Pacific Symposium on Biocomputing 15:269-280(2010)

STRUCTURAL PREDICTION OF PROTEIN-RNA INTERACTION BY COMPUTATIONAL DOCKING WITH PROPENSITY-BASED STATISTICAL POTENTIALS

LAURA PÉREZ-CANO[1], ALBERT SOLERNOU[1], CARLES PONS[1,2], JUAN FERNÁNDEZ-RECIO[1]

[1]*Life Sciences Department, Barcelona Supercomputing Center (BSC),*

[2] *Computational Bioinformatics, National Institute of Bioinformatics (INB),*

Jordi Girona 29, Barcelona, 08034, Spain

Despite the importance of protein-RNA interactions in the cellular context, the number of available protein-RNA complex structures is still much lower than those of other biomolecules. As a consequence, few computational studies have been addressed towards protein-RNA complexes, and to our knowledge, no systematic benchmarking of protein-RNA docking has been reported. In this study we have extracted new pairwise residue-ribonucleotide interface propensities for protein-RNA, which can be used as statistical potentials for scoring of protein-RNA docking poses. We show here a new protein-RNA docking approach based on FTDock generation of rigid-body docking poses, which are later scored by these statistical residue-ribonucleotide potentials. The method has been successfully benchmarked in a set of 12 protein-RNA cases. The results show that FTDock is able to generate near-native solutions in more than half of the cases, and that it can rank near-native solutions significantly above random. In practically all these cases, our propensity-based scoring helps to improve the docking results, finding a near-native solution within rank 100 in 43% of them. In a remarkable case, the near-native solution was ranked 1 after the propensity-based scoring. Other previously described propensity potentials can also be used for scoring, with slightly worse performance. This new protein-RNA docking protocol permits a fast scoring of rigid-body docking poses in order to select a small number of docking orientations, which can be later evaluated with more sophisticated energy-based scoring functions.

1. Introduction

Understanding the molecular mechanism of protein-RNA recognition in order to understand and predict such interactions is one of the grand challenges in structural biology. In recent years, the growing awareness for the importance of RNA in the context of protein-RNA interactions, together with the publication of the 50S and 30S ribosome subunits,[1,2] have increased the volume of data on these complexes. However, in spite of this, the level of structural knowledge of protein-RNA association is quite poor in comparison to that of other biomolecules.

Given that experimental determination of protein-RNA complexes at high resolution is challenging (being crystallization one of the main bottle-necks), computational approaches for structural modeling at different resolution levels are increasingly needed in order to complement existing experimental data on protein-RNA interactions of interest. One promising tool is computational protein-RNA docking, which can provide structural models at residue and even atomic resolution level, but there are still very few reported methods as compared to protein-protein or protein-ligand docking, and certainly there are no systematic studies on large data sets. To our knowledge, the largest reported benchmark set for protein-RNA docking consisted on five cases.[3] In this context, the Critical Assessment of PRediction of Interactions (CAPRI) experiment (http://www.ebi.ac.uk/msd-srv/capri), a blind international docking competition to evaluate the performances of protein-protein computational docking methods, proposed recently the first protein-RNA complex target. The experiment nicely showed how some docking methods can be adapted to predict the tridimensional structure of a protein-RNA complex. Indeed, our pyDock scoring protocol,[4] which achieved excellent results for protein-protein docking in past CAPRI tests,[5] identified the second best model among all participants in the scorer experiment, with excellent ligand RMSD of 3.8 Å with respect to the X-ray structure of the complex. However, this experiment also highlighted the limitations of current methods. In addition to better treatment of flexibility, new scoring parameters specifically adapted for protein-RNA binding are needed.

