Rapid enzymatic synthesis of long RNA polymers: A simple protocol to generate RNA libraries with random sequences

Irina Randrianjatovo-Gbalou, Marc Delarue

Rapid enzymatic synthesis of long RNA polymers: a simple protocol to generate RNA libraries with random sequences

Irina Randrianjatovo-Gbalou[1] and Marc Delarue

Unit of Structural Dynamics of Biological Macromolecules and CNRS UMR 3528,

Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France.

[1]corresponding author, irina.gbalou@gmail.com

Abstract

RNA aptamers have several advantages over DNA aptamers due to their propensity to fold into three-dimensional structures. However, the synthesis of large RNA libraries remains a challenge as it requires more precautions to conserve their functional integrity, especially when such libraries are intended for aptamers or ribozymes selection. Here, we present an enzymatic method that enables the rapid synthesis of RNA polymers thanks to the efficient incorporation of ribonucleotides (NTPs) as well as chemically modified ribonucleotides by human DNA polymerase Theta (θ) mutants. These mutants have the ability to generate long single-stranded RNA polynucleotides of random sequences due to their improved template-free terminal nucleotidyltransferase activity. Here we describe the detailed protocols to produce large and diverse libraries of RNA, to make them ready to use in repeated cycles of Systematic Evolution of Ligands by Exponential enrichment (SELEX) and to synthesize C2'-modified nucleic acids with higher nuclease resistance.

Keywords: RNA ; functional RNA synthesis ; DNA polymerase theta; randomized library; 2'-O-methyl library; C2'-modified oligonucleotides.

1. Introduction

The discovery of numerous classes of ribonucleic acids (RNAs) and their functions in a wide spectrum of biological processes has revolutionized the field of RNA biology. Functional RNAs encompass aptamers [1]–[4], ribozymes [5], antisense RNAs [6]–[8], small interfering RNAs [9] or microRNAs.

During the last decade, the selection of aptamers has gained more and more interest as they proved useful affinity reagents for cellular imaging [10]–[14] and are potentially promising therapeutics compared to antibodies or other molecules [15], [16]. Aptamers are single-stranded RNA or DNA oligonucleotides that can mimic the functional properties of antibodies and are selected in vitro from large libraries for their high affinity to a target molecule. The first choice in the aptamer discovery process is to decide whether to select for DNA or RNA libraries. The generation of RNA libraries is more laborious and needs to first generate libraries of cDNA and to use RT-PCR. Also, RNA is significantly unstable because of the presence of nucleases in biological solutions [17], and, last but not least, producing RNAs is more expensive than DNA. Compared to custom DNA synthesis, which is easily realized chemically using solid-phase processes from classical suppliers, the production of RNA oligonucleotides, with or without special modifications, is done at a much higher cost. As an example, the synthesis of 2'-Fluoro-RNA of 100 nucleotides costs up to $2500 per milligram [3], which is an important limitation to RNA aptamer development. On the other hand, RNA molecules are more versatile in forming 3D conformations due to the propensity of the 2'hydroxyl group of the sugar moiety to form hydrogen bonds and numerous therapeutics drugs have been discovered from RNA aptamers [18].

In this work, we propose a rapid and low-cost enzymatic method for producing large RNA libraries that might be used for RNA aptamer selection and RNA functionalization with chemical modifications that confer nuclease resistance and better pharmacokinetics and pharmacodynamics [16], [19]. For that purpose, we designed mutants of the A-family Human DNA polymerase θ, whose terminal transferase properties have been described previously [20], so that the selected variants now accept the incorporation of natural and modified ribonucleotides [21] and elongate single-stranded oligonucleotides from their 3'-end in a template-free manner. For the record, the active site of DNA polymerases display highly conserved sequence motifs from which it is possible to classify them into families [22]. The human DNA polymerase θ (polθ) crystal structure has been recently solved [23] and indeed shares the three expected strictly conserved motifs (A, B and C), all located in the palm domain (see Figure 1). During the DNA primer elongation process, the conserved aspartate (D2330) in motif A (amino acid sequence: DYSQLELR) interacts with the incoming deoxyribonucleotide (dNTP) and the magnesium ion to stabilize the transition state that leads to the formation of a new phosphodiester bond. In a previous study [21], several mutants were generated by single or multiple substitutions of residues around the incoming nucleotide (L2334, E2335, Q2384, Y2387) to increase the acceptance of ribonucleotides (NTPs) instead of dNTPs.

After having optimized the artificial RNA synthesis protocol by these mutants, we present here in detail the procedures that enable 1) to generate large and diverse libraries of RNA, 2) to make the fragments ready to use for RT-PCR and RNA aptamer selection, and finally 3) to produce fully modified RNA or DNA oligonucleotides with enhanced resistance to nuclease degradation.

2. Enzymatic RNA synthesis of random sequences and library generation

2.1. Materials and reagents

All the reagents used in the following procedures were RNAse-free or were beforehand DEPC-treated. The main steps of RNA synthesis and library generation are summarized in Figure 2.

- RNAse-free water not DEPC-treated (Thermofisher)
- Diethyl pyrocarbonate (DEPC) (Sigma Aldrich)
- 500 µM ribonucleotides, resuspended in RNAse-free water (ATP, GTP, CTP, UTP, mix of the four NTPs) (Trilink Biotechnologies).
- Fluorescent RNA 15-mer primer (RP-HPLC purified) : 5'AlexaFluor488-UACGCAUUAGCAAUG (Eurogentec)
- 5 mM Manganese chloride solution, $MnCl_2$ (Sigma Aldrich)
- 5 µM Human DNA polymerase θ (WT or mutant prepared by site-directed mutagenesis)
- RNAse inhibitor RNAsin® (Promega)
- Human DNA polymerase θ buffer: 20 mM Tris-HCl pH 8 (Sigma Aldrich), 10% glycerol (Sigma Aldrich), 0.01% IGEPAL C6-30 (Sigma Aldrich), 0.1 $mg.ml^{-1}$ BSA (Sigma Aldrich).
- RNAse-free EDTA solution (ThermoFischer)
- T4 RNA ligase 1 High concentration kit (New England Biolabs)
- Fluorescent RNA oligonucleotide, ddC-blocked and phosphorylated at the 5'-end and conjugated with Cy5 dye, (sequence of ligRNA: 5'-pCUAUGC**U**(Cy5)AAUGUCCddC) (Eurogentec).
- Formamide blue solution: 10 mM EDTA (Sigma Aldrich), 98% formamide (Sigma Aldrich), 0.1% Bromophenol blue (Sigma Aldrich).
- Denaturing polyacrylamide gel: prepared at final concentrations of 15% acrylamide/bisacrylamide (from 30% acrylamide/bisacrylamide 29:1, Bio-RAD), 8 M Urea (Sigma), Tris-Borate-EDTA buffer 1X (from a 10X stock solution that contains 0.89 M Trizma with 0.02 M EDTA adjusted to pH 8.3 with boric acid, Sigma-Aldrich).
- Electrophoresis migrating buffer: Tris-Borate-EDTA buffer 1X
- RNA clean & concentrator™-5 kit (Zymo Research).
- MAXIMA Reverse Transcriptase kit, with dNTPs mix and RT buffer (ThermoFisher)
- RT_rev-primer sequence: 5'-AAAATAAGGGACATTAGCATAA, RP-HPLC purification (Eurogentec)
- Phusion High-fidelity DNA polymerase (NEB)

2.2. Random RNA fragments synthesis

2.2.1. Kinetics of terminal transferase activity of polθ variants for the incorporation of rNTPs

The random RNA fragments synthesis is performed according to the protocol described by Randrianjatovo-Gbalou et al. [21]. Here, the primer extension is performed by the DNA polymerase θ mutant DW9 which carries the double substitution L2334M-E2335G. The enzyme displays an increased terminal transferase activity and can add randomly each of the four natural ribonucleotides to the 3'-end of the AlexaFluor488-RNA primer (15-mer) used to initiate the elongation reaction (sequence: 5'-UACGCAUUAGCAAUG).

