W502–W506 *Nucleic Acids Research, 2015, Vol. 43, Web Server issue* *Published online 11 June 2015*
doi: 10.1093/nar/gkv557

RNAssess—a web server for quality assessment of RNA 3D structures

Piotr Lukasiak[1,2,*], Maciej Antczak[1], Tomasz Ratajczak[1], Marta Szachniuk[1,2], Mariusz Popenda[2], Ryszard W. Adamiak[1,2] and Jacek Blazewicz[1,2]

[1]Institute of Computing Science, Poznan University of Technology, Piotrowo 2, 60-965 Poznan, Poland and [2]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Received March 9, 2015; Revised May 15, 2015; Accepted May 16, 2015

ABSTRACT

Nowadays, various methodologies can be applied to model RNA 3D structure. Thus, the plausible quality assessment of 3D models has a fundamental impact on the progress of structural bioinformatics. Here, we present RNAssess server, a novel tool dedicated to visual evaluation of RNA 3D models in the context of the known reference structure for a wide range of accuracy levels (from atomic to the whole molecule perspective). The proposed server is based on the concept of local neighborhood, defined as a set of atoms observed within a sphere localized around a central atom of a particular residue. A distinctive feature of our server is the ability to perform simultaneous visual analysis of the model-reference structure coherence. RNAssess supports the quality assessment through delivering both static and interactive visualizations that allows an easy identification of native-like models and/or chosen structural regions of the analyzed molecule. A combination of results provided by RNAssess allows us to rank analyzed models. RNAssess offers new route to a fast and efficient 3D model evaluation suitable for the RNA-Puzzles challenge. The proposed automated tool is implemented as a free and open to all users web server with an user-friendly interface and can be accessed at: http://rnassess.cs.put.poznan.pl/

INTRODUCTION

RNAs are one of the most important molecules from a biological and medical point of view, since they are involved in a wide range of biochemical reactions in cells, as well as in playing regulatory roles. The activity of RNA is strongly dependent on its 3D structure. Various experimental methods, such as X-ray crystallography, nuclear magnetic resonance or cryo-microscopy have been applied to determine RNA 3D structure. Unfortunately, experimental determination of a high-resolution 3D structure of RNA is often difficult, because of the nature of RNA molecules. Most of functional RNAs have not been determined. Thus, several computational methods to predict RNA 3D structure have been introduced, namely RNAComposer (1,2), MC-Fold/MC-Sym (3), MMB (4), NAST (5), DMD Vfold (6), Swa rna loop (7), iFoldRNA (8), ModeRNA (9). Computational approaches can be divided into various categories according to: requirements of expert knowledge (i.e. manual, semi-automated and automated) and characteristics of input data needed for prediction (i.e. sequence only, sequence and secondary structure, and sequence and structural template).

The quality evaluation of models in the context of the reference structure can be performed in various ways using wide range of measures. RNA 3D structures are most often evaluated by numerical measures, namely root-mean-square deviation (RMSD), which is a measure that uses a rigid body translation and rotation of one atoms set with respect to the other set by a least-squares superposition procedure (10), interaction network fidelity (INF) that is a balanced measure of binary classification of the base-stacking and base-pairing annotations between a model and a reference structure (11), Deformation Index (DI), which measures dissimilarities between structures at the nucleotide scale for both intra-domain and inter-domain interactions (11), mean of circular quantities (MCQ) that estimates structural similarity based on a torsion angle space representation of 3D structure (12), CAD-score, which measures contact area distance (13), *P*-value that estimates the probability that a given structure prediction is better than that expected by chance (14).

Moreover, visual approaches such as deformation profile (11), which depicts the conformation differences between compared 3D structures in local, inter-domain, and intra-domain scales, and visualizations provided by RNAlyzer (15) showing both global and local coherence between a particular model and a reference structure (e.g. 2D map and 3D landscape) and cutoff, which determines the number of local neighborhood areas observed around residues for a

*To whom correspondence should be addressed. Tel: +48 61 6653033; Fax: +48 61 8771525; Email: piotr.lukasiak@cs.put.poznan.pl

fixed sphere radius predicted below a certain threshold can be also applied.

