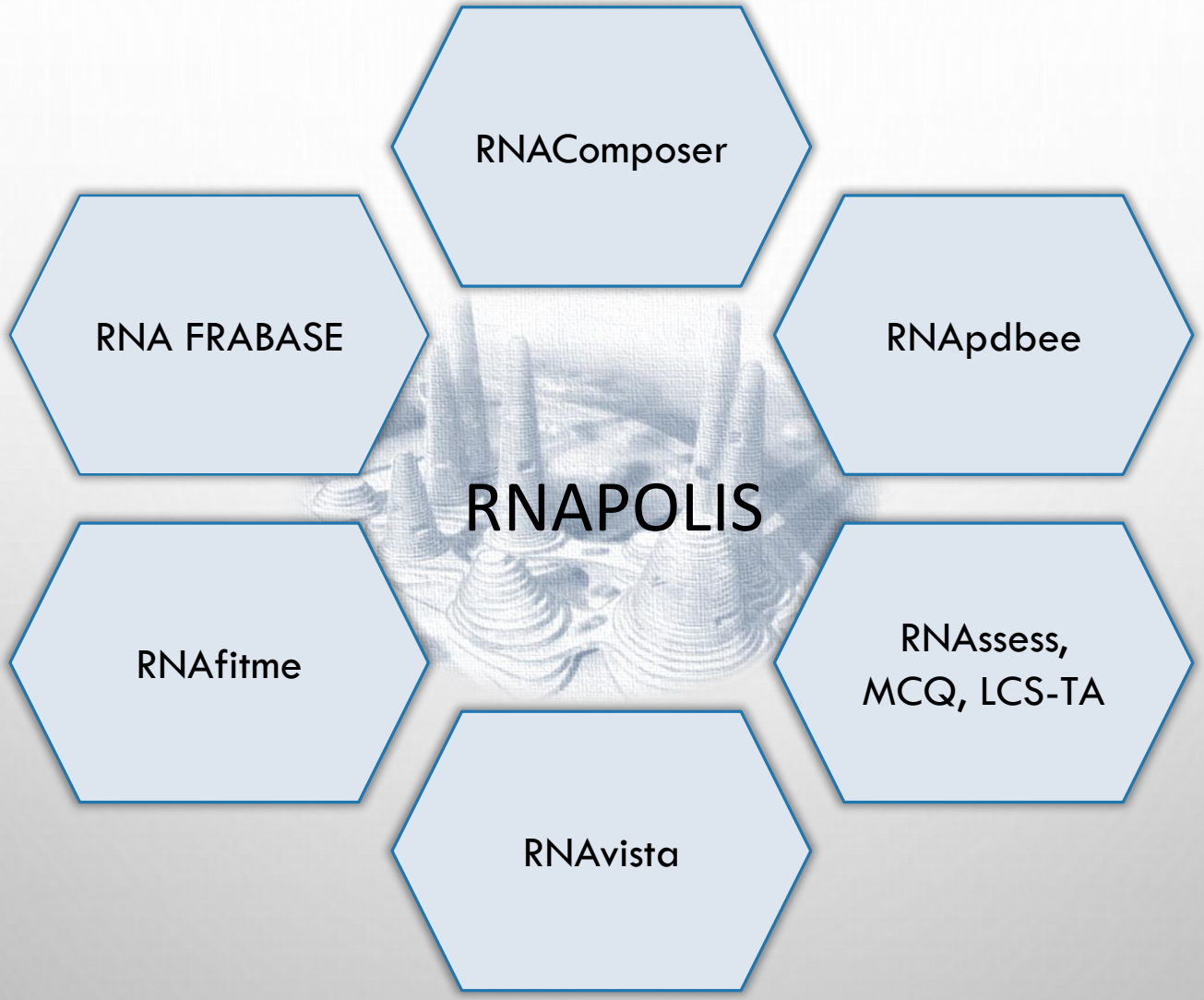
RNA 2D/3D STRUCTURE IN RNAPOLIS. THERE AND BACK AGAIN.

MARTA SZACHNIUK

MACIEJ ANTCZAK & TOMASZ ZOK

Institute of Computing Science,
Poznan University of Technology,
Poland

Institute of Bioorganic Chemistry
Polish Academy of Sciences,
Poznan, Poland

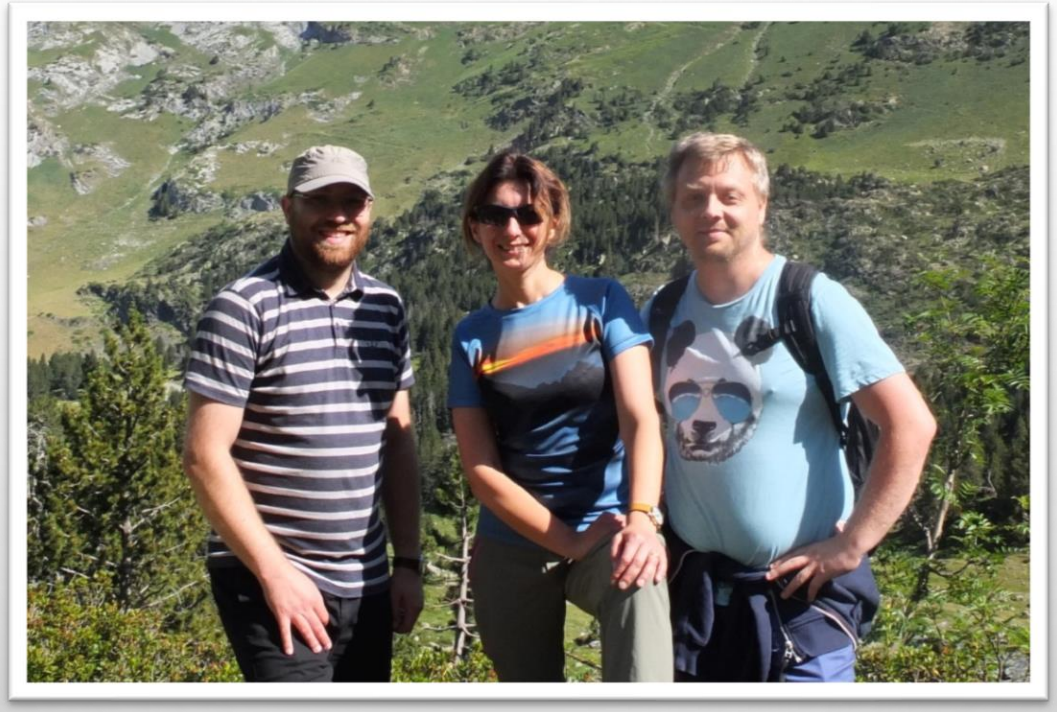

RNAComposer
RNA FRABASE
RNApdbee
RNAPOLIS
RNAfitme
RNAssess,
MCQ, LCS-TA
RNAvista