In this sense, a number of studies have used available structural data of real protein-RNA interfaces in order to understand this type of interaction and extract better parameters for predictions.[6-16] Some of these studies reported individual propensities and all-atom statistical potentials for the characterization of modeled protein-RNA interactions at atomic level. For instance, all-atom hydrogen-bond statistical potentials have been applied to identify

near-native docking solutions.[3] Other all-atom statistical potentials found interesting details of the protein-RNA interaction.[12-16] However, coarse-grained statistical potentials at residue-nucleotide level have the advantage of low computational cost for their application to larger benchmark sets in order to develop new docking methods (something that is needed before considering more detailed functions). In a previous work,[6] we extracted and successfully used single residue interface propensities for protein-RNA to identify RNA-binding sites in proteins. In the present study, we have extracted new pairwise residue-ribonucleotide propensities for protein-RNA with the goal of being of predictive value for docking. For that, we have used a standard FFT-based approach to generate protein-RNA rigid-body docking poses, and then we have used the new propensities to successfully score these docking poses. We have also tested other reported single residue or residue-ribonucleotide propensities,[6,9,15,16] in order to check the capability of these statistical potentials for the scoring of rigid-body docking poses.

2. Methods

2.1. *Pairwise residue-ribonucleotide interface propensities from protein-RNA structural data*

We extracted protein-RNA pairwise interface propensities using the same training data set that we previously used to extract individual interface propensities, which was composed of 282 non-redundant protein-RNA interactions.[6] These propensities can be calculated from the observed frequency of the specific residue-ribounucleotide pairs of type *pq* (p = 1...20 for amino acid residues; and q = 1...4 for ribonucleotides) at the protein-RNA interfaces, as compared with the expected frequency of these pairs according to the protein and ribonucleotide surface composition, as it is shown by equations 1-4:

$$P^I_{pq} = \frac{N^I_{pq} / \sum_{pq} N^I_{pq}}{N^S_p / \sum_p N^S_p \times N^S_q / \sum_q N^S_q} \qquad (1)$$

where N^I_{pq} is the number of pairs between residue type p and ribonucleotide type q at the protein-RNA interfaces (the pairs were defined by having at least one atom within 4 Å distance from each other), ΣN^I_{pq} the total number of residue-ribonucleotide pairs at protein-RNA interfaces, N^S_p and N^S_q the number of surface residues and ribonucleotides of type p and q respectively (surface residue or ribonucleotide were defined as those with accessible surface area ASA > 0.1 Å^2), and ΣN^S_p and ΣN^S_q the total number of surface residues and ribonucleotides, respectively.

Then, propensities P^I_{pq} were easily converted to free-energy estimates or statistical potentials by equation 2:

$$\Delta G^{stat}_{pq} = -RT\ln(P^I_{pq}) \qquad (2)$$

The statistical potential ΔG^{stat}_{pq} thus represents the empirical energy of forming a pair between a residue of type p and a ribonucleotide of type q at the interface, given their frequencies at the protein and ribonucleotide surfaces, being R the gas constant and T the absolute temperature (we have used here RT as 0.59 kcal/mol). Therefore, negative statistical potential values indicate favorable binding energies.

2.2. *Protein-RNA rigid-body docking and scoring by propensity-based statistical potentials*

We used FTDock[17] to generate 10.000 protein-RNA docking poses. We used the same FTDock version as we previously used for testing our pyDock method for scoring of protein-protein docking, that is, with no electrostatics and 1.2 Å grid resolution.

We evaluated all generated docking poses by a very fast algorithm that scored solutions based on the existing contacts at interface. For every residue-ribonucleotide pair at the interface of the docking pose (that is, those that have at least one atom within 4 Å distance from each other), the corresponding propensity value according to its type was assigned. The propensity-based values of all pairs were summed and formed the final score of the given docking pose *i*, as in equation 3:

$$\Delta G_i^{stat} = \sum_{pq} \Delta G_{pq}^{stat} \qquad (3)$$

Finally, all docking solutions were ranked according to these propensity-based scores. For comparison, we also tested other previously described propensity values, either pairwise residue-ribonucleotide or single residue propensities. In the case of pairwise residue-ribonucleotide propensities, we converted the reported values (usually observed and expected frequencies) to statistical potentials as above described. In the case of single residue propensities, we summed the corresponding values of all interface residues according to their types.

