

## RNA dactyloscopy

Marta Magdalena Gabryelska<sup>1,2</sup>✉, Eliza Wyszko<sup>1</sup> and Jan Barciszewski<sup>1</sup>

<sup>1</sup>Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland; <sup>2</sup>Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, United Kingdom

**Despite the wealth of data on RNA secondary structure, conformational dynamics and tertiary structure *in vitro* and *in vivo*, predicting RNA biological activity in cellular environments remains difficult. Here, we present a comparison between *in silico* RNA fingerprinting and published experimental data that sheds light on efficient design of the hammerhead ribozyme molecules with a high intracellular efficiency. Our method, which we call RNA dactyloscopy, is a reliable tool for assessing the catalytic properties, modeling and design of RNA.**

**Key words:** hammerhead ribozyme; RNA tertiary structure; fingerprint

**Received:** 31 May, 2016; **revised:** 21 June, 2016; **accepted:** 05 August, 2016; **available on-line:** 01 December, 2016

### INTRODUCTION

In criminology, dactyloscopy refers to an analysis and classification of patterns observed in individual fingerprints. It is obviously associated with detective stories. We thus could be recognized as private investigators that want to figure out what happened within a cell. RNAs have a variety of secrets and many of them remain hidden. Therefore, assuming that every RNA molecule is unique, we came up with an idea of the RNA fingerprint (Gabryelska *et al.*, 2013). This allowed us to make the first step towards understanding how a distinctive geometry of a molecule can reveal its future path. Here we coin the term of RNA dactyloscopy and discuss its applicability.

The ubiquity and diverse functionality of ribonucleic acids (RNA) within the cell suggest its ability to act both, as a catalyst and storage of genetic information. This is the basis of the RNA world. In molecular biology, the building blocks of proteins or nucleic acids allow for assembly of cellular macromolecules into biologically active structures based on modeling and designing of new objects. This would make predictions of the RNA tertiary structure with a biological activity possible. *In silico* design of the definite structure and function of a molecule presents a big challenge for molecular biology.

It is known that the catalytic potential of RNA (ribozymes) is very wide. Hammerhead ribozymes (HH) catalyze transesterification (hydrolysis, degradation) of messenger RNAs and are a promising, versatile tool for down-regulation of gene expression *in vivo*. Due to their small size and a relatively high activity in protein-free environment in comparison to other catalytic RNAs, HH can be used as a model for RNA structure-function relationship studies (Perrault *et al.*, 2011).

### MATERIALS AND METHODS

**Structure modeling.** A selected substrate strand was joined with the ribozyme 3' end by the GGG sequence that allowed building a model of a type I cis-acting hammerhead ribozyme, necessary for modeling of an entire complex.

The RNA secondary structure predictions were performed through the Vienna RNAfold server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>, Lorenz *et al.*, 2011) with standard parameters to predict the minimum free energy structures (Turner model, 1999). Catalytic core sequence (CUGAUGA and GAA) were turned into NNNNNNN and NNN respectively, so these nucleotides did not affect the MFE structure prediction. The resulted dot-bracket notation was enriched with a catalytic core of C3-G5 base pairing to resemble the ribozyme catalytic core active conformation.

The RNA tertiary structure modeling was performed through the RNAComposer system (<http://rnacomposer.cs.put.poznan.pl>) (Popenda *et al.*, 2012). The program generates up to 10 structural models, where only the first one is the same in all repeats. Additional models are chosen by chance. For each secondary structure, 10 tertiary structures were given, 5 of which possessing the lowest CHARMM force field energy were selected for further calculations. Measurements were done with PyMol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). We analyzed different tertiary structure models of substrate:ribozyme complexes and measured two distances, D1: G12(N1) (general base) – C17(2'O) (nucleophile) and D2: G8(2'O) (general acid) – C1.1(5'O) (the leaving group). Also, the D1/D2 ratio was calculated which turned out to be the most promising factor.

