# AROMATIC Π−NETWORKS IN SM/LSM PROTEIN INTERFACES

### *UDC 547.96 : 541.6*

# Luka M. Breberina[1], Milan R. Nikolić[1], Srđan Đ. Stojanović[2]

[1]Faculty of Chemistry, University of Belgrade, Serbia
[2]ICTM - Department of Chemistry, University of Belgrade, Serbia

**Abstract**. *In this work, we have analyzed the influence of π–π interactions on stability and properties of Sm/LSm assemblies. Phe residues were found to be involved in π–π interactions much more frequently than Tyr or His. Similarly, the Phe–Phe π–π interacting pair had the highest frequency of occurrence. Furthermore, a significant number of π-networks were observed at the interface of Sm/LSm proteins. Generally speaking, the distance between the interacting pairs was in the range of 5–6 Å. 3π and 7π-networks were found to frequently have plane–plane angles less than 60º. Solvent accessibility pattern of Sm/LSm proteins revealed that all of the interacting residues were from buried areas. Moreover, most of the π–π interacting residues of Sm/LSm proteins were evolutionary conserved and were in the strand regions. A high percentage of these residues could be considered as stabilization centers that (significantly) contribute to the net stability of Sm/LSm proteins.*

**Key words**: *aromatic π–networks, Sm/LSm proteins, interfaces, stabilization centers, conservation score*

## 1. INTRODUCTION

Noncovalent inter- and intramolecular interactions involving aromatic rings are ubiquitous in chemical and biological systems, and span from molecular recognition to self-assembly, and to catalysis and transport [1]. Interaction between the arene systems (π–π) has been recognized as a key stabilizing force in wide-ranging fields spanning molecular biology, crystal engineering, and material design [1-3]. Aromatic π–π interactions not only determine biological structures but also modulate the physical properties of residues at enzyme active sites [4]. The nature of π–π interaction was primarily thought to be dispersive with notable electrostatic contribution depending on the system in question [5]. At the supramolecular level, the aromatic rings can interact in different ways: stacked arrangement

(face-to-face, perfect alignment, offset, slipped, parallel displaced) and edge-to-face, T-shaped conformation [6]. The calculated $\pi$–$\pi$ interaction energies of the parallel, edge–face (T-shaped) and offset stacked are -1.48, -2.46 and -2.48 kcal/mol, respectively [7], and the major source of attraction is not short-range (such as charge-transfer), but long-range interactions (quadrupole–quadrupole electrostatic and dispersion) [8]. It has been suggested that the perpendicular and the parallel-displaced configurations are more common than the sandwich geometry as these, especially the former one exposes three aromatic faces to the outside, offering greater possibility for additional interactions with other groups [9]. Aromatic residues show a high tendency towards forming clusters beyond the dimer which has a significant influence on protein folding, structure, and stability [10, 11].

The Sm family of proteins, encompassing the Sm and Sm-like (LSm) proteins [12], are common participants in RNA metabolism in Eubacteria [13], Archaea [14, 15], and eukaryotes [16, 17]. Sm proteins primarily occur as small (~9–29 kDa) stand-alone proteins lacking other domains [18] that assemble to form characteristic homomorphic or heteromorphic rings containing six or seven proteins. Members of the family are characterized by the conserved bipartite Sm domain or ''Sm fold'' which functions, at least in part, in binding to neighboring Sm proteins within such rings [12, 19, 20]. All Sm proteins form structures of a higher order which can be defined or none defined. In general they are very stable and sometimes the presence of chaotropic agent is necessary for their disruption [14, 17]. We have previously reported contribution of hydrogen bonds, salt bridges and non-canonical interactions to the stability of Sm oligomers [21, 22]. In our work [23], we showed that the hot spots of Sm proteins are located within densely packed regions; these are highly conserved and have large energy contributions to the interface interactions.

In an effort to search for the factors that contribute to the affinity and specificity of protein–protein interactions, many previous studies were aimed at the analysis of the properties of protein–protein interfaces. This manuscript expands on our previous work on the non-canonical interactions of Sm/LSm proteins [21, 22] by analyzing the same class of proteins with respect to $\pi$–$\pi$ interactions. We have systematically analyzed the influence of $\pi$–$\pi$ interactions to the stability of Sm/LSm proteins. We have focused our study at the protein interface and hence the $\pi$–$\pi$ interactions within a protein are not considered. Results from this study might be used for understanding of structure-function relationships, and can provide a new dimension of molecular recognition and self-assembly.