2.3. *Benchmarking the method on known protein-RNA complex structures*

In order to benchmark our method, we compiled a set of non-redundant protein-RNA complexes of known structure, in which there is an available unbound structure for at least one of the two components. This produced a total of 12 cases, two of which had available structure for both unbound protein and RNA molecules, five had available only the unbound protein structure, and the remaining five had available structure for only the unbound RNA (Table 1). In order to avoid redundancy we ensured there was no more than 70% of sequence identity between any pair of proteins within the data set. On the other hand, we considered as unbound proteins those with more than 95% of sequence identity with respect to the bound protein, and as unbound RNAs those with more than 85% of sequence identity with respect to the bound RNA structures in the protein-RNA complexes.

Table 1. Structural data set of protein-RNA interactions used in this study. For each molecule the PDB and chain identifiers are shown. The RMSD in Å between the receptor or ligand used here and the bound structure is also shown in brackets.

Name	Complex PDB	Receptor PDB (RMSD)	Ligand PDB (RMSD)
Tyrosyl-tRNA synthetase splicing factor / group I intron RNA	2RKJ_a:c	1Y42_x (0.9)	1Y0Q_a (3.0)
Ct domain of elongation factor SelB / SECIS RNA	1WSU_a:e	1LVA_a (0.7)	1MFK_a (3.1)
NF-Kb / anti-NFKb RNA aptamer	1OOA_a:c	1OOA_a (0.0)	2JWV_a (5.4)
Stnthetic Fab / P4-P6 ribozyme domain	2R8S_l:r	2R8S_l (0.0)	1HR2_a (4.3)
Elongation factor SelB from E.Coli / SECIS RNA	2PJP_a:b	2PJP_a (0.0)	1MFK_a (3.1)
SRP 19 / 7S.S SRP RNA	1LNG_a:b	1LNG_a (0.0)	1Z43_a (2.1)
SRP ribonucleoprotein core Variant 6 / RNA	2PXV_a:b	2PXV_a (0.0)	1CQL_a (8.1)
RNA-binding protein 15.5 K complexed / RNA	1E7K_a:c	2JNB_a (3.2)	1EK7_c (0.0)
HutP / Hut mRNA	1WPU_a:c	1WPV_a (0.2)	1WPU_c (0.0)
Norwakl Virus Polymerase with CTP / RNA	3BSO_a:p	1SH0_b (1.3)	3BSO_p (0.0)
Pp7 Coat protein dimer in complex / RNA hairpin	2QUX_a:c	2QUD_a (0.8)	2QUX_c (0.0)
Structure of 9-subunit archaeal exosome / RNA	2JEA_a:c	2JE6_a (0.4)	2JEA_c (0.0)

This set of structures was used to benchmark the docking results. We compared the predicted docking poses with the real protein-RNA complex structures by superimposing protein alpha-carbons of predicted and real complexes, and then calculating the RMSD between the RNA molecules (considering all atoms). A near-native solution was defined as a docking pose with RNA RMSD, calculated as described above, smaller than 10 Å, which is in line with the criteria used in the CAPRI experiment. This was calculated for the 10,000 docking poses generated by FTDock,[17] and we computed success rates as the percentage of cases in which at least a near-native solution was found within a given number of docking poses as scored by the docking algorithm. The success rates expected by random were calculated by randomly shuffling the scores of the docking solutions (the process was repeated 100 times and the average was calculated).

