Towards *in vivo* origami: bio-orthogonal scaffolded RNA nanoribbons self-assembled via cotranscriptional folding

E. Torelli[1], J. W. Kozyra[1,§], B. Shirt-Ediss[1], L. Piantanida[2,^], K. Voïtchovsky[2], N. Krasnogor[1*]

[1]ICOS, CSBB, Newcastle University, Newcastle upon Tyne, NE1 7RX, UK
[2]Department of Physics, Durham University, Durham, DH1 3LE, UK
[§]Present Address: Nanovery Ltd, London, W1W 7LT, UK
[^]Present Address: Micron School of Materials Science & Engineering, Boise State University, Boise, ID 83725, USA

*Corresponding authors. Emails: natalio.krasnogor@newcastle.ac.uk, emanuela.torelli@newcastle.ac.uk

Background

Our main goal is the synthesis of RNA nanostructures as "bio-circuit board" in different engineered bacterial genera, combining portability, bio-orthogonality and specific functionalities tailored to different applications (e.g. control of metabolic pathways or cell's rheology). Towards this perspective, we successfully demonstrate the cotranscriptional folding of scaffolded RNA origami as a further step along the *in vivo* origami expression to precisely control biomolecular processes.

Several self-assembly strategies have been presented to design and synthesize RNA nanostructures, from RNA architectonics[1-5] to single-stranded self-assembly[6]. Nonetheless, the scaffolded RNA origami approach is not yet fully explored. Drawing from DNA origami technique in which a long scaffold sequence (commonly of viral origin) is folded into a desired nanostructure using several short staple strands[7], here we present a full isothermal protocol for bio-orthogonal scaffolded RNA origami assembly via cotranscriptional folding at 37°C[8].

In detail, bio-orthogonal De Bruijn sequence (DBS) scaffold[8,9] and seven staple strands, for the first time, are *in vitro* transcribed by bacteriophage T7 polymerase from dsDNA templates and folded into a nanoribbon shape in a one-pot reaction at physiologically compatible temperature. Scaffold DBS is characterized by lack of genetic information, restriction enzyme sites and reduced ambiguity in the addressability[9]. Staple strands sequences are designed in order to promote a reasonable transcriptional yield, allow an efficient control of the 5' transcript sequence content and reduce RNA byproducts, like aberrant sequences. The RNA nanostructure self-assembly is successfully confirmed by atomic force microscopy[8].

Furthermore, our new split Broccoli aptamer system[10] is included into the nanoribbon

design in order to demonstrate the folding through its functional activation and fluorescence emission upon binding to a specific dye, DFHBI-1T. Broccoli RNA aptamer[11] is dissected into two nonfunctional fragments each of which is integrated into the 5' or 3' end of two staple strands complementary to the DBS scaffold. When the RNA origami self-assembled into the target configuration, the split sequences are in closed proximity and the transcribed RNA nanoribbons are able to capture the specific dye showing a clear green fluorescent emission.

Finally, we simulate the RNA origami *in silico* using the nucleotide-level coarse grained model oxRNA to investigate the thermodynamic stability of the assembled nanostructure in isothermal conditions over a period of time (Figure 1).