## 2. MATERIALS AND METHODS

### 2.1. Dataset

For this study we used the Protein Data Bank (PDB) 08 January 2014 list of 97746 structures [24]. The selection criteria for a Sm/LSm proteins to be included in the dataset were based on the following criteria: (1) no theoretical model structures and no NMR structures were accepted; these structures were not included since it was difficult to define the accuracy of the ensemble of structures in terms of displacement that was directly comparable to the X-ray diffraction studies. (2) only crystal structures with the resolution of 3.0 Å or better and a crystallographic R-factor of 25.0% or lower were accepted, and (3) crystal structures of proteins containing Sm-like fold (SCOP Classification, version 1.75) [25] without RNA binding were accepted. If not already present, all hydrogen atoms were added and optimized using the program REDUCE [26] with default settings. When

multiple alternative conformations of certain residues were present, as indicated by the altLoc field in the PDB file, only the first conformation was considered. Using these criteria, we created a dataset of 17 Sm/LSm proteins. The PDB IDs are as follows: 1b34, 1d3b, 1h64, 1hk9, 1i4k, 1i8f, 1jbm, 1kq1, 1m5q, 1mgq, 1n9r, 1th7, 1u1s, 1ycy, 2vgn, 3bdu, and 3pgg. To reduce biased statistics, caused by the lack of hetero-oligomer proteins in the dataset, we did not divide dataset into homo and hetero sub-datasets. In order to have a non-redundant set of interfaces, we used PISCES sequence culling server [27]. We excluded the interfaces that contain more than 35% sequence identity. After the interface dataset had been assembled, several interfaces that contained ligands were rejected, leaving 215 interfaces that were actually used as the dataset in our analysis.

### 2.2. π–π interaction analysis

Aromatic-Aromatic Interactions Database ($A^2ID$) [10] was used for the calculation of various types of π–π interactions and their geometrical features with default settings (Fig. 1). To represent the π–π network, three parameters have been considered: centroid to centroid distance between aromatic ring pair ($R$), interplanar angle ($\phi$) and the centre of the aromatic ring (one aromatic ring surrounded by number of aromatic rings). In a given protein structure the program scans from N-terminal to C-terminal and identifies the first aromatic residue. A virtual sphere is created from the centroid of the π-ring of aromatic residue with a radius defined by the cutoff distance ($R$). If π-ring of other aromatic residues falls within this virtual sphere, they were considered as a part of that network and so on. Thus the nearest aromatic neighbors for each aromatic residue were calculated based on distance criteria. The aromatic residue present in one π-network was not considered in other π-networks in the protein. A lower cut-off distance of 2.5 Å is taken to differentiate the centroid of two rings of tryptophan in the calculations and is treated individually as a π-system. The network was defined as 2π if the program could locate the centroid of another aromatic residue within the cutoff distance. Similarly, the network was defined as 3π if the centroid of a third aromatic residue was also located within the cutoff distance from the centroid of the first aromatic ring or the second aromatic ring and so on for 4π and higher π-network. Thus, a ring in an nπ network will have at least a single connectivity with any of the n−1 aromatic networks. In case no new ring is found within the cutoff distance of a certain nπ networks then the network is called nπ network. Another important geometrical parameter between a pair of aromatic residues is the angle between two π-planes ($\phi$) as defined in Fig. 1. The $\phi$ angles close to 0º and 180º correspond to a situation where the two aromatic rings of neighboring residues are parallel to each other. The aromatic rings of the amino acid side chains of His, Phe, Tyr and Trp are considered to be π-systems.
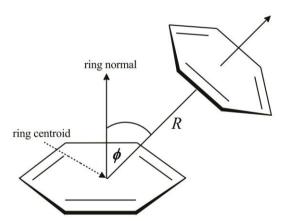


**Fig. 1** Parameters for π–π interactions: $R$ is the distance between the two π-systems; $\phi$ is the interplanar angle between normals to the planes of π-systems.

### 2.3. Secondary structure and solvent accessibility studies

The secondary structure and solvent accessibility (ASA) of the amino acid residues were among the key factors that were essential for understanding the environmental and structure–function relationship of proteins. Hence, a systematic analysis of each interaction forming a residue was performed based on its location in different secondary structures of Sm/LSm proteins and their solvent accessibility. We used the program DSSP [28] to obtain information about secondary structures and solvent accessibility. The secondary structures have been classified into alpha helix, beta turn, beta strand and coil, as suggested by the DSSP output. Solvent accessibility is the ratio between the solvent accessible surface area of a residue in a 3D structure and in an extended tripeptide conformation. Solvent accessibility was divided into three classes: buried (0–20%), partially buried (20–50%), and exposed (>50%), indicating respectively; the least, moderate and high accessibility of the amino acid residues to the solvent.

