International Journal of Oceans and Oceanography
ISSN 0973-2667 Volume 5, Number 1 (2011), pp. 43-51

http://www.ripublication.com/ijoo.htm

Settlement Inducing Protein Complex of *Balanus amphitrite:* An Analysis of Canonical and Non-Canonical Interactions in their Structural Stability

N. Marimuthu[1]*, J. Jerald Wilson[2] and A.K. Kumaraguru[2]

[1]Andaman & Nicobar Centre for Ocean Science and Technology, National Institute of Ocean Technology (Ministry of Earth Sciences, Govt. of India), Dollygunj, Port Blair 744103, Andamans, India.
[2]Department of Marine and Coastal Studies, School of Energy, Environment and Natural Resources, Madurai Kamaraj University, Madurai 625021, Tamil Nadu, India.
**Corresponding Author E-mail: marinemari@hotmail.com*

Abstract

Settlement inducing protein complex from *Balanus amphitrite* plays a vital role in their settlement. Recently the protein sequence of this complex has been elucidated. However the three dimensional structure of this protein complex has not been elucidated yet. In this study, the three dimensional structure of a modeled SIPC using bioinformatics tools was discussed. Further the importance of Canonical and non-canonical interactions in the structural stability of this modeled protein was also studied. The study reveals that the number of conventional hydrogen bonding is higher than non-conventional hydrogen bonding to the structural stability of the proteins. Among all the amino acids in CH…OC, NH…OC and OH…O interactions, percentage contribution is higher in main chain-main chain interactions only. Among the cation-π interactions of modeled SIPC, Lys-Phe pair shows the highest cation-π interaction energy of –5.56 kcal/mol, however more number cation-π interactions are formed between Arg-Phe pairs only. Overall this study may help to understand the strong interactions to the structural stability of modeled SIPC complex.

Keywords: SIPC, Canonical interaction, Non-canonical interaction, Cation-π interaction.

Introduction

Proteins are the machinery of life. The functional properties of proteins depend upon their three dimensional structures. The sequence of amino acids in the polypeptide chains folds to generate, compact domains with specific three dimensional structures (Branden and Tooze, 1998). These unique three dimensional structures are determined by the interactions of amino acid residues along the polypeptide chain as well as with the surrounding medium. In proteins, both the covalent and noncovalent forces determine the three dimensional structure. The various noncovalent interactions that influence the protein structural stability are electrostatic interactions, hydrogen bonding and hydrophobic forces (Voet and Voet, 1999). These interactions could be both intermolecular as well as intramolecular and a balance between the attractive and repulsive forces does exist. Apart from these conventional hydrogen bonds, it is now generally accepted that other weak electrostatic interactions termed non-canonical interactions (NCI), such as C–H···O interaction, contribute to structural stability of both small molecules and biological macromolecules. More over, cation-π interactions between amino acid side-chains are increasingly being recognized as important structural and functional features of proteins and other biomolecules (Gallivan and Dougherty, 1999). Cation-π interactions can occur between cationic side-chain of either lysine or arginine and the aromatic side-chain of phenylalanine, tyrosine or tryptophan.

Settlement Inducing Protein Complex (SIPC) from the adult barnacle tissues were believed to be a chemical cue for their settlement on marine structures i.e., SIPC induced the cypris settlement of the barnacle, *Balanus amphitrite.* According to a report (Matsumura *et al.,* 1998), it was synthesized in larval development and accumulated in the cyprid stage. SIPC comprises three complex proteins viz., 76, 88 and 98 KDa. Recently reported the amino acid composition of Settlement Inducing Protein Complex (Dreanno *et al.,* 2006). Therefore the bioinformatics study of protein interactions involved in the modeled SIPC are vital.

Materials and Methods

The canonical hydrogen bonds and non-canonical weak interactions in the entire modeled-SIPC were identified using the program Protein Structure Analysis Package (PSAP) (Balamurugan *et al.,* 2007) is available at http://iris.physics.iisc.ernet.in/cgi-bin/psap/index.pl and was developed to calculate the number of canonical (strong) and non-canonical weak interactions. The coordinate file for modeled SIPC in a PDB format is necessary to find these interactions.

