D386–D391 *Nucleic Acids Research, 2008, Vol. 36, Database issue* *Published online 5 October 2007*
doi:10.1093/nar/gkm786

RNA FRABASE version 1.0: an engine with a database to search for the three-dimensional fragments within RNA structures

Mariusz Popenda[1], Marek Błażewicz[2], Marta Szachniuk[2,3] and Ryszard W. Adamiak[1,*]

[1]Laboratory of Structural Chemistry of Nucleic Acids, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, [2]Institute of Computing Science, Poznań University of Technology, Piotrowo 2, 60-965 Poznań and [3]Laboratory of Bioinformatics, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland

Received August 12, 2007; Revised September 15, 2007; Accepted September 17, 2007

ABSTRACT

The RNA FRABASE is a web-accessible engine with a relational database, which allows for the automatic search of user-defined, 3D RNA fragments within a set of RNA structures. This is a new tool to search and analyse RNA structures, directed at the 3D structure modelling. The user needs to input either RNA sequence(s) and/or secondary structure(s) given in a 'dot-bracket' notation. The algorithm searching for the requested 3D RNA fragments is very efficient. As of August 2007, the database contains: (i) RNA sequences and secondary structures, in the 'dot-bracket' notation, derived from 1065 protein data bank (PDB)-deposited RNA structures and their complexes, (ii) a collection of atom coordinates of unmodified and modified nucleotide residues occurring in RNA structures, (iii) calculated RNA torsion angles and sugar pucker parameters and (iv) information about base pairs. Advanced query involves filters sensitive to: modified residue contents, experimental method used and limits of conformational parameters. The output list of query-matching RNA fragments gives access to their coordinates in the PDB-format files, ready for direct download and visualization, conformational parameters and information about base pairs. The RNA FRABASE is automatically, monthly updated and is freely accessible at http://rnafrabase.ibch.poznan.pl (mirror at http://cerber.cs.put.poznan.pl/rnadb).

INTRODUCTION

The multitude of recent discoveries regarding novel functions of RNA, including the RNAi mechanism and the involvement of regulatory RNAs in cancer and infectious and neurodegradative diseases, make the RNA studies a field of growing practical importance (1).

In this respect, the knowledge of complex 3D folds of the RNA structures is essential to understand the increasing number of their biological functions (2,3). In recent years, we have observed a continuous growth in the statistics of protein data bank (PDB)-deposited RNA structures, although their number (above a thousand structures) is still much smaller than nearly 40 000 protein structures available (4). Undoubtedly, solving the structure of the ribosome (5,6) gave access to a wealth of structural information (7), which made the above statistics more favourable.

Most of the studies on the RNA structure start from the analysis on the secondary structure level. From this viewpoint, programs for secondary structure prediction (8–11) comparative analysis (9), secondary structure alignment (10,12) and visualization (13–15) are indispensable, often offered as the web-accessible tools.

Understanding the RNA structure–function relationship with the use of the computer-aided modelling requires methods to search for the RNA motifs and fragments and to analyse and manipulate tertiary structures. Despite the increasing demand, the computational tools to perform those functions are much less developed than in the case of proteins (16). Until presently only a small number of methods have been designed for the 3D structure-based RNA search, e.g. using a reduced representation of the RNA conformational space to pseudotorsional angles η and θ (17), dihedral angles (18) or phosphorus atoms (19). These programs allow for alignments of the RNA tertiary structures and their fragments. Others are offered for the analysis of RNA torsional angles and helical parameters (20,21), identification and classification of canonical and non-canonical base pairs (22–24) as well as drawing secondary structure from atom coordinates (23).

*To whom correspondence should be addressed. Tel: +48 61 8528503; Fax: +48 61 8520532; Email: adamiakr@ibch.poznan.pl