### 2.4. Computation of conservation of amino acid residues

We computed the conservation of amino acid residues in each protein using the ConSurf server [29]. This server computes the conservation based on the comparison of the sequence of a PDB chain with the proteins deposited in Swiss–Prot [30] 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 were evolutionary related to each one of the proteins in the dataset 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 to be highly variable and residues with a score of 9 are considered to be highly conserved.

### 2.5. Computation of stabilization centers

Stabilization centers are the clusters of residues that make cooperative, non-covalent and long-range interactions [31]. Thus, they are likely to play an important role in maintaining the stability of protein structures. Residues can be considered as parts of stabilization centers if they are involved in medium or long-range interactions and if two supporting residues can be selected from their C and N terminal flanking tetrapeptides, which together with the central residues form at least seven out of the nine possible contacts. We used an online server, available at http://www.enzim.hu/scide [32], to analyze the stabilization centers of interaction–forming residues. This server defines the stabilization center based on the following criteria: (1) two residues are in contact if there is at least one heavy atom–atom distance smaller than the sum of their van der Waals radii plus 1 Å; (2) a contact is recognized as "long-range" interaction if the interacting residues are at least ten amino acids apart; (3) two residues form a stabilization center if they are in long-range interaction, and if it is possible to select one–one residues from both flanking tetrapeptides of these two residues that make at least seven contacts between these two triplets [32].

Figures 2 and 3 were prepared using the program Discovery Studio Visualizer 4.0 [33].

3. RESULTS AND DISCUSSION

In this study we have focused on the analysis of π–π interactions at Sm/LSm protein interfaces. We have studied: (1) preference of aromatic residues to take part in π–π interactions, (2) interaction geometry, (3) solvent accessibility and secondary structure preferences, (4) conservation score of interacting residues and (5) stabilization center residues.

### 3.1. Preference of aromatic residues to take part in π–π interactions

There were a total of 713 π–π interactions between aromatic amino acid residues (Phe, Tyr and His), observable at the interfaces of Sm/LSm proteins, included in the studied dataset (215 in total). In the case of most of the considered interfaces (e.g. the structures with PBD ID codes 1d3b, 1i4k, 1kq1, 1m5q, 1mgq, 1n9r and 1th7), a large number π–π interactions was found. Nonetheless, in some cases, no π–π interactions could be detected. An illustrative example of a typical π–π interaction involving two aromatic side groups is shown in Fig. 2.

The frequencies of occurrence of Phe, Tyr and His residues at Sm/LSm protein interfaces as well as the frequency of π–π interaction involving specific aromatic amino acid residues is summarized in Table 1. The most abundant π residues at the Sm/LSm interface were Phe and Tyr, which made up for about 8% of the total interface area. The contribution of His was somewhat smaller. This result is consistent with earlier report on Sm/LSm proteins [23].

**Table 1** Frequency of occurrence of residues from Sm/LSm protein interfaces, involved in π–π interactions

| Amino acid | N | % | $N_{\pi-\pi}$ | $\%_{\pi-\pi}$ | π-networks (%) |
|---|---|---|---|---|---|
| His | 312 | 3.3 | 29 | 4.1 | 3π (100) |
| Phe | 408 | 4.4 | 391 | 54.9 | 2π (91); 3π (9) |
| Tyr | 389 | 4.2 | 293 | 41.0 | 2π (84); 3π (8); 7π (8) |
| Total | 1109 | 11.9 | 713 | 100 | – |
| *Pair (π–π)* | | | | | |
| His–His | – | – | 0 | 0 | – |
| His–Phe | – | – | 7 | 1.5 | 3π (100) |
| His–Tyr | – | – | 16 | 3.4 | 3π (100) |
| Phe–Phe | – | – | 328 | 69.9 | 2π (97); 3π (3) |
| Phe–Tyr | – | – | 39 | 8.4 | 2π (82); 3π (18) |
| Tyr–Tyr | – | – | 79 | 16.8 | 2π (51); 3π (11); 7π (38) |
| Total | – | – | 469 | 100 | – |

N, the number of occurrences of the specific amino acid in whole database; %, percent of occurrence of the specific amino acid in whole database; $N_{\pi-\pi}$, number of π–π interactions in Sm/LSm protein interfaces; $\%_{\pi-\pi}$, percent of π–π interactions in Sm/LSm protein interfaces; π-networks, occurrence of aromatic residues in different π-networks.