2.2.2. Operating procedures

- Mix the following in a 1.5 ml-microcentrifuge tube:

5 x DNA polymerase θ buffer	40 µL
50 mM $MnCl_2$	20 µL
5 mM NTPs mix	20 µL
5 µM Alexa-Fluor488 RNA primer	20 µL
40 U/µL RNAsin	5 µL
RNase-free water	75 µL
50 µM PolθDW9	20 µL
	Total 200 µL

- Incubate at 42°C for 5 s, 30 s, 1 min, 5 min and 15 min.
- Stop the reaction by adding 25 mM EDTA (final concentration 12.5 mM).
- Clean-up the synthesis products by using RNA clean & concentrator™-5 kit according to the supplier purification protocol for small fragments (17-200 nt fragments). Elute the purified RNA fragments (synRNA) with 20 µL RNAse-free water.
- Sample 10 µL of synthesis product and add 20 µL of formamide blue to stop the reaction.
- Heat the tubes at 95°C for 5 min and immediately store on ice.
- Load the samples on a 15% denaturant polyacrylamide gel and migrate for 4 h (2000V, 40W).

2.3. Ligation of a fixed oligonucleotide at the 3' end of the random fragments

In a second step, the 3'-end constant region has been added by the ligation of a 15-mer RNA oligonucleotide phosphorylated at the 5'-end, conjugated with CY5 dye and ddC-blocked. The ligation has been performed using T4 RNA ligase I (from NEB).

- Mix the following in a 1.5 ml-microcentrifuge tube:

synRNA	100 pmol
ligRNA	20 pmol
10 x T4 RNA ligase buffer	2 µL
10 mM ATP	2 µL
PEG 8000 50%	5 µL
RNAsin (40 U/µL)	1 µL (2 U/µL)
RNAse-free water	1 µL
T4 RNA ligase 1	1 µL (10 Units)
	Total 20 µL

- Incubate the mix overnight at 18°C.
- Heat the mix at 65°C to inactivate the ligase.
- To verify the ligation efficiency, sample 10 µL of the ligation product and add 20 µL of formamide blue to stop the reaction.
- Heat the tubes at 95°C for 5 min and immediately store on ice.
- Load the samples on a 15% denaturant polyacrylamide gel and migrate for 4h (2000V, 40W).
- Scan the gel through TYPHOON™ FLA 9500 biomolecular imager (GE Healthcare Life science) by first measuring the Cy5 fluorescence emission after 635 nm red laser excitation, then the AlexaFluor488 fluorescence emission after 473 nm laser excitation.

2.4. Reverse transcription using MAXIMA Reverse transcriptase

In order to check that the synthetic RNA fragments do possess the two constant regions at the 5' and 3'ends and a variable region in between, a reverse transcription followed by a cDNA amplification has been done. The reverse transcription (RT) assay has been performed by using the MAXIMA Reverse Transcriptase and the customized reverse primer (RT_rev-primer sequence: 5'AAAATAAGGGACATTAGCATAA) as described in Figure 2.

- ***RT-premix composition***

Template RNA (0,01 pg – 500 ng)	2 µL
RT-rev-primer (1 µM)	1 µL
dNTPs mix (0.5 mM)	4 µL
RNase-free H_2O	qsp 14.5 µL

Incubate at 65°C for 5min, then chill on ice.

- ***RT reaction***

Add to the RT premix:

5X RT buffer	4 µL

RNase inhibitor (20 U)	0.5 µL
MAXIMA reverse transcriptase (200 U)	1 µL
	Total RT reaction volume 20 µL

Mix gently, centrifuge and incubate at 55°C for 60 min. To stop the reaction, incubate at 85°C for 5 min, then chill on ice.

- ***RT controls***

Negative control (NC): all the reagents described above were added except the MAXIMA reverse transcriptase.

No template control (NT): All reagents described above were added except the RNA template.

2.5. cDNA amplification

cDNA amplification has been completed by Phusion High-fidelity DNA polymerase.

- ***PCR mix reaction***

5 x Phusion buffer	10 µL
2.5 mM dNTPs mix	4 µL
20 µM Forward primer	1.25 µL
20 µM reverse primer	1.25 µL
Template DNA (≈100 ng) or RT controls	2 µL
DMSO	1.5 µL
Phusion DNA polymerase	0.5 µL
RNase-free H_2O	29.5 µL
	Total volume 50 µL

- ***Temperature cycles***

Initial denaturation	30s – 98°C	
Denaturation	10s – 98°C	30 cycles
Annealing	1min 30 – 56°C	
Extension	30s – 72°C	
Final extension	1min 30 – 72°C	

Analyze samples after migration on a 3% agarose gel and TAE buffer.

3. C2'-modified nucleotides incorporation by DNA polymerase θ mutants

The ability of DNA polymerase θ mutants to incorporate C2'-modified has been evaluated. Two mutants were tested, DW9 (L2334M-E2335G) and CS13 (E2335G). The template-free terminal

transferase activity of both enzymes has been verified through the elongation of a fluorescent or P32-radiolabelled DNA and with 2'-Fluoro- or 2'-O-methyl-modified nucleotides as substrates.

3.1. Reagents

- 2'-Fluoro-dATP, dCTP, dGTP, dTTP, dUTP (Trilink Biotechnologies)
- 2'-O-methyl-ATP, CTP, GTP, UTP (Trilink Biotechnologies)
- Fluorescent ATTO488-DNA primer 14mer (sequence: 5'TACGCATTAGCATA, Eurogentec)
- DNA primer 14mer 5'TACGCATTAGCATA (Eurogentec)
- T4 Polynucleotide kinase (NEB)
- ATP, [γ-^{32}P], 3000 mCi/mmol (PerkinElmer)
- DNA polymerase θ-DW9 and CS13 mutants, prepared according to the protocol described in [21].
- Formamide blue solution: 10 mM EDTA (Sigma Aldrich), 98% formamide (Sigma Aldrich), 0.1% Bromophenol blue (Sigma Aldrich).
- Denaturing polyacrylamide gel: 8% acrylamide/bisacrylamide 29:1 (Bio-RAD), 8 M Urea (Sigma), Tris-Borate-EDTA buffer 1X (Sigma Aldrich).
- Electrophoresis migrating buffer: Tris-Borate-EDTA buffer 1X

3.2. Radiolabeling of ssDNA primer

The oligonucleotides used for the radioactive template-free nucleotidyltransferase activity of DNA polymerase θ mutants were labelled following the protocol described in [21].

3.3. Incorporation of C2'-modified nucleotides by template-free nucleotidylterminal transferase activity

- ***For non-radioactive reaction, mix the following in a 1.5 ml-microcentrifuge tube***

5 x DNA polymerase θ buffer	2 µL
50 mM $MnCl_2$	1 µL
5 mM 2'Fluoro-dNTP or 2'-O-methyl-NTP	1µL
5 µM ATTO488 ssDNA primer	1 µL
RNase-free water	4 µL
50 µM PolθDW9 or PolθCS13	1 µL
	Total 10 µL

- Incubate at 42°C for 30 min
- Stop the reaction by adding 20 µL of Formamide blue.
- Heat the tubes at 95°C for 5 min, store in ice until loading on denaturing 8% polyacrylamide gel.

- ***For radioactive reaction, mix the following in a 1.5 ml-microcentrifuge tube***

5 x DNA polymerase θ buffer	2 µL
50 mM $MnCl_2$	1 µL
5 mM 2'Fluoro-dNTP or 2'-O-methyl-NTP	1µL
500 nM [γ-^{32}P]ATP-ssDNA primer	1 µL
RNase-free water	4 µL
50 µM PolθDW9 or PolθCS13	1 µL
	Total 10 µL

- Incubate at 42°C for 30 min
- Stop the reaction by adding 20 µL of Formamide blue.
- Heat the tubes at 95°C for 5 min, store in ice until loading on denaturing 8% polyacrylamide gel.