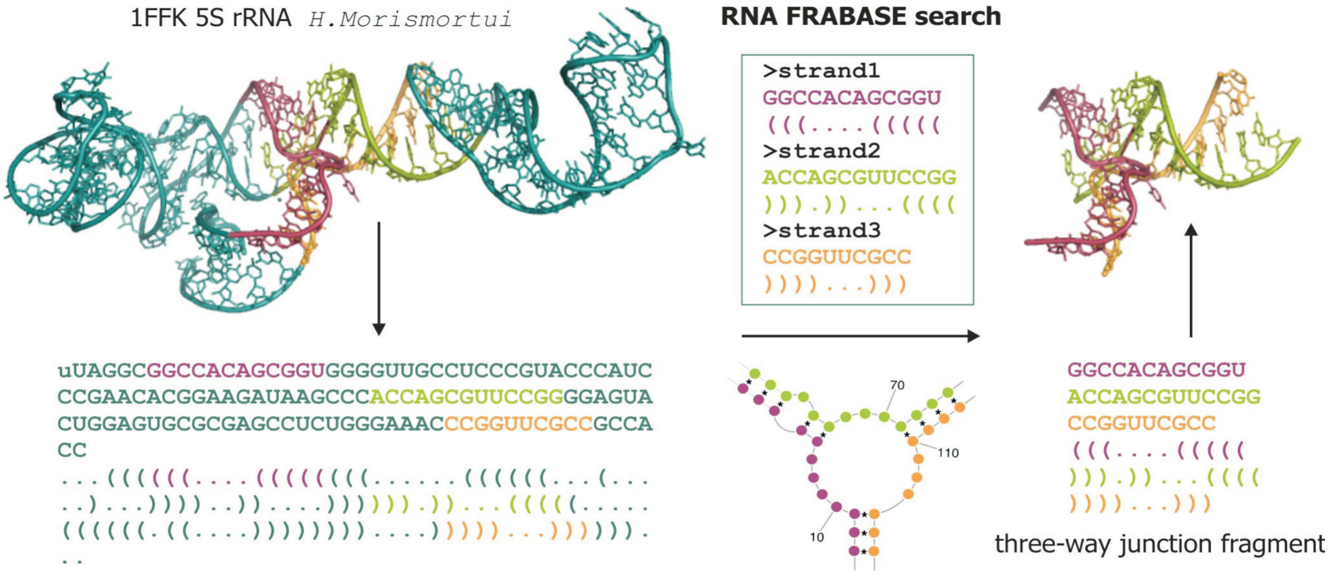


Figure 1. The encoding and searching concept of the RNA FRABASE presented for the *Haloarcula marismortui* 5S rRNA structure (PDB code 1FFK).

The RNA databases play an important role in the RNA analysis on the tertiary structure level. Both the RCSB PDB (4) and nucleic acids database (NDB) (25) hold a collection of RNA structures. Deposited files, apart from atom coordinates, provide a wealth of structural data, give supportive information and allow for the structure visualization. The RNABase (26) consolidates all biomolecular structures containing RNA both from the RCSB PDB and NDB and makes complete conformational maps and analytical data available for each structure. The MeRNA database (27) offers information and classification of metal ion-binding sites in the RNA structures. The structural classification of RNA (SCOR) database (28), which surveys the 3D RNA motifs within the PDB- and NBD-deposited RNA structures, is also frequently visited. This elaborated, RNA architecture hierarchy-based database, gives in its 2.0.1 version not only the access to over 8000 RNA structural elements, mostly internal and hairpins loops, but also to less frequent tertiary interactions motifs. The SCOR provides also a glossary of the RNA structural motifs. Unfortunately, this database is not updated on the regular basis.

Here, we present the RNA FRABASE, the web-accessible engine with relational database, which allows for the automatic search of the user-defined, 3D fragments within the RNA structures available in this database. The search based on regular expressions and the pattern-matching method is very efficient. The pattern includes the RNA sequence(s) and/or secondary structure(s). The output list of query-matching RNA fragments gives access to their coordinates in the standard PDB-format files, which are ready for direct download and visualization at the client site. Additional information about their conformational parameters and canonical base pairs is given. The RNA FRABASE is automatically, monthly updated and is freely accessible at http://rnafrabase.ibch.poznan.pl. We hope that this new tool will be of general interest for researchers working in the field of RNA structural biology.

METHOD OUTLINE

Searching for 3D RNA structural motifs in a conformational space is a process that is much more complex than searching the databases for the primary and secondary structural patterns with the use of regular expressions and the pattern-matching method. A number of algorithms as structural pattern finders have been developed (29–31). Their efficiency comes from the simplicity of the data format and the RNA structure descriptors used to encode the primary and the secondary RNA structures.

Therefore, in order to design a method for the 3D RNA fragments automatic search, we have decided to extract structural data from the 3D RNA structures, which would enable us to take advantage of efficient searches on the level of sequences and secondary structures.

As presented in Figure 1, the encoding and searching concept of RNA FRABASE consists of three major stages.

Stage 1. The information about base pairing and other structural data is extracted from the PDB-deposited RNA structures and transformed to the primary and secondary RNA structure string formats. The RNA FRABASE holds the RNA sequences in the one-letter code and the secondary structures encoded in the 'dot-bracket' notation (9). This transformation has been conducted with our own script employing a list of base pairs for every RNA structure delivered by the RNAView software (23).