Although the frequency of occurrence of Phe and Tyr was comparable, Phe was more frequently involved in π–π interactions. There are only a few cases histidine residues were involved in π–π interactions. Based on this, it seems that Phe residues are more important for π–π interactions in Sm/LSm proteins than other aromatic residues. The greater electronegativity of sp$^2$ C relative to H in aromatic (benzene) core produces substantial C−H$^+$ dipole. It seems that C−H dipole accounts well for π–π interaction in phenylalanine [34].

Of six possible π−π interacting pairs (Table 1), Phe−Phe pair had the highest frequency of occurrence (69.9%), followed by Tyr−Tyr (17%) and Phe−Tyr (8%). Interestingly, there were no His−His interactions.
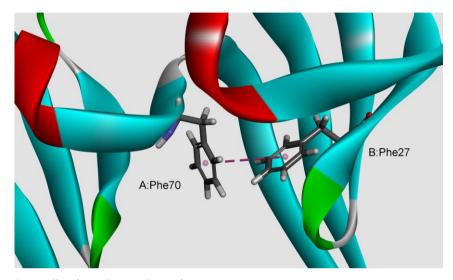


**Fig. 2** Details of π−π interactions of the human small nuclear ribonucleoprotein (snRNP) (PDB ID code 1d3b). The π−π interaction is marked with the pink dashed line (A:Phe70 − B:Phe27).
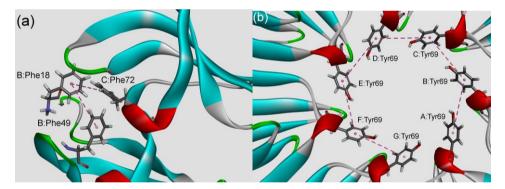


**Fig. 3** Details of π-networks in Sm/LSm proteins. a) 3π-network (triangular motif) in 1n9r protein (B:Phe49 − B:Phe18 − C:Phe72). b) 7π-network (2222211 motif) in 1m5q protein (A:Tyr69 − B:Tyr69 − C:Tyr69 − D:Tyr69 − E:Tyr69 − F:Tyr69 − G:Tyr69). π−π interactions are marked with pink dashed lines.

The frequency of Phe−Phe and Tyr−Tyr π−π interactions implicate their importance for the stability of Sm/LSm proteins. The specific arrangement or connectivity of π-clusters in proteins could significantly influence their structural stability. π-Clusters found at the surface of thermophilic proteins or clusters of aromatic residues, buried in the

globular proteins, may serve as examples of this. It has also been shown that addition of an aromatic pair on the protein surface increases its stability [35]. For that reason, we have also sought for the possible π-networks at the interfaces of Sm/LSm proteins (Table 1).
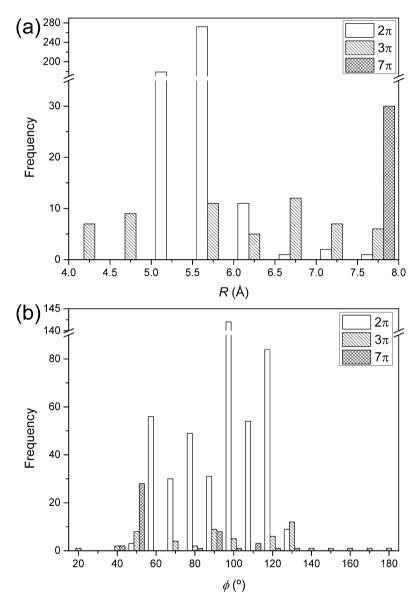


**Fig. 4** Interaction geometry. a) Distance distribution of π–π interactions. b) Angle distribution of π–π interactions.

In Sm/LSm proteins, small π-networks (2π) are preferred over large π-networks illustrating that π-networks are closely packed in the proteins (Table 1). Relatively high preference for triangular architecture (3π) further reduces probability for higher membered ring formation. However in some cases, Tyr residue is involved in large π–π networks (7π). A large π-network can enhance the stability of a protein conformation and can have a considerable influence on protein–ligand interactions. An illustrative example of 3π and 7π-networks in the dataset of proteins is shown in Fig. 3.