3. Results

3.1. *New pairwise residue-ribonucleotide interface propensities for protein-RNA*

We computed pairwise residue-ribonucleotide interface propensities from a set of protein-RNA complex structures, and then we converted them to statistical binding potentials (see Methods). The resulting values for all residue-ribonucleotide types are shown in Figure 1. We found in this analysis that the most populated residue-ribonucleotide pairs at protein-RNA interfaces are those composed of the amino acid residues arginine (R), lysine (K) and histidine (H). On the contrary, the least favored pairs are composed of the following residues: aspartic acid (D), glutamic acid (E), cysteine (C), valine (V), leucine (L) and isoleucine (I). In most of the cases, for a given residue type, we do not see significant differences in the pairwise propensity values with regards to the four ribonucleotide types. This can be clearly seen in Figure 1, in which the major differences can be found among the residue types and not among the ribonucleotide types. These results are consistent with our previously reported single residue propensities for protein-RNA,[6] and show the important role of electrostatic forces in protein-RNA binding, with negative RNA charge playing a determinant role in RNA-binding areas in proteins. Interestingly, the important role of electrostatics in protein-RNA binding underlines a major difference with protein-protein association, where desolvation and hydrophobic effect seem much more important.

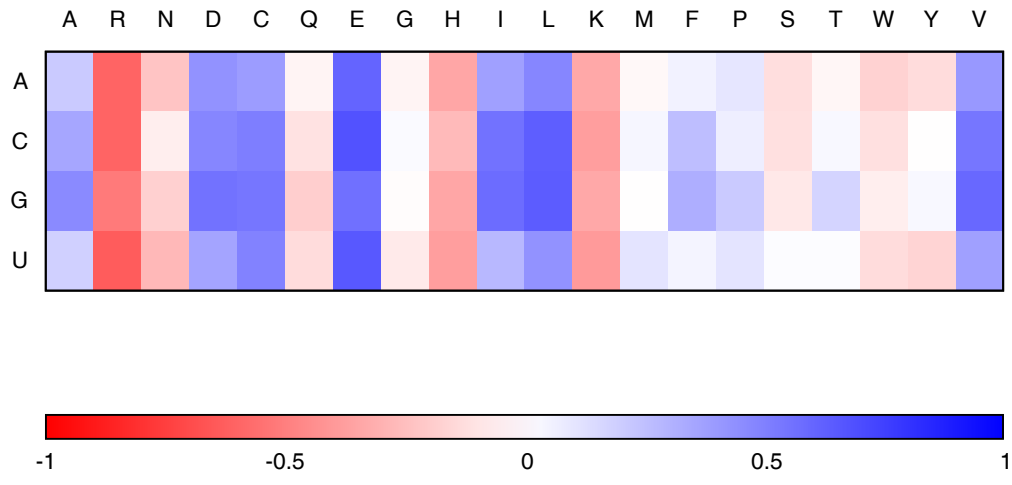

Figure 1. Pairwise protein-RNA statistical potentials (favorable pairs are in red; disfavored in blue).

3.2. *Protein-RNA rigid-body docking and scoring by propensity-based statistical potentials*

For seven out of the 12 cases, the FTDock rigid-body docking generated at least a near-native solution within the total 10,000 docking poses. The results are shown in Table 2. In two cases, the near-native solutions are ranked below 10 by FTDock, but the rest of cases there is no near-native solution ranked below 100. We can also see that the propensity-based scoring alone improves the best rank of a near-native solution in three of the cases, as compared with the original FTDock scores. As a consequence, in three cases a near-native solution is found with rank below 100. Moreover, when the propensity-based scores are used in combination with the original FTDock scores (simply by adding the scores, no weight optimization has been attempted here to avoid overfitting given the small number of cases), in practically all cases the near-native rank values improve with respect to the original FTDock scores (except 1WPU, which slightly worsens from 159 to 170). Actually, in four cases a near-native solution is found with rank below 100. This indicates a clear predictive value of the pairwise propensity scores. Interestingly, the FTDock and the propensity values are quite complementary: in three of the cases, a near-native solution is found with rank below 10 by either FTDock or propensity-based scoring.