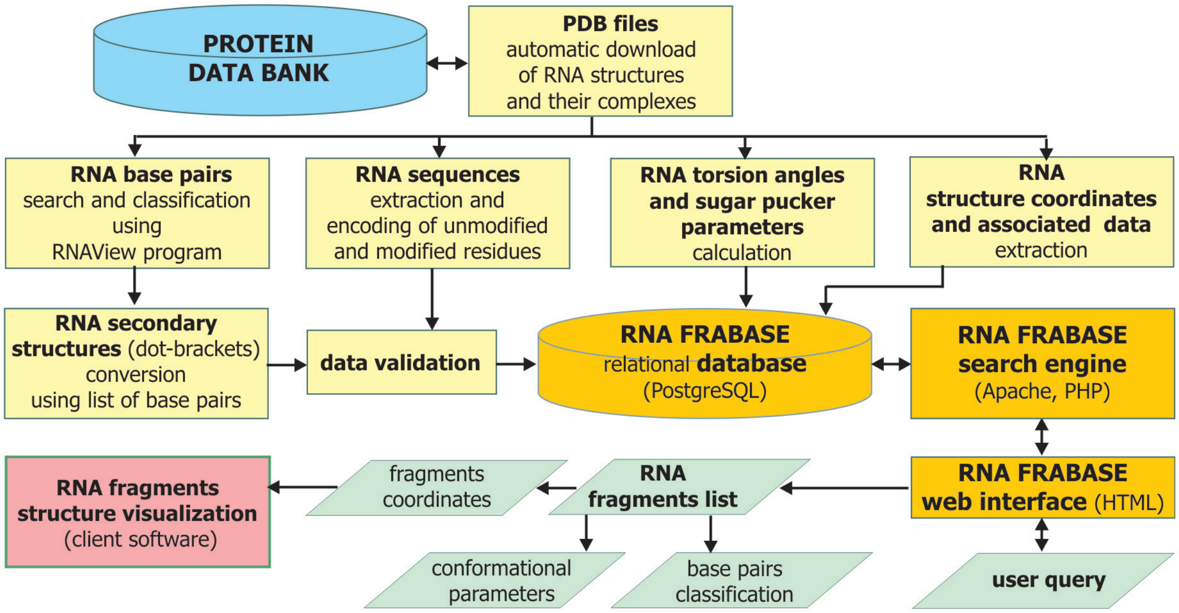

PROTEIN DATA BANK
PDB files automatic download of RNA structures and their complexes
RNA base pairs search and classification using RNAView program
RNA sequences extraction and encoding of unmodified and modified residues
RNA torsion angles and sugar pucker parameters calculation
RNA structure coordinates and associated data extraction
RNA secondary structures (dot-brackets) conversion using list of base pairs
data validation
RNA FRABASE relational database (PostgreSQL)
RNA FRABASE search engine (Apache, PHP)
RNA fragments structure visualization (client software)
fragments coordinates
RNA fragments list
RNA FRABASE web interface (HTML)
conformational parameters
base pairs classification
user query

Figure 2. The RNA FRABASE flowchart.

Stage 2. To start searching for the 3D RNA fragments, the user needs to provide an input pattern defining either RNA sequence(s) and/or secondary structure(s) in the RNA FRABASE format.

Stage 3. The search engine scans all RNA structures placed in the relational database to find those, which exactly match the input pattern. Subsequently, the engine links the resulting fragment locations in the primary and secondary structures to the atom coordinates of defined residues within 3D RNA structures.

IMPLEMENTATION

The core of RNA FRABASE consists of the database, the search engine and the web interface. They are directly associated with all the actions undertaken by the user. The remaining part of the system, concerning the data collection and computation, consists of several assistant scripts performing in the background. These scripts are responsible for sugar, backbone and χ torsion angles calculation, sugar pucker calculation and identification of the base pairs using the RNAView software (23).

The flowchart representation of RNA FRABASE showing the relations between all of its components is presented in Figure 2. The RNA FRABASE runs in SUSE Linux environment. It has been developed as the relational database in PostgreSQL. The search engine is served by the Apache http daemon along with the PHP scripts. The interface component consists of the web pages designed and implemented in HTML. The assistant programs have been encoded in AWK and PHP. The RNA FRABASE has been tested on Windows and Unix/Linux platforms. It can be operated through many web browsers, like Mozilla, Firefox, Internet Explorer and Opera. The service is hosted and maintained by the Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland.