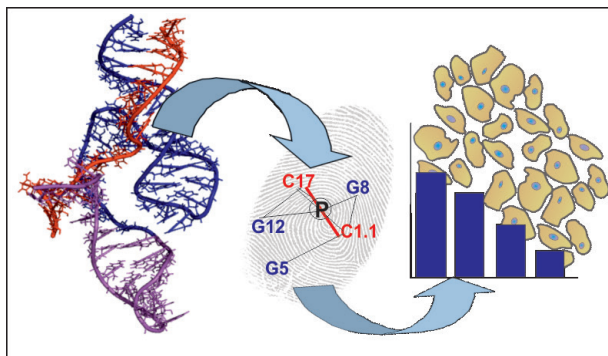
**Statistics.** Statistical analysis of comparison of *in silico* and *ex vivo* data and calculated correlations was performed with the GraphPad Prism (GraphPad Software) using the Kruskal-Wallis test.

### RESULTS AND DISCUSSION

We looked at the HH activity using the tertiary structure models of substrate:ribozyme complexes in active conformational states (Gabryelska *et al.*, 2013). To do this, we applied two methods: minimal free energy (MFE) calculation of secondary structures and the RNAComposer (Popenda *et al.*, 2012) application. We have shown a distinctive geometry of six parameters, particu-

✉ e-mail: [mgabryel@staffmail.ed.ac.uk](mailto:mgabryel@staffmail.ed.ac.uk)

**Abbreviations:** HH, hammerhead ribozyme; MFE, minimal free energy



**Figure 1. Hammerhead ribozyme dactyloscopy.** Tertiary structure of the ribozyme in complex with a substrate is used to obtain the catalytic core geometry parameters which than allow predicting the molecule's activity within cells.

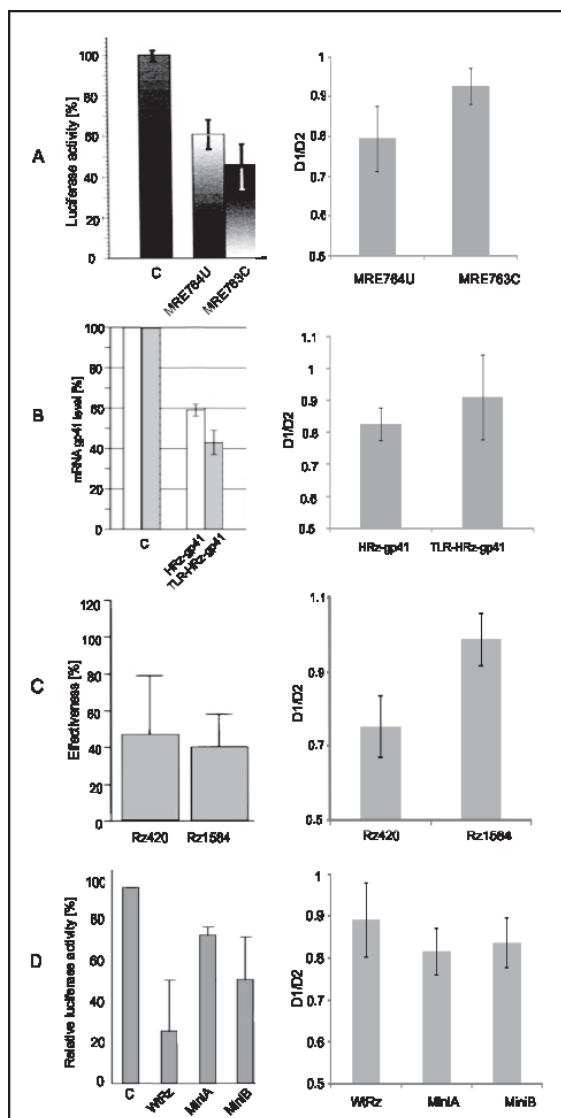
larly the distances and angles between the interacting atoms within the catalytic center. As such unique distribution of traits is specific for each ribozyme sequence, we called it the ribozyme fingerprint. This specific set of data characterizes the catalytic core very well. The RNA is an object, geometrical shape made of points, distances and angles. This leads us to a point of thinking of each and every RNA molecule as an individual set of computational data that hide functional information inside.