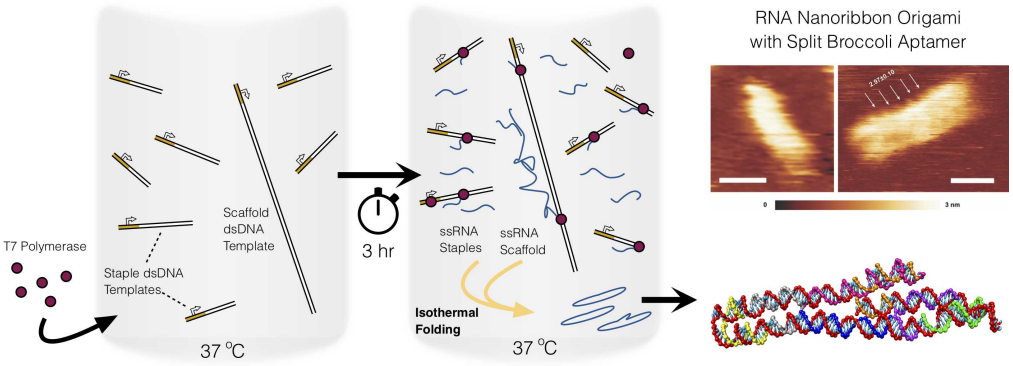

Figure 1. Schematic summary of the cotranscriptionally folded bio-orthogonal light-up RNA origami.

Results

OxRNA simulation

The RNA origami is visualized and simulated *in silico* at equilibrium in isothermal condition using oxRNA coarse-grained model[12,13]. The simulated RNA nanoribbon/split aptamer system stays at equilibrium and the split aptamer is preserved at 37 °C (Figure 2).

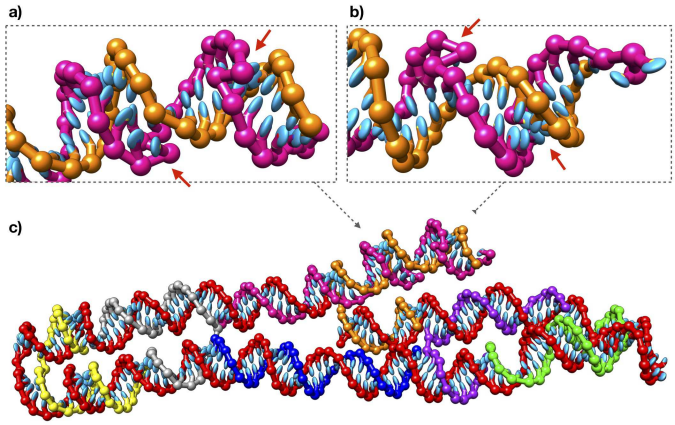

Figure 2. oxRNA simulation of a pre-assembled RNA ribbon that is formed of the scaffold (red) which is bound by 5 staples (various colours). The two split aptamer strands (pink and orange) are located at the 5' and 3' ends of the scaffold, accordingly. Enlarged views show the front (a) and the back (b) of the aptamer structure in respect to the RNA ribbon (c). Red arrows indicate the location of bulge loops[8].

Cotranscriptional folding and characterization by atomic force microscopy (AFM)

An *in silico* designed bio-orthogonal DBS (212 nt) is used as scaffold together with a set of seven staple RNA strands. The DBS scaffold bio-orthogonality was demonstrated in *E. coli* cells[9].

All staple strands sequences are designed in order to start with -GG or -GA and taking into account the lack of self-complementarity at the 3' end and of pairing between the 3' end adenine and uracil, 9 bases upstream. The 5' -GG or -GA content can improve the transcription in terms of yield and the control of 5' sequence content, while the 3' end characteristics can reduce the presence of transcription byproducts derived from a *trans* or a *cis* self-primer mechanism.

Each optimized sequence is placed downstream from the T7 RNA polymerase promoter and each dsDNA template is transcribed by T7 polymerase in an optimized transcription/folding buffer: *in vitro* transcribed RNA sequences are analyzed by denaturing PAGE[8]. The transcribed scaffold and staple strands show a distinct main band corresponding to the full-length transcript of the expected size. In addition, higher molecular weight products are observed as previously reported[14]: as RNA staples are designed to not support RNA-primed RNA extension, we postulate that, after the production of a certain amount of transcript, T7 polymerase accepts template RNA sequences that are not folded in 3' end stable secondary structures (e.g. harpin).

In our design to facilitate origami folding, all the sequences are optimized to weaken secondary structures and avoid hairpin formation[9]. We then hypothesize that during the cotranscriptional folding the 3' ends of the RNA transcripts are sequestered by the origami self-assembly.

