Structural Characterization of RNA-Binding Sites of Proteins: Preliminary Results

Fadi Towfic
Bioinformatics and Computational-
Biology Graduate Program
Iowa State University
Ames, IA 50011-1040,USA
ftowfic@cs.iastate.edu

David C. Gemperline
Department of Biology
Department of Chemistry
Carthage College
2001 Alford Park Drive
Kenosha, WI 53140-1994, USA
dcgemperline@gmail.com

Cornelia Caragea
Department of Computer Science
Iowa State University
Ames, IA 50011-1040,USA
cornelia@cs.iastate.edu

Feihong Wu
Bioinformatics and Computational-
Biology Graduate Program
Iowa State University
Ames, IA 50011-1040,USA
wuflyh@cs.iastate.edu

Drena Dobbs
Department of Genetics,
Development and Cell Biology
Bioinformatics and Computational-
Biology Graduate Program
Iowa State University
Ames, IA 50011-1040,USA
ddobbs@iastate.edu

Vasant Honavar
Department of Computer Science
Bioinformatics and Computational-
Biology Graduate Program
Iowa State University
Ames, IA 50011-1040,USA
honavar@cs.iastate.edu

Abstract

We explore whether protein-RNA interfaces differ from non-interfaces in terms of their structural features and whether structural features vary according to the type of the bound RNA (e.g., mRNA, siRNA...etc.), using a non-redundant dataset of 147 protein chains extracted from protein-RNA complexes in the protein data bank. Our analysis of surface roughness, solid angle and CX value of amino acid residues for each of the protein chains in the dataset shows that: The protein-RNA interface residues tend to be protruding compared to non-interface residues and tend to have higher surface roughness and exhibit moderately convex or concave solid angles. Furthermore, the protein chains in protein-RNA interfaces that contain Viral RNA and rRNA significantly differ from those that contain dsRNA, mRNA siRNA, snRNA, SRP RNA and tRNA with respect to their CX values. The results of this analysis suggests the possibility of using such structural features to reliably identify protein-RNA interface residues when the structure of the protein is available but the structures of complexes formed by the protein with RNA are not.

1 Introduction

Protein-RNA interactions play a vital role in RNA splicing, translation, replication of many viruses as well as many other processes in the cell. The prediction of protein-RNA interfaces can aid in the design of drug-inhibitors for viruses, down-regulation of unwanted genes as well as contributing to our basic understanding of the mechanisms involved in protein-RNA recognition [16, 17, 11, 7]. At least nine families of RNA-binding proteins have been identified using sequence-based analyses of the major groups of RNA-binding proteins, together with functional characterization of mutations that affect the specificity or affinity of RNA

binding (for review, see [3]). In contrast, the number of experimentally determined structures for protein-RNA complexes is still relatively small and heavily biased (ribosomal proteins represent 50% of all RNA binding proteins in the Protein Data Bank [PDB] [2]).

Because of the importance of protein-RNA interactions in biological regulation and the considerable effort required to identify RNA binding residues through biophysical analyses of protein-RNA complexes or *in vitro* binding studies, there is an urgent need for computational methods to identify RNA binding sites given a protein's primary amino acid sequence, and when available, its 3-dimensional structure. Several recent studies have focused on the development of machine learning approaches to amino acid sequence-based prediction of RNA-binding residues in proteins [21, 20, 8, 9]. The predictions obtained using such methods have already contributed to the design of wet-lab experiments to decipher mechanisms of protein-RNA recognition [19, 1]. However, the machine learning approaches to prediction of RNA-binding residues of proteins have focused largely on the analysis of amino acid sequence as opposed to the structural features of the protein chain. Other analyses of protein-RNA interfaces [10, 15, 13] have focused on the analysis of hydrogen bonds or van der Waals contacts in between the protein and the RNA. There has been relatively little attention paid to structural features of the interface (e.g protrusion or roughness) rather than the atomic forces.

