Stability changes upon mutation of solvent-accessible residues in proteins evaluated by database-derived potentials. J Mol Biol 1996; 257: 1112-26.
[38] 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. J Mol Biol 1997; 272: 276-90.
[39] Glaser F, Pupko T, Paz I, *et al*. 2003. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 2003; 19: 163-4.
[40] Boeckman B, Bairoch A, Apweiler R, *et al.* The SWISS-PROT Protein knowledgebase and its supplement TrEMBL. Nucleic Acids Res 2003; 31: 365-70.
[41] Dostanyi ZS, Fiser A, Simon I. Stabilization centers in proteins: identification, characterization and predictions. J Mol Biol 1997; 272: 597-612.
[42] Dosztanyi ZS, Magyar CS, Tusnady EG, Simon I. SCide: identification of stabilization centers in proteins. Bioinformatics 2003; 19: 899-900.
[43] Hoffman MM, Khrapov MA, Cox JC, Yao J, Tong L, Ellington AD. The amino-acid nucleotide interaction database (AANT) Nucleic Acids Res 2004; 32: D174-181.
[44] Malkov SN, Zivkovic VM, Beljanski MV, Hall MB, Zaric SD. A reexamination of the propensities of amino acids towards a particular secondary structure: classification of amino acids based on their chemical structure. J Mol Model 2008; 14: 769-75.

Received: June 20, 2008 Revised: September 19, 2008 Accepted: September 25, 2008