Table 2. Results of protein-RNA docking and scoring. The total number of near-native solutions in the docking set is shown. The best rank of a near-native solution (RMSD < 10Å) is shown, after scoring by FTDock, by propensity-based potentials, and by combined score (in brackets is given the RMSD in Å of the near-native solution with respect to the x-ray complex structure, in addition to the fraction of native f_{nat} and non-native $f_{non\text{-}nat}$ contacts as defined in CAPRI).

Complex PDB	Number of near-native solutions	Best near-native rank by FTDock scoring (RMSD) (f_{nat}; $f_{non\text{-}nat}$)	Best near-native rank by propensity-based scoring (RMSD) (f_{nat}; $f_{non\text{-}nat}$)	Best near-native rank by FTDock + propensity scoring (RMSD) (f_{nat}; $f_{non\text{-}nat}$)
1WSU	5	2015 (8.9) (0.23; 0.84)	70 (9.0) (0.23; 0.88)	1049 (9.0) (0.23; 0.88)
2PJP	9	1590 (9.3) (0.43; 0.75)	213 (9.7) (0.43; 0.75)	763 (8.9) (0.43; 0.75)
1LNG	4	131 (5.3) (0.55; 0.22)	660 (5.3) (0.55; 0.22)	92 (5.3) (0.55; 0.22)
1E7K	46	7 (9.6) (0.05; 0.90)	778 (8.7) (0.18; 0.67)	7 (9.6) (0.05; 0.90)
1WPU	44	159 (8.1) (0.35; 0.57)	1989 (9.8) (0.19; 0.74)	170 (8.1) (0.35; 0.57)
2QUX	14	157 (8.1) (0.45; 0.70)	1 (8.5) (0.45; 0.67)	60 (8.1) (0.45; 0.70)
2JEA	17	9 (9.1) (0.18; 0.95)	61 (8.3) (0.18; 0.96)	7 (9.1) (0.18; 0.95)

3.3. *Examples of successful predictions*

It is remarkable that for the 2QUX case (unbound protein vs. bound RNA), the scoring by the new pairwise propensities is able to find a near-native solution ranked 1. As can be seen in Figure 2, the predicted RNA orientation on the protein surface is very close to that in the x-ray structure. The best rank obtained by FTDock scoring was 157, so the effect of using the new propensity-based potentials on the final scoring is dramatic in this case.

Another example of successful application of protein-RNA docking can be found in a recent CAPRI blind experiment. Targets 33 and 34 were a protein-RNA case. In target 33, both molecules (protein and RNA) needed to be modeled since there was no available x-ray structure. That was an extremely difficult case for which no group was able to submit an acceptable model. Target 34 was the same complex, but with the bound structure of the RNA provided by the organizers (although with random orientation). For this target, we generated docking poses with FTDock and with RotBUS (and in-house developed program for rigid-body docking; see upcoming publication). The docking poses were scored by our protein-protein scoring function pyDock,[4] with desolvation parameters for

RNA adapted from those used by us in protein-protein docking. In addition, we applied distance restraints to one residue-ribonucleotide pair and six ribonucleotides that, according to literature, were likely to be at the interface. The result was an acceptable model within the ten submitted models (it was ranked 885 before applying restraints, and rank 3 after restraints). Moreover, in the scorers experiment, our method identified the second best model among all participants, with excellent ligand RMSD of 3.8 Å with respect to the X-ray structure of the complex.