4. Results

4.1. Enzymatic synthesis of RNA libraries for aptamer selection

The main goal of this work was to propose a rapid method to generate a pool of randomized sequences of RNA fragments by exploiting the properties of a versatile DNA polymerase that belongs to the A-family, the Human DNA polymerase θ. As previously described [21], this DNA polymerase has been evolved to accept natural NTPs and to incorporate them in a template-free fashion at the 3'-end of a single-stranded DNA or RNA primer. Creating pools of oligonucleotides for SELEX involves the ability to generate fragments with constant regions at 5'- and 3'-ends on either side of the randomized central sequence. During the SELEX procedure it is critical to control the amplification step of nucleic acids aptamers that show large affinities to the molecule target. These regions are needed for the reverse transcription reaction in order to obtain the complementary DNA and to perform PCR to enrich the pool for the next cycle. The same regions must also not interact or form undesirable secondary structures nor influence the 3D conformation of the future aptamer. That is what we have accomplished and report in this study.

The RNA synthesis was stopped by the addition of 12.5 mM EDTA at the appropriate time to achieve the desired average length and the RNAs were cleaned-up. The T4 RNA ligase I was then used to ligate a 5'-phosphorylated ssRNA fragment labelled with Cy5 and preferably blocked at its 3'-end to avoid auto-ligation to the newly synthesized RNA fragment (labelled with AlexaFluor488). The products of the reaction were resolved by gel electrophoresis on a 15% acrylamide gel and 8 M urea. A double imaging has been performed i) by using the Alexa488 filter for the detection of the acceptor RNA length and ii) by using the Cy5 filter for the detection of the ligation of the Cy5-ligRNA.

The double fluorescence detection performed in the same gel allowed the observation of both the quality of the primer elongation (green fluorescence) and the efficiency of the ligation reaction (red fluorescence). The Figure 3 presents the primer elongation rate of the DW9 mutant by adding natural ribonucleotides at 3'-end of a RNA primer. The 6 first lanes of the gel display the evolution of the

elongation of the primer as a function of the incubation time (5 seconds to 15 minutes) and demonstrate an exponential increase of the length of the RNA with the duration of the incubation. The following lanes (7 to 11) correspond to the same newly synthesized RNA now ligated to the 15-mer constant region. The efficiency of the ligation has been evaluated by the presence of both green and red fluorescence, resulting in a yellow color along the gel and displaying variable sizes. In addition, thick bands are observed at exactly +15-added nucleotides, which correspond to the ligation of the constant region to the non-elongated RNA primers that have remained in the reaction bulk.

Finally, the lane 12 shows the absence of auto-ligation when the ligRNA fragments are mixed with T4 RNA ligase. Indeed, no band is visible except the "+0" one that corresponds to the length of the ligRNA itself. Taking together these results and the successful cDNA amplification presented in Figure 4, we can confirm that it is possible to add a constant region to the synthetic RNA pool and that the RNA library can hence be directly screened and amplified for different applications, in particular for aptamer selection.

4.2. Conferring higher nucleases resistance to artificial oligonucleotides

It is now admitted that the development of both DNA and RNA aptamers has several advantages compared to antibodies, as they can be inexpensive and rapidly selected. However, their application in biological media is strongly limited by the sensitivity of natural oligonucleotides to nuclease degradation [17]. One of the alternatives that can be applied to avoid this unwanted effect is the chemical modification of these molecules using nucleotide analogs. In this work, we have studied the ability of two DNA polymerase θ mutants to form long polymers of 2'-Fluoro and 2'-OMe modified nucleotides. Both compounds enable better nucleases resistance to nucleic acids [19].

The Figure 5 shows the improved template-free terminal transferase activity of polθ-CS13 mutant in incorporating 2'-Fluoro-dNTPs compared to the wild-type polymerase. Polθ-CS13 demonstrated a high efficiency to form long polymers, regardless of the type of nucleotides (2'-Fluoro-dATP, -dTTP, -dCTP, -dGTP and -dUTP alone or all mixed). Conversely, this mutant failed to incorporate 2'-OMe modified nucleotides while polθ-DW9 did accept 2'-OMe modified nucleotides (Figure 6), and especially generated long polymers when the four 2'-OMe-NTPs were mixed together (Figure 6, lane 5). The same mutant succeeded as well to incorporate 2'-Fluoro-modified nucleotides (data not shown).

5. Biotechnological applications and futures directions

Human DNA pol θ mutants demonstrated a reliable ability to perform RNA random synthesis and to incorporate a large panel of modified nucleotides[20], [21], [24]. The randomness of its nucleotidyltransferase activity has been assessed quantitatively and this new enzymatic method to produce long RNA polymers should be useful for therapeutics applications. In particular, the continuous quest for ultra-effective, selective and non-toxic nucleic acids-based drugs made of aptamers may become more

effective [21]. Indeed, this work offers a viable biological alternative to the initial generation of RNA random sequences by chemical synthesis and library design. By establishing an efficient enzymatic SELEX procedure, the assay costs will be reduced at the same time as the duration of the selection cycle, as a 1-to-5min reaction is sufficient to generate a 40 nucleotides-length RNA oligos. The next step to this method would be the capability to realize the full RNA aptamer selection in RNA space and avoiding the enrichment step by RT-PCR. This would require the direct amplification of RNA into RNA by a (thermostable) RNA-dependent RNA polymerase.

We also showed that the polθ-DW9 mutant was able to produce a randomized pool of long 2'-OMe polymers that could be used for selection of 2'-OMe aptamers. In the same logic as this work, Liu et al.[25] managed to evolve two thermostable DNA polymerases SFM4-6 and SFM4-9 which enable the transcription and reverse transcription of 2'-OMe oligonucleotides, that lead them to announce, for the first time, that the selection of fully 2'-OMe modified aptamers is now a reality. With a rapid method to generate large pools of 2'-OMe fragments and a robust mean to select and amplify 2'-OMe aptamers, important progresses may potentially arise and result in the discovery of functional molecules with higher nuclease resistance and better target affinity[19], [25].

The possibility to rapidly synthesize RNAs broadens the scope of RNA biology. In addition, it allows to imagine that large libraries of messenger RNAs (mRNAs) could be used for generating libraries of proteins of random sequences by directly translating them in a cell-free protein synthesis (CFPS) system. Such biotechnological application could greatly facilitate the development of therapeutic molecules and new biocatalysts.

6. Fundings

This work was funded by Action Incitative Concertée of Institut Pasteur granted to M.D. The authors declare that they have no conflict of interest. A provisional patent application no. 62/560693, titled DNA POLYMERASE THETA MUTANTS, THE METHODS OF PRODUCING THESE MUTANTS, AND THEIRS USES, has been filed and deposited to the UPSTO on 20th of September, 2017.

7. Acknowledgements

We would like to thank Sandrine Rosario, Institut Pasteur, for her help with the modified nucleotides incorporation experiments (Figures 5 and 6), and Dr. Michel Ryckelynck from IBMC of CNRS and Strasbourg University for his advices in the SELEX library preparation.

References

[1] G. Mayer, "The Chemical Biology of Aptamers," *Angewandte Chemie International Edition*, vol. 48, no. 15, pp. 2672–2689, Mar. 2009.

[2] M. Famulok, J. S. Hartig, and G. Mayer, "Functional Aptamers and Aptazymes in Biotechnology, Diagnostics, and Therapy," 2007.

[3] J. Zhou and J. Rossi, "Aptamers as targeted therapeutics: current potential and challenges," *Nat Rev Drug Discov*, vol. 16, no. 6, p. 440, Jun. 2017.

