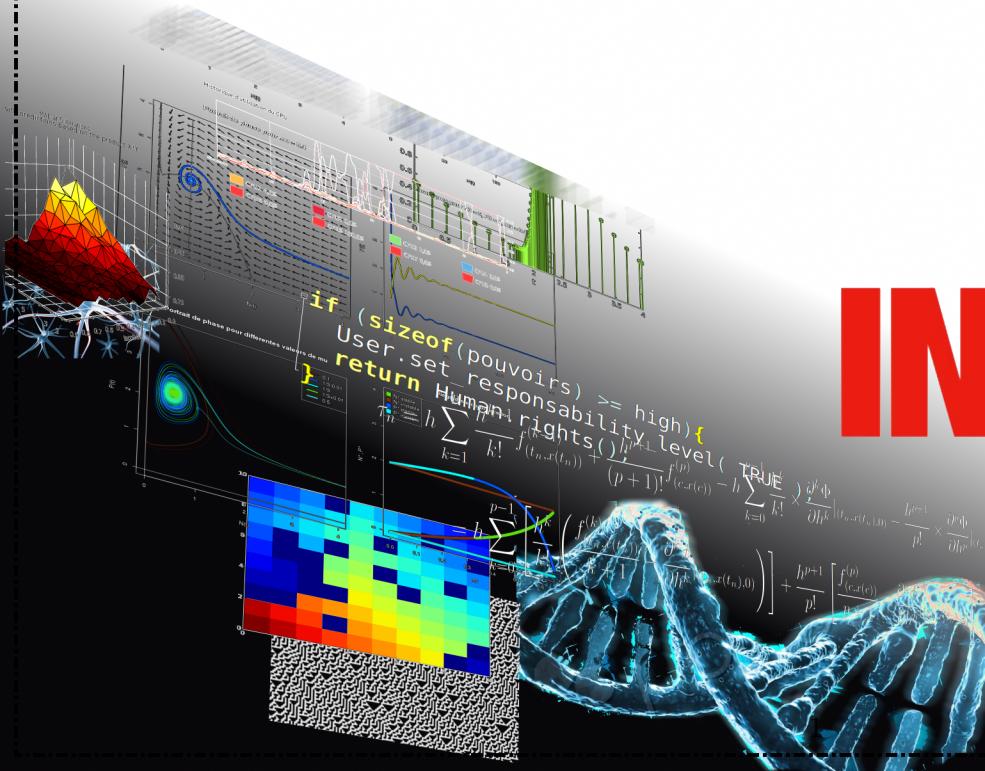


# Aptamer 3D-Structure Computation Tool

INSA DNA Concept Club - INSA Lyon

September 3, 2016



BIO  
INFORMATIQUE  
&  
MODELISATION

## Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Intro</b>  | <b>2</b> |
| <b>2</b> | <b>Reliability</b>                                    | <b>3</b> |
| <b>3</b> | <b>How it Works</b>                                   | <b>4</b> |
| 3.1      | If needed: transcription to RNA . . . . .             | 4        |
| 3.2      | Get the secondary structure . . . . .                 | 4        |
| 3.3      | Computation of the 3D structure from the 2D . . . . . | 4        |
| 3.4      | Minimization . . . . .                                | 4        |
| 3.5      | Selection of the best ones . . . . .                  | 4        |
| 3.6      | Creation of PDB files from the results . . . . .      | 4        |
| 3.7      | Re-converting to DNA . . . . .                        | 4        |
| 3.8      | Add hydrogens, and re-minimize . . . . .              | 5        |

## 1 Intro

### About:

This web tool has been developped by students of the INSA DNA Concept Club at the National Institute of Applied Sciences in Lyon, France.

It is a part of the team's larger project, the developpement of a simple, fast, and sensitive biosensor for several sexually transmitted diseases at the same time, that the team wil present to the iGEM competition.

See <https://2016.igem.org/Team:INSA-Lyon> for more information.

### Goals:

We deal with short DNA or RNA sequences called *aptamers*.

This tool has been made to predict the 5 most probable spatial 3D structures the molecule would have *in vivo*.

It should help you visualizing what changes happen when you substitue a nucleotide, etc...

The tool **outputs PDB files**, so you can use those aptamers for docking later.

### License:

The computation is mostly an automatisation of several prediction procedures described in articles (see **References** below) using software like **Biopython**, **ViennaRNA** and **Rosetta**.

Please notice Rosetta is used under an adapted **Academic Licence**.

By using this tool, you recognize you use this service for non-commercial purposes and you agree to Rosetta's license agreement at <https://els.comotion.uw.edu/licenses/86>.

### Concerned students & acknowledgements:

Louis Becquey & Juliens Orlans.

Acknowledgements to Pedro da Silva, Hubert Charles and the Biosciences Department of the Institute, and Etienne Fachaix for the web interface.



INSTITUT NATIONAL  
DES SCIENCES  
APPLIQUÉES  
LYON

## 2 Reliability

The structures you get as output may differ ones from each other. That's why we give five instead of only one. Most often, we can observe variable domains and stable ones.