Finally, all dsDNA templates encoding scaffold and staples strands are mixed in a one-pot reaction to transcribe each RNA sequence able to fold into the target origami configuration. The comparison between nondenaturing polyacrylamide gel images of fully folded sample, partially folded products and conventionally folded origami using synthetic RNAs, suggests the correct cotranscriptional folding of the RNA origami nanoribbon. The unpurified RNA origami sample is well-characterized by AFM confirming the correct folding of the origami with average lengths of 27.2 ± 3.5 nm and 8.4 ± 1.8 nm (Figure 3). During AFM imaging, interactions between RNA nanostructures and agglomerates are also observed and can be explained respectively as interactions between complementary unpaired ssRNA scaffold regions and undesired assemblies of transcribed RNA byproducts.

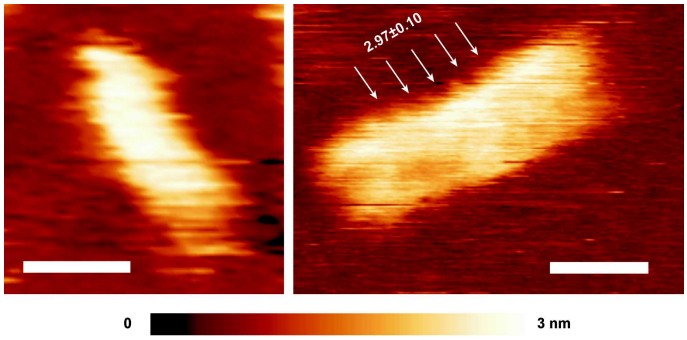

Figure 3. High-resolution AFM images of co-transcriptional folded RNA origami. The white arrows show the dsRNA helical pitch (3.2 ± 0.3 nm). Scale bar: 10 nm[8].

Light-up RNA origami: characterization by in-gel imaging

Broccoli aptamer has compact size, robust folding under physiological conditions and it selectively binds to a noncytotoxic and cell permeable dye[11]. For these reasons, we chose Broccoli aptamer among other light-up aptamers: Broccoli is divided into two afunctional sequences which are integrated into two staple strands. When the origami self-assembled, the split sequences are in close proximity and a fluorescent signal is detected[10]. After cotranscriptional folding and PAGE, the transcription mix is

analyzed by an optimized in-gel imaging protocol using the specific DFHBI-1T dye. Gel image clearly shows a main fluorescent band (lane 7, approx. 6 times higher compared to each of the two lower bands) corresponding to the RNA origami band as revealed by the following staining with SYBR Gold (Figure 4). In lane 7 is also detectable a faint fluorescent band characterized by higher mobility compared to the scaffold: this band could represent unknown transcribed sequences. Lane 5 and 6 correspond to partially folded samples: the fluorescent bands are not visible in the fully assembled origami sample case. The double staining using DFHBI-1T dye (Figure 4, left) and SYBR Gold (Figure 4, right) allow the confirmation of the proper self-assembly considering the specific migration distance and fluorescence of the RNA origami band[8].