### 3.2. Interaction geometry

The geometrical properties of residues involved in π–π interactions are quantified in terms of the parameters ($R$, $\phi$) described in the Materials and Methods section. We have analyzed the distribution of the distance ($R$) and angle ($\phi$) parameters of π–π interacting pairs (Fig. 4). The distribution of distances (Fig. 4a) shows that in larger π-networks the highest frequency of occurrence shifts towards higher R values: the most frequent R values observed in 2π, 3π and 7π-networks were 5.5 Å, 6.75 Å and 7.75 Å, respectively. The number of small 2π-networks decreases as the centroid distance increases (>6 Å) which show that the small π-networks are compact and play an important role in the stability of proteins. Fig. 4b provides an overview of the plane–plane angles ($\phi$) in the Sm/LSm proteins. An analysis of 2π aromatic packaging showed a preference for nearly T-shaped (perpendicular) orientation, i.e. $\phi$ angle between 60° and 120°. However, in 3π and 7π-networks, plane–plane angles less than 60°–these indicate coplanarity that could maximize π–π stacking and packing [36]–were observed in relatively high number of cases. From this, one may infer that as the length of the π–π networks increases, the occurrence of the π–π stacking becomes more frequent.

### 3.3. Solvent accessibility and secondary structure preferences

We have carried out a systematic analysis of the solvent accessibility patterns for the aromatic residues in Sm/LSm proteins using DSSP as described in Subsection 2.3. We observed that all aromatic residues, involved in π–π interactions, preferred to be in solvent buried region (ASA <20%). Hence, π–π interacting residues stabilize the interface regions in these proteins.

To understand the interactions that confer stability of one specific secondary structure in protein, it is important to know the conformational preferences of amino acids involved. The propensity of the amino acid residues to favor a particular conformation is well described. Such a conformational preference is not only dependent on the amino acid alone, but also on the local amino acid sequence [37]. Thus, we have conducted a systematic analysis of the preference and pattern for Phe, Tyr and His residues (which were involved in π–π interactions) from Sm/LSm interfaces. The analysis was based on the occurrence of mentioned π–π interacting residues in different secondary structures (helix, strand and turn).

According to the results summarized in Table 2, Phe, Tyr and His predominantly prefer strand conformations. On the other hand, the frequency of occurrence of π–π interacting His and Phe in helix/strand secondary structures was similar. Hence, the preference of an amino acid to form π–π interactions in particular secondary structure is not the same as the preference of the amino acid for a particular secondary structure [38].

This analysis indicates that the π−π interactions do not occur at random but have a residue-specific preference for a particular secondary structure.

**Table 2** Frequency of occurrence of π−π interaction forming residues
in different secondary structures

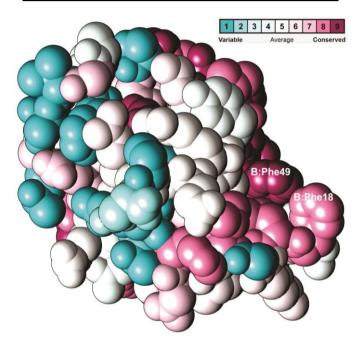| Amino acid | Helix (%) | Strand (%) | Turn (%) |
|---|---|---|---|
| His | 9.5 | 80.2 | 10.3 |
| Phe | 11.4 | 86.1 | 2.5 |
| Tyr | 0.0 | 100.0 | 0.0 |



**Fig. 5** Conservation pattern of small nuclear ribonucleoprotein F (SmF) from
*Saccharomyces cerevisiae* (PDB ID code 1n9r; Chain B) using Chimera.
Conservation score of B:Phe18 and B:Phe49 residues is 8 and 9, respectively.

### 3.4. Conservation score of interacting residues

The level of evolutionary conservation was often used as an indicator for the importance of certain position in maintaining the protein's structure and/or function [39]. Hence, we used the ConSurf server to compute the conservation score of residues involved in π−π interactions in Sm/LSm protein interfaces. Among the π−π interacting residues, 78.3% of them had a conservation score of higher or equal to 6; this is the cutoff value used to identify the stabilizing residues. 22.7% of the residues had the highest score of 9. The most of other residues comprising mentioned interfaces also show a great degree of conservation. Analysis of the conservation patterns of π−π interactions have shown that the 7π-network interactions have been conserved more than 2π or 3π interactions.

**Table 3** Involvement of stabilizing center residues in π–π interactions of Sm/LSm interfaces

| Amino acid | $N_{\pi-\pi}$ | SC | $SC_{\%}$ |
|---|---|---|---|
| His | 29 | 1 | 3.4 |
| Phe | 391 | 166 | 42.4 |
| Tyr | 293 | 105 | 35.8 |
| Total | 713 | 272 | 38.2 |

$N_{\pi-\pi}$, number of π–π interactions in Sm/LSm protein interfaces; SC, number of SC residues involved in π–π interactions; $SC_{\%}$, % of SC residues involved in π–π interactions.

Thus, it seems the majority of the residues involved in π–π interactions is evolutionarily conserved and might have a significant contribution towards the stability of Sm/LSm proteins. The conservation grade of π–π interacting residues in small nuclear ribonucleoprotein F (SmF) from *Saccharomyces cerevisiae* (PDB ID code 1n9r; Chain B) is shown in Fig. 5 (this is generated using Chimera [40]).