CURRENT RESEARCH & THE NEAREST PLANS

- Modeling and annotation of pseudoknots
- Pseudoknot order and hierarchy of RNA folding
- Modeling of circular RNA 3D structures
- Analysis and classification of quadruplex structures
- Algorithms for RNA 3D structure quality assessment
- RNA FRABASE renewal and adaptation to mmCIF data

RNAPOLIS TEAM

Tomasz Zok Marta Szachniuk Maciej Antczak



Mariusz Popenda Ryszard W. Adamiak

PhD students
Joanna Miskiewicz
Jakub Wiedemann

MSc students
Marcin Zablocki
Michal Zurkowski

RNACOMPOSER 1.1
2D ⇒ 3D STRUCTURE

MACIEJ ANTCZAK

Institute of Computing Science,
Poznan University of Technology,
Poland

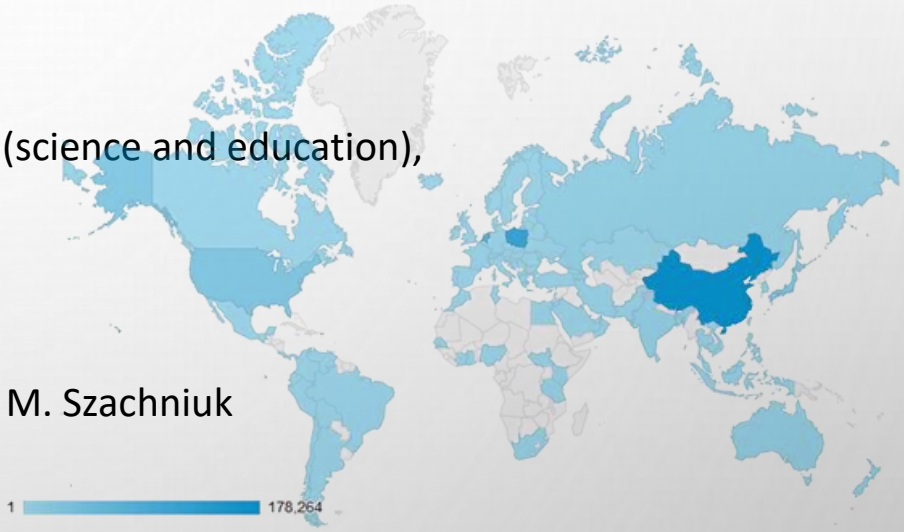
Institute of Bioorganic Chemistry
Polish Academy of Sciences,
Poznan, Poland

RNACOMPOSER

RNAComposer webserver was made available (and published) in 2012 and can be accessed at:

http://rnacomposer.cs.put.poznan.pl
http://rnacomposer.ibch.poznan.pl

Until 21.07.2018, RNAComposer has collected:
~ **3,352,700** sessions performed,
~ **10,000** returning users from all over the world (science and education),
~ **200** citations of the 1st paper,
~ **2** scientific awards.

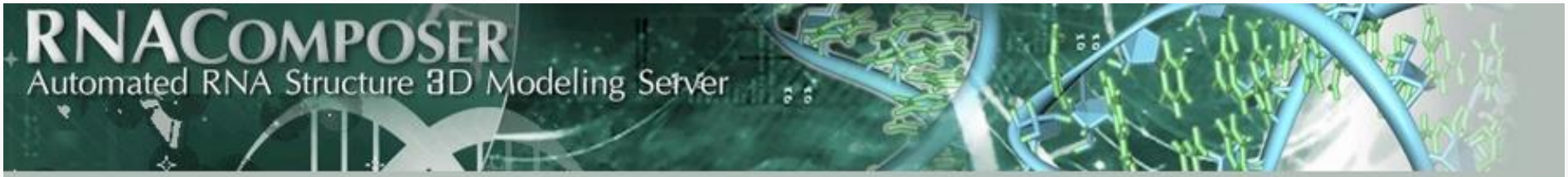


Team (currently):
M. Antczak, M. Popenda, T. Zok, R.W. Adamiak & M. Szachniuk

References (main):

- M. Popenda, M. Szachniuk, M. Antczak, K.J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz, R.W. Adamiak. Automated 3D structure composition for large RNAs, *Nucleic Acids Research*, 2012, 40(14), e112.
- M. Antczak, M. Popenda, T. Zok, J. Sarzynska, T. Ratajczak, M. Tomczyk, R.W. Adamiak, M. Szachniuk. New functionality of RNAComposer: an application to shape the axis of miR160 precursor structure, *Acta Biochimica Polonica*, 2016, 63(4), 737-744.

http://rnacomposer.cs.put.poznan.pl/



Home
Tools
Help
About
References
Links
Contact us

User ID:
Password:
Log in
Forgot your password?
Create an account

You are **823,989** visitor.
Visitors online: **11**

Supported by:
Department of Structural Chemistry and Biology of Nucleic Acids & Department of Bioinformatics, Institute of Bioorganic Chemistry, Polish Academy of Sciences
Institute of Computing Science, Poznan University of Technology

Welcome to RNAComposer, a fully automated RNA structure modeling server. (Mirror site: rnacomposer.ibch.poznan.pl)

The RNAComposer system offers a new user-friendly approach to the fully automated prediction of large RNA 3D structures. The method is based on the machine translation principle and operates on the RNA FRABASE database acting as the dictionary relating RNA secondary structure and tertiary structure elements.

RNAComposer works in two modes:

- **interactive mode** - allows to work on one RNA molecule of interest at a time; its use is limited up to 500 nt residues and results in a single 3D-RNA structure model. Input your RNA sequence and secondary structure (Example 1 and Example 2) or sequence only (Example 3). Example 3 is offered for introductory purposes.
- **batch mode** - is designed for large-scale automated modeling of RNA structures up to 500 nt residues, based on user-defined RNA secondary structures. As an input a set of up to 10 RNA sequences can be used. This mode is available only for registered users.

Allows to influence the prediction process!