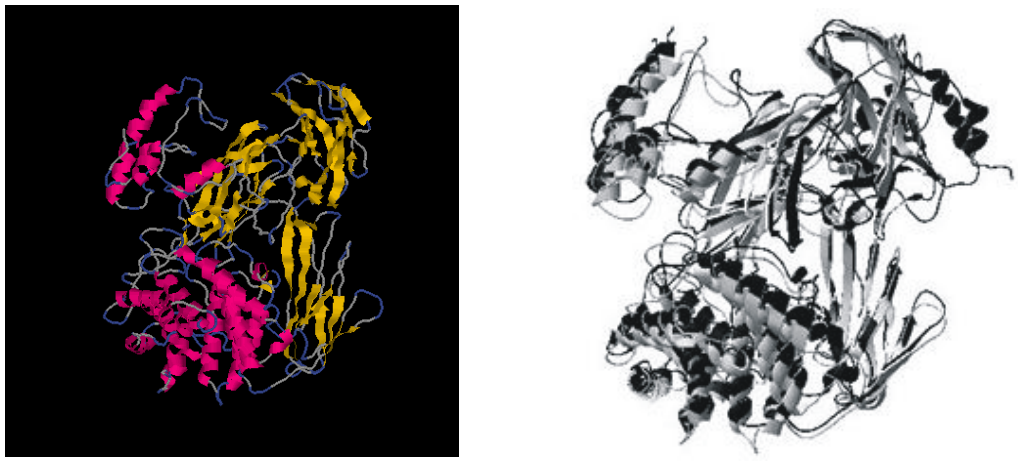
Cation–π interactions are found to be common among structures in the Protein Data Bank (PDB). To search the cation–π interaction in PDB a computer program CAPTURE (Cation–π Trends Using Realistic Electrostatics) is available at http://capture.caltech.edu (Gallivan and Dougherty, 1999) and was developed to calculate the distance between the cationic group ammonium nitrogen (NZ) in Lys or the guanidinium carbon (CZ) in Arg and the centers of all aromatic rings.

Results and Discussion

The SIPC (from Arginine-837 to Tyrosine-1541) was modeled using ESyPred3D-Web Server 1.0 (Lambert *et al.,* 2002). The coordinate file (prot_21412244731543.pdb) of modeled SIPC was retrieved from ESyPred3D-Web Server 1.0 (Lambert *et al.,* 2002), with the help of Human Complement Component C3-B (Protein Id-2A73_B) as a template. The identities and similarities of amino acid sequences between the SIPC and Human Complement Component C3-B were also studied in the BLAST. In this study, the canonical, non-canonical and cation-π interactions to the structural stability of modeled SIPC are discussed.

Homology modeling of SIPC by using Human Complement Component C3-B as a template

Settlement Inducing Protein Complex (SIPC) comprises three complex proteins with the molecular weight of 76, 88 and 98KDa. The total amino acids present in the FASTA-sequence of SIPC were 1547 (Dreanno *et al.,* 2006). Homology-based studies revealed that the 98, 88 and 76-KDa of SIPC peptide fragments shared significant homology with members of the α-2-macroglobulin (A2M), complement factor, and insect thioester-containing protein (TEP) families (Dreanno *et al.,* 2006). Among these, Human Complement C3-B was used as a template to study the homology modeling of SIPC in this study. The structure of the modeled SIPC was visualized using RASMOL (ExPASy Proteomics tools) (Gasteiger *et al.,* 2003). Figure-1 shows the structure of modeled SIPC by using Human Complement Component C3-B (Protein Id-2A73_B) as a template and the Structure of Superimposed model of modeled SIPC along with Human Complement Component C3-B (Protein Id-2A73_B).



A: Modeled SIPC

B: Superimposed model of modeled SIPC along with Human Complement Component C3-B (Protein Id-2A73_B

Figure 1: Structure of modeled and superimposed model of SIPC.

Relative abundance of amino acids in modeled SIPC

The percentage composition of all amino acids in the modeled SIPC was listed in Table-1. The composition was determined with the help of Protein Structure Analysis Package (PSAP) (Balamurugan *et al.*, 2007). This modeled protein has 792 amino acids of which Leucine has higher composition of 9.09% and Tryptophan has the lower composition (0.63%).

Table 1: Percentage composition of amino acid in both protein sequence and modeled structure of SIPC.