Against this background, it is natural to ask: *Do protein-RNA interfaces differ from non-interfaces in terms of their structural features? Do the structural features vary according to the type of the bound RNA (e.g., mRNA, siRNA etc.)?* If we find that the protein-RNA interfaces differ from noninterfaces in terms of their structural features, then the structural features can be exploited by machine learning approaches to predict protein-RNA interface residues when the structure of the protein is available but the structures of the complexes formed by the protein with RNA are not. If the different classes of protein-RNA interfaces significantly differ from each other with respect to their structural features, it might be possible to improve the specificity and sensitivity of protein-RNA interface residue prediction by training separate classifiers for each type of bound RNA.

We describe an analysis of the structural features of protein chains from RNA-binding proteins that explores this question using a non-redundant dataset of 147 protein chains from the RB147 dataset [21]. We focus on three of the six structural properties of amino acid residues used in a recent analysis of protein-protein interfaces by Wu *et al.* [22], namely, surface roughness [14], solid angle [4] and CX value [18]. The results of our analysis show that protein-RNA interface residues tend to be protruding compared to non-interface residues. Furthermore, interface residues tend to have rough surfaces and have moderately convex or concave solid angles. Our analysis also shows that the protein chains in protein-RNA interfaces containing Viral-RNA and rRNA significantly differ from those that contain dsRNA, mRNA siRNA, snRNA, SRP RNA and tRNA with respect to their CX values.

The rest of the paper is organized as follows: Section 2 describes the RB147 dataset and each of 3 properties of amino acid residues examined in this study. Section 3 presents the results of our analysis, comparing interface and noninterface residues based on these 3 properties. Section 4 concludes with a summary and an outline of some directions for further research.

2 Materials and Methods

2.1 Dataset

The RB147 dataset [21] used in this study contains protein chains extracted from structures of protein-RNA complexes in the PDB solved by X-ray crystallography, after eliminating protein chains from structures with resolution worse than 3.5Å and protein chains sharing a sequence identity greater than 30% with one or more other protein chains. The RB147 dataset contains 147 non-redundant protein chains and a total of 32,324 amino acids. The RNA-binding residues are defined as follows: an RNA-binding residue is an amino acid containing at least one atom within 5Å of any atom in the bound RNA. According to this definition, RB147 contains a total of 6,157 RNA-binding residues and 26,167 non-binding residues.

2.2 Classification of the Protein Chains Based on the Type of the Bound RNA

The protein chains in the RB147 dataset were classified into 9 classes according to the type(s) of RNA that was found in the corresponding protein-RNA complex based on a taxonomy of RNA types used previously by Ellis *et al.* [6]: dsRNA, mRNA, rRNA, siRNA, snRNA, SRP RNA, tRNA, Viral RNA or "other" (which denoted synthetic RNAs or pre-mRNAs or a class of RNAs not included in any of the other categories). The classification for each PDB id and chain in the dataset is shown in table 1. Over half of the protein chains belong to complexes with rRNA, with tRNA, mRNA and viral RNA being the other dominant groups (in that order).

2.3 Analysis of Structural Properties

Each chain in the dataset was analyzed in terms of its surface roughness [14], solid angle [4], and CX value [18].