You are in interactive mode

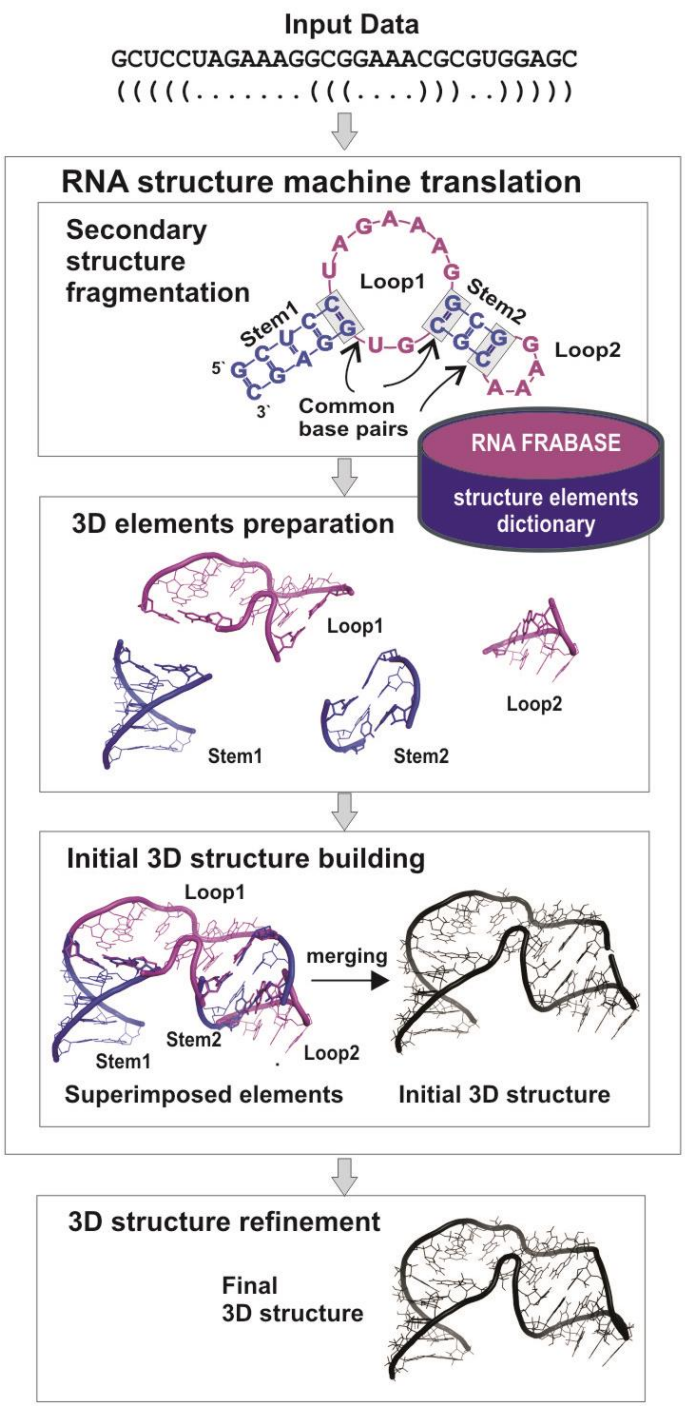
Enter RNA sequence and secondary structure in dot-bracket format (Example 1 and Example 2) or sequence only (Example 3). A maximum sequence length is limited to 500 residues.

Load example: 1 2 3 Reset

Which one is the most popular among users?

```
#HIV-2 DIS RNA hairpin
>example1
CCCCAAGCUAUGCAACCUUAUGGUCCCUGUGGGAAGGGGA
((((..((...))..((....)).(((...)))..)))).
```

Compose Email results to:



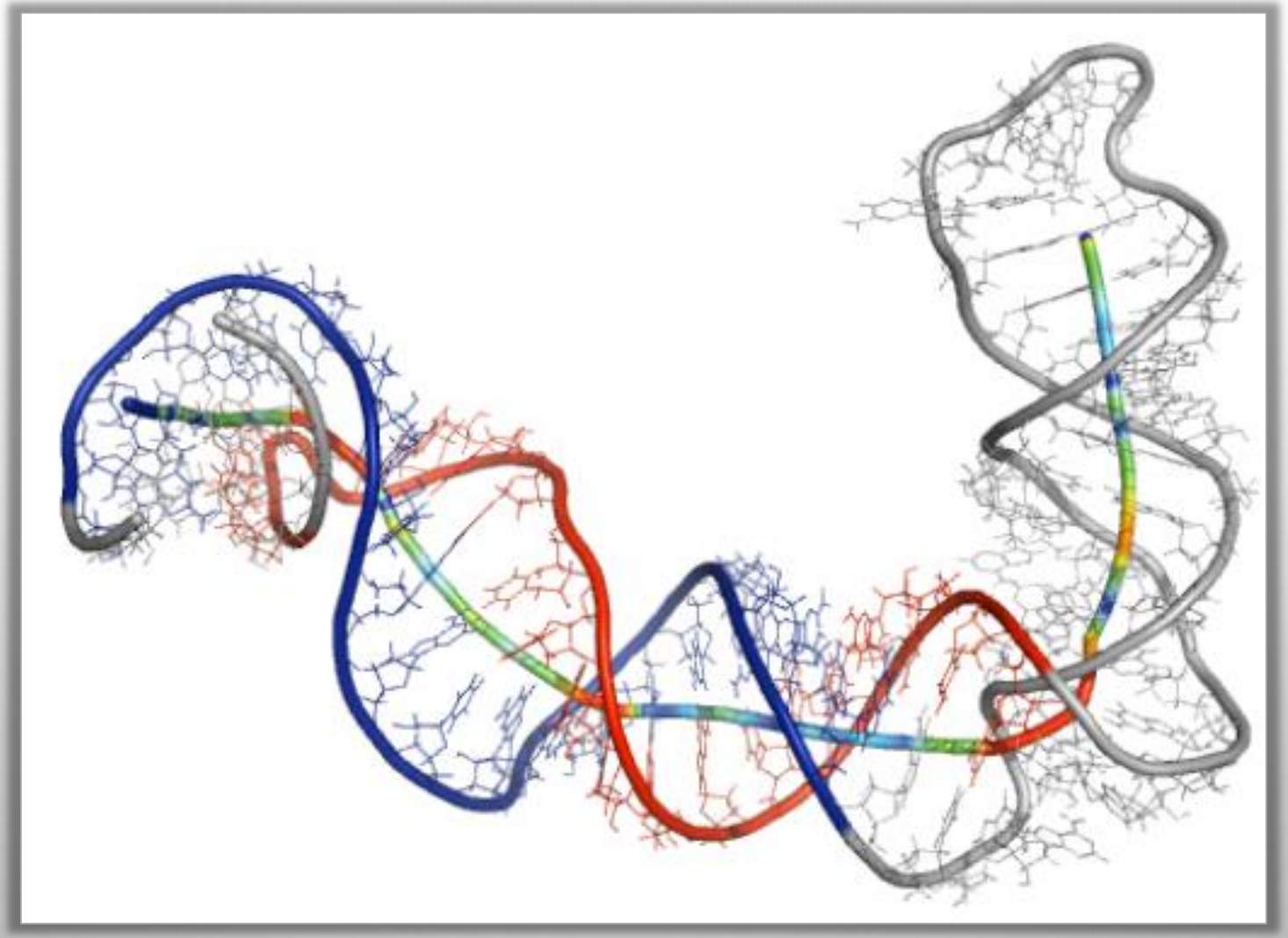
RNACOMPOSER FLOW

Input: *sequence and – optionally – secondary structure in dot-bracket notation.*