[4] D. H. Bunka, O. Platonova, and P. G. Stockley, "Development of aptamer therapeutics," *Current Opinion in Pharmacology*, vol. 10, no. 5, pp. 557–562, Oct. 2010.

[5] N. G. Walter and D. R. Engelke, "Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs.," *Biologist (London, England)*, vol. 49, no. 5, pp. 199–203, Oct. 2002.

[6] S. L. DeVos and T. M. Miller, "Antisense oligonucleotides: treating neurodegeneration at the level of RNA.," *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics*, vol. 10, no. 3, pp. 486–97, Jul. 2013.

[7] S. T. Crooke, *Antisense drug technology : principles, strategies, and applications*. CRC Press, 2008.

[8] D. Campoccia, L. Montanaro, and C. R. Arciola, "A review of the biomaterials technologies for infection-resistant surfaces.," *Biomaterials*, vol. 34, no. 34, pp. 8533–54, Nov. 2013.

[9] S. Shukla, C. S. Sumaria, and P. I. Pradeepkumar, "Exploring Chemical Modifications for siRNA Therapeutics: A Structural and Functional Outlook," *ChemMedChem*, vol. 5, pp. 328–349, 2010.

[10] A. Autour, E. Westhof, and M. Ryckelynck, "iSpinach: a fluorogenic RNA aptamer optimized for in vitro applications," *Nucleic Acids Research*, vol. 44, no. 6, pp. 2491–2500, 2016.

[11] C. T. Farrar, C. M. William, E. Hudry, T. Hashimoto, and B. T. Hyman, "RNA Aptamer Probes as Optical Imaging Agents for the Detection of Amyloid Plaques," *PLoS ONE*, vol. 9, no. 2, 2014.

[12] R. J. Trachman, A. Abdolahzadeh, A. Andreoni, R. Cojocaru, J. R. Knutson, M. Ryckelynck, P. J. Unrau, and A. R. Ferré-D'Amaré, "Crystal Structures of the Mango-II RNA Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness," *Biochemistry*, vol. 57, no. 26, pp. 3544–3548, Jul. 2018.

[13] A. Autour, S. C. Y. Jeng, A. D. Cawte, A. Abdolahzadeh, A. Galli, S. S. S. Panchapakesan, D. Rueda, M. Ryckelynck, and P. J. Unrau, "Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells," *Nature Communications*, vol. 9, no. 1, p. 656, Dec. 2018.

[14] P. Fernandez-Millan, A. Autour, E. Ennifar, E. Westhof, and M. Ryckelynck, "Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI.," *RNA (New York, N.Y.)*, vol. 23, no. 12, pp. 1788–1795, Dec. 2017.

[15] Y. Morita, M. Leslie, H. Kameyama, D. E. Volk, and T. Tanaka, "Aptamer Therapeutics in Cancer: Current and Future.," *Cancers*, vol. 10, no. 3, Mar. 2018.

[16] K. D. Kovacevic, J. C. Gilbert, and B. Jilma, "Pharmacokinetics, pharmacodynamics and safety of aptamers," *Advanced Drug Delivery Reviews*, vol. 134, pp. 36–50, Sep. 2018.

[17] A. V Lakhin, V. Z. Tarantul, and L. V Gening, "Aptamers: problems, solutions and prospects.," *Acta naturae*, vol. 5, no. 4, pp. 34–43, Oct. 2013.

[18] J. Zhou, M. L. Bobbin, J. C. Burnett, and J. J. Rossi, "Current progress of RNA aptamer-based

therapeutics.," *Frontiers in genetics*, vol. 3, p. 234, 2012.

[19] C. Kratschmer and M. Levy, "Effect of Chemical Modifications on Aptamer Stability in Serum," *Nucleic Acid Therapeutics*, vol. 27, no. 6, pp. 335–344, Dec. 2017.

[20] T. Kent, P. A. Mateos-Gomez, A. Sfeir, and R. T. Pomerantz, "Polymerase θ is a robust terminal transferase that oscillates between three different mechanisms during end-joining," *eLife*, vol. 5, no. JUN2016, pp. 16203–16208, Jun. 2016.

[21] I. Randrianjatovo-Gbalou, S. Rosario, O. Sismeiro, H. Varet, R. Legendre, J.-Y. Coppée, V. Huteau, S. Pochet, and M. Delarue, "Enzymatic synthesis of random sequences of RNA and RNA analogues by DNA polymerase theta mutants for the generation of aptamer libraries," *Nucleic Acids Research*, vol. 46, no. 12, pp. 6271–6284, Jul. 2018.

[22] M. Delarue, O. Poch, N. Tordo, D. Moras, and P. Argos, "An attempt to unify the structure of polymerases," *"Protein Engineering, Design and Selection,"* vol. 3, no. 6, pp. 461–467, 1990.

[23] K. E. Zahn, A. M. Averill, P. Aller, R. D. Wood, and S. Doublié, "Human DNA polymerase θ grasps the primer terminus to mediate DNA repair," *Nature Structural & Molecular Biology*, vol. 22, no. 4, pp. 304–311, Mar. 2015.

[24] T. Kent, T. D. Rusanov, T. M. Hoang, W. A. Velema, A. T. Krueger, W. C. Copeland, E. T. Kool, and R. T. Pomerantz, "DNA polymerase θ specializes in incorporating synthetic expanded-size (xDNA) nucleotides.," *Nucleic acids research*, vol. 44, no. 19, pp. 9381–9392, Nov. 2016.

[25] Z. Liu, T. Chen, and F. E. Romesberg, "Evolved polymerases facilitate selection of fully 2'-OMe-modified aptamers.," *Chemical science*, vol. 8, no. 12, pp. 8179–8182, Dec. 2017.

Figures captions

Figure 1. Structure of the catalytic site of Human DNA Pol θ. Ribbon representation of the C-terminal catalytic domain of the Human DNA polymerase θ, showing the binding of the incoming ddGTP in the active site (PDB id. 4X0Q). The three strictly conserved carboxylates (D2330, D2540 and E2541) are displayed as sticks: they coordinate the Mg2+cation (green) and are facing the incoming nucleotide. The E2335, L2334, Y2387 and Q2384 residues are also displayed as sticks to highlight that they are possible targets to generate mutants that would accept ribonucleotides. The same residues are located in the three motifs A, B and C that are respectively distinguished with the following color code : yellow (D2330, L2334, E2335), pink (Q2384, Y2387) and green (D2540, E2541).

Figure 2. Enzymatic RNA synthesis and library generation procedures. This diagram summarizes the main steps of the preparation of RNA library for SELEX.

Figure 3. Dual-color visualization of the catalytically synthesized of RNA pool by pol θ-DW9 mutant and the subsequent ligation of a constant region to the 3'-end using T4 RNA ligase. Time-course of the elongation of a 15-mer ssRNA primer by DW9 mutant in the presence of a stoichiometric mix of the four NTPs (0.5 mM each) and separated in a denaturing 8% acrylamide gel. At each indicated time (lanes 2 to 6), the reaction was stopped by the addition of formamide blue. (AlexaFluor488) The green fluorescence signal reveals the extent of the elongation of all RNA fragments containing the 5'AlexaFluor-labelled RNA primer. (CY5) The red fluorescence signal reveals the fragments that have been ligated with the constant region at 3'-end by T4 RNA ligase. (Merged) The green and red signals have been merged to evaluate the quality of both reactions (elongation + ligation).

Figure 4. Amplification of cDNA after reverse transcription of the enzymatically synthesized RNAs. Four samples were analyzed on a 3% agarose gel after amplification of RNA libraries obtained after 5 min-synthesis (duplicates 1 and 2) and 1 min-synthesis (duplicates 3 and 4). NC: reverse transcription negative control where all the reagents were present except the MAXIMA reverse transcriptase. NT: No template control where the RT mix has been prepared without the RNA template.