2

$RNA\ Type$	$PDBIDs$
dsRNA	$1DI2_A, 1UVJ_A, 1YZ9_A$
mRNA	$1AV6_A, 1G2E_A, 1GTF_Q, 1KNZ_A$ $1KQ2_A, 1M8X_A, 1WPU_A, 1WSU_A$ $2A1R_A, 2ASB_A$
rRNA	$1APG_A, 1DFU_P, 1FEU_A, 1FJG_B$ $1FJG_C, 1FJG_D, 1FJG_E, 1FJG_G$ $1FJG_I, 1FJG_J, 1FJG_K, 1FJG_L$ $1FJG_M, 1FJG_N, 1FJG_P, 1FJG_Q$ $1FJG_S, 1FJG_T, 1FJG_V, 1G1X_A$ $1HRO_W, 1I6U_A, 1JBR_A, 1MZP_A$ $1SDS_A, 1T0K_B, 1UN6_B, 1VQO_1$ $1VQO_2, 1VQO_3, 1VQO_A, 1VQO_B$ $1VQO_C, 1VQO_D, 1VQO_E, 1VQO_G$ $1VQO_H, 1VQO_I, 1VQO_J, 1VQO_K$ $1VQO_L, 1VQO_M, 1VQO_N, 1VQO_P$ $1VQO_Q, 1VQO_R, 1VQO_S, 1VQO_T$ $1VQO_U, 1VQO_V, 1VQO_W, 1VQO_X$ $1VQO_Y, 1VQO_Z, 1W2B_5, 1Y69_8$ $1Y69_K, 1Y69_U, 2AVY_F, 2AVY_U$ $2AW4_0, 2AW4_1, 2AW4_2, 2AW4_3$ $2AW4_D, 2AW4_E, 2AW4_G, 2AW4_H$ $2AW4_J, 2AW4_L, 2AW4_N, 2AW4_P$ $2AW4_Q, 2AW4_R, 2AW4_S, 2AW4_Y$ $2AW4_Z, 2BH2_A, 2D3O_1, 2D3O_S$ $1G1X_B, 1G1X_C$
siRNA	$1RPU_A, 1SI3_A, 2BGG_A$
snRNA	$1A9N_A, 1EC6_A, 1LNG_A, 1M8V_A$ $1OOA_A$
SRP RNA	$1E8O_A, 1HQ1_A$
tRNA	$1ASY_A, 1B23_P, 1C0A_A, 1EIY_B$ $1F7U_A, 1FFY_A, 1H3E_A, 1H4S_A$ $1J1U_A, 1J2B_A, 1K8W_A, 1N78_A$ $1Q2S_A, 1QF6_A, 1QTQ_A, 1R3E_A$ $1SER_A, 1TFW_A, 1U0B_B, 1VFG_A$ $1WZ2_A, 2BTE_A, 2CT8_A, 2FMT_A$
Viral RNA	$1A34_A, 1DDL_A, 1H2C_A, 1LAJ_A$ $1N35_A, 1NB7_A, 1PGL_2, 1RMV_A$ $1WNE_A, 2AZ0_A, 2BU1_A$
Other	$1B2M_A, 1JID_A, 1M5O_C, 1YVP_A$ $1ZH5_A, 2A8V_A, 2BX2_L$

Table 1. Classification for each of the 147 **protein chains in the dataset. The four letter PDB ids are subscripted by the chain. As can be seen from the table, over half (**55.7%**) of the RNAs are rRNAs.**

The analysis was repeated on subsets of the dataset corresponding to the classification based on the type of the RNA found in the interface (see table 1). We implemented a program in Java, Structure-Analyzer 1.0 (available at http://www.public.iastate.edu/~ftowfic) for this analysis. The Java package has an easy-to-use API to allow its use in other applications. The program generates a standard tab-delimited output file with the PDBID, chain name, residue name (three letter abbreviation), residue number, a + or - indicating whether or not the residue is part of the interface, a score derived from the structural property being examined (roughness, cx, solid angle) and a + or - denoting whether or not the residue is part of the surface of the protein (the definition of a surface residue can be varied within the class as desired). In our analysis, surface residues are defined as residues that have a solvent accessible surface area that is at least 5% of their total surface area [22, 5].

2.4 Roughness Calculation

The roughness value for a residue denotes the degree of irregularity of that point at the surface as outlined by Lewis *et al.* and Lee *et al.* [14, 12]. The surface roughness value (D) is given by:

$$D = 2 - \frac{\partial \log A_s}{\partial \log R}$$

The roughness calculation requires a molecule surface area (A_s), which is obtained by rolling a sphere with radius R against the protein and calculating the area of the resulting surface as implemented in the MSP software package [5]. The radius R is varied from 0.2 to 4.0Å in 0.1 increments, and the resulting points are used to calculate the roughness values according to the previous equation. For a perfectly smooth surface $D = 2$ whereas for a rough surface, $D > 2$.