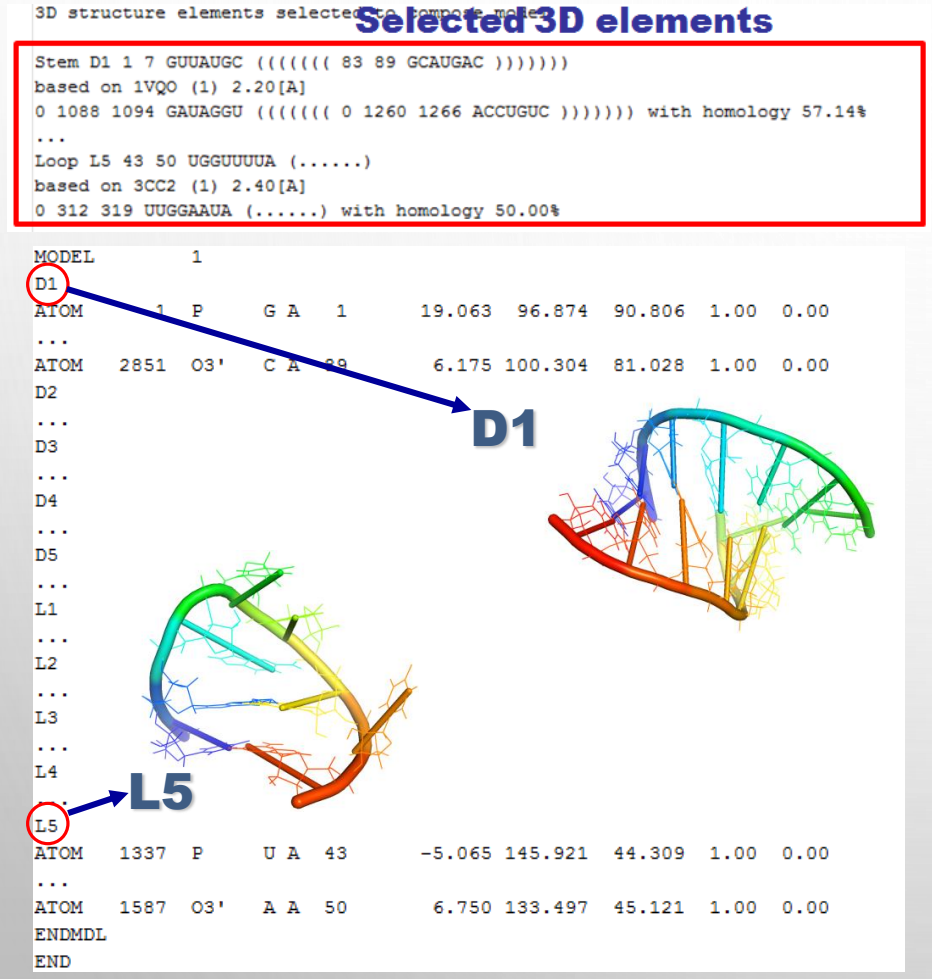
0. Secondary structure prediction (if needed).
1. Input data **validation**.
2. Secondary structure **decomposition into topologically common fragments**.
3. **Selecting** the most appropriate **3D elements** for all identified fragments from the database of known RNA 3D structures.
4. 3D element **adjustment** (if needed).
5. **Assembling** considered 3D elements into an initial 3D model.
6. 3D model **minimization** in both torsion and Euclidean space.

Output: *up to 10 RNA three-dimensional models.*

IF THE FIRST 3D MODEL DOESN'T MEET EXPECTATIONS…



CAN YOU INFLUENCE THE 3D STRUCTURE PREDICTION?



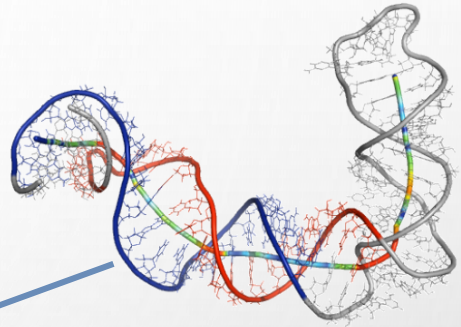
1. **Introduce own restraints for the inter-atomic distances and torsion angles** applied during 3D model refinement.
2. **Introduce own 3D elements** at the input of the prediction process.
3. **Filter 3D elements** originated from **the particular set of RNAs** described by PDB ids.
4. **Consider 3D elements** originated from **high-resolution RNAs only**.
5. **Generate double helices as well as single strands** based on A-RNA template.

INFLUENCING RNA 3D STRUCTURE PREDICTION PROCESS

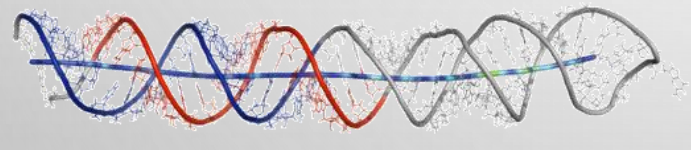
Modeling of miRNA precursor structure: miR160c (Arabidopsis thaliana).

```
D1      L1 D2    L2       D3       L3 D4 L4     D5      L5
        C        CC-               A  AAAAC          AG
GUCGUGC UGGCU    CUGUAUGCCAC AG       AUCGAUUU   U
||||||| |||||    |||||||||||| ||       ||||||||   U
CAGUACG ACUGA    GACAUGCGGUG UC       UAGCUAAA   U
        A        UGA               A  ---AC          AC
```

RNAComposer

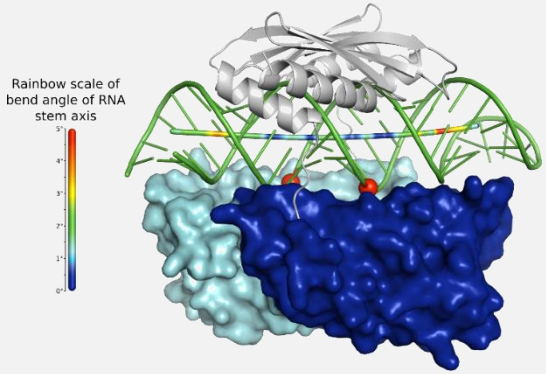


How to refine the 3D model?



miRNA – red miRNA* – blue

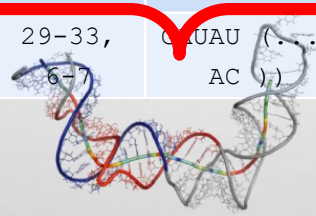
miRNA structure should adapt to the shape of DLC1 catalytic valley:

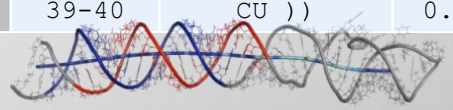


Reference:
M. Antczak, M. Popenda, T. Zok, J. Sarzynska, T. Ratajczak, K. Tomczyk, R.W. Adamiak, M. Szachniuk, New functionality of RNAComposer: application to shape the axis of miR160 precursor structure, *Acta Biochimica Polonica*, 2016, 63(4), 737–744.