Amino acid	SIPC complex protein		Modeled SIPC	
	Nos.	Sequence composition	Nos.	Sequence composition
ALA	99	6.40	56	7.07
CYS	20	1.29	10	1.26
ASP	88	5.69	46	5.81
GLU	126	8.14	70	8.84
PHE	71	4.59	38	4.8
GLY	112	7.24	66	8.33
HIS	17	1.10	7	0.88
ILE	76	4.91	40	5.05
LYS	84	5.43	36	4.55
LEU	133	8.60	72	9.09
MET	24	1.55	16	2.02
ASN	61	3.94	29	3.66
PRO	77	4.98	34	4.29
GLN	53	3.43	26	3.28
ARG	74	4.78	46	5.81
SER	130	8.40	62	7.83
THR	121	7.82	53	6.69
VAL	120	7.76	56	7.07
TRP	12	0.78	5	0.63
TYR	49	3.17	24	3.03

Strong and Weak interactions in the modeled-SIPC

The contribution of strong (conventional) and weak interactions (non-conventional) in the modeled SIPC was listed in Table-2. The composition was determined with the help of Protein Structure Analysis Package (PSAP) as a bioinformatics tool. By using Protein Structure Analysis Package (PSAP), the amino acids involved in Canonical interaction were determined in the modeled SIPC. The percentage contribution of Canonical interactions between the amino acids in the modeled SIPC was listed in Table-3. The number of Strong (conventional or canonical) and weak interactions between the amino acids in the modeled-SIPC is shown in Figure-2.

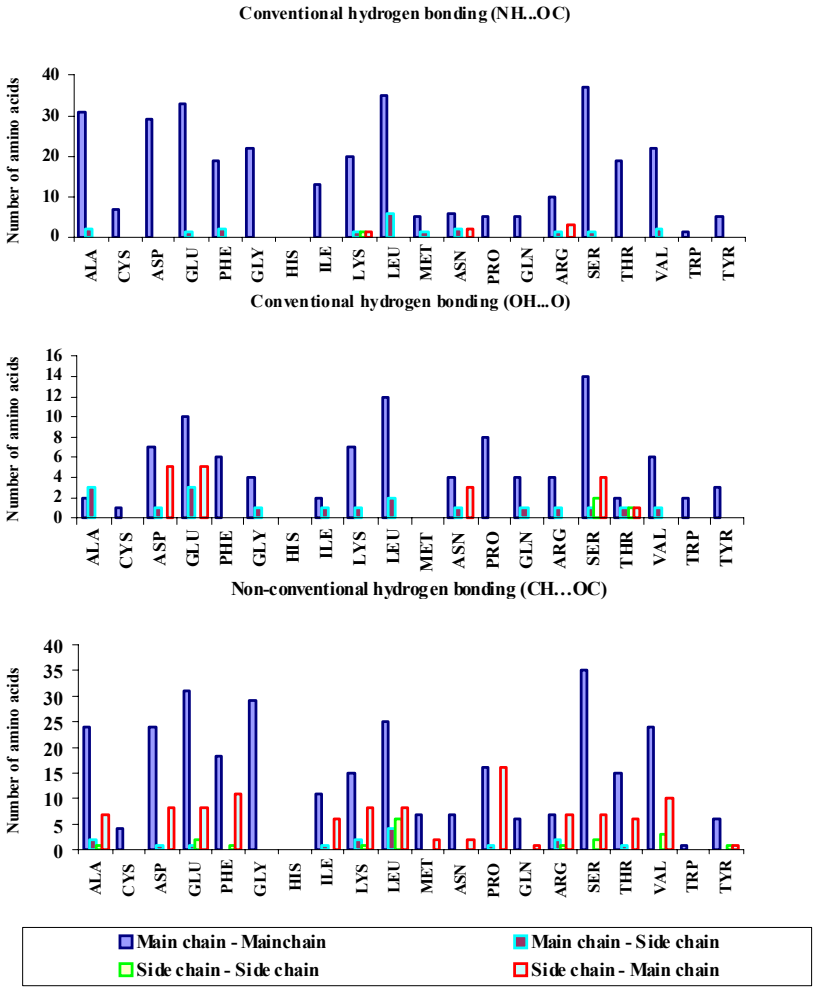


Figure 2: Contribution of modeled-SIPC amino acids in conventional hydrogen bonding (NH…OC & OH…O) and non-conventional hydrogen bonding (CH…OC).

Table 2: Contribution of strong and weak interaction in modeled SIPC.