Database

As of August 2007, the database contains (i) RNA sequences in one-letter code and secondary structures, in the 'dot-bracket' notation, derived from 1065 PDB-deposited RNA structures and their complexes, (ii) a collection of atom coordinates of unmodified and modified nucleotide residues occurring in RNA structures, (iii) calculated RNA torsion angles and sugar pucker parameters, (iv) information about base pairs given by the RNAView software (23) and (v) PDB identification codes with links to deposited information about RNA structures, experimental method used, resolution and short description of RNA structures. The RNA FRABASE holds not only RNA structures composed of unmodified nucleotides, naturally occurring modified nucleotides (e.g. typical for tRNAs) or synthetic nucleotide analogues, but also the RNA–DNA chimeras and RNA–DNA hybrids. Every database entry is automatically checked for conformational errors.

RNA sequences in the RNA FRABASE database are coded in the one-letter format. The unmodified RNA monomer residues are coded in capital letters, i.e. A, C, G, U. All modified RNA and non-RNA units, like DNA units, and 5′-dephosphorylated residues are recorded in small letters giving the closest analogy to parent nucleoside. The sequences are stored and searched in the 5′–3′ direction.

RNA secondary structures are coded using the 'dot-bracket' notation (9). In the 'dot-bracket' notation an unpaired nucleotide is represented as a dot and a base pair is represented as a pair of opening and closing brackets. In the RNA FRABASE brackets are used only in case of an explicit Watson–Crick AU, UA, CG, GC and wobble GU or UG base pairing. Bracket notation is extended to represent secondary structure of pseudoknots and kissing loops, by inserting squared '[''and'']' brackets. The curly brackets '{''and''}' as well as angle brackets '<''and''>' are used for the most complicated structures such as higher order pseudoknots in RNAs.

Search engine

The RNA structure is formed by a single, folded oligoribonucleotide strand. The RNA FRABASE engine is designed to search for user-specified RNA fragments which might be single- or multi-stranded (Figure 1). The number of strands, which compose a fragment, is unlimited (e.g. the seven-way junction, Figure S1). The user is asked to provide an input pattern describing those strands. To ensure that the input pattern is legal, one should follow the RNA FRABASE format, which is described subsequently. Every strand should be defined in one separate section. The section concerning the strand is composed of two or three significant lines:

- Line 1 is obligatory—contains '>' followed by a unique strand identifier
- Line 2 is optional—contains the RNA sequence in the 5′–3′ direction and in the one-letter format coherent with IUPAC–IUB codes
- Line 3 is optional for single-stranded fragments—contains the secondary structure in the 'dot-bracket' notation

In any place of the entry, a user can insert a comment line starting from '#'. The following convention should be used when querying for the fragments containing two or more strands: each strand must be associated with the subsequent one at least by one W–C base pair or GU, the last strand must be paired with the first one, and the number of opening and closing brackets should be the same for all the strands to define the whole fragment (Figure 1).

Moreover, each input pattern is scored by an objective-weighted function, which assesses the level of precision of the fragment's definition. The pattern is accepted for searching if its score goes beyond a system threshold. Both the function and the threshold values have been optimized experimentally. The input pattern, when accepted, is passed along to the main search engine procedure. The algorithm based on regular expression-matching principles is used. In its first step, the most informative strand of the RNA fragment is located within all the entities stored in the database. Next, the remaining strands of the fragment are matched one by one within the subset of structures selected in the previous step. Each resulting fragment is checked for conformity with the secondary structure defined in the input pattern. The efficiency of the algorithm depends mainly on the complexity of the query, i.e. number of strands, their lengths and the level of sequence description.

The search engine can work in two modes. In the default mode, both capital and small characters in the sequence(s) pattern are treated equally. In the second mode, those characters are recognized and the 'modified residues sensitive search' can be executed.