INFLUENCING RNA 3D STRUCTURE PREDICTION PROCESS (2)

Secondary structure element		RNA FRABASE search pattern	3D element selected automatically by RNAComposer				3D element selected manually			
Id	Topology		PDB ID Chain	Position	Topology	Res. [Å] PSI [%]	PDB ID chain	Position	Topology	Res. [Å] PSI [%]
L1	CCU (.(AAG).)	NCN (.(NAN).)	3U5H BA	2338-2340, 1895-1897	CCU (.(AAG).)	3.00 100.0	4FE5 B	25-27, 43-45	UCG (.(CAA).)	1.32 33.33
L2	CCU (.(AAG).)	NCN (.(NAN).)	3U5H BA	2338-2340, 1895-1897	CCU (.(AAG).)	3.00 100.0	4FE5 B	25-27, 43-45	UCG (.(CAA).)	1.32 33.33
L3	CGAG (..(CCGG)..)	NRRN (..(NYRN)..)	3J3F AB	490-493, 574-577	CGUG (..(CCGG)..)	5.00 87.50	2ZJR X	1491-1494, 1530-1533	CAAG (..(UCAG)..)	2.91 50.00
L4	GAUAC (...(GC))	(...( ))	1U 3K BA	29-33, 6-7	GAUAU (...(AC))	NMR 71.43	397D AB	22-26, 39-40	ACUG (...(CU))	1.30 0.00





Reference:

M. Antczak, M. Popenda, T. Zok, J. Sarzynska, T. Ratajczak, K. Tomczyk, R.W. Adamiak, M. Szachniuk, New functionality of RNAComposer: application to shape the axis of miR160 precursor structure, *Acta Biochimica Polonica*, 2016, 63(4), 737–744.

FUTURE PLANS

1. Take into consideration **recurrent motifs predicted from sequence** in RNA 3D structure prediction process.

2. Allow user to **introduce own planarity restraints** applied during 3D model refinement.

3. Develop **rigid body minimization protocol** allowing for treating of long-range interactions.

4. Extend **secondary structure decomposition procedure and 3D elements database** to address specific motifs like **pseudoknots**, and **G-quadruplexes**.

5. Develop **fully-automated method for 3D structure prediction** of extremely large RNA 3D structures (**up to 1000 nts**).

6. Refine **3D models ranking protocol**.

13

RNAPDBEE 2.0
3D ⇒ 2D STRUCTURE

TOMASZ ZOK

Institute of Computing Science,
Poznan University of Technology,
Poland

Institute of Bioorganic Chemistry
Polish Academy of Sciences,
Poznan, Poland

WHY 3D → 2D?

- For RNAComposer to have its building blocks obviously ☺
- 3D data is very complex to analyze: non-canonical interactions, base-(ribose/phosphate), etc.
- Extraction of 2D data from 3D coordinates is not easy:
  - Technically – PDB format quirks, base modifications, missing atoms, etc.
  - Logically – pseudoknots introduce ambiguity!
- 2D visualization is a big help in understanding 3D structure

RNAPDBEE

RNApdbee webserver was made available (and published) in 2014 and can be accessed at:

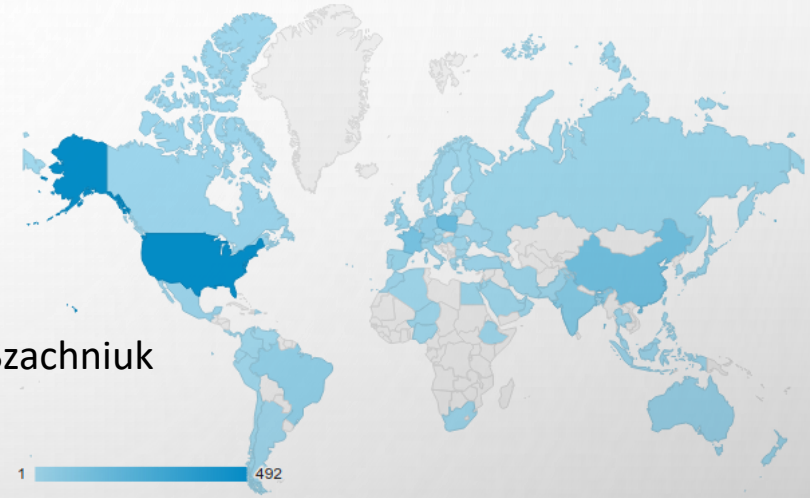
http://rnapdbee.cs.put.poznan.pl

Until July 2018, RNApdbee has collected:

~ **11,000** visitors

~ **2,700** sessions performed

~ **19** citations of the 1st paper



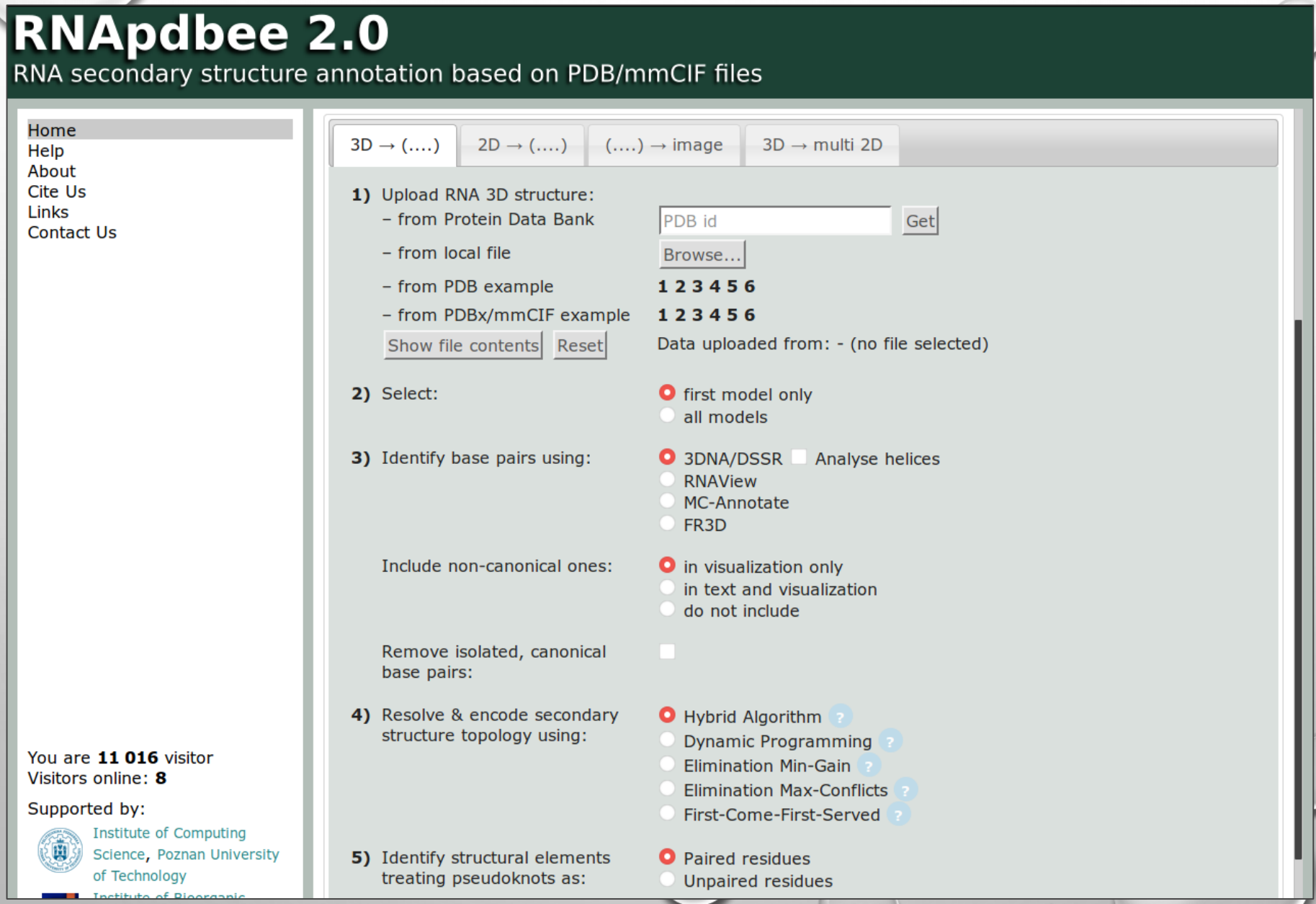
Team (currently):