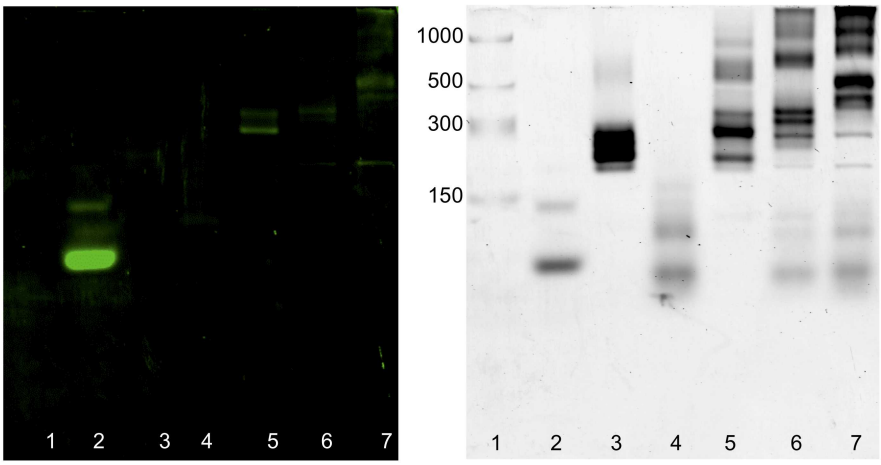

Figure 4. In-gel imaging of co-transcriptional folded light-up RNA origami. 10% TBE gel electrophoresis after DFHBI-1T (a) and after SYBR® Gold (b) staining. Lanes: 1: low range ssRNA ladder; 2: Broccoli aptamer; 3: transcribed RNA scaffold; 4: transcribed RNA staples; 5: transcribed scaffold, s1 and s2 staples; 6: transcribed scaffold, s1, s2, l1 and r1 staples; 7: transcribed RNA origami. Molecular size in nucleotides are indicated[8].

Conclusion

Here we combine three important aspects: i) bio-orthogonality, ii) RNA origami assembly monitoring using our Split aptamer system, and iii) origami cotranscriptional folding in a one-pot reaction at 37 °C from double-stranded DNA templates.

The Split aptamer introduced as a tag into the RNA origami design can be used as a fluorescence detection system to monitor the self-assembly inside a living cell without

using fluorescent reporter proteins and thus reducing the burden to the cell. The in-gel imaging assay using a cell permeable and compatible dye can be applied as a rapid and specific pre-screening of the well folded nanostructures[8]. Furthermore, the fluorogenic aptamer can be replaced with other functional sequences required for different purposes.

The cotranscriptionally folded RNA origami can be compatible with expression in bacterial cells and could act as a small "bio-circuit board" allowing the precise molecular arrangement through protruding staple strands. The resulting spatial organization can represent a strategy to direct metabolic processes in engineered bacterial cells[15].

Fundings

This work was supported by Engineering and Physical Sciences Research Council grant EP/N031962/1. Prof. Krasnogor is supported by a Royal Academy of Engineering Chair in Emerging Technology award. KV and LP acknowledge funding from the Biotechnology and Biological Sciences Research Council (grant BB/M024830/1).

References

[1] Shu, D. et al. *Nano Lett*. 2004, **4**, 1717-1723.

[2] Chworos, A. et al. *Science* 2004, **306**, 2068-2072.

[3] Severcan, I. et al. *Nat. Chem.* 2010, **2**, 772-779.

[4] Shu, Y. et al. *Methods* 2011, **54**, 204-214.

[5] Ishikawa, J. et al. *Wiley Interdiscip. Rev.: RNA* 2013, **4**, 651-664.

[6] Afonin, A. K. et al. *Nat. Nanotech.* 2010, **5**, 676-682.

[7] Rothemund, P. W. K. *Nature* 2006, **440**, 297-302.

[8] Torelli, E. et al. *ACS Synth. Biol.* 2020*,* **9**, 1682-1692.

[9] Kozyra, J. et al. *ACS Synth. Biol.* 2017, **6**, 1140-1149.

[10] Torelli, E. et al. *Sci. Rep.* 2018, **8**, 6989.

[11] Filonov, G. S. et al. *J. Am. Chem. Soc.* **2014***, 136*, 16299-16308.

[12] Šulc, P. et al. *J. Chem. Phys.* 2014, **140**, 235102.

[13] Matek, C. et al. *J. Chem. Phys.* 2015, **143**, 243122.

[14] Oesinghaus, L. et al. *Nat. Comm.* 2019*,* **10**, 2092.

[15] Delebecque, C. J. et al. *Science* 2011, **333**, 470-474.