To our knowledge, several applications have been presented to deal with different evaluation aspects that do not require a reference RNA 3D structure to carry out the quality assessment and are available online: wwPDB Deposition Tool (16), MolProbity (17), CAD-score (13), Rclick (18), Setter (19) and WebRasp (20). However, there is no single, online available tool allowing user to perform models evaluation against a reference structure. To fill this gap RNAssess server is proposed. This tool was developed on the RNAlyzer (15) framework and is equipped with a wide range of new functionalities allowing for easy online analysis of RNA 3D models in the reference structure context. It allows users to analyze RNA 3D structures quality using well-established measures for both global and local quality assessment and visualization of RNA 3D structure regions of interest in atomic perspective. RNAssess offers novel opportunities for those working on RNA 3D structure prediction methods and their evaluation (RNA-Puzzle contest (21)). It should be also of interest of the rapidly growing community of RNA 3D modelers. Since RNAssess can compare any kind of RNA 3D structure of the same sequence, we envisage its future application in RNA molecular dynamics simulation.

METHOD OUTLINE

The proposed server integrates a set of aforementioned measures, namely RMSD (a general and adjusted version is considered; the latter is calibrated by the particular sphere atoms coverage), deformation index, INF (considering various base–base-pairing interactions such as canonical, non-canonical, stacking and all), solving their disadvantage of sensitivity to local deviations by changing the accuracy level perspective of an analyzed molecule in an interactive way. Moreover, *P*-value and cut-off can be applied for the global quality assessment only. The latter is a percentage of spheres built around certain nucleotide residues for the particular radius (considering 12 and 16Å), with the RMSD value below the threshold of 2Å. The implemented web server incorporates a set of web services and through the combination of information obtained provides: (i) ranking of multiple 3D models, which reflects the coherence in the context of the reference structure, (ii) visualization of prediction accuracy of atoms located in the neighborhood of every nucleotide for a fixed accuracy level (i.e. sphere visualization) (iii) the ability to change the presentation of average values of a used measure with an increasing structural neighborhood (sphere radius) (iv) colored 2D map and 3D landscape that allow to observe structural regions where the prediction is inaccurate for a particular model and all considered accuracy levels (v) cut-off plot that shows how accurate the prediction of a particular model is from a local point of view, e.g. the recognition of structural regions of models predicted correctly (below a certain threshold). A global quality assessment of each 3D model is extended by preparation of well-known Deformation Profile plot (11).

RNAssess allows us to perform the analysis using a wide range of measures and ranks considered models globally (the results are presented in the global quality analysis section) and locally (based on the concept of spheres introduced by RNAlyzer tool (15)). The user can evaluate and visualize the coherence between 3D models and the reference structure of RNA molecule by several spatial comparisons conducted between the sets of atoms located inside a series of spheres built around residues through an analyzed molecule chain. A sphere is built for a fixed radius around the central atom selected by the user (C1′, O5′, O3′ or P) of every residue of the reference structure. As a result of the sphere-building process, a set of all atoms located in the particular sphere is obtained. Next, for every sphere built of the reference structure residue, a corresponding set of atoms from the analyzed model is identified. Finally, substructures located inside corresponding spheres are compared with each other using a measure selected by the user. The scheme of a computational process of RNAssess is presented in Figure 1.

Structural regions built for selected accuracy level and residue (i.e. sphere view) can be visualized in the context of the reference structure using JSmol in a wireframe or cartoon view. RNAssess supports two interactive plots, namely multimodel and cutoff, where the user can dynamically adjust the accuracy level (i.e. sphere radius), window width of the chart presentation and threshold. A sphere visualization of corresponding superimposed structural regions of the considered models is displayed in the case of selecting of both sphere radius and residue.

IMPLEMENTATION

RNAssess system is designed in a two-layer architecture (depicted in Supplementary Figure S1). The first one is a user-friendly web application stacked on PHP. The back-end layer is represented by RESTful web services. The system allows users to efficiently analyze data using interactive charts and visualize interesting structural regions on demand. The computational component aggregates selected functionalities of RNAView (22), MC-Annotate (23) (to identify base pairs in the 3D structure of RNA when INF-based measure is used), Molprobity3 (17) (to use ClashScore routine), Gnuplot (to efficiently generate high-quality plots) and Deformation Profile Scripts (11) (to produce a graphical representation of a discrepancy metric between a pair of RNA 3D models for the same molecule). RNAssess can be operated through the most popular web browsers, such as Google Chrome, Mozilla Firefox, Opera and Internet Explorer, running under MacOS, Unix/Linux and Windows operating systems. To ensure efficiency and fault tolerance of RNAssess system, the performance experiments have been conducted. The analyzed models have been generated by RNAComposer web server (1) as well as downloaded from RSCB PDB (16). Secondary structures were verified using RNApdbee (24). A summary of the obtained results is presented in Supplementary Tables S1 and S2.