Figure 5. Nucleotidyltransferase activity of human pol θ in presence of 2'-Fluoro nucleotides. Denaturing 15% acrylamide gel showing the efficiency of the pol θ-CS13 in forming long homo- and heteropolymers of 2'-fluoro-nucleotides compared to pol θ-WT. A ssDNA of 14-mer has been used as a primer and elongated for 30 min at 42°C.

Figure 6. Nucleotidyltransferase activity of human pol θ in presence of 2'-O-Methyl nucleotides. Denaturing 15% acrylamide gel showing the efficiency of the pol θ-DW9 in forming long homo- and heteropolymers of 2'-O-Me-nucleotides compared to pol θ-WT. A ssDNA of 14-mer has been used as a

primer and elongated for 30 min at 42°C.

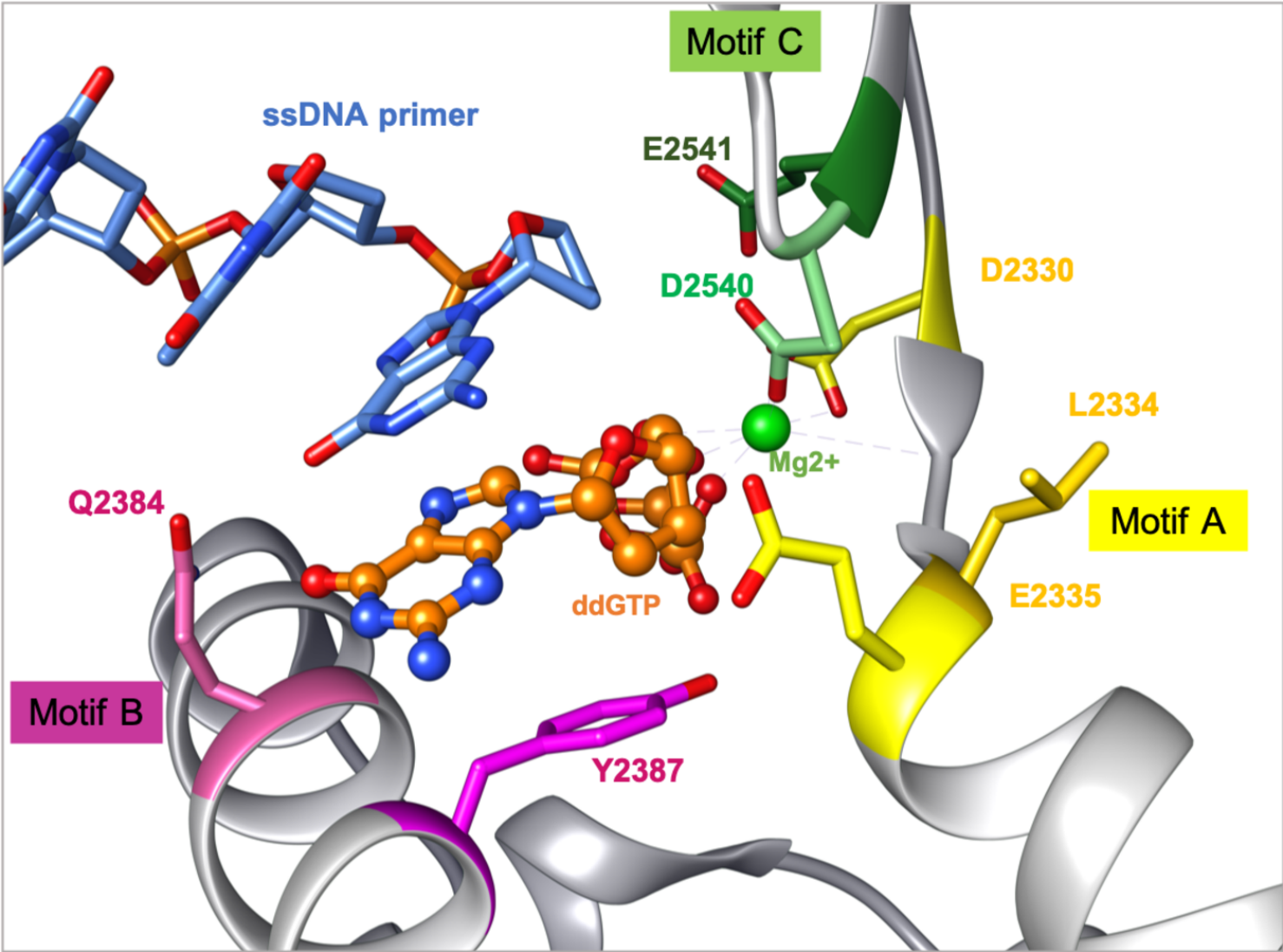
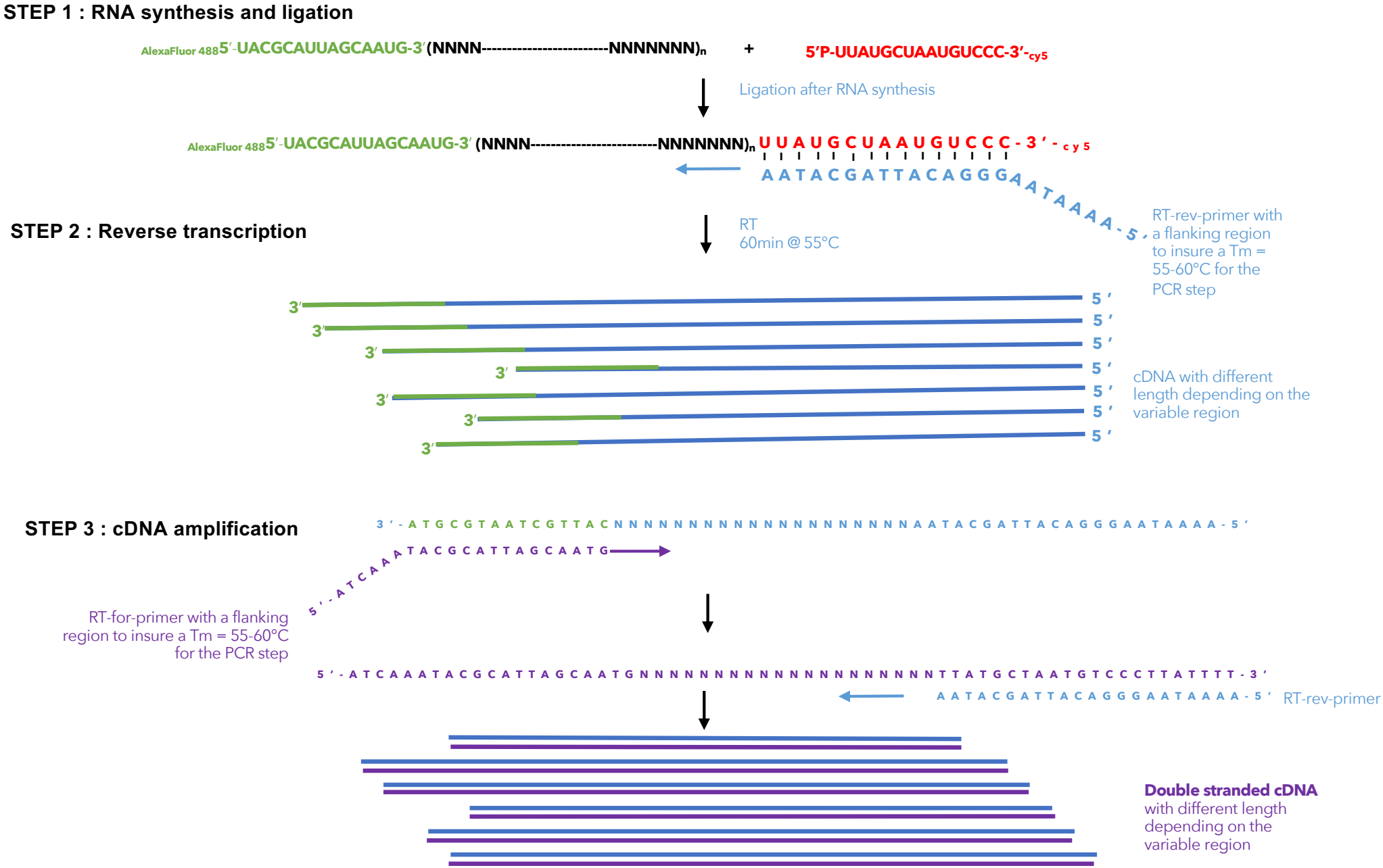
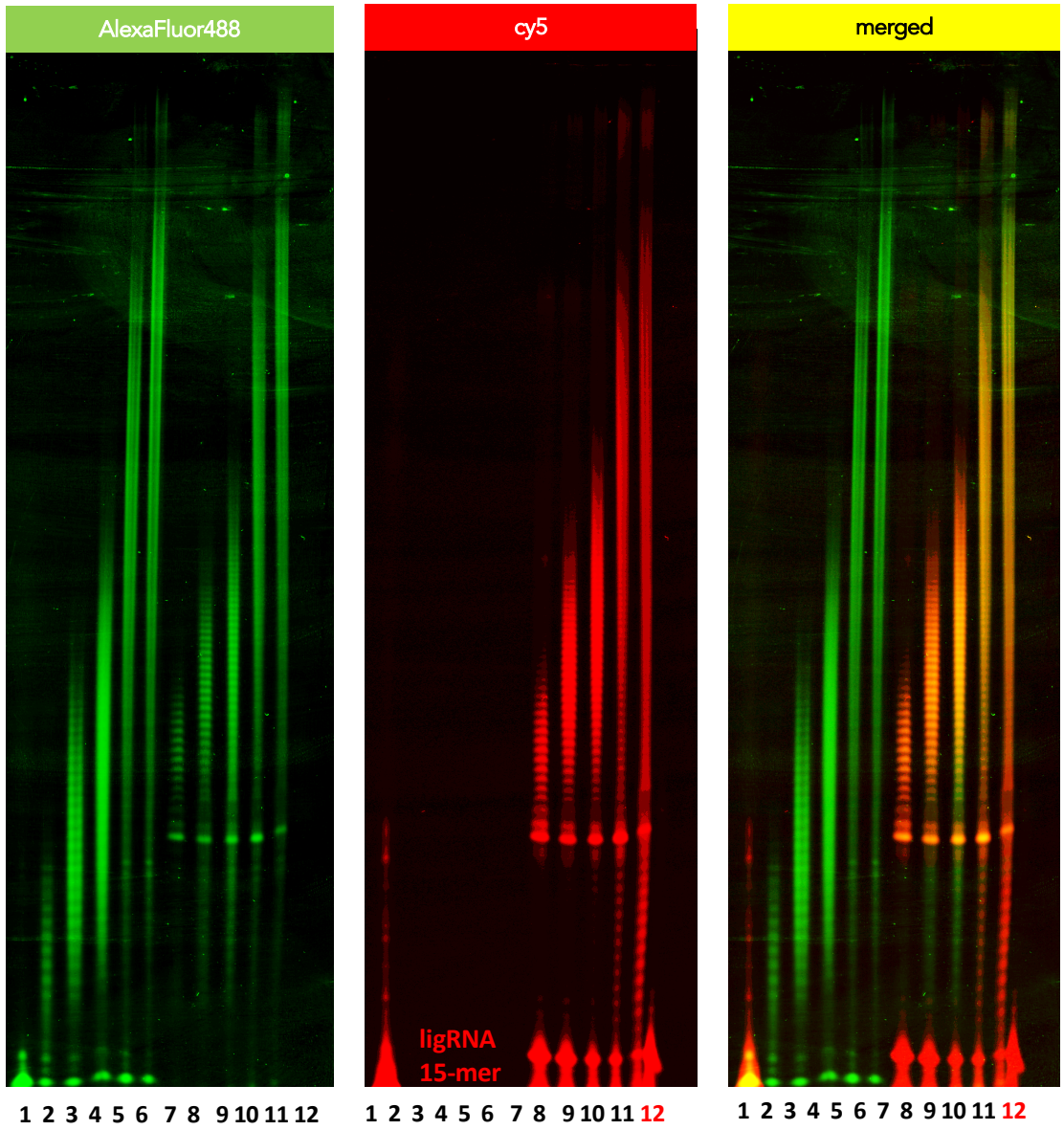