Interactions	Non-canonical or Non-conventional Interations				Canonical or Conventional Interations								Cation-π
	CH…OC				NH…OC				OH…O				
Interactions between the chain	MM	MS	SS	SM	MM	MS	SS	SM	MM	MS	SS	SM	--
Contribution in Nos.	305	15	18	108	324	19	1	6	98	18	3	18	8
Contribution in %	32.41	1.59	1.91	11.48	34.43	2.02	0.11	0.64	10.41	1.91	0.32	1.91	0.85

MM- Interactions inside the Main chain amino acids of modeled SIPC; MS- Interactions between the amino acids of Main chain and Side chain; SS- Interactions inside the Side chain amino acids of modeled SIPC; SM- Interactions between the amino acids of Side chain and Main chain.

Table 3: Percentage contribution of amino acids in conventional and non-conventional hydrogen bonds of modeled SIPC.

Amino acid	Cation-π %	*CH…OC*				NH…OC				OH…O			
		MM%	MS%	SS%	SM%	MM%	MS%	SS%	SM%	MM%	MS%	SS%	SM%
ALA	0.00	42.86	3.57	1.79	12.50	55.36	3.57	0.00	0.00	3.57	5.36	0.00	0.00
CYS	0.00	40.00	0.00	0.00	0.00	70.00	0.00	0.00	0.00	10.00	0.00	0.00	0.00
ASP	0.00	52.17	2.17	0.00	17.39	63.04	0.00	0.00	0.00	15.22	2.17	0.00	10.87
GLU	0.00	44.29	1.43	2.86	11.43	47.14	1.43	0.00	0.00	14.29	4.29	0.00	7.14
PHE	13.16	47.37	0.00	2.63	28.95	50.00	5.26	0.00	0.00	15.79	0.00	0.00	0.00
GLY	0.00	43.94	0.00	0.00	0.00	33.33	0.00	0.00	0.00	6.06	1.52	0.00	0.00
HIS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ILE	0.00	27.50	2.50	0.00	15.00	32.50	0.00	0.00	0.00	5.00	2.50	0.00	0.00
LYS	5.56	41.67	5.56	2.78	22.22	55.56	2.78	2.78	2.78	19.44	2.78	0.00	0.00
LEU	0.00	34.72	5.56	8.33	11.11	48.61	8.33	0.00	0.00	16.67	2.78	0.00	0.00
MET	0.00	43.75	0.00	0.00	12.50	31.25	6.25	0.00	0.00	0.00	0.00	0.00	0.00
ASN	0.00	24.14	0.00	0.00	6.90	20.69	6.90	0.00	6.90	13.79	3.45	0.00	10.34
PRO	0.00	47.06	2.94	0.00	47.06	14.71	0.00	0.00	0.00	23.53	0.00	0.00	0.00
GLN	0.00	23.08	0.00	0.00	3.85	19.23	0.00	0.00	0.00	15.38	3.85	0.00	0.00
ARG	13.04	15.22	4.35	2.17	15.22	21.74	2.17	0.00	6.52	8.70	2.17	0.00	0.00
SER	0.00	56.45	0.00	3.23	11.29	59.68	1.61	0.00	0.00	22.58	1.61	3.23	6.45
THR	0.00	28.30	1.89	0.00	11.32	35.85	0.00	0.00	0.00	3.77	1.89	1.89	1.89
VAL	0.00	42.86	0.00	5.36	17.86	39.29	3.57	0.00	0.00	10.71	1.79	0.00	0.00
TRP	0.00	20.00	0.00	0.00	0.00	20.00	0.00	0.00	0.00	40.00	0.00	0.00	0.00
TYR	8.33	25.00	0.00	4.17	4.17	20.83	0.00	0.00	0.00	12.50	0.00	0.00	0.00

MM- Interactions inside the Main chain aminoacids of modeled SIPC; MS- Interactions between the aminoacids of Main chain and Side chain; SS- Interactions inside the Side chain aminoacids of modeled SIPC; SM- Interactions between the amino acids of Side chain and Main chain.

The number of Conventional H_2 bonding (Strong interaction – NH…OC and OH…O) between the amino acids in the modeled SIPC was recorded as 487. In this study, these interactions recorded were classified into four types, i.e., between the main chains, between the main chain and side chain, between the side chain and main chain and between the side chains of the modeled SIPC. Among 487 strong interactions present in this modeled protein, 422 interactions are present between the main chains. Thirty seven strong interactions are present between the main chain and side chain. In between the side chain and main chain, twenty four interactions were recorded. Only four strong interactions were recorded between the side chains of this modeled protein.