2.5 Solid Angle Calculation

The solid angle value for a residue measures the gross shape of a local region of a protein by calculating the intersection area of a sphere (centered around a point on the protein surface, with radius 6Å) with the protein [4]. The solid angle has a range of 0 to 4π and is calculated for a residue by averaging the solid angles for all points that are part of the residue. A point with solid angle $< 2\pi$ lies on a surface that is locally convex and a point with $> 2\pi$ lies on surface that is locally concave.

2.6 CX Value Calculation

The CX value measures the ratio of the number of atoms that occupy a 6Å sphere compared to the empty volume within the sphere [18]. The analysis for CX was conducted

3

on surface residues. The CX score may be calculated according to three different methods: First, the CX score can be extracted from the alpha-carbon atom and that score is used for the whole residue (*alphacarbon* method). Another method is to simply average the CX score across all atoms for the residue and use the average score as the CX score for the residue (*averagecx* method). Finally, the CX scores for the atoms in the R-group of the residue can be averaged and that average can be used as the CX score for the residue (*rgroup* method).

2.7 Interface Propensity Calculations

Consider a residue-based property (such as residue roughness) with k discrete values: $(v_1, v_2, ..., v_k)$. Each surface residue is assigned to one of k disjoint subsets $S_1, S_2, ..., S_k$ based on the value of the residue property. Let I_i and N_i respectively be the *fractions* of interface residues and non-interface residues in the set S_i. Let I and N respectively denote the *fractions* of interface and non-interface residues in the entire dataset (over all of the score ranges). The log-propensity for the interface can then be expressed according to the following equation

$$\log_2(Propensity_i) = log_2(\frac{\frac{I_i}{I}}{\frac{I_i + N_i}{I + N}})$$

The interface propensity I_i of the property at value v_i is a measure of the preference for the value (or a range of values) v_i among the interface residues (relative to the entire set of surface residues). $I_i > 0$ denotes that the specified property value (or a range of values) v_i tends to be more preferred among the interface residues relative to the surface residues. Similarly, $I_i < 0$ denotes that the specified property value v_i tends to be less preferred among the interface residues relative to the entire set of surface residues.

3 Results

Now we proceed to explore the questions: *Do protein-RNA interfaces differ from non-interfaces in terms of their structural features? Do the structural features vary with the type of bound RNA?*

3.1 CX Protrusion Index

Figure 1 shows the relative CX score propensities of the interface residues, based on three different calculations of CX score for each residue: the average of the CX scores for all atoms (*averagecx*); use the average of the CX scores atoms in the R-group only (*rgroup*); or the CX score for the alpha carbon atom as the score for the corresponding
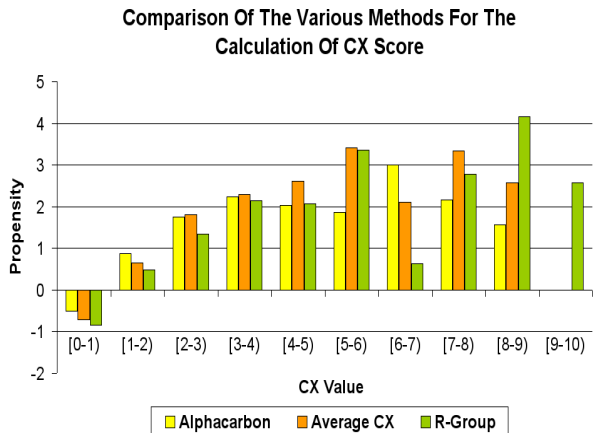

Figure 1. Comparison of various methods to obtain CX values for residues on the protein surface. The figure shows that averaging the CX score across all atoms in the residue produces similar results to averaging across the R-group atoms alone.

residue (*alphacarbon*). The figure shows that averaging the CX score across all atoms in the residue produces similar results to averaging across the R-group atoms alone. Figures 2, 3 and 4 (respectively) show the residue propensities based on the types of the bound RNA using the *averagecx*, *alphacarbon* and *rgroup* methods respectively. The observed CX value residue propensities of interfaces involving different types of RNA appear to be sensitive to the method used to calculate the CX values.