A high D1/D2 ratio, representing a close distance between stems I and II, was found to correlate with high efficiency of the ribozyme. This turned out to be a very helpful method for *in silico* selection of highly active molecules against a specific target. It is an easy and reliable alternative approach to traditional ribozyme design based on the RNA tectonics (Westhof *et al.*, 1996). A direct correlation between structure modeling and results of an intracellular experiment shows that the RNA design supported by computational analysis is a way to predict its cellular activity (Fig. 1) (Gabryelska *et al.*, 2013).

Unique set of computational data can be obtained really fast. If so, one can use them to predict the RNA intracellular activity just before it is synthesized. To validate our findings, we used data for hammerhead ribozymes acting *in trans*, e.g. preferably against the GUC or AUC sites, in animal cell lines, modeled as type I *in cis* ribozymes described earlier (Scherr *et al.*, 1997; Kato *et al.*, 2001; Kim *et al.*, 2003; Fedoruk-Wyszomirska *et al.*, 2009). We compared the activity and *in silico* properties of these molecules, particularly the data on downregulation of gene expression and the proposed ribozymes' secondary structures.

The hammerhead ribozymes targeting N-ras oncogene have been evaluated *ex vivo* in HeLa cells (Scherr *et al.*, 1997). The catalytic activity of two minimal variants (MRE763C and MRE764U) was measured as the loss of the reporter luciferase activity in 48 hrs post transfection. MRE763C and MRE764U displayed higher activity, which lead to loss of luciferase activity in about 54% and 39%, respectively. This overlaps with D1/D2 parameter of 0.8 and 0.93 for MRE764U and MRE763C, respectively (Fig. 2A). These results suggest that MRE763C should be more efficient in animal cells.

In another study, two hammerhead ribozymes have been designed for downregulation of HIV-1 gp41 mRNA (Fedoruk-Wyszomirska *et al.*, 2009). One was shortened, a minimal HHRz (HRz-gp41) and the second one was TLR-extended (TLR-HRz-gp41). Experiments in HeLa cells indicated a decrease in gp-41 mRNA ex-



**Figure 2. Comparison of the ribozymes' activity in the cells and *in silico* properties described in different papers A, B, C and D (Scherr *et al.*, 1997; Fedoruk-Wyszomirska *et al.*, 2009; Kim *et al.*, 2003; Kato *et al.*, 2001, respectively).**

On the left, original data of downregulating gene expression in the cell lines. On the right, results of D1/D2 calculations. C – control.

pression by 58% and 41% caused by TLR-HRz-gp41 and HRz-gp41, respectively. Computational analysis revealed the D1/D2 parameter of 0.83 and 0.91 for HRz-gp41 and TLR-HRz-gp41, respectively (Fig. 2B). Our data confirmed that TLR-HRz-gp41 should be more efficient in animal cells.

A different example is hammerhead ribozymes used to downregulate the expression of E6AP gene (Kim *et al.*, 2003). Two of these ribozymes fulfilled criteria presented in the previous work (Gabryelska *et al.*, 2013). The effectiveness of the ribozymes was estimated at the protein level. Rz420 and Rz1584 decrease the level of target by 51% and 60%, respectively. D1 to D2 ratio was 0.75 and 0.99 for R420 and Rz1584, respectively, which confirmed the second ribozyme to be more efficient in downregulation of gene expression in an animal cell line (Fig. 2C).