The contributions of weak interactions (CH…OC) were recorded as 446 which were lower than the strong interactions (NH…OC and OH…O). Similarly, the non-

conventional H_2 bonding is existing in larger amount inside the main chain of the modeled SIPC (Figure-2). Among the ninteen different canonical interactions (NH…OC and OH…O) present in the modeled SIPC, 44.84% of interactions were recorded in the Main chain (Table-2). However in RNA binding proteins, this interaction was observed in the range of 64% (Anbarasu *et al.,* 2007). Further analysis indicated that the main-chain atoms of {Main chain-Main chain (MM)-[C-H…O=C], (MM)-[N-H…O] and (MM)-[O-H…O]} interactions contribute significantly (about 77%) either as donor or acceptor or both. This analysis on the dataset of modeled SIPC domain indicates that the distribution of atoms involved in the Canonical and non-canonical interactions occurred in the main chain. Non-conventional interactions between Main chain and Main chain (MM), Main chain and Side chain (MS), Side chain and Side chain (SS) and Side chain and Main chain (SM) contributed to 47.39% which was significantly lower than the canonical interactions.

Moreover, they are more oriented towards the high incidence in the main-chain atoms. This suggests that the conventional or canonical interactions may contribute significantly to the stability of the modeled SIPC. Relative contribution of the non-canonical and canonical hydrogen bonds to the stability of SIPC was also evaluated. The correlation between the number of conventional (Canonical interactions) hydrogen bonding and non-conventional (Non-canonical interactions) hydrogen bonding was recorded as +0.82.

Cation-π interactions between amino acid side-chains of SIPC

Cation-π interactions can occur between cationic side-chain of either lysine or arginine and the aromatic side-chain of phenylalanine, tyrosine or tryptophan. The stabilization energy originates in part from electrostatic attraction between the cations (of the basic amino acid residue) and regions of high electron density in π-orbital of the aromatic group, leading to cation-π interaction. This interaction in the modeled SIPC was determined with the help of web-based version of the CaPTURE program (Gallivan and Dougherty, 1999). Table-4 shows the energy contribution between amino acids during cation-π interaction.

Table 4: Cation-π interaction energy contribution of modeled SIPC.

Protein	Cation	Res	Aromatic	Res	Ees	Evdw	Ecat-pi
Modeled SIPC	ARG	886	PHE	889	-2.16	5.11	2.95
	ARG	1008	PHE	1010	-1.67	-1.64	-3.31
	ARG	1369	PHE	1066	-1.73	-1.81	-3.54
	ARG	1505	PHE	1579	-3.14	-1.61	-4.75
	ARG	1027	TYR	979	-2.32	-3.08	-5.4
	ARG	1170	TYR	1171	-1.34	-3.13	-4.47
	LYS	1221	PHE	1231	-4.39	-1.17	-5.56
	LYS	1237	TYR	1287	-3.55	-0.71	-4.26

In this modeled SIPC, there were a total of 8 energetically significant cation-π interactions involved. Even though 17 pairs of cationic interaction (between Arginine and Phenylalanine) present in the modeled protein, 4 pairs were energetically significant. They were ARG (886) – PHE (889), ARG (1008) – PHE (1010), ARG (1369) – PHE (1066) and ARG (1505) – PHE (1579).

Two pairs [ARG (1027) – TYR (979) and ARG (1170) – TYR (1171)] out of 13 Arginine and Tyrosine pairs and one pair between Lysine and Phenylalanine LYS (1221) – PHE (1231) showed energetically significant cation-π interactions. Similarly another pair between Lysine and Tyrosine also existed. It was LYS (1237) – TYR (1287). There were no energetically significant Arg-Trp and Lys-Trp. The lesser regression coefficient (R^2=0.0001) in the case of cation-π interactions may be due to low incidence of aromatic amino acid residues in proteins.

There are reports of this interaction for their role in the enhancement of stability of thermophilic proteins (Chakravarthy and Varadarajan, 2000; Gromiha *et al.,* 2002), folding of polypeptides (Shi *et al.,* 2002), and the stability of membrane protein (Gromiha, 2003). The stability and specificity of both protein DNA (Gromiha *et al.,* 2004a; Gromiha *et al.,* 2004b) and protein RNA (Anbarasu *et al.,* 2007; Chakkaravarthi and Gromiha, 2006) complexes are also reported on the basis of these cation-π interactions.