M. Antczak, M. Popenda, T. Zok, R.W. Adamiak & M. Szachniuk

References:

- Antczak, M., Zok, T., Popenda, M., Lukasiak, P., Adamiak, R. W., Blazewicz, J., & Szachniuk, M. (2014). **RNApdbee--a webserver to derive secondary structures from pdb files of knotted and unknotted RNAs**. *Nucleic Acids Research*, 42(W1), W368–W372.
- Zok, T., Antczak, M., Zurkowski, M., Popenda, M., Blazewicz, J., Adamiak, R. W., & Szachniuk, M. (2018). **RNApdbee 2.0: multifunctional tool for RNA structure annotation**. *Nucleic Acids Research*, 46(W1), W30–W35.

http://rnapdbee.cs.put.poznan.pl



RNAPDBEE 2.0: A MULTIFUNCTIONAL TOOL

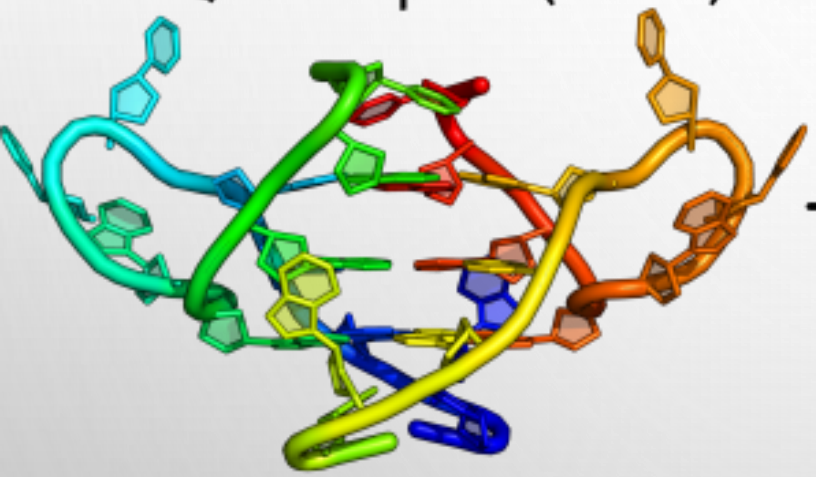
- Supports formats: PDB, **PDBx/mmCIF**, BPSEQ, CT, dot-bracket
- Integrates tools: RNAView **(with PDBx/mmCIF support)**, MC-Annotate **(vide supra)**, DSSR, **FR3D**, PseudoViewer, VARNA **(with own patches)**, **R-chie**
- Analyses:
  - Canonical and non-canonical base pairs
  - Stacking and **base-ribose** / **base-phosphate** interactions
  - **2D motifs (loops, stems, single strands)**
  - Pseudoknots using **one of five algorithms**
- Allows to run a **3D → multi 2D** scenario to seek a consensus based on different pseudoknot dedicated algorithms

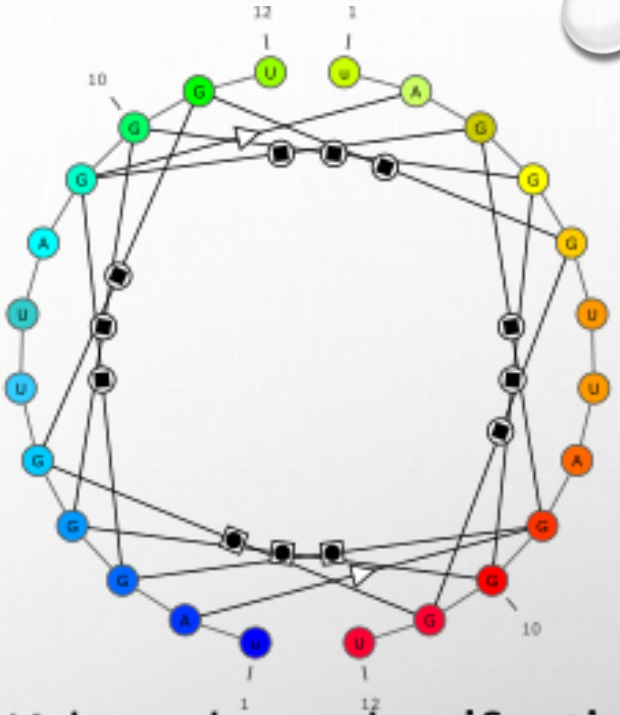
RNAPDBEE 2.0: PROPERTIES

- Input may come from **PDB servers** or **user supplied 3D/2D structures**
- The same structure can be easily analysed with **varying parameters**
- All results can be **downloaded in an aggregated archive**
- All images are **publication-quality vector graphics**
- **Missing residues** (i.e. lack of corrdinates in PDB) are present in sequence and marked with minus sign '-' in dot-bracket
- **Modified residues** are shown as lowercase in sequence (including self-detection of modifications)
- **RNA-RNA interactions** are analysed (both canonical and non-canonical)

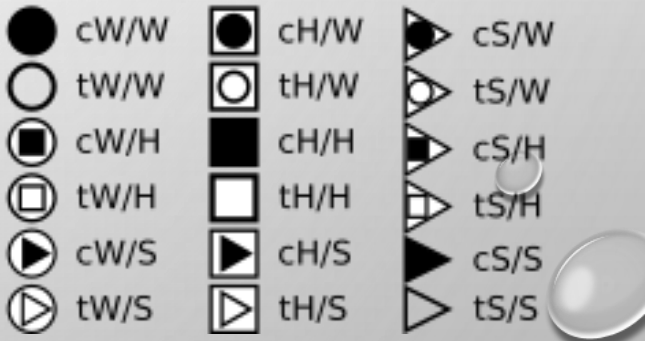
RNAPDBEE 2.0: VISUALIZATION

Quadruplex (2KBP)