Motif C
ssDNA primer
E2541
D2540
D2330
L2334
Mg2+
Motif A
Q2384
ddGTP
E2335
Motif B
Y2387

Figure 2.


STEP 1 : RNA synthesis and ligation
AlexaFluor 488 5′-UACGCAUUAGCAAUG-3′ (NNNN----------NNNNNNN)n + 5′P-UUAUGCUAAUGUCCC-3′-cy5
Ligation after RNA synthesis
AlexaFluor 488 5′-UACGCAUUAGCAAUG-3′ (NNNN----------NNNNNNN)n U U A U G C U A A U G U C C C - 3 ′ - c y 5
A A T A C G A T T A C A G G G A A T A A A A - 5 ′
STEP 2 : Reverse transcription
RT
60min @ 55°C
RT-rev-primer with a flanking region to insure a Tm = 55-60°C for the PCR step
3′
5′
cDNA with different length depending on the variable region
STEP 3 : cDNA amplification
3 ′ - A T G C G T A A T C G T T A C N N N N N N N N N N N N N N N N N N N N N N N N N N A A T A C G A T T A C A G G G A A T A A A A - 5 ′
5 ′ - A T C A A A T A C G C A T T A G C A A T G
RT-for-primer with a flanking region to insure a Tm = 55-60°C for the PCR step
5 ′ - A T C A A A T A C G C A T T A G C A A T G N N N N N N N N N N N N N N N N N N N N N N N N N N N T T A T G C T A A T G T C C C T T A T T T T - 3 ′
A A T A C G A T T A C A G G G A A T A A A A - 5 ′
RT-rev-primer
Double stranded cDNA with different length depending on the variable region