### 3.5. Stabilization center residues

Stabilization centers (SC) are composed of certain clusters of residues, involved in the cooperative long range interaction of proteins that regulate flexibility, rigidity and stability of protein structures. Stabilization centers are important in regulating the turnover of certain proteins by preventing their decay with cooperative long range interactions. The most frequent stabilization center residues are usually found at buried positions and have hydrophobic or aromatic side-chains, but some polar or charged residues may also play an important role in stabilization. The stabilization centers show a significant difference in the composition and in the type of linked secondary structural elements, when compared with the rest of the residues. The performed structural and sequential conservation analysis showed a higher conservation of stabilization centers over protein families [31, 41]. In addition, sequence and structure motifs have an application in drug design [42].

We have computed the stabilization centers for all π–π interaction forming residues in Sm/LSm interfaces. Table 3 shows the percentage contribution of the individual amino acid residue which is part of the stabilizing center involved in π–π interactions.

Considering the whole data set, 272 (38.2%) of all stabilizing residues are involved in building π–π interactions. It was interesting to note that all residues involved in π–π interactions were included in at least one stabilization center. Phe was included in more Sm/LSm interface stabilization centers than Tyr residues. Among the stabilization centers involving π residues, His showed the least contribution (3.4%). These observations strongly suggest that the mentioned residues may contribute significantly to the structural stability of studied proteins in addition to participating in π–π interactions.

## 4. CONCLUSIONS

We have systematically analyzed the influence of π–π interactions on the stability of Sm/LSm assemblies. We have found that most of the Phe, Tyr and His Sm/LSm interface residues exhibit π–π interactions. Although the frequency of occurrence of Phe and Tyr was comparable, Phe was more frequently involved in π–π interactions. Involvement of His

residues in π–π interactions was the least frequent. Phe–Phe π–π interacting pair had the highest frequency of occurrence (69.9%), followed by Tyr–Tyr (17%) and Phe–Tyr (8%). There were no His–His interactions. The significant number of π–π interacting residues identified in the dataset is involved in the formation of π-networks. Most of the π–π interacting pairs of Sm/LSm proteins prefer to be in distance range of 5-6 Å. At the plane–plane angles less than 60º (indication of coplanarity, more probably related to maximization of π–π stacking and packing) 3π and 7π-networks were more frequent. The most of the π–π interacting residues preferred to be in strand secondary structure. Since, most of the π–π interacting residues have the tendency of being buried, these interactions might be important in stabilizing the interface regions of these proteins. Moreover, the majority of the residues involved in π–π interactions were evolutionarily conserved; all residues involved in π–π interactions were included in at least one stabilization center, providing an additional stabilization of the Sm/LSm proteins. On the whole, the results obtained from this study might be very helpful in further understanding of the structural stability and functions of Sm/LSm proteins.

REFERENCES

1. L.M. Salonen, M. Ellermann and F. Diederich, Aromatic rings in chemical and biological recognition: energetics and structures, *Angewandte Chemie International Edition,* **50** (21), 4808-4842 (2011).
2. K.L. Copeland, S.J. Pellock, J.R. Cox, M.L. Cafiero and G.S. Tschumper, Examination of tyrosine/adenine stacking interactions in protein complexes, *Journal of Physical Chemistry B,* **117** (45), 14001-14008 (2013).
3. K.A. Wilson, J.L. Kellie and S.D. Wetmore, DNA-protein π-interactions in nature: abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar, *Nucleic Acids Research,* **42** (10), 6726-6741 (2014).
4. S. Yanagisawa, P.B. Crowley, S.J. Firbank, A.T. Lawler, D.M. Hunter, W. McFarlane, C. Li, T. Kohzuma, M.J. Banfield and C. Dennison, π-interaction tuning of the active site properties of metalloproteins, *Journal of the American Chemical Society,* **130** (46), 15420-15428 (2008).
5. M.O. Sinnokrot and C.D. Sherrill, Substituent effects in π-π interactions: sandwich and T-shaped configurations, *Journal of the American Chemical Society,* **126** (24), 7690-7697 (2004).
6. G.B. McGaughey, M. Gagne and A.K. Rappe, π-stacking interactions. Alive and well in proteins, *Journal of Biological Chemistry,* **273** (25), 15458-15463 (1998).
7. S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami and K. Tanabe, Origin of attraction and directionality of the π/π interaction: model chemistry calculations of benzene dimer interaction, *Journal of the American Chemical Society,* **124** (1), 104-112 (2002).
8. A.V. Morozov, K.M.S. Misura, K. Tsemekhman and D. Baker, Comparison of quantum mechanics and molecular mechanics dimerization energy landscapes for pairs of ring-containing amino acids in proteins, *Journal of Physical Chemistry B,* **108**, 8489-8496 (2004).
9. C. Chipot, R. Jaffe, B. Maigret, D.A. Pearlman and P.A. Kollman, Benzene dimer: a good model for π–π interactions in proteins? A comparison between the benzene and the toluene dimers in the gas phase and in an aqueous solution, *Journal of the American Chemical Society,* **118**, 11217-11224 (1996).
10. M. Chourasia, G.M. Sastry and G.N. Sastry, Aromatic-aromatic interactions database, A(2)ID: An analysis of aromatic π-networks in proteins, *International Journal of Biological Macromolecules,* **48** (4), 540-552 (2011).