Conclusion

Among three types of interactions, greater contribution observed in the modeled SIPC was Canonical Interactions. Since these interactions between Main chain and Main chain (MM), Main chain and Side chain (MS), Side chain and Side chain (SS) and Side chain and Main chain (SM) contributed to 51.75% which was significantly higher than the non-canonical interactions. The stabilization energy results from Cation-π interaction reveal the importance of this interaction to the modeled SIPC protein stability (Table-4).

Acknowledgements

Thanks are due to Dr. S. Chakkaravarthi, Postdoctoral Research Fellow, Molecular Systems Biology Group (Chemistry), University of Coimbra, Portugal for his technical support and critical review of this manuscript. We acknowledge the Department of Ocean Development (Ministry of Earth Sciences, Govt. of India) for the financial support.

References

[1] Branden, C., and Tooze, J., 1998, Introduction to protein structures, Garland Publishing, Inc, New York.
[2] Voet, D., and Voet, J. G., 1999, Biochemistry, Wiley and Sons, New York.

[3] Gallivan, J. P., and Dougherty, D. A., 1999, "Cation-π interactions in structural biology", Proc. Natl. Acad. Sci. USA., 96, pp. 9459-9464.
[4] Matsumura, K., Nagano, M., Yoshinaga, Y. K., Yamazaki, M., Clare, A. S., and Fusetani, N., 1998, "Immunological Studies on the Settlement-Inducing Protein Complex (SIPC) of the Barnacle Balanus amphitrite and Its Possible Involvement in Larva-Larva Interactions", Proceedings: Biological Sciences, 265(1408), pp. 1825-1830.
[5] Dreanno, C., Matsumura, K., Dohmae, N., Takio, K., Hirota, H., Kirby, R. R., and Clare, A. S., 2006, "An α2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite", PNAS., 103(39), pp. 14396–14401.
[6] Balamurugan, B., Md. Roshan, M. N. A., Hameed, B. S., Sumathi, K., Senthilkumar, R., Udayakumar, A., Babu, K. H. V., Kalaivani, M., Sowmiya, G., Sivasankari, P., Saravanan, S., Ranjani, C. V., Gopalakrishnan, K., Selvakumar, K. N., Jaikumar, M., Brindha, T., Michael, D., and Sekar, K., 2007, "PSAP: protein structure analysis package", J. Appl. Cryst., 40, pp. 773-777.
[7] Lambert, C., Leonard, N., De Bolle, X., and Depiereux, E., 2002, "ESyPred3D: Prediction of proteins 3D structures", Bioinformatics, 18(9), pp. 1250-1256.
[8] Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R., and Bairoch, A., 2003, "ExPASy: the proteomics server for in-depth protein knowledge and analysis", Nucl. Acids Res., 31(13), pp. 3784-3788.
[9] Anbarasu, A., Anand, S., Mathew, L., and Sethumadhavan, R., 2007, "Influence of cation-π interactions on RNA binding proteins", Int. J. Biol. Macromol., 40, pp. 479-483.
[10] Chakravarthy, S., and Varadarajan, R., 2000, "Elucidation of factors responsible for enhanced thermal stability of proteins: a structural genomics based study". Biochemistry, 41, pp. 8152-8161.
[11] Gromiha, M. M., Thomas, S., and Santhosh, C., 2002, "Role of cation-π interaction to the stability of thermophilic proteins" Prep. Biochem. Biotech., 2, pp. 355-362.
[12] Shi, Z., Olson, C. A., and Kallenbach, N. R., 2002, "Cation-π interaction in model α-helical peptides", J. Am. Chem. Soc., 124, pp. 3284-3291.
[13] Gromiha, M. M., 2003, "Influence of cation-π interaction in different folding types of membrane proteins", Biophys. Chem., 103, pp. 251-258.
[14] Gromiha, M. M., Santhosh, C., and Suwa, M., 2004a, "Influence of cation-π interactions in protein–DNA complexes", Polymer, 45, pp. 633-639.
[15] Gromiha, M. M., Pujadas, G., Magyar, C., Selvaraj, S., Simon, I., 2004b, "Locating the stabilizing residues in (α/β)8 barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation", Proteins, 55, pp. 316-329.
[16] Chakkaravarthi, S., and Gromiha, M. M., 2006, Analysis of cation-π interaction to the structural stability of RNA binding proteins, Polymer, 47, pp.709-721.