ANOVA analysis for the *alphacarbon* method (ANOVA p-value = 0.056, cutoff = 0.05) shows that tRNA and mRNA cluster together with variances around 0.2. SnRNA, rRNA, dsRNA and "other" cluster together with variances around 0.3. Finally, siRNA, SRP RNA and Viral RNA cluster together with variances around 0.65. ANOVA analysis for the *averagecx* method (ANOVA p-value = 0.00001, cutoff = 0.05) shows that Viral RNA, mRNA, "other", and snRNA cluster together with variances around 0.2. DsRNA, rRNA siRNA and SRP RNA cluster together with variance around 0.35 and tRNA is the only RNA type with variance around 0.5. The results of ANOVA analysis for the *rgroup* method (ANOVA p-value = 0.0002, cutoff = 0.05) are similar to those of *averagecx* with mRNA, snRNA and "other" clustering together with variances around 0.13, whereas dsRNA, SRP RNA, siRNA, rRNA and Viral RNA cluster together with variances around 0.37. Finally, the tRNA group is isolated with a variance of 0.5.

Regardless of the method used to calculate the CX scores, we can observe some general trends: The rRNA and tRNA propensities are always negative at CX score range [0,1) and have an increasing propensity as the CX
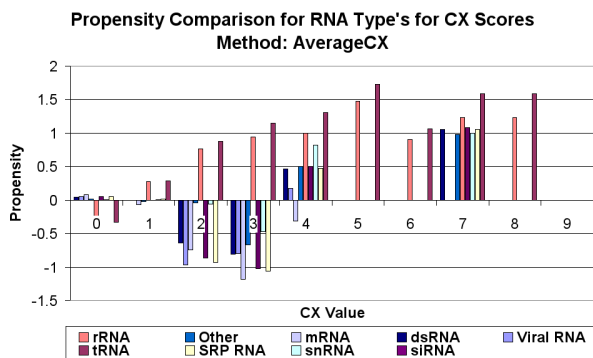
Figure 2. Propensity scores for CX values (Y-axis) calculated using the *averagecx* for different ranges of CX values (X-axis). Propensities for CX values 0-4 tend to vary across different RNA types as compared to propensities for higher CX scores. Different colors correspond to different RNA types.
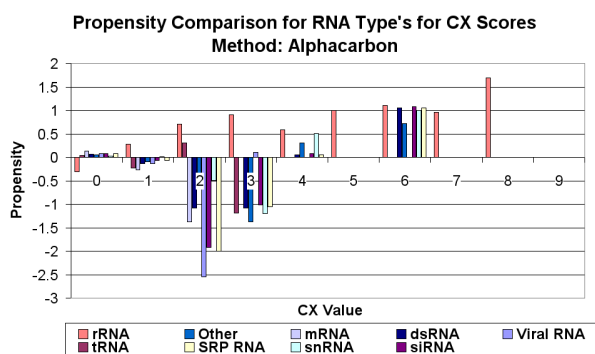

Figure 3. Propensity scores for CX values (Y-axis) calculated using the *alphacarbon* for different ranges of CX values (X-axis). Different colors correspond to different RNA types.
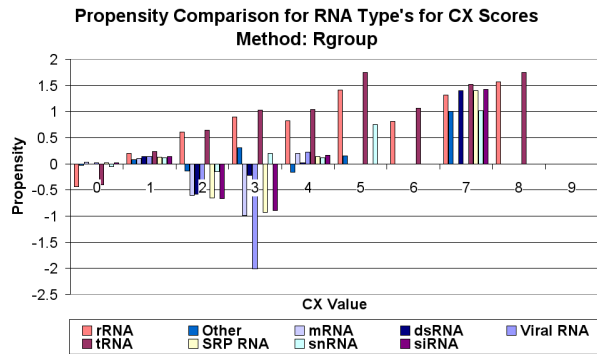