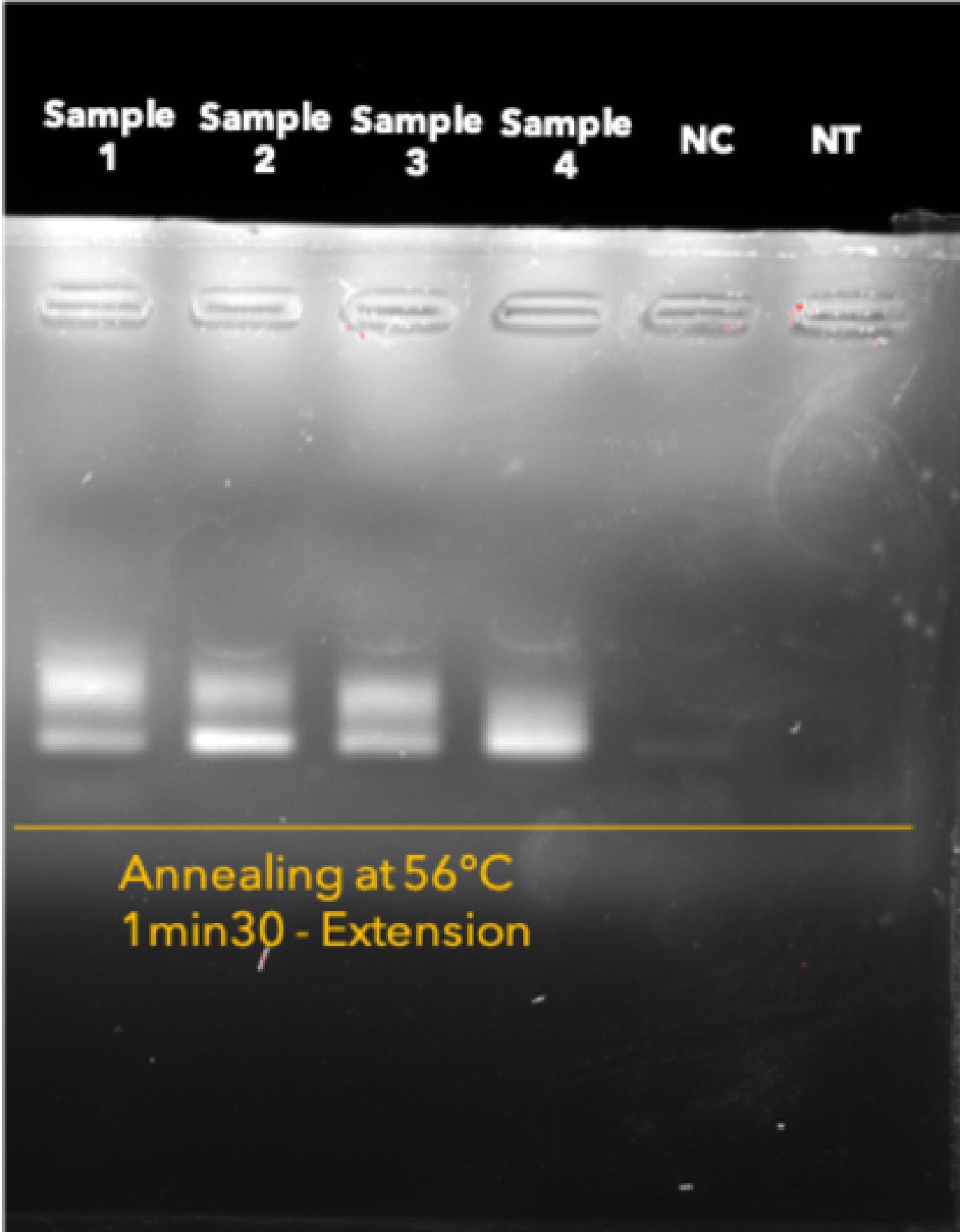


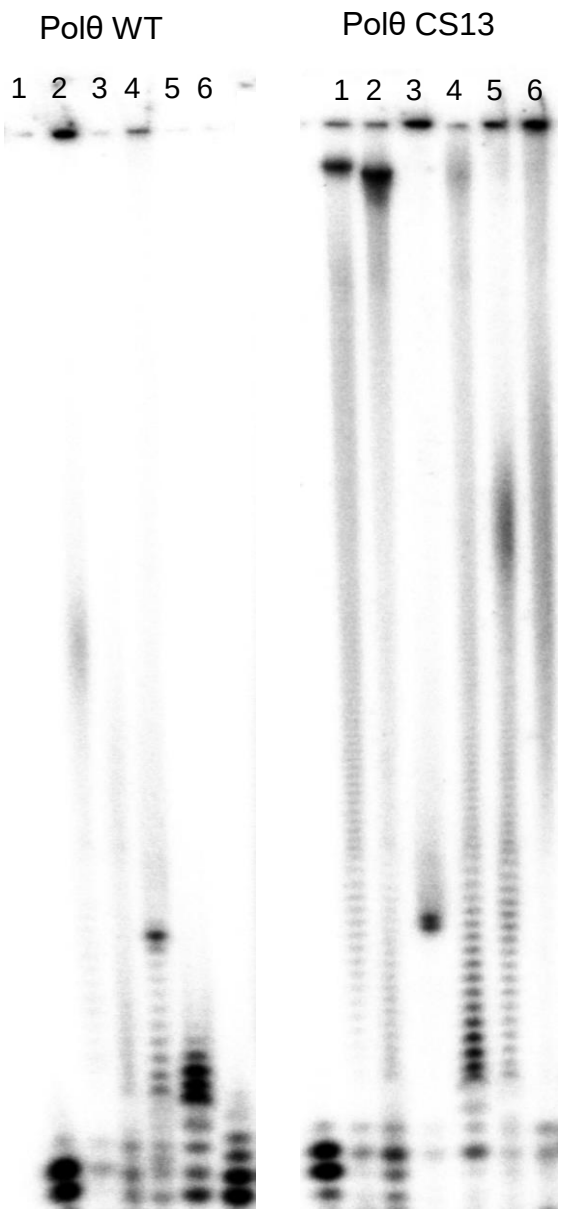
1. Control : RNA primer + ligRNA w/out enzyme

2. Elongation control after 5 sec
3. Elongation control after 30 sec
4. Elongation control after 1 min
5. Elongation control after 5 min
6. Elongation control after 15 min
7. Ligation after 5 sec-elongation (+15)
8. Ligation after 30 sec-elongation (+15)
9. Ligation after 1 min-elongation (+15)
10. Ligation after 5 min-elongation (+15)
11. Ligation after 15 min-elongation (+15)

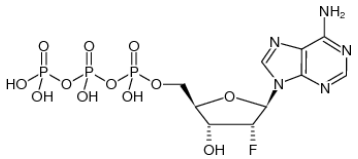
12. Auto-ligation control

AlexaFluor 488 5'-UACGCAUUAGCAAUG-3' (NNNN----------NNNNNNN)$_n$ 5'P-UUAUGCUAAUGUCCC-3'-$_{cy5}$