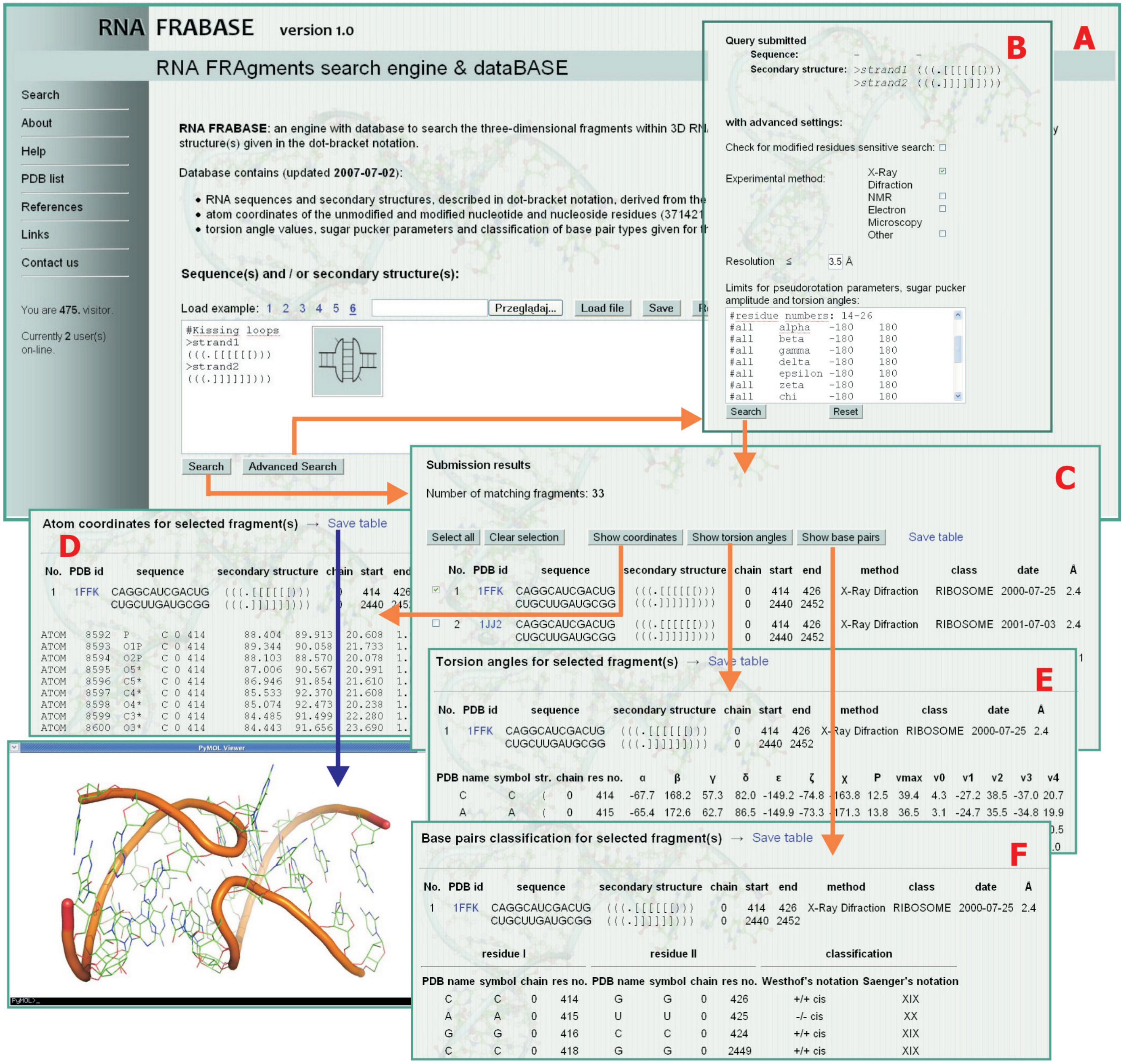
User interface

Figure 3 presents snapshots of different information and results given by the RNA FRABASE. Upon entry, the RNA FRABASE home page http://rnafrabase.ibch.poznan.pl gives direct access to the 'Search' window (panel A). Panel A contains the sequence/structure entry box in which a user provides the input data, i.e. either RNA sequence(s) and/or secondary structure(s) in the RNA FRABASE format. This data might be directly loaded from the user's own file(s) using browser option. The user query might also be saved in the text format. As a help for the user, six example entries are given concerning: tRNA, RNA hairpin, bulge, internal loop, three-way junction and kissing loops.

Alternatively, the user can enter the 'Advanced Search' (panel B), which provides filters to narrow the search using the following options:

- The search for RNA structures containing modified nucleotide residues. To activate it, the 'modified residues sensitive search' box should be checked.
- The experimental method used to determine RNA structure and its resolution. To activate this filter, the appropriate box should be checked and the resolution value (Å) typed in.
- The ranges of conformational parameters (torsion angles, sugar pucker parameters). Limits for conformational parameters should be typed in the appropriate format, which is described in detail in the 'Help' section of the RNA FRABASE.