Figure 4. Propensity scores for CX values (Y-axis) calculated using the *rgroup* for different ranges of CX values (X-axis). Different colors correspond to different RNA types.
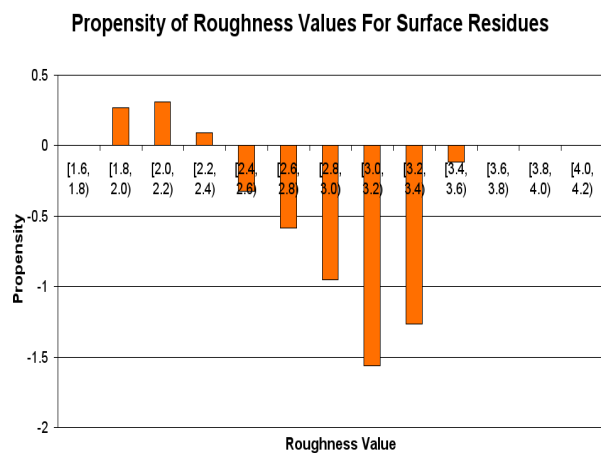

Figure 5. The propensity values for various roughness scores on the surface residues.

3.2 Roughness Value

Figure 5 shows the propensities for the roughness score. Residues with rough surfaces are preferred in protein-RNA interfaces. Figure 6 shows the distribution of the propensities for each roughness score range classified by the type of RNA. The figure indicates that, unlike the CX scores, the roughness scores do not vary significantly among the different types of RNA at each score range as all RNA types behave almost identically for each roughness score range (ANOVA p-value = 0.32 with cutoff 0.05). Using the variances calculated by ANOVA, tRNA, SRP RNA, snRNA, rRNA and dsRNA seem to cluster well with each other (all have variances close to 0.2). The remaining RNA types (Viral RNA, siRNA, mRNA and "other") have variances around 0.1.
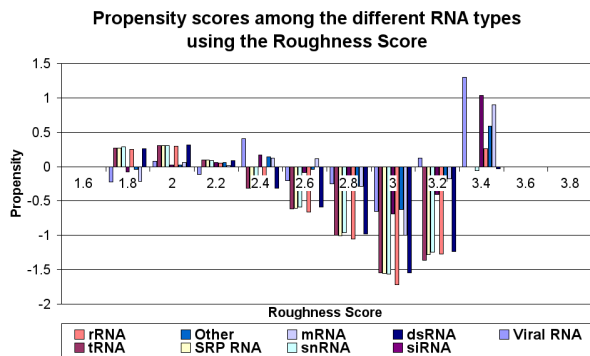
scores increase. However, all other types of RNA (mRNA, snRNA...etc) tend to have low (negative) propensity values from [0,4), and the propensities for all types then tend to rise after CX range [4,5). One interesting exception is the score range [0,1), which tends to have slightly positive (albeit very small) interface propensities in the case of RNAs other than tRNA and rRNA, suggesting that in protein-RNA interfaces containing such RNAs non-protruding residues can be interface residues. The CX value residue propensities of interface residues bound to Viral RNAs and rRNAs appear to differ significantly from that of residues that bind to other types of RNAs figures 2, 3 and 4.

Figure 6. Propensity scores for roughness values classified by RNA type. Roughness propensity is similar across different types of RNA interfaces

3.3   Solid Angle

Figure 7 shows the solid angle propensities. Interface residues seem to prefer residues whose surfaces are not too concave and not too convex. Figure 8 shows the solid angle propensities across the various RNA types in table 1. ANOVA analysis across all RNA types (p-value = 0.998 with cutoff = 0.05) shows that the tRNA, SRP RNA, snRNA, rRNA and dsRNA types cluster well with each other (variances for each of these types is around 0.2) while Viral RNA, siRNA, "other" and mRNA types cluster together (variances around 0.05 each). It is interesting to note that Viral RNA, mRNA and siRNA-binding amino acid residues appear to differ slightly with respect to solid angle propensities from those that bind other types of RNA (see figure 7). Specifically, interface residues that bind to mRNA and siRNA appear to have a preference for more concave surfaces.