RNAssess server supports several usage scenarios. In a typical scenario, the user uploads pdb files of 3D models that will be assessed in the context of the reference structure, selects the measure, comparison mode (all atoms or one particular atom per each residue only), atom treated as the center of the sphere, considered accuracy levels and submits the request to the server. The correctness of the input

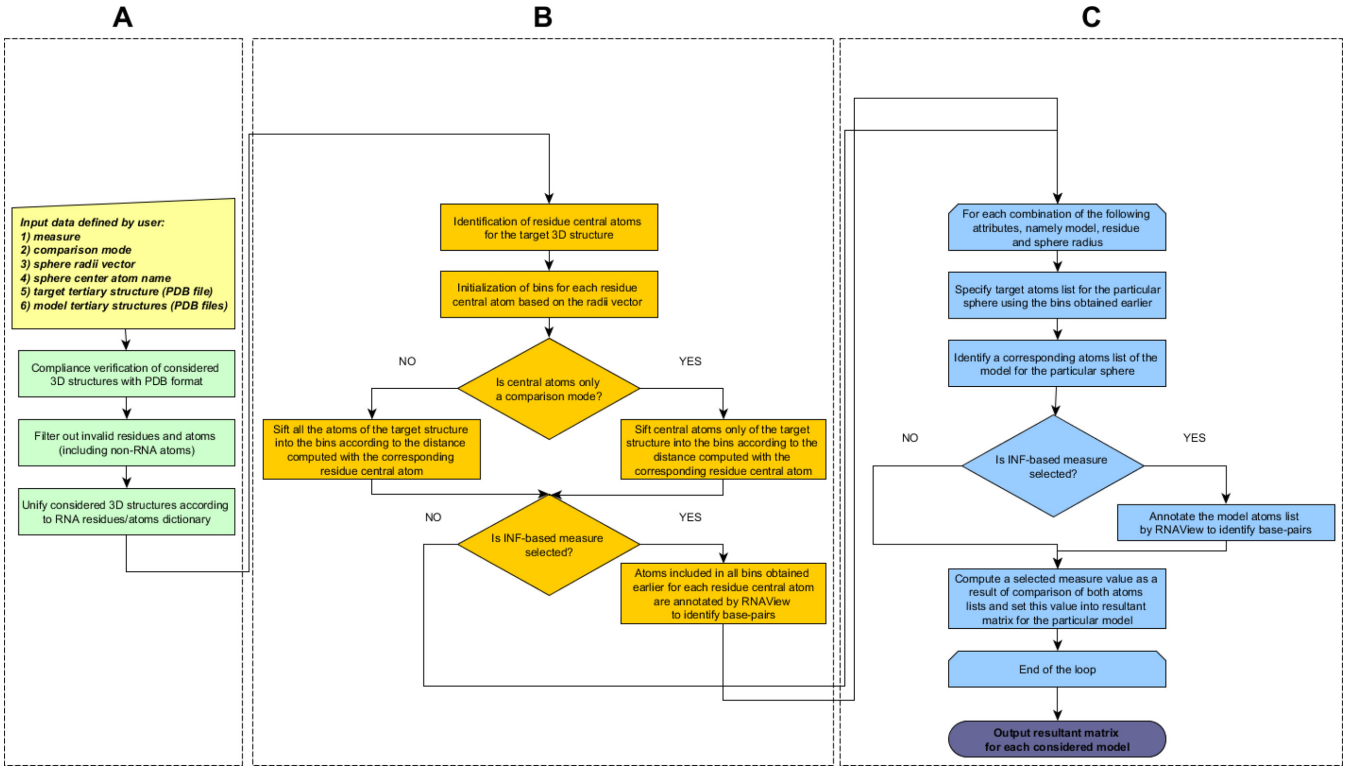

Figure 1. Workflow scheme of RNAssess computational process. (**A**) Input data reading, verification and unification, (**B**) a reference 3D RNA structure analysis involving computation of the atoms set of spheres built for every residue of the reference structure and every sphere radius depicted by the user, (**C**) Quality assessment of analyzed 3D RNA models involving measurement of convergence between considered models and the reference structure.

of data is validated. When any of the submitted files (i.e. any model or the reference structure) is inconsistent with the pdb file format (16) a log window will be displayed including a list of encountered errors. The models submitted for evaluation should be aligned exactly to the reference structure, otherwise only the subset of atoms being coherent with the reference structure will be analyzed. An output consists of three components, namely a global quality analysis section extended by Deformation Profile, static and interactive linear plots and section including 2D map and 3D landscape plots.