11. E. Lanzarotti, R.R. Biekofsky, D.A. Estrin, M.A. Marti and A.G. Turjanski, Aromatic-aromatic interactions in proteins: beyond the dimer, *Journal of Chemical Information and Modeling,* **51** (7), 1623-1633 (2011).

12. B. Seraphin, Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs, *EMBO Journal,* **14** (9), 2089-2098 (1995).

13. P. Valentin-Hansen, M. Eriksen and C. Udesen, The bacterial Sm-like protein Hfq: a key player in RNA transactions, *Molecular Microbiology,* **51** (6), 1525-1533 (2004).

14. S. Fischer, J. Benz, B. Spath, L.K. Maier, J. Straub, M. Granzow, M. Raabe, H. Urlaub, J. Hoffmann, B. Brutschy, T. Allers, J. Soppa and A. Marchfelder, The archaeal Lsm protein binds to small RNAs, *Journal of Biological Chemistry,* **285** (45), 34429-34438 (2010).

15. C. Mura, D. Cascio, M.R. Sawaya and D.S. Eisenberg, The crystal structure of a heptameric archaeal Sm protein: Implications for the eukaryotic snRNP core, *Proceedings of the National Academy of Sciences of the United States of America,* **98** (10), 5532-5537 (2001).

16. I.W. Mattaj and E.M. De Robertis, Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins, *Cell,* **40** (1), 111-118 (1985).

17. B. Zaric, M. Chami, H. Remigy, A. Engel, K. Ballmer-Hofer, F.K. Winkler and C. Kambach, Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function, *Journal of Biological Chemistry,* **280** (16), 16066-16075 (2005).

18. V. Anantharaman, E.V. Koonin and L. Aravind, Comparative genomics and evolution of proteins involved in RNA metabolism, *Nucleic Acids Research,* **30** (7), 1427-1464 (2002).

19. H. Hermann, P. Fabrizio, V.A. Raker, K. Foulaki, H. Hornig, H. Brahms and R. Luhrmann, snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions, *EMBO Journal,* **14** (9), 2076-2088 (1995).

20. P. Khusial, R. Plaag and G.W. Zieve, LSm proteins form heptameric rings that bind to RNA via repeating motifs, *Trends in Biochemical Science,* **30** (9), 522-528 (2005).

21. S.Đ. Stojanović, E.R. Isenović and B.L. Zarić, Contribution of non-canonical interactions to the stability of Sm/LSm oligomeric assemblies, *Molecular Informatics,* **30** (5), 430-442 (2011).

22. B.L. Zarić, V.B. Jovanović and S.Đ. Stojanović, Non-covalent interactions across subunit interfaces in Sm proteins, *Journal of Theoretical Biology,* **271** (1), 18-26 (2011).

23. S.Đ. Stojanović, B.L. Zarić and S.D. Zarić, Protein subunit interfaces: A statistical analysis of hot spots in Sm proteins, *Journal of Molecular Modelling,* **16** (11), 1743-1751 (2010).

24. P.W. Rose, B. Beran, C. Bi, W.F. Bluhm, D. Dimitropoulos, D.S. Goodsell, A. Prlic, M. Quesada, G.B. Quinn, J.D. Westbrook, J. Young, B. Yukich, C. Zardecki, H.M. Berman and P.E. Bourne, The RCSB Protein Data Bank: redesigned web site and web services, *Nucleic Acids Research,* **39** (Database issue), D392-D401 (2011).

25. A.G. Murzin, S.E. Brenner, T. Hubbard and C. Chothia, SCOP: a structural classification of proteins database for the investigation of sequences and structures, *Journal of Molecular Biology,* **247** (4), 536-540 (1995).