As the output, a list of query-matching RNA fragments is generated (panel C). The list can be saved in the CSV file format. The list contains the following information concerning each RNA fragment: the PDB ID of the parent RNA structure (directly linked to the original PDB file), sequence, secondary structure in the 'dot-bracket' notation, identifier of the chain in which RNA fragment has been found, position of the fragment within the chain(s), experimental method used, functional class, date of structure deposition and structure resolution. Upon marking the structure(s) on the search result list, major structural information is available: atom coordinates (panel D), torsional angles (panel E) and the canonical base pairs classification (panel F). Clicking 'show coordinates' button gives access to the fragment(s) coordinates in the PDB-format which can be saved and downloaded for direct visualization at the client site as exemplified by the kissing loop motif structure (example 6, PDB ID 1FFK) using e.g. PyMol software (32). Clicking 'show torsion angles' gives the complete list of sugar and backbone parameters: α, β, γ, δ, ε, ζ, χ, P, $\nu_{\max}$, ν_0, ν_1, ν_2, ν_3 and ν_4 to be saved in the CSV file format. The third option lists the Watson–Crick and GU base pairs in the Saenger's and Westhof's notations.



RNA FRABASE: an engine with database to search the three-dimensional fragments within 3D RN
structure(s) given in the dot-bracket notation.

Database contains (updated **2007-07-02**):

- RNA sequences and secondary structures, described in dot-bracket notation, derived from the
- atom coordinates of the unmodified and modified nucleotide and nucleoside residues (371421
- torsion angle values, sugar pucker parameters and classification of base pair types given for tl

You are **475.** visitor.

Currently **2** user(s) on-line.

Sequence(s) and / or secondary structure(s):

Load example: 1 2 3 4 5 6

```
#Kissing loops
>strand1
(((.[[[[[[)))
>strand2
(((.]]]]]])))
```

Query submitted

Sequence: – –

Secondary structure: *>strand1* (((.[[[[[[)))
>strand2 (((.]]]]]])))

with advanced settings:

Check for modified residues sensitive search: □

Experimental method: X-Ray Difraction ☑ NMR □ Electron Microscopy □ Other □

Resolution ≤ 3.5 Å

Limits for pseudorotation parameters, sugar pucker amplitude and torsion angles:

```
#residue numbers: 14-26
#all    alpha   -180    180
#all    beta    -180    180
#all    gamma   -180    180
#all    delta   -180    180
#all    epsilon -180    180
#all    zeta    -180    180
#all    chi     -180    180
```

Submission results

Number of matching fragments: 33

	No.	PDB id	sequence	secondary structure	chain	start	end	method	class	date	Å
☑	1	1FFK	CAGGCAUCGACUG CUGCUUGAUGCGG	(((.[[[[[[))) (((.]]]]]])))	0 0	414 2440	426 2452	X-Ray Difraction	RIBOSOME	2000-07-25	2.4
□	2	1JJ2	CAGGCAUCGACUG CUGCUUGAUGCGG	(((.[[[[[[))) (((.]]]]]])))	0 0	414 2440	426 2452	X-Ray Difraction	RIBOSOME	2001-07-03	2.4

Atom coordinates for selected fragment(s) → Save table

No.	PDB id	sequence	secondary structure	chain	start	end
1	1FFK	CAGGCAUCGACUG CUGCUUGAUGCGG	(((.[[[[[[))) (((.]]]]]])))	0 0	414 2440	426 245

```
ATOM    8592  P     C 0 414      88.404  89.913  20.608  1.
ATOM    8593  O1P   C 0 414      89.344  90.058  21.733  1.
ATOM    8594  O2P   C 0 414      88.103  88.570  20.078  1.
ATOM    8595  O5*   C 0 414      87.006  90.567  20.991  1.
ATOM    8596  C5*   C 0 414      86.946  91.854  21.610  1.
ATOM    8597  C4*   C 0 414      85.533  92.370  21.608  1.
ATOM    8598  O4*   C 0 414      85.074  92.473  20.238  1.
ATOM    8599  C3*   C 0 414      84.485  91.499  22.280  1.
ATOM    8600  O3*   C 0 414      84.443  91.656  23.690  1.
```

Torsion angles for selected fragment(s) → Save table

No.	PDB id	sequence	secondary structure	chain	start	end	method	class	date	Å
1	1FFK	CAGGCAUCGACUG CUGCUUGAUGCGG	(((.[[[[[[))) (((.]]]]]])))	0 0	414 2440	426 2452	X-Ray Difraction	RIBOSOME	2000-07-25	2.4