The Supplementary Figure S2 presents snapshot of the global quality analysis table, where the results of the analysis performed for all considered models using aforementioned measures are included. The models can be ranked and filtered on the user's demand according to the value of the particular measure and model names, respectively. Moreover, a user can display the deformation profile of each model by enlarging a corresponding image. The submitted models can also be analyzed using a graphical mode. Each non-interactive graph can be enlarged simply by clicking on it. Multimodel and cutoff plots (interactive ones) can be adjusted on the user demand by changing the value of the sphere radius or threshold, respectively with a slider at the top of plot area. The range selector located at the bottom of the plot area can be used to change the window width of the chart presentation. The values of the used measure as well as the position of nucleotide will be presented on the right hand side of the corresponding model name. The visualization of spheres corresponding to the analyzed models built on the particular residue for a fixed accuracy level, will be presented using the JSmol tool.

RNAssess server provides four input data examples, where three of them came from RNA-Puzzles contest (21) and the fourth one was introduced by the authors to present significant advantages of the proposed approach. By selecting one of them, RNAssess proceeds with computation and takes the user directly to the results page. The user can download all static charts presented on results page.

RNAssess server is free and publicly available to all users under http://rnassess.cs.put.poznan.pl.

RESULTS AND DISCUSSION

RNAssess was developed to assess the quality of RNA 3D models against the reference structure. The system provides a set of web services that allow the user to perform a comprehensive analysis in a numerical and graphical manner of RNA 3D models. In contrast to the other applications currently available, RNAssess is equipped with the set of well-established measures enriched by visualization tools allowing the user to depict results of the analysis together with an easy access to visualizations of the corresponding RNA 3D regions.

To present the advantages of the proposed server, we give a short description of the analysis performed on four RNA 3D models generated by RNAComposer system (1). RNA of *Tetrahymena* ribozyme (PDB Id: 1X8W) was chosen as a reference molecule. A model denoted by 1X8W1 is gener-

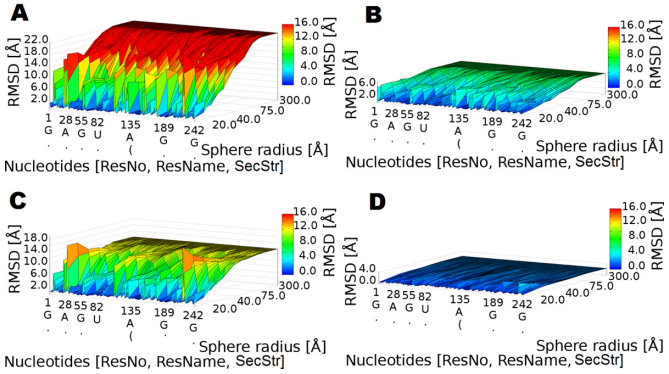

Figure 2. 3D landscapes of predicted 3D models (**A**: 1X8W1, **B**: 1X8W2, **C**: 1X8W3, **D**: 1X8W4) for Tetrahymena ribozyme (PDB Id: 1X8W) (X, Y and Z axes correspond to the nucleotide sequence, RMSD value and sphere radius value respectively). Looking at them along Z–axis, one can observe changes in the considered measure values that appear within the increasing surrounding of the considered nucleotide residues in question, from local to global perspective. Here, model A presents substantial discrepancies, which are highly propagated with the increasing sphere radii, while model D closely resembles the reference structure. In models B and C, respectively, structural discrepancies observed from a local perspective are compensated in the global perspective.