26. J.M. Word, S.C. Lovell, J.S. Richardson and D.C. Richardson, Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation, *Journal of Molecular Biology,* **285** (4), 1735-1747 (1999).

27. G. Wang and R.L. Dunbrack, Jr., PISCES: a protein sequence culling server, *Bioinformatics,* **19** (12), 1589-1591 (2003).

28. W. Kabsch and C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers,* **22** (12), 2577-2637 (1983).

29. W.L. DeLano, Unraveling hot spots in binding interfaces: progress and challenges, *Current Opinion in Structural Biology,* **12** (1), 14-20 (2002).

30. B. Boeckmann, A. Bairoch, R. Apweiler, M.C. Blatter, A. Estreicher, E. Gasteiger, M.J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout and M. Schneider, The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003, *Nucleic Acids Research,* **31** (1), 365-370 (2003).

31. Z. Dosztanyi, A. Fiser and I. Simon, Stabilization centers in proteins: identification, characterization and predictions, *Journal of Molecular Biology,* **272** (4), 597-612 (1997).

32. Z. Dosztanyi, C. Magyar, G. Tusnady and I. Simon, SCide: identification of stabilization centers in proteins, *Bioinformatics,* **19** (7), 899-900 (2003).

33. Accelrys Software Inc., 2013. Discovery Studio Visualizer, Release 4.0, San Diego: Accelrys Software Inc.

34.  C.A. Hunter and J.K.M. Sanders, The nature of π-π interactions, *Journal of the American Chemical Society,* **112** (14), 5525-5534 (1990).

35.  B. Ma, T. Elkayam, H. Wolfson and R. Nussinov, Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces, *Proceedings of the National Academy of Sciences of the United States of America,* **100** (10), 5772-5777 (2003).

36.  S. Marsili, R. Chelli, V. Schettino and P. Procacci, Thermodynamics of stacking interactions in proteins, *Physical Chemistry Chemical Physics,* **10** (19), 2673-2685 (2008).

37.  S. Chakkaravarthi, M.M. Babu, M.M. Gromiha, G. Jayaraman and R. Sethumadhavan, Exploring the environmental preference of weak interactions in (α/β)₈ barrel proteins, *Proteins,* **65** (1), 75-86 (2006).

38.  S.N. Malkov, M.V. Zivkovic, M.V. Beljanski, M.B. Hall and S.D. Zaric, A reexamination of the propensities of amino acids towards a particular secondary structure: classification of amino acids based on their chemical structure, *Journal of Molecular Modelling,* **14** (8), 769-775 (2008).

39.  M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures, *Nucleic Acids Research,* **33** (Web Server issue), W299-W302 (2005).

40.  E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin, UCSF Chimera-A visualization system for exploratory research and analysis, *Journal of Computational Chemistry,* **25** (13), 1605-1612 (2004).

41.  C. Magyar, M.M. Gromiha, G. Pujadas, G.E. Tusnady and I. Simon, SRide: a server for identifying stabilizing residues in proteins, *Nucleic Acids Research,* **33** (Web Server issue), W303-W305 (2005).

42.  D.J. Craik, N.L. Daly and C. Waine, The cystine knot motif in toxins and implications for drug design, *Toxicon,* **39** (1), 43-60 (2001).

## AROMATIČNA Π-MREŽA U INTERFEJSIMA SM/LSM PROTEINA

*U ovom radu smo analizirali uticaj π–π interakcija na stabilnost i osobine Sm/LSm proteinskih agregata. Ostatak fenilalanina znatno češće uzima učešće u π–π interakcijama u odnosu na His i Tyr. Slično, Phe–Phe π–π interagujući parovi su najučestaliji. Prepoznat je značajan broj π-mreža u interfejsima Sm/LSm proteinima. U većini slučajeva, rastojanje između interagujućih parova aminokiselina bilo je u opsegu 5–6 Å. Za 3π i 7π-mreže, prsten–prsten uglovi manji od 60° su bili učestaliji. Razmatrajući delove Sm/LSm proteina dostupne rastvaraču, može se zaključiti da se svi interagujući parovi nalaze u unutrašnjim regionima. Pored toga, većina π–π interagujućih aminokiselinskih ostataka je evoluciono konzervativan i nalazi se u regionima sa nabranom strukturom. Veliki broj ovih ostataka se može smatrati stabilizacionim centrima, koji (značajno) doprinose ukupnoj stabilnost Sm/LSm proteina.*

Ključne reči: *aromatična π–mreža, Sm/LSm proteini, interfejs, stabilizacioni centri, skor konzervativnosti*