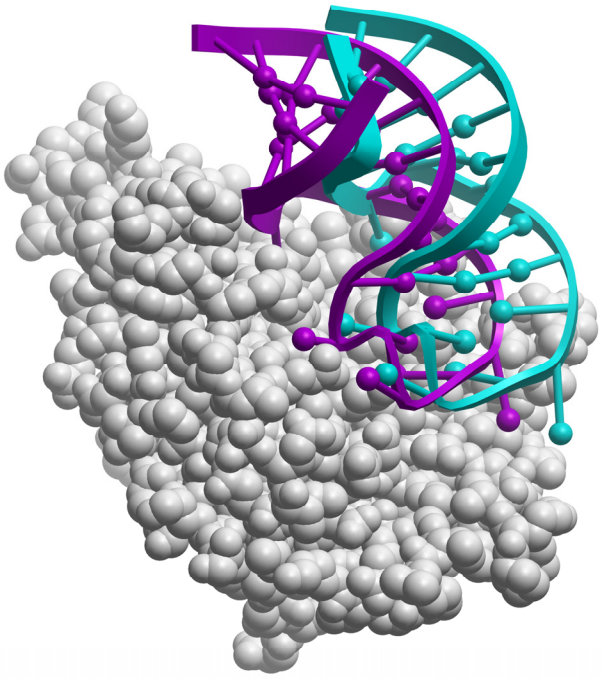

Figure 2. Solution ranked 1 after the new propensity-based scoring in the docking of Pp7 Coat protein dimer with RNA hairpin (protein in white CPK; RNA in cyan ribbon). For comparison, the x-ray structure (PDB 2QUX) of the protein-RNA complex is shown (RNA in magenta ribbon, after superimposing protein molecules).

4. Discussion

4.1. *New pairwise propensities for protein-RNA interaction. Comparison to other reported propensities*

We can compare our new propensities with other reported protein-RNA propensities, all of them derived from smaller data sets.[9,15,16] In practically all of the studies, the favored residues are R and K, but their pairwise preferences for the ribonucleotides change. In our work here and in others,[9,16] there are no significant differences on the preferences of R and K for any ribonucleotide, but in some studies they see preferred pairs as R-U,[11,12,15] or K-A.[11,12] Interestingly, in our study we can see H residues with higher pairwise propensities, while this is not observed in any of the previous studies. Histidine residue can act as positively charged, depending on the environment, so this can explain its high propensities. This was already discussed in our previous study on single residue interface propensities.[6] On the other side, while in our study the aromatic residues are not amongst the most preferred pairs, in other studies they are.[9,16] We have seen before that this can strongly depend on the data set used for deriving the propensities. For instance, we checked that there are more aromatic residues involved in the interaction with single chain RNA molecules, so the proportion of this type of RNA in the data set could modify the propensity values.[6] Interestingly, we see that although the global propensities of pairs involving W and Y residues are small (favorable,

but small), they have slightly higher preferences for A and U ribonucleotides (especially in the case of Y residue). While all the reported propensities may reflect different characteristics of the protein-RNA interaction, we are more confident in the general applicability of the propensities described in this work, given that they were extracted from the largest data set so far (282 non-redundant protein-RNA interactions).[6]

4.2. *Use of other reported propensities (single and pairwise) for scoring*

We have seen above that our new pairwise residue-ribonucleotide propensities improve the scoring of rigid-body protein-RNA docking poses. In order to check whether this represents an advance over other reported protein-RNA propensities, we have also used these other propensities in the same conditions to check their results in docking. For this, we selected different protein-RNA propensities from bibliography[9,15,16] and from an own previous work.[6] We used both pairwise residue-ribonucleotide interface propensities[6,9,15,16] as well as single residue interface propensities.[6,9,16] We compiled a total number of seven different protein-RNA propensity matrices, five of which were pairwise propensities, and two single residue interface propensities. Because the propensity definition varied among the different studies, we considered for all of them the reported expected (F_{obs}) and observed (F_{exp}) frequency values, which we transformed into propensities ($P = F_{obs} / F_{exp}$) and then into their respective statistical potentials following our definition (eq. 2 and 3).

The results of using different propensity values can be seen in Figure 3. In general, the pairwise propensities described in this work, especially when is combined with FTDock score, give the best success rates. Actually, most of the other propensities give success rates that are not significantly above random. However, the pairwise propensities from Westhof laboratory[16] had reasonable good results, indeed significantly above random. Moreover, it is interesting that single residue propensities from our previous work[6] and from Westhof laboratory[16] obtained quite good success rates. Two important conclusions can be derived from this: *i)* the determinant for protein-RNA specificity lies mostly on the protein residues, and *ii)* the predictive value of the statistical potentials depend strongly on the size and composition of the database used to derive the propensities.