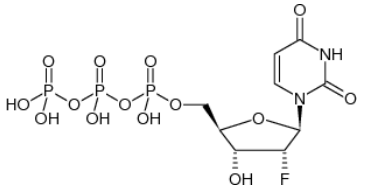

Sample 1
Sample 2
Sample 3
Sample 4
NC
NT
Annealing at 56°C
1min30 - Extension



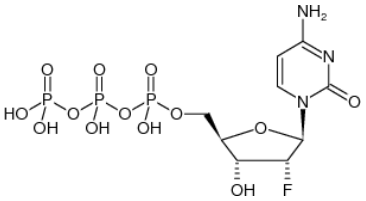
2’-fluoro-dATP



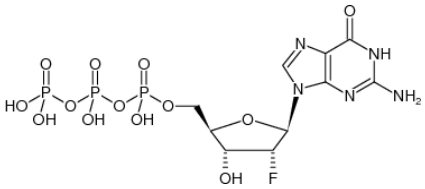
2’-fluoro-dUTP



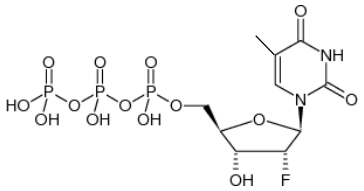
2’-fluoro-dCTP



2’-fluoro-dGTP

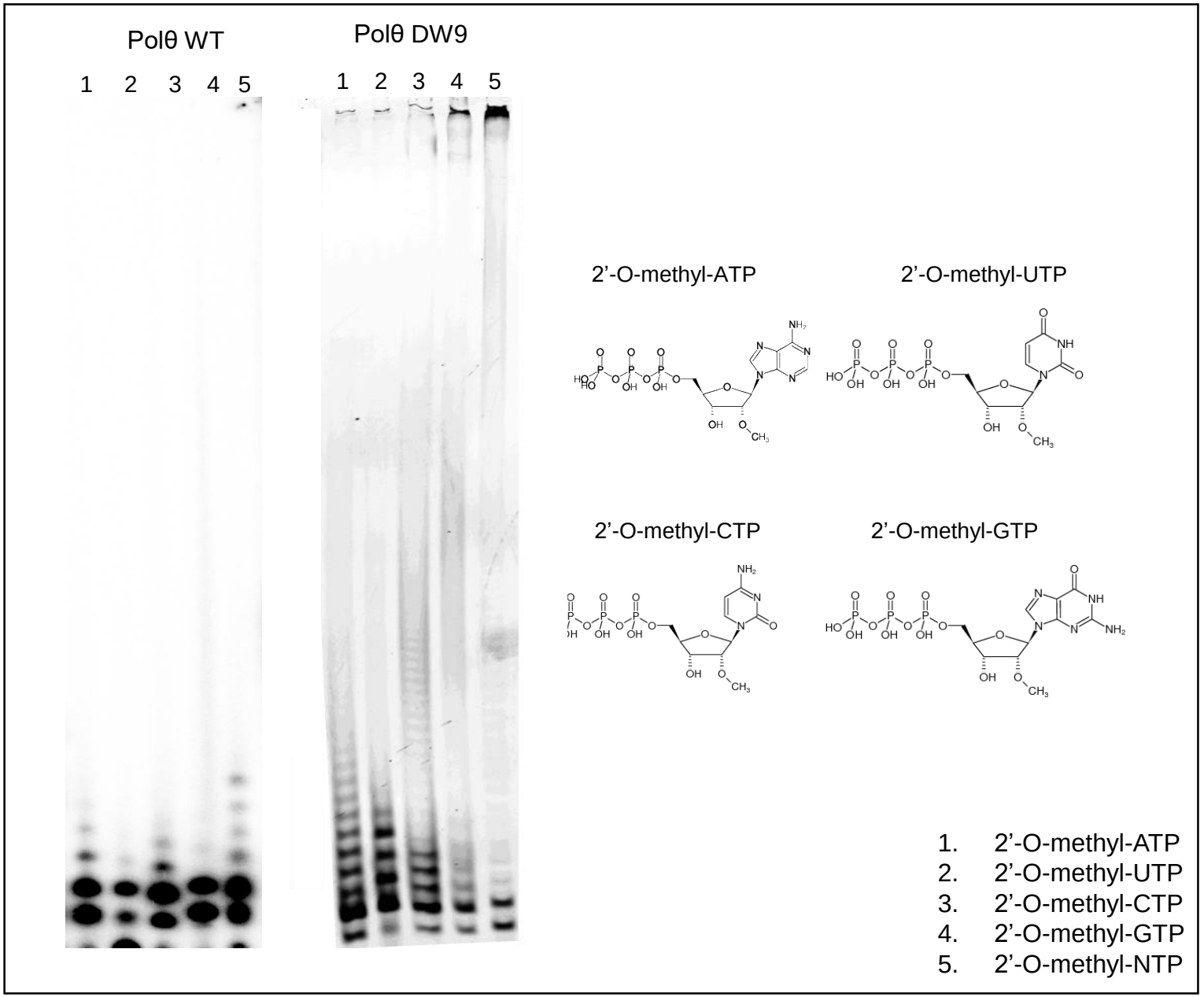


2’-fluoro-dTTP

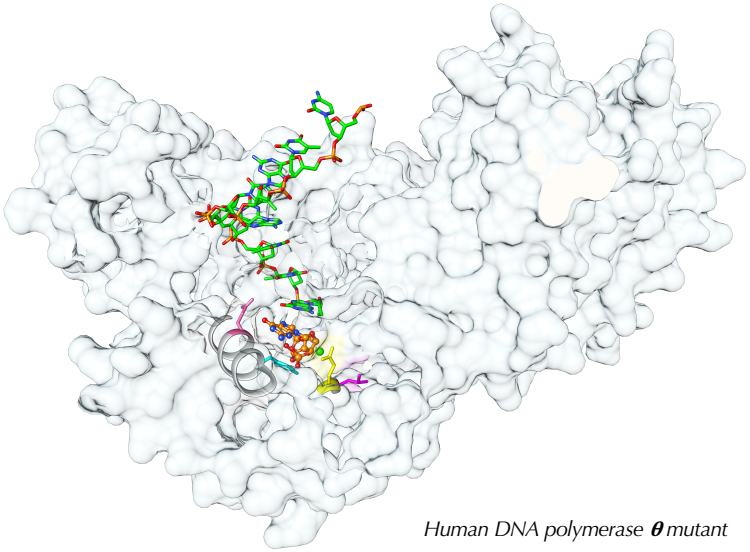


1. 2’-Fluoro-dATP
2. 2’-Fluoro-dGTP
3. 2’-Fluoro-dCTP
4. 2’-Fluoro-dTTP
5. 2’-Fluoro-dUTP
6. 2’-Fluoro-dNTP

Figure 6.


Polθ WT
1 2 3 4 5
Polθ DW9
1 2 3 4 5
2’-O-methyl-ATP
2’-O-methyl-UTP
2’-O-methyl-CTP
2’-O-methyl-GTP
1. 2’-O-methyl-ATP
2. 2’-O-methyl-UTP
3. 2’-O-methyl-CTP
4. 2’-O-methyl-GTP
5. 2’-O-methyl-NTP

Rational design of polymerases

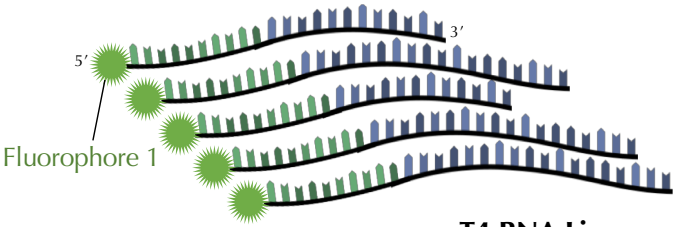


*Human DNA polymerase **θ** mutant*

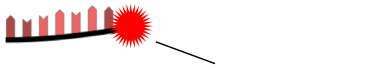
Artificial RNAs synthesis

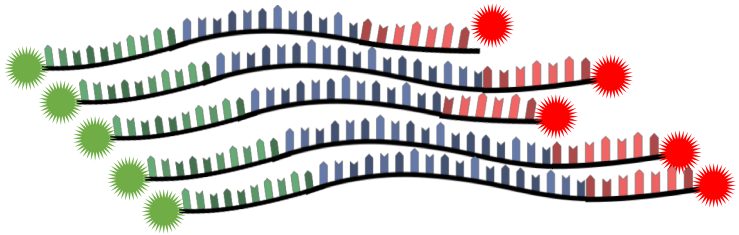
Fixed region
DNA or RNA

40-200 N random sequence

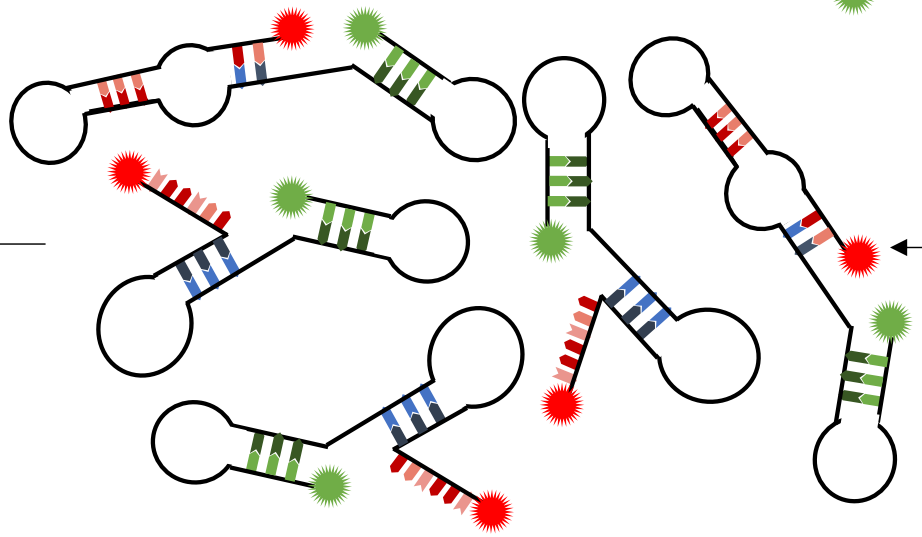


T4 RNA Ligase
+





3D folding



Biotechnological applications

1- Rapid Aptamer selection
2- HTP Cell-free expression system
3- RNA analogs library
4- Artificial mRNA synthesis