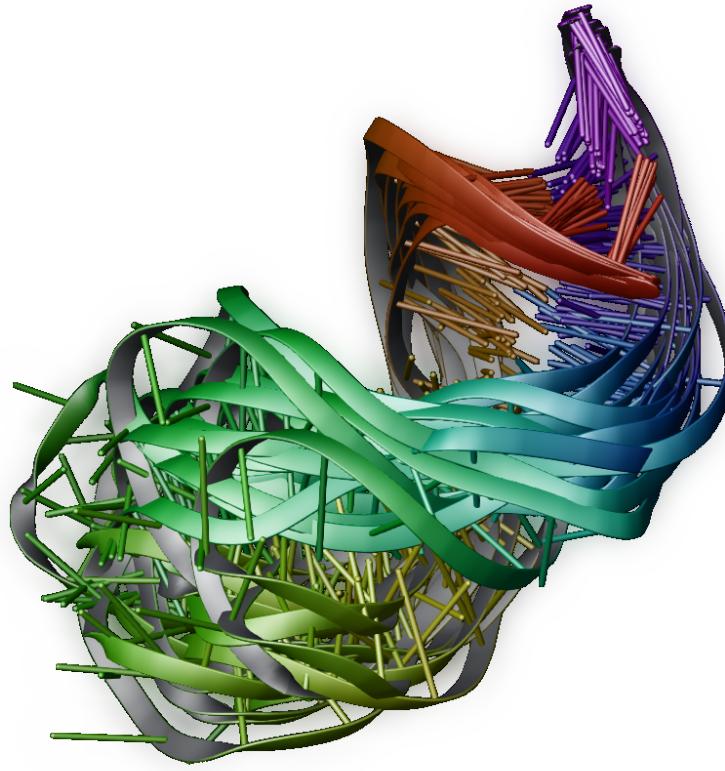


Figure 1: 20 most probable structures of the ATP aptamer

Obviously, we can't guarantee the truth of the results and we do not disclaim any responsibility for or liability related to them.

You can find more info about the scoring of structures in Rosetta's documentation.

### 3 How it Works

We consider that you start with your DNA or RNA sequence.

#### 3.1 If needed: transcription to RNA

The procedure has been thought for RNAs. So we transcript. But we have read this is not wrong<sup>[2]</sup>.

**Tech:** We use the Biopython package to do so.

#### 3.2 Get the secondary structure

If you did not give it yourself, we compute the secondary structure with ViennaRNA's RNAFold algorithm.

As always, if you know the secondary structure of your RNA, because you found it in an article or somewhat, prefer trusting what you read. This step is the most uncertain part of the procedure, because it is highly dependant of the environmental conditions.

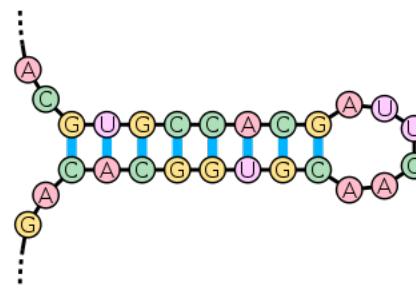


Figure 2: Example of RNA secondary structure.

**Important remark:** For now, we are unable to predict **pseudoknots** or **G-quadruplexes**, that may be frequent in aptamers. Giving your own secondary structure may be a way to explicit thoses structures.

#### 3.3 Computation of the 3D structure from the 2D

We use Rosetta's `rna_denoovo` tool on the given 2D data, with flags `-ignore_zero_occupancy` and `-no_minimize`, and let it run for 20000 cycles (default).

#### 3.4 Minimization

Rosetta will now score the files and minimize only the 100 best ones.

(I mean moving the atoms in space to reach a local minimum of potential energy.)

**Remark:** it is a delicate step, because we select 100 best structures precisely before having minimized them...

#### 3.5 Selection of the best ones

We ask Rosetta to score the files, so we can keep only the 5 best ones. This time, comparing really makes sense.

#### 3.6 Creation of PDB files from the results

Rosetta exports its results in PDB format.

#### 3.7 Re-converting to DNA

Then, we use a hand-made python script to modify the PDF files to change the uraciles in thymines and desoxygenate the riboses.

### 3.8 Add hydrogens, and re-minimize

Finally, Rosetta corrects the PDB (atom numbers, missing hydrogens) and re-minimizes the structure for having the most stable atom positions for DNA.

## References

- [1] Clarence Yu Cheng, Fang-Chieh Chou, Rhiju Das. **Modeling Complex RNA Tertiary Folds with Rosetta.** Methods in Enzymology, 553, 35-64, 2015, ISSN 15577988
- [2] Chin H Lin, Dinshaw J Patel. Structural basis of DNA folding and recognition in an AMP-DNA **aptamer complex : distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP.** Chemistry & biology, 4, 817-832, 1997, ISSN 10745521.
- [3] Rosetta's Documentation,  
<http://rosettacommons.org/docs/latest/Home>
- [4] ViennaRNA software,  
<https://www.tbi.univie.ac.at/RNA/>