PDB name	symbol	str.	chain	res no.	α	β	γ	δ	ε	ζ	χ	P	vmax	v0	v1	v2	v3	v4
C	C	(	0	414	-67.7	168.2	57.3	82.0	-149.2	-74.8	-163.8	12.5	39.4	4.3	-27.2	38.5	-37.0	20.7
A	A	(	0	415	-65.4	172.6	62.7	86.5	-149.9	-73.3	-171.3	13.8	36.5	3.1	-24.7	35.5	-34.8	19.9

Base pairs classification for selected fragment(s) → Save table

No.	PDB id	sequence	secondary structure	chain	start	end	method	class	date	Å
1	1FFK	CAGGCAUCGACUG CUGCUUGAUGCGG	(((.[[[[[[))) (((.]]]]]])))	0 0	414 2440	426 2452	X-Ray Difraction	RIBOSOME	2000-07-25	2.4

residue I				residue II				classification	
PDB name	symbol	chain	res no.	PDB name	symbol	chain	res no.	Westhof's notation	Saenger's notation
C	C	0	414	G	G	0	426	+/+ cis	XIX
A	A	0	415	U	U	0	425	-/- cis	XX
G	G	0	416	C	C	0	424	+/+ cis	XIX
C	C	0	418	G	G	0	2449	+/+ cis	XIX

Figure 3. The RNA FRABASE interface snapshots: panel A, front page with the 'Search' window; panel B, 'Advanced Search' with provided filters; panel C, a list of query-matching RNA fragments with structural information—upon checking the structure(s) on the list, major structural information about 3D RNA fragments is available concerning coordinates (panel D), torsional angles (panel E) and canonical base pairs classification (panel F). The RNA fragment(s) coordinates saved in the PDB-format files are ready for direct download and visualization.

CONCLUSIONS

We have presented the RNA FRABASE, which is a web-accessible engine with a database, which allows for the automatic search of the user-defined, 3D fragments within a set of RNA structures. To the best of our knowledge, this is the first web-accessible tool that performs this operation. Until now, this kind of search has required tedious manipulations. The algorithm behind the engine is highly efficient, making a typical search through all the RNA structures in our database possible in a short time on a user's standard PC. The output list of the query-matching RNA fragments gives access to their coordinates in the PDB-format files, ready for direct download and visualization. Within the advanced query option, the use of a filter defining the limits for RNA torsion angles should be very useful to compare the conformational details of the user's own structures with those deposited within RCSB PDB database. Several applications of the RNA FRABASE might be envisaged. The one pursued in our laboratory is the 3D RNA structure modelling directed towards the RNA therapeutics. Currently, we apply the RNA FRABASE in connection with the 3D RNApredict program (33) for the high-throughput prediction of a large set of RNA structures (pre-miRNA) using the RNA templates approach.

In future, we intend to enlarge the searching capabilities of the engine by applying the RNA structural descriptor format (31), include the NDB-deposited RNA structures and introduce the list of non-canonical base pairs.

AVAILABILITY AND CITATION

The RNA FRABASE database in its 1.0 version is continuously maintained and automatically updated every month. The database is freely available at http://rnafrabase.ibch.poznan.pl. Users are invited to contact us through the 'Contact us' link to give critical comments and suggestions. If you use the RNA FRABASE in a published report, please cite this article.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by the Foundation for Polish Science (SP 01/04) and the Ministry of Education and Science (3T09A014 29) grants to R.W.A. The generous support of the Poznań Supercomputing and Networking Centre is acknowledged. The funding to pay the Open Access publication charges for this article was provided by the Ministry of Education and Science (grant 3 T09A 014 29).

Conflict of interest statement. None declared.

