Available online at www.sciencedirect.com
SCIENCE DIRECT
Polymer 47 (2006) 709–721
polymer
www.elsevier.com/locate/polymer

Analysis of cation–π interactions to the structural stability of RNA binding proteins

S. Chakkaravarthi [a,*], M. Michael Gromiha [b]

[a] School of Biotechnology and Chemical Engineering, Vellore Institute of Technology, Deemed University, Vellore 632 014, Tamilnadu, India
[b] Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), AIST Tokyo Waterfront Bio-IT Research Building, 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan

Received 3 August 2005; received in revised form 17 November 2005; accepted 17 November 2005
Available online 7 December 2005

Abstract

Cation–π interactions play an important role to the stability of protein structures. In this work, we have analyzed the influence of cation–π interactions in RNA binding proteins. We observed cation–π interactions in 32 out of 51 RNA binding proteins and there is a strong correlation between the number of amino acid residues and number of cation–π interactions. The analysis on the influence of short ($<\pm3$ residues), medium (±3 or ±4 residues) and long range contacts ($>\pm4$ residues) showed that the cation–π interactions are mainly formed by long-range contacts. The cation–π interaction energy for Arg–Trp is found to be the strongest among all interacting pairs. Analysis on the preferred secondary structural conformation of the residues involved in cation–π interaction indicates that the cationic Lys and Arg prefer to be in α-helices and β-strands, respectively, whereas the aromatic residues prefer to be in strand and coil regions. Most of the cation–π interactions forming residues in RNA binding proteins are conserved among homologous sequences. Further, the cation–π interactions have distinct roles to the stability of RNA binding proteins in addition to other conventional non-covalent interactions. The results observed in the present study will be useful in understanding the contribution of cation–π interactions to the stability of RNA binding proteins.

Keywords: Cation–π interactions; RNA binding proteins; Accessible surface area

1. Introduction

Selective binding of proteins to specific sites on nucleic acids has been a challenging and interesting problem since the earliest days of molecular biology. The first protein-nucleic acid recognition problem to be defined was the enzymatic linking of an amino acid with its correct tRNA [1,2], a process whose specificity was seen as crucial for accurate gene expression. Protein recognition of specific RNA sites was also implicit in early studies of ribosome assembly [3,4]. Since then, the participation of specific protein-RNA complexes in a large number of cellular processes has become evident. RNA structures are flexible molecules that display complex secondary and tertiary structures including short lengths of double helices (A-form), hairpin loops, bulges and pseudo-knots. Proteins tend to interact with the complex secondary structure elements such as stem-loops and bulges in RNA [5]. In addition, non-Watson-Crick base pairing can occur in loop regions of RNA structures and such features can also be preferentially identified by proteins [6]. There are several types of interactions, which give an effect to macromolecular structure and interactions. Ion–ion bonds, hydrogen bonds and hydrophobic interactions are often important for both recognition and binding specificity in protein-DNA/RNA interactions. A growing number of experimental and theoretical studies have emphasized the existence of favorable interactions between positively charged groups and π-aromatic systems [7–9]. Both intermolecular and intramolecular cation–π interactions are recognized to play an important role in the stability of protein-DNA complexes [10]. This type of non-covalent binding force is assumed to be significant in protein structure [11] as well as in biomolecular association processes such as antigen–antibody binding [12,13] and receptor–ligand interaction [14,15]. There are reports of this interaction for their role in the enhancement of stability of thermophilic proteins [16,17], folding of polypeptides [18] and the stability of membrane proteins [19,20]. The stability and specificity of

* Corresponding author. Tel.: +91 416 2202616; fax: +91 416 2243092.
E-mail address: chakkaravarthi77@gmail.com (S. Chakkaravarthi).

0032-3861/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.polymer.2005.11.059

protein-DNA complexes are also reported on the basis of these cation–π interactions [21,22]. Although the structural studies of protein-RNA complexes are mostly focused on discovering the specific mechanisms of protein-RNA interactions by analyzing intra and inter-molecular interactions in diverse aspects, importance of the cation–π interaction in the structural stability of RNA binding proteins has not yet been elucidated.