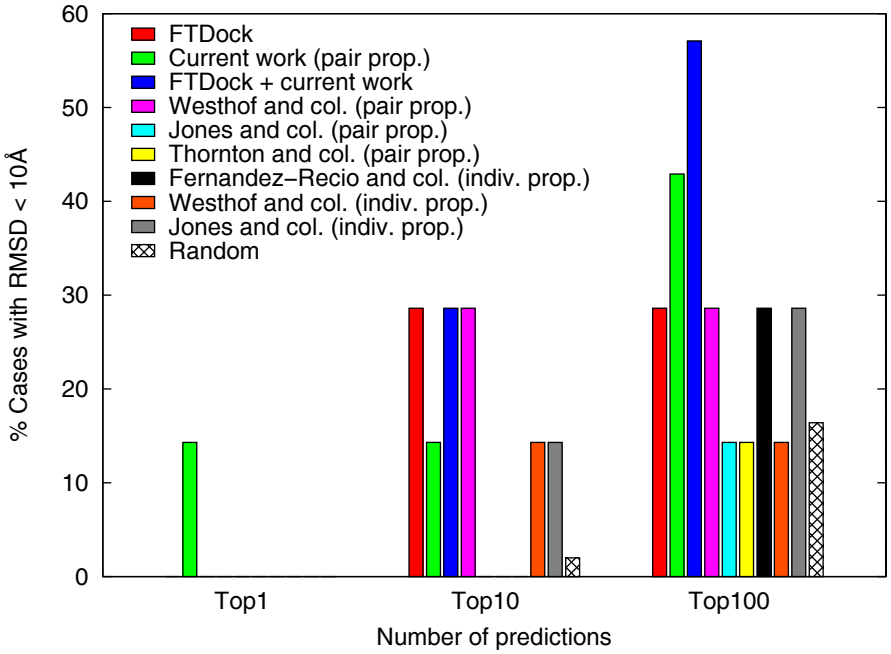

Figure 3. Success rates of protein-RNA docking after scoring by different propensity-based statistical potentials.

4.3. *FTDock docking and scoring performance is better than in protein-protein docking*

While the number of protein-RNA cases for which a near-native solution is found by FTDock (58%) is inferior to that in protein-protein docking (where a near-native solution is found in 77% of the cases), it is interesting that for some cases, the number of near-native solutions is quite high (Table 2). Especially in the cases of 1EK7 and 1WPU, more than 40 near-native solutions are found, a number clearly above the performance in protein-protein docking. Moreover, the success rate of FTDock scoring for the top 10 solutions is well above random (in protein-protein, success rates for FTDock scoring, run in the same conditions as here, was always similar to random).[4] This could indicate that geometry complementarity is more important in protein-RNA than in protein-protein (electrostatics should also be important in the interaction, but the FTDock version we are running here has no electrostatics, so this should not affect the scoring). This higher importance of surface complementarity is consistent with the fact that no statistical preferences are found within the ribonucleotides. That is, while protein residues have different preferences derived from their different chemico-physical characteristics, the ribonucleotides are more similar in thermodynamic behaviour, and perhaps the contribution of RNA to specificity lies more on conformational aspects. Actually, in the cases where FTDock has the best results (1EK7, 2JEA), the RNA molecule is in the bound conformation, and thus the rigid-body approach can take full advantage of the geometry complementarity.

5. Conclusions

In summary, we have proposed here a new protocol for protein-RNA docking, based on FFT rigid-body docking followed by scoring with new pairwise residue-ribonucleotide interface propensities derived from protein-RNA complex structures. The docking approach and new propensities have been tested in the largest protein-RNA benchmark, to our knowledge. The results show that FTDock can be successful if RNA conformation is in the bound conformation, and that the new propensities help to improve the rank of the near-native docking poses in virtually all cases.