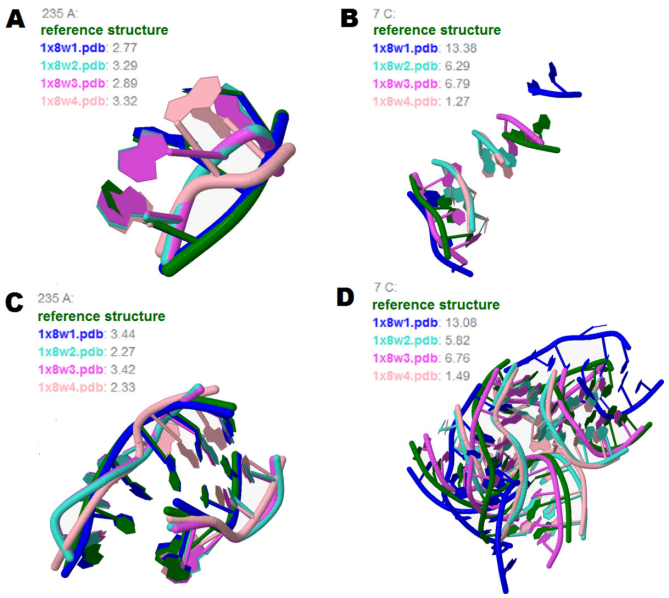

Figure 3. Visualization of atoms located in the neighborhood of the adenosine residue 235 (**A** and **C**) as well as the cytydine residue 7 (**B** and **D**) for the considered models in the reference structure context (PDB Id: 1X8W, sphere radii A, B—10 Å; C, D—20 Å). The reference structure is marked in green. The above type of analysis allows the user to recognize local discrepancies of RNA 3D models that are globally well predicted and to recognize correctly predicted local conformations of RNA 3D models characterized by a low level of prediction accuracy from the global perspective.

ated as a result of default predicion procedure performed with RNAComposer system. The other models, denoted by 1X8W2, 1X8W3 and 1X8W4, were generated with extended RNAComposer structural fragments dictionary including refined counterparts defined by the user. It can be clearly seen that the model 1X8W4 (depicted in Figure 2D) contains the smallest deviations from the reference structure and the model 1X8W1 (depicted in Figure 2A) the largest deviations from global perspective. The other two models 1X8W2 and 1X8W3 (depicted in Figure 2B and C) are located in the middle of the RMSD measure ranking.

The analysis of RNA 3D models that are globally most similar to the reference structure, allows us to identify local regions predicted with lower accuracy in comparison to corresponding regions of the models globally not similar to the reference one. In our example, the surrounding of the adenosine residue 235 (for a sphere radius equal to 10 Å) for the worst model from a global perspective (1X8W4) is the most similar to the corresponding region located in the reference structure among the four models that were analyzed (Figure 3A). The increase of the sphere radius allows us to observe that the prediction accuracy of model 1X8W4 (being globally the best among all considered RNA 3D models) outperforms the others (for a sphere built with a radius equal to 20 Å depicted in Figure 3C). The Figure 3B and D visualize the corresponding regions surrounding cytydine residue 7 observed within the RNA 3D models for the aforementioned sphere radii. The models 1X8W4 and 1X8W1 are evaluated as the most accurately predicted and considerably deviated, respectively, from local as well as global perspective. The above analysis can be easily performed using the interactive multimodel plot where accuracy levels can be customized during the analysis (depicted in Supplementary Figures S2 and S3, respectively).

The modeling accuracy with respect to the reference structure can be also analyzed with a cut-off plot (depicted in Supplementary Figure S4). The best model from the global perspective (1X8W4) is characterized by a high level of prediction accuracy (below 2 Å threshold), but even for this model the percentage of correctly predicted spheres built around nucleotide residues decreases below 50% for the radius higher than 40 Å. Similar analysis can be performed using other measures e.g. Deformation Index (example results are depicted in Supplementary Figures S6–S8).

Our web server is equipped with new functionalities and commonly applied measures. It allows the user to effectively assess the quality of RNA 3D models in the reference structure context. In the future, RNAssess will be enriched by new features such as user-defined alignments constructed between the considered models and the reference structure. Moreover, the quality assessment approach when dealing with unknown reference structures will be developed. We hope that the presented approach will have an impact on the field of RNA 3D structure modeling where even an inexperienced user would be able to upload RNA 3D models obtained for the same input sequence and evaluate their structure quality.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Eric Westhof for the discussion and suggestions. This study has been conducted in the European Centre for Bioinformatics and Genomics, Poznan, Poland.

FUNDING

National Science Center, Poland [2012/06/A/ST6/00384, 2012/05/B/ST6/03026]. Funding for open access charge: National Science Centre, Poland [2012/06/A/ST6/00384].
Conflict of interest statement. None declared.