RNA base-base classification

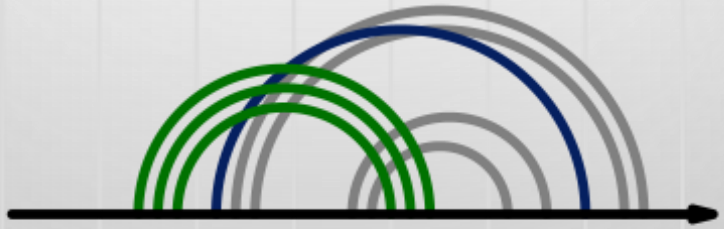


RNAPDBEE 2.0: VISUALIZATION

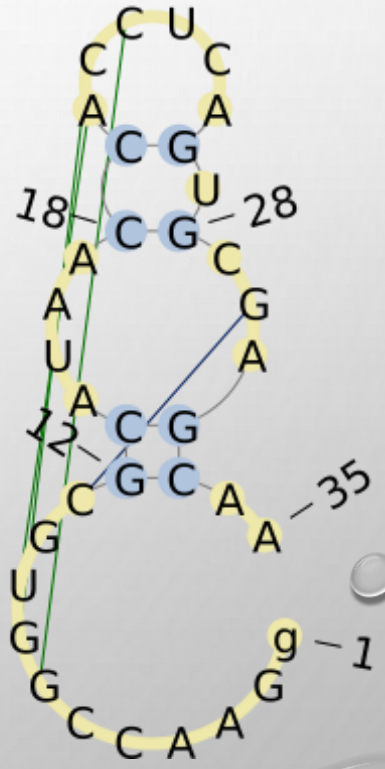
```
>strand_A
gGAACCGGUGCGCAUAACCACCUCAGUGCGAGCAA
......[[[.{((....((]]]...).).}.))..
```