From the results shown here and our experience in the blind CAPRI experiment, we can envisage a possible strategy for protein-RNA docking. First, rigid-body docking between protein and RNA based on FFT algorithms or on any other efficient approach. Then, we can use a combination of geometry-based scoring and propensity statistical potentials as a filter to select a few hundred docking poses, which later will be evaluated with more complete energy function. Finally, the use of a minimal information that can be integrated as distance restraints can dramatically improve the results. One of the main bottlenecks for continuing development in protein-RNA docking is the lack of cases in which both the unbound protein and RNA structures are simultaneously available. Therefore, it seems that future developments will need to focus on the use of homology-based models of RNA in docking.

Although beyond the scope of current work, it would be interesting in the future to check the capabilities of our protein-RNA statistical potentials for the prediction of protein-DNA interactions. However, reported data on protein-RNA and protein-DNA propensities highlight the specific differences between these types of interfaces, specifically that in DNA-binding the phosphate group is determinant for the interaction, so protein charged residues are preferred, while in RNA-binding the ribose and nucleotides are more relevant, being the protein aromatic residues key for specificity.[7,8,10,12,14] These findings, together with preliminary tests (data not shown) indicate that our protein-RNA propensities are specific for protein-RNA binding. As future work, derivation of new protein-DNA propensities, in the same fashion as that described here for protein-RNA, could be successfully applied for DNA-binding interface prediction and docking studies.

Acknowledgments

This work is supported by grant BIO2008-02882 from Plan Nacional I+D+i (Spanish Ministry of Science). LP-C is recipient of an FPU fellowship from the Spanish Ministry of Science.

References

1. N. Ban, P. Nissen, J. Hansen, P.B. Moore and T.A. Steitz, Science 289:905-920 (2000).
2. B.T. Wimberly, D.E. Brodersen, W.M. Clemons, R.J. Morgan-Warren, A.P Carter, C. Vonrhein, T. Hartsch and V. Ramakrishnan, Nature 407:327-339 (2000).
3. Y. Chen, T. Kortemme, T. Robertson, D. Baker and G. Varani, Nucleic Acid Res. 32:5147-5162 (2004).
4. T.M Cheng, T.L Blundell and J. Fernández-Recio, Proteins 68:503-15 (2007).
5. S. Grosdidier, C. Pons, A. Solernou and J. Fernández-Recio, Proteins 69:852-8 (2007).
6. L. Pérez-Cano and J. Fernández-Recio, Proteins, *in press* (2009) *[published online: 7 Jul 2009]*.
7. R.P. Bahadur, M. Zacharias and J. Janin, Nucleic Acids Res 36:2705-2716 (2008).
8. C.M. Baker and G.H. Grant, Biopolymers 85:456-470 (2007).
9. J.J. Ellis, M. Broom and S. Jones, Proteins 66:903-911 (2007).
10. N. Morozova, J. Allers, J. Myers and Y. Shamoo, Bioinformatics 22:2746-2752 (2006).
11. O.T. Kim, K. Yura and N. Go, Nucleic Acid Res 34:6450-6460 (2006).
12. D. Leujene, N, Delsaux, B. Charloteaux, A. Thomas and R. Brasseur, Proteins 61:258-271 (2005).
13. E. Jeong, H. Kim, S. Lee and K. Han, Mol. Cells 16:161-167 (2003).
14. J. Allers and Y. Shamoo, J. Mol. Biol. 311:75-86 (2001).
15. S. Jones, D. Daley, N. Luscombe, H. Berman and J. Thornton, Nucleic Acid Res. 29:943-954 (2001).
16. M. Treger and E. Westhof, J. Mol. Recogn. 14:199-214 (2001).
17. H.A. Gabb, R.M. Jackson, M.J. Stemberg, J. Mol. Biol. 272:106-120 (1997).