4   Summary and Discussion

We have analyzed a non-redundant dataset of protein-RNA interfaces in terms of three structural properties of amino acid residues, namely, CX score, roughness, and solid angle. The results of our analysis show that:

- Amino acid residues in protein-RNA interfaces tend to be more protruding (as measured by CX values) compared with surface residues.
- Amino acid residues in protein-RNA interface tend to have more rough surfaces compared with surface residues.
- Amino acid residues in protein-RNA interfaces tend to have moderately convex or concave solid angles.
- Amino acid residues in protein-RNA interfaces containing Viral RNA and rRNA significantly differ from
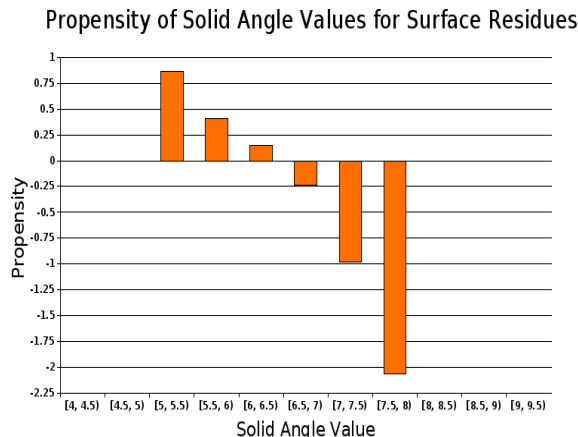

Figure 7. Solid angle propensities. Smaller solid angles appear to be preferred by interface residues

those that contain dsRNA, mRNA siRNA, snRNA, SRP RNA and tRNA with respect to their CX values.

It is possible that the general trends observed across all RNA types is biased by rRNA-binding proteins, which make up over half of the protein-RNA complexes in PDB. One way to determine whether this is indeed the case is to repeat the protein-RNA interface analysis separately for each RNA-type. Also of concern is the relatively small size of the RB147 dataset protein-RNA dataset. Terribilini *et al.* [21] have noted that PDB included only 198 proteinnucleic acid complexes in 1996, but by April 2007, this number had grown to 1,734, of which 529 were proteinRNA complexes. The resulting availability of larger and more diverse datasets can be expected to significantly improve the quality and quantity of data available for performing the analysis of the sort reported here.

Work in progress is aimed at:

- Assembling a comprehensive of protein-RNA interface database (PRIDB), and associated services for querying, analysis, and visualization of protein-RNA interfaces.
- Developing machine learning approaches for reliable identification of putative RNA-binding residues in proteins that improve upon the state-of-the-art sequence-based methods [20] by taking advantage of structural information when it is available.
- Analysis of sequence and structural properties of protein-binding residues in RNA and the development of machine learning approaches to reliable identification of putative protein-binding RNA residues.
- Characterization of the sequence and structural correlates of of protein-RNA interfaces and the similarities
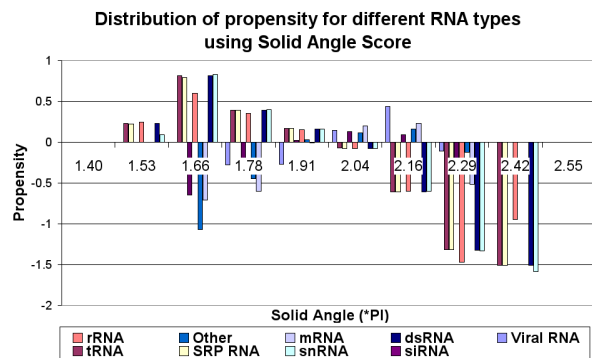
Figure 8. Propensity scores for solid angle values classified by RNA type. The scores for solid angle seem to indicate that the solid angle propensities vary across RNA types less than than CX value propensities, but more than roughness propensities. Propensities for solid angles in the range 1.66π to 2π range appear to vary substantially across different RNA types.

and differences among different types of protein-RNA interfaces, and between protein-protein, protein-DNA, and protein-RNA interfaces.