In this study we have analyzed the cation–π interaction in 51 RNA binding proteins. The energetic contribution due to cation–π interactions have been brought out for each of the 51 proteins and for all six pairs of residues (Arg–Phe, Arg–Tyr, Arg–Trp, Lys–Phe, Lys–Tyr and Lys–Trp) involved in such interactions. The percentage composition of specific amino acid residues contributing to cation–π interactions was calculated. Further, the characteristic features of residues involved in cation–π interactions have been evaluated in terms of secondary structure, solvent accessibility and sequential separation of residues involved in cation–π interactions. We observed that the cation–π interaction energy for the pairs with Arg is stronger than that with Lys. Sequential separation of cation–π interactions in RNA binding proteins shows that most of the interactions are formed due to long range interactions. Cation–π interaction forming residues Lys and Arg prefer to be in α-helices and β-strands, respectively, whereas aromatic residues prefer β-strands and coil regions. Further, most of the residues contributing for cation–π interactions are not involved in binding with RNA.

2. Materials and methods

2.1. Data set

We have considered a set of 51 RNA-binding proteins from the information available in literature [23] for the present study. This set has been obtained with the following conditions: (i) the three dimensional structures of these proteins have been solved with ≤ 3.0 Å resolution, (ii) the similarity search using PSI-BLAST yielded the *e*-value of less than 0.001 and (iii) the sequence identity is less than 80%. The complexes, whose proteins were homologous but recognized different nucleotide sequences, were included in the data set.

The PDB [24] codes of the proteins are: 1b23, 1b2m, 1b7f, 1c0a, 1c9s, 1cx0, 1dfu, 1di2, 1dk1, 1e7x, 1ec6, 1efw, 1f7u, 1f8v, 1feu, 1ffy, 1fxl, 1g59, 1gax, 1gtf, 1gtn, 1g2e, 1h4q, 1h4s, 1hc8, 1hdw, 1he0, 1he6, 1hq1, 1i6u, 1il2, 1jbr, 1jbs, 1jid, 1k8w, 1knz, 1kq2, 1l9a, 1lng, 1mms, 1qf6, 1qtq, 1ser, 1urn, 1zdh, 1zdi, 2bbv, 2fmt, 5msf, 6msf and 7msf.

2.2. Computation of amino acid composition

The amino acid composition for each amino acid residue that are involved in cation–π interactions (Lys, Arg, Phe, Trp and Tyr) was computed using the standard formula,

$$\mathrm{comp(i)} = \frac{n(\mathrm{i})}{N} \tag{1}$$

where n(i) is the number of amino acids of type i and N is the total number of amino acids in a protein.

2.3. Occurrence and energetic contribution due to cation–π interactions

The number of cation–π interaction in each protein has been calculated using the program CAPTURE [25] available at http://capture.caltech.edu. In the present study only energetically significant interactions ($E_{\mathrm{cat}-\pi} \leq 2$ kcal/mol) were considered. The percentage composition of a specific amino acid residue contributing to cation–π interactions is obtained by the equation,

$$\mathrm{comp}_{\mathrm{cat}-\pi}(\mathrm{i}) = n_{\mathrm{cat}-\pi}(\mathrm{i}) \times \frac{100}{n(\mathrm{i})} \tag{2}$$

where i stands for the five residues, Lys, Arg, Phe, Trp and Tyr, $n_{\mathrm{cat}-\pi}$ is the number of residues involved in cation–π interactions and n(i) is the number of residues of type i in the considered protein structures.

We have computed the energetic contribution of cation–π interactions for each RNA binding protein in the data set and for all possible pairs of positively charged and aromatic amino acids. The total cation–π interaction energy ($E_{\mathrm{cat}-\pi}$) has been divided into electrostatic (E_{es}) and van der Waals energy (E_{vw}) and were computed using the program CAPTURE, which has implemented a subset of OPLS force field [26] to calculate the energies. The electrostatic energy (E_{es}) is calculated using the equation

$$E_{\mathrm{es}} = \frac{q_{\mathrm{i}} q_{\mathrm{j}} e^2}{r_{\mathrm{ij}}} \tag{3}$$

where q_{i} and q_{j} are the charges for the atoms i and j, respectively, and r_{ij} is the distance between them. The van der

Table 1
Composition of aromatic and positively charged residues in RNA binding proteins

Proteins	Lys%	Arg%	Phe%	Tyr%	Trp%
RNA binding protein	7.30±3.70	5.98±2.65	3.79±1.49	3.27±1.55	1.03±0.90
TMH	2.29±1.68	2.97±0.95	7.98±1.67	4.14±0.79	4.19±1.38
TMS	4.73±1.75	3.48±0.85	4.40±1.36	6.56±1.93	1.85±1.33
Globular	5.83	4.74	3.97	3.60	1.48

TMS, transmembrane strand; TMH, transmembrane helical.