REFERENCES

1. Zamore,P.D. and Haley,B. (2005) Ribo-gnome: the big world of small RNAs. *Science*, **309**, 1519–1524.
2. Tinoco,J.Jr. and Bustamante,C. (1999) How RNA folds. *J. Mol. Biol.*, **293**, 271–281.
3. Leontis,N.B., Lescoute,A. and Westhof,E. (2006) The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.*, **16**, 279–287.
4. 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.
5. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905–920.
6. Wimberly,T., Brodersen,D.E., Clemons,W.M.Jr, Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
7. Noller,H.F. (2005) RNA structure: reading the ribosome. *Science*, **309**, 1508–1514.
8. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
9. Hofacker,I.L., Fontana,W., Stadler,P.F., Bonhoeffer,S., Tacker,M. and Schuster,P. (1994) Fast folding and comparison of RNA secondary structures. *Monatshefte f. Chemie*, **125**, 167–188.
10. Mathews,D.H. and Turner,D.H. (2002) Dynalign: an algorithm for finding the secondary structure common to two RNA sequences. *J. Mol. Biol.*, **317**, 191–203.
11. Reeder,J., Steffen,P. and Giegerich,R. (2007) pknotsRG: RNA pseudoknot folding including near-optimal structures and sliding windows. *Nucleic Acids Res.*, **35**, W320–W324.
12. Ji,Y., Xu,X. and Stormo,G.D. (2004) A graph theoretical approach for predicting common RNA secondary structure motifs including pseudoknots in unaligned sequences. *Bioinformatics*, **20**, 1591–1602.
13. De Rijk,P, Wuyts,J. and De Wachter,R. (2003) RnaViz 2: an improved representation of RNA secondary structure. *Bioinformatics*, **19**, 299–300.
14. Byun,Y. and Han,K. (2006) PseudoViewer: web application and web service for visualizing RNA pseudoknots and secondary structures. *Nucleic Acids Res.*, **34**, W416–W422.
15. Kaiser,A., Kruger,J. and Evers,D.J. (2007) RNA movies 2: sequential animation of RNA secondary structures. *Nucleic Acids Res.*, **35**, W330–W334.
16. Wolfson,H.J., Shatsky,M., Schneidman-Duhovny,D., Dror,O., Shulman-Peleg,A., Ma,B. and Nussinov,R. (2005) From structure to function: methods and applications. *Curr. Protein Pept. Sci.*, **6**, 171–183.
17. Duarte,C.M., Wadley,L.M. and Pyle,A.M. (2003) RNA structure comparison, motif search and discovery using a reduced representation of RNA conformational space. *Nucleic Acids Res.*, **31**, 4755–4761.
18. Ferre,F., Ponty,Y., Lorenz,W.A. and Clote,P. (2007) DIAL: a web server for the pairwise alignment of two RNA three-dimensional structures using nucleotide, dihedral angle and base-pairing similarities. *Nucleic Acids Res.*, **35**, W659–W668.
19. Dror,O., Nussinov,R. and Wolfson,H.J. (2006) The ARTS web server for aligning RNA tertiary structures. *Nucleic Acids Res.*, **34**, W412–W415.
20. Gendron,P., Lemieuxs,S. and Major,F. (2001) Quantitative analysis of nucleic acid three-dimensional structures. *J. Mol. Biol.*, **308**, 919–936.
21. Lu,X.-J. and Olson,W.K. (2003) 3 DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
22. Lemieux,S. and Major,F. (2002) RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res.*, **30**, 4250–4263.
23. 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.
24. Das,J., Mukherjee,S., Mitra,A. and Bhattacharyya,D. (2006) Non-canonical base pairs and higher order structures in nucleic acids: crystal structure database analysis. *J. Biomol. Struct. Dynamics*, **24**, 149–161.
25. Berman,H.M., Westbrook,J., Feng,Z., Iype,L., Schneider,B. and Zardecki,C. (2003) The nucleic acid database. *Methods Biochem. Anal.*, **44**, 199–216.
26. Murthy,V.L. and Rose,G.D. (2003) RNABase: an annotated database of RNA structures. *Nucleic Acids Res.*, **31**, D502–D504.
27. Stefan,L.R., Zhang,R., Levitan,A.G., Hendrix,D.K., Brenner,S.E. and Holbrook,S.R. (2006) MeRNA: a database of metal ion binding sites in RNA structures. *Nucleic Acids Res.*, **34**, D131–D134.
28. Klosterman,P.S., Hendrix,D.K., Tamura,M., Holbrook,S.R. and Brenner,S.E. (2004) Three dimensional motifs from the SCOR, structural classification of RNA database: extruded strands, base triples, tetraloops and U-turns. *Nucleic Acids Res.*, **32**, 2342–2352.
29. Laferriere,A., Gautheret,D. and Cedergren,R. (1994) An RNA pattern matching program with enhanced performance and portability. *Comput. Appl. Biosci.*, **10**, 211–212.
30. Macke,T.J., Ecker,D.J., Gutell,R.R., Gautheret,D., Case,D.A. and Sampath,R. (2001) RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Res.*, **29**, 4724–4735.
31. Chang,T.H., Huang,H.D., Chuang,T.N., Shien,D.M. and Horng,J.T. (2006) RNAMST: efficient and flexible approach for identifying RNA structural homologs. *Nucleic Acids Res.*, **34**, W423–W428.
32. DeLano,W.L.T. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
33. Popenda,M., Bielecki,L. and Adamiak,R.W. (2006) High-throughput method for the prediction of low-resolution, three-dimensional RNA structures. *Nucleic Acids Symp. Ser.*, **50**, 67–68.