References

[1] E. Bechara, L. Davidovic, M. Melko, M. Bensaid, S. Tremblay, J. Grosgeorge, E. W. Khandjian, E. Lalli, and B. Bardoni. Fragile x related protein 1 isoforms differentially modulate the affinity of fragile x mental retardation protein for g-quartet rna structure. *Nucleic Acids Research*, 35:299–306, 2007.

[2] H. Berman, J. Westbrook, Z. Feng, and et al. The protein data bank. *Nucleic Acids Res*, 28:235–242, 2000.

[3] Y. Chen and G. Varani. Protein families and rna recognition. *FEBS Journal*, 272(9):2088–2097, 2005.

[4] M. L. Connolly. Measurement of protein surface shape by solid angles. *Journal of Molecular Graphics*, 4(1):3–6, 1986.

[5] M. L. Connolly. The molecular surface package. *J Mol Graph*, 11(2):139–41, 1993.

[6] J. Ellis, M. Broom, and S. Jones. Protein-rna interactions: structural analysis and functional classes. *Proteins*, 66(4):903–11, 2007.

[7] E. O. Freed and A. J. Mouland. The cell biology of hiv-1 and other retroviruses. *Retrovirology*, 3(77), 2006.

[8] E. Jeong, I.-F. Chung, and S. Miyano. A neural network method for identification of rna-interacting residues in protein. *Genome Informatics*, 15(1):105–116, 2004.

[9] E. Jeong and S. Miyano. A weighted profile method for protein-rna interaction prediction. *Trans. On Comput. Syst. Biol.*, IV:123–139, 2006.

[10] S. Jones, D. T. A. Daley, N. M. Luscombe, H. M. Berman, and J. M. Thornton. Protein-rna interactions: a structural analysis. *Nucleic Acids Research*, 29(4):943–954, 2001.

[11] M. S. Jurica and M. J. Moore. Pre-mrna splicing: awash in a sea of proteins. *Mol. Cell*, 12:5–14, 2003.

[12] B. Lee and F. M. Richards. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol*, 55:379–400, 1971.

[13] D. Lejeune, N. Delsaux, B. Charloteaux, A. Thomas, and R. Brasseur. Protein-nucleic acid recognition: statistical analysis of atomic interactions and influence of dna structure. *Proteins*, 61(2):258–71, 2005.

[14] M. Lewis and D. Rees. Fractal surfaces of proteins. *Science*, 230(4730):1163–1165, 1985.

[15] T. M and W. E. Statistical analysis of atomic contacts at rna-protein interfaces. *J Mol Recognit*, 14(4):199–214, 2001.

[16] M. J. Moore. From birth to death: the complex lives of eukaryotic mrnas. *Science*, 309:15141518, 2005.

[17] H. F. Noller. Rna structure: reading the ribosome. *Science*, 309:15081514, 2005.

[18] A. Pintar, O. Carugo, and S. Pongor. Cx, an algorithm that identifies protruding atoms in proteins. *Bioinformatics*, 18(7):980–4, 2002.

[19] M. Terribilini, J.-H. Lee, C. Yan, R. L. Jernigan, S. Carpenter, V. Honavar, and D. Dobbs. Identifying interaction sites in "recalcitrant" proteins: Predicted protein and rna binding sites in rev proteins of hiv-1 and eiav agree with experimental data. In *Pacific Symposium on Biocomputing*, volume 11, pages 415–426, 2006.

[20] M. Terribilini, J.-H. Lee, C. Yan, R. L. Jernigan, V. Honavar, and D. Dobbs. Prediction of rna binding sites in proteins from amino acid sequence. *Bioinformatics*, 12:1450–1462, 2006.

[21] M. Terribilini, J. D. Sander, J.-H. Lee, P. Zaback, R. L. Jernigan, V. Honavar, and D. Dobbs. Rnabindr: a server for analyzing and predicting rna-binding sites in proteins. *Nucleic Acids Research*, 35(2):1–7, 2007.

[22] F. Wu, F. Towfic, D. Dobbs, and V. Honavar. Analysis of protein protein dimeric interfaces. *IEEE International Conference on Bioinformatics and Biomedicine proceedings*, 2007. In Press.