REFERENCES

1. Popenda,M., Szachniuk,M., Antczak,M., Purzycka,K.J., Lukasiak,P., Bartol,N., Blazewicz,J. and Adamiak,R.W. (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res.*, **40**, e112.
2. Purzycka,K.J., Popenda,M., Szachniuk,M., Antczak,M., Lukasiak,P., Blazewicz,J. and Adamiak,R.W. (2015) Automated 3D RNA structure prediction using the RNAComposer method for riboswitches(1.). *Methods Enzymol.*, **553**, 3–34.
3. Parisien,M. and Major,F. (2008) The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature*, **452**, 51–55.
4. Flores,S.C., Sherman,M.A., Bruns,C.M., Eastman,P. and Altman,R.B. (2011) Fast flexible modeling of RNA structure using internal coordinates. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, **8**, 1247–1257.
5. Jonikas,M.A., Radmer,R.J., Laederach,A., Das,R., Pearlman,S., Herschlag,D. and Altman,R.B. (2009) Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA*, **15**, 189–199.
6. Cao,S., Giedroc,D.P. and Chen,S.-J. (2010) Predicting loop-helix tertiary structural contacts in RNA pseudoknots. *RNA*, **16**, 538–552.
7. Sripakdeevong,P., Kladwang,W. and Das,R. (2011) An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 20573–20578.
8. Sharma,S., Ding,F. and Dokholyan,N.V. (2008) iFoldRNA: three-dimensional RNA structure prediction and folding. *Bioinformatics*, **24**, 1951–1952.
9. Rother,M., Milanowska,K., Puton,T., Jeleniewicz,J., Rother,K. and Bujnicki,J.M. (2011) ModeRNA server: an online tool for modeling RNA 3D structures. *Bioinformatics*, **27**, 2441–2442.
10. Kabsch,W. (1976) A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr.*, **A32**, 922–923.
11. Parisien,M., Cruz,J.A., Westhof,E. and Major,F. (2009) New metrics for comparing and assessing discrepancies between RNA 3D structures and models. *RNA*, **15**, 1875–1885.
12. Zok,T., Popenda,M. and Szachniuk,M. (2013) MCQ4Structures to compute similarity of molecule structures. *Central European Journal ofOper ations Research*, **22**, 457–474.
13. Olechnovic,K. and Venclovas,C. (2014) The CAD-score web server: contact area-based comparison of structures and interfaces of proteins, nucleic acids and their complexes. *Nucleic Acids Res.*, **42**, W259–W263.
14. Hajdin,C.E., Ding,F., Dokholyan,N.V. and Weeks,K.M. (2010) On the significance of an RNA tertiary structure prediction. *RNA*, **16**, 1340–1349.
15. Lukasiak,P., Antczak,M., Ratajczak,T., Bujnicki,J.M., Szachniuk,M., Adamiak,R.W., Popenda,M. and Blazewicz,J. (2013) RNAlyzer–novel approach for quality analysis of RNA structural models. *Nucleic Acids Res.*, **41**, 5978–5990.
16. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
17. Davis,I.W., Leaver-Fay,A., Chen,V.B., Block,J.N., Kapral,G.J., Wang,X., Murray,L.W., Arendall,W.B. III, Snoeyink,J., Richardson,J.S. *et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.*, **35**, W375–W383.
18. Nguyen,M.N. and Verma,C. (2014) Rclick: a web server for comparison of RNA 3D structures. *Bioinformatics*, **31**, 966–968.
19. Hoksza,D. and Svozil,D. (2012) Efficient RNA pairwise structure comparison by SETTER method. *Bioinformatics*, **28**, 1858–1864.
20. Capriotti,E., Norambuena,T., Marti-Renom,M.A. and Melo,F. (2011) All-atom knowledge-based potential for RNA structure prediction and assessment. *Bioinformatics*, **27**, 1086–1093.
21. Cruz,J.A., Blanchet,M.F., Boniecki,M., Bujnicki,J.M., Chen,S.J., Cao,S., Das,R., Ding,F., Dokholyan,N.V., Flores,S.C. *et al.* (2012) RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction. *RNA*, **18**, 610–625.
22. Yang,H., Jossinet,F., Leontis,N., Chen,L., Westbrook,J., Berman,H. and Westhof,E. (2003) Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.*, **31**, 3450–3460.
23. Gendron,P., Lemieux,S. and Major,F. (2001) Quantitative analysis of nucleic acid three-dimensional structures. *J. Mol. Biol.*, **308**, 919–936.
24. Antczak,M., Zok,T., Popenda,M., Lukasiak,P., Adamiak,R.W., Blazewicz,J. and Szachniuk,M. (2014) RNApdbee–a webserver to derive secondary structures from pdb files of knotted and unknotted RNAs. *Nucleic Acids Res.*, **42**, W368–W372.