R-chie

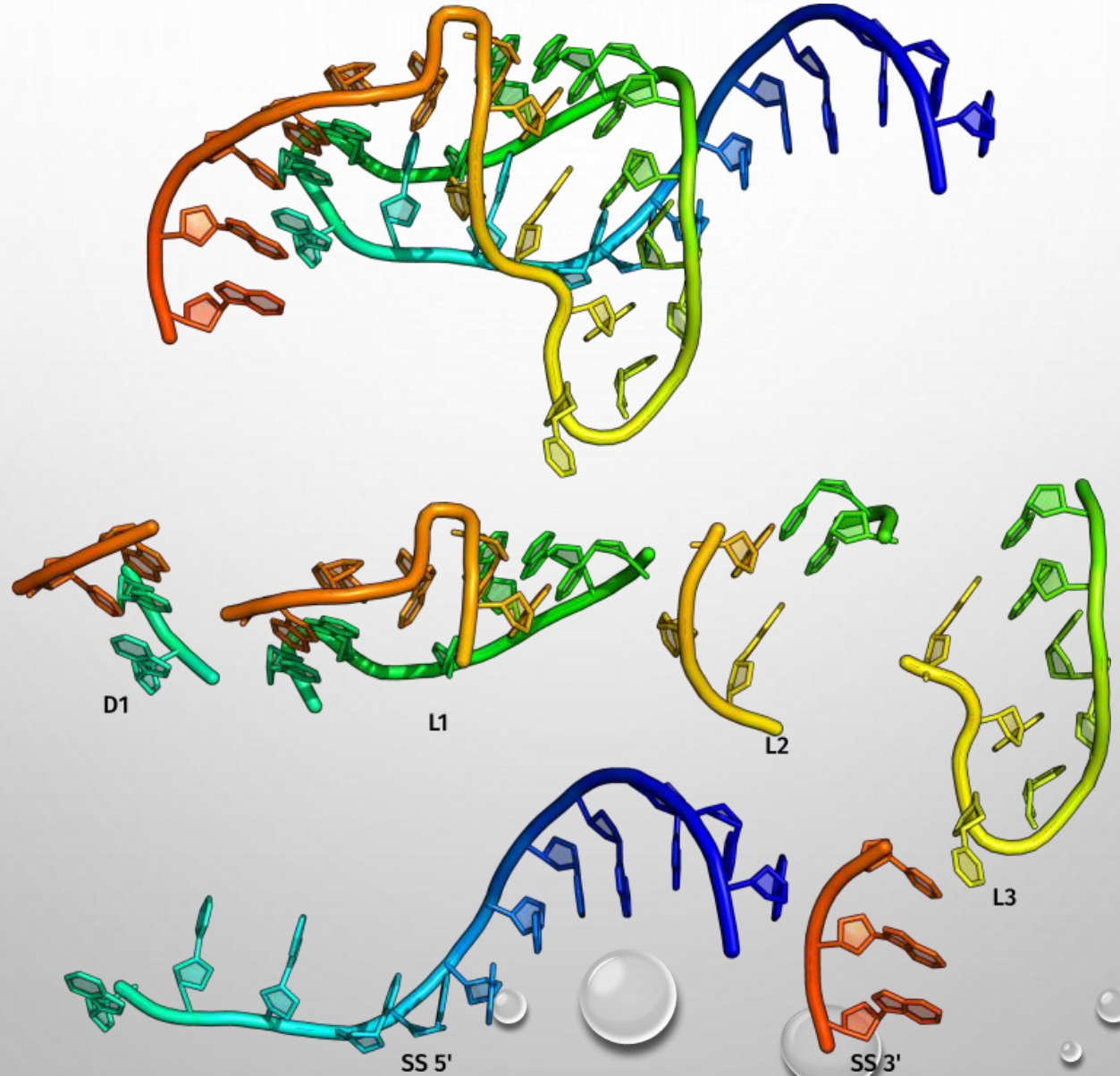
PseudoViewer



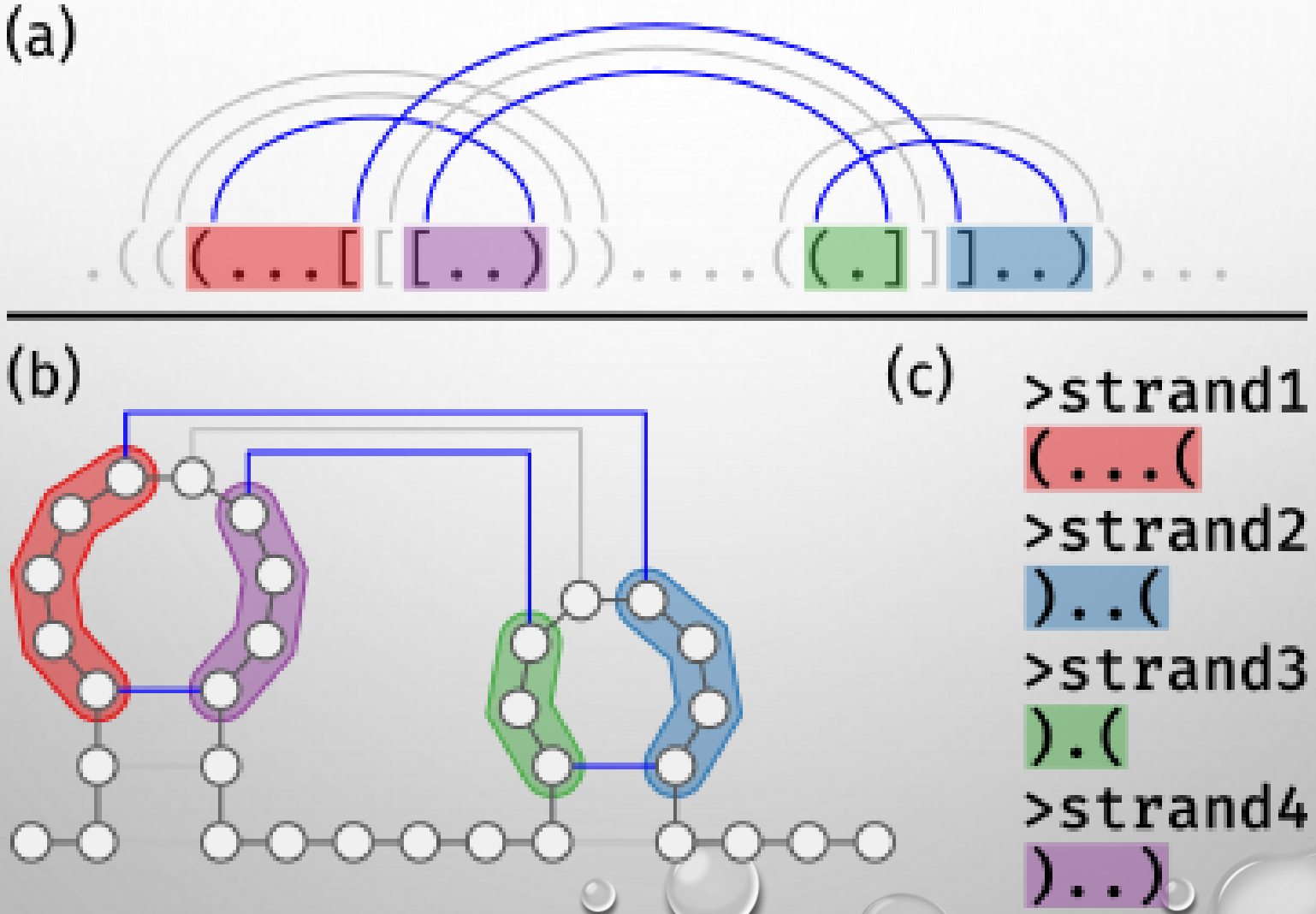
www.e–rna.org



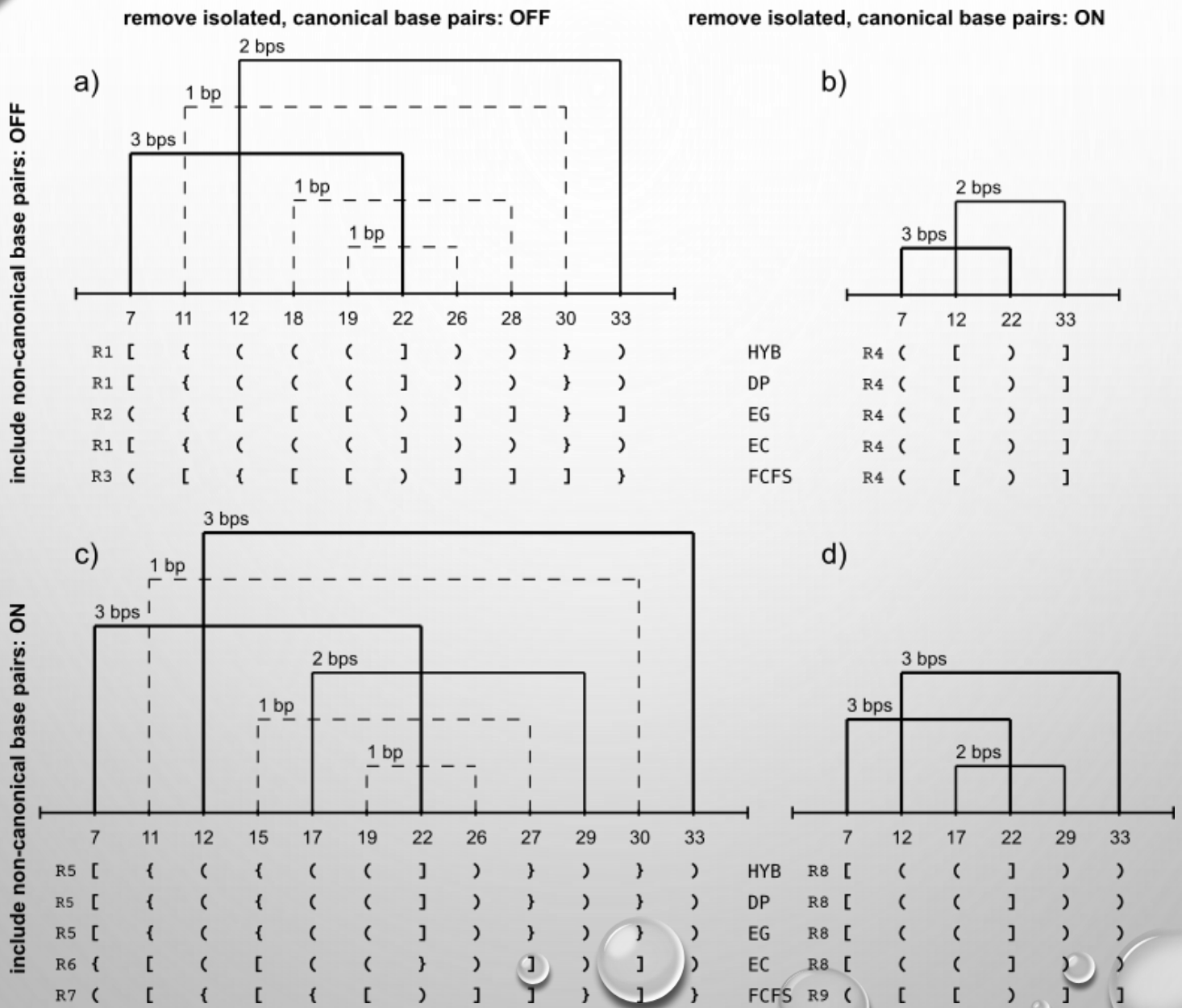
RNAPDBEE 2.0:
2D MOTIF ANALYSIS



RNAPDBEE 2.0:
2D MOTIF ANALYSIS



RNAPDBEE 2.0:
3D → MULTI 2D



RNAPDBEE 2.0: FUTURE PLANS

- Find and annotate **3D motifs,** both local and long-range
- Analyse new type of interactions: **RNA-protein** and **RNA-ligand**
- Assert **ions** roles for local RNA structure

You are welcome to use RNApdbee at:

http://rnapdbee.cs.put.poznan.pl

And let us know your opinions and feature requests ☺