Out of three ribozymes against the BCR-ABL fusion gene, one was a minimal hammerhead ribozyme (WtRz)

and the others were shortened variants with only one base pair in the helix II (MiniA and MiniB) (Kato *et al.*, 2001). WtRz was the most active, leading to the decrease of the reporter luciferase activity by around 70%, in comparison to 45% (MiniA) and 25% (MiniB) exerted by other constructs. The D1/D2 ratio turned out to be 0.816437 (MiniA), 0.837608 (MiniB) and 0.89273 for WtRz, confirming the last one to be the most active under given conditions (Fig. 2D).

To our knowledge, there is no similar approach for predicting *in vivo* activity of ribozymes. It could be widely applied to characterize any RNAs with specified function, or as a selection system, making experiments less expensive and time-consuming. There are no other similar methods available. RNA dactyloscopy (generation of RNA fingerprints) is an attractive and efficient method of predicting experimental results. The concept of uniqueness of every RNA molecule at the atomic level gives broader perspective on the use and design of nucleic acids in general. Better structural prediction tools and automatic calculation algorithms would make such analysis more accurate and reliable. Every functional nucleic acid could be described in a similar way and thus correlations could be found, as it is all encoded in the common laws of mathematics and chemistry.

#### Author contributions

M.M.G., E.W. and J.B. designed research and analyzed data; M.M.G. performed research; M.M.G. and J.B. wrote the manuscript.

#### Competing financial interests

The authors declare no competing financial interests.

#### Funding source statement

This study was supported by The Polish Ministry of Science and Higher Education (grant numbers: N N302 220535, N N301 411238). M.M.G. was a scholarship holder within the project “Scholarship support for Ph.D. students specializing in majors strategic for Wielkopolska’s development”, Sub-measure 8.2.2 Human Capital Operational Programme, co-financed by the European Union under the European Social Fund. Supported by the Polish Ministry of Science and Higher Education, under the KNOW program.

#### REFERENCES

- Fedoruk-Wyszomirska A, Szymański M, Wyszko E, Barciszewska MZ, Barciszewski J (2009) Highly active low magnesium hammerhead ribozyme. *J Biochem* **145**: 451–459. doi: 10.1093/jb/mvn182
- Gabryelska MM, Szymański M, Wyszko E, Popenda P, Barciszewski J (2013) Prediction of hammerhead ribozyme intracellular activity with the catalytic core fingerprint. *Biochem J* **451**: 439–451. doi: 10.1042/BJ20121761
- Kato Y, Kuwabara T, Warashina M, Toda H, Taira K (2001) Relationships between the activities *in vitro* and *in vivo* of various kinds of ribozyme and their intracellular localization in mammalian cells. *J Biol Chem* **276**: 15378–15385. doi: 10.1074/jbc.M010570200
- Kim Y, Cairns MJ, Marouga R, Sun L (2003) E6AP gene suppression and characterization with *in vitro* selected hammerhead ribozymes. *Cancer Gene Ther* **10**: 707–716. doi: 10.1038/sj.cgt.7700623
- Lorenz R, Bernhart SH, Hoener zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL (2011) ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**: 26. doi: 10.1186/1748-7188-6-26
- Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G, Breaker RR (2011) Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comp Biol* **7**: e1002031. doi: 10.1371/journal.pcbi.1002031
- Popenda M, Szachniuk M, Antczak M, Purzycka KJ, Lukasik P, Bartol N, Blazewicz J, Adamiak RW (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res* **40**: e112. doi: 10.1093/nar/gks339
- Scherr M, Grez M, Ganser A, Engels JW (1997) Specific hammerhead ribozyme-mediated cleavage of mutant N-ras mRNA *in vitro* and *ex vivo*. Oligoribonucleotides as therapeutic agents. *J Biol Chem* **272**: 14304–14313. doi: 10.1074/jbc.272.22.14304
- Westhof E, Masquida B, Jaeger L (1996) RNA tectonics: towards RNA design. *Fold Des* **1**: R78–R88. doi: 10.1016/S1359-0278(96)00037-5