

Evolving Thermostable Reverse Transcriptase Activity in a DNA Polymerase Scaffold**

Katharina B. M. Sauter and Andreas Marx*

DNA polymerases are involved in all DNA synthesis that occurs in nature.^[1] Furthermore, DNA polymerases are the workhorses in numerous important molecular biological core technologies like the widely applied polymerase chain reaction (PCR), cDNA cloning, genome sequencing, and diagnostics based on nucleic acids.^[2] In each of these processes the DNA polymerase requires a DNA strand as a template for directing DNA synthesis. In contrast, RNA templates are processed with marginal activity.^[1] Nevertheless, it has been reported that some DNA polymerases like *Thermus thermophilus* (*Tth*) exhibit increased reverse transcriptase (RT) activity exclusively in the presence of Mn^{2+} ions.^[3] However, for many applications in mRNA diagnosis, for example, in pathogen detection or gene expression analysis, employment of Mn^{2+} is not suited.^[4] Thus, we asked whether one is able to generate a DNA polymerase that is proficient to first reverse transcribe RNA into DNA and subsequently amplify the DNA by PCR without the requirement of Mn^{2+} ions. Through this approach, new insights into DNA polymerase substrate recognition and valuable tools for applications should be gained.

Methods of directed evolution have been shown to be promising to engineer nucleic acid polymerases with altered properties.^[5] Alterations have mainly been achieved by directed molecular evolution by using genetic complementation and/or screening,^[6] phage display,^[7] or in vitro compartmentalization.^[8] Herein, we show that the generation of a new DNA polymerase function, that is, RT PCR activity, is achievable through iterative screening of small libraries of DNA polymerases. Our example is one of the rare examples where using an arbitrary randomized (nonfocused) library was sufficient to identify a new DNA polymerase activity.

To evolve reverse transcriptase activity in an N terminal, shortened form of DNA polymerase from *Thermus aquaticus* (*Klentaq*)^[9] we randomized the open reading frame by error prone PCR (epPCR).^[10] Protein expression was conducted in 96 well plates, and enzymes were screened for PCR activity

[*] Dipl. Chem. K. B. M. Sauter, Prof. Dr. A. Marx
Fachbereich Chemie
Universität Konstanz
Universitätsstrasse 10, M 726, 78457 Konstanz (Germany)
Fax: (+49) 7531 88 5140
E mail: andreas.marx@uni-konstanz.de

[**] We gratefully acknowledge financial support by the DFG and Volkswagen Foundation. Samuel Weisbrod is gratefully acknowledged for his assistance in the preparation of the manuscript.

on DNA templates directly after heat denaturation of host proteins and lysis without the requirement of further purification steps. All procedures were performed by automated liquid handling, thus enabling high throughput if required. DNA polymerase activities were monitored using SYBR green I for parallel quantification of PCR product formation and fluorescent readout in 384 well plates. Subsequently, the active variants were tested for their RT PCR activity by employing a natural RNA target from bacteriophage MS2. RT PCR activities were quantified with SYBRgreen I by melting curve analysis of PCR product using a conventional real time PCR cyclor in 96 well plates.

Already after screening 768 PCR active variants derived from about 2000 initially collected clones, we identified a mutant (termed M1) that shows promising features in RT PCR and was therefore further studied. M1 was subsequently purified to homogeneity (see the Supporting Information) for functional characterization, and we found that this mutant exhibits strikingly increased RT PCR activity in comparison to the wild type enzyme (Figure 1a). While the parental

the desired product exclusively (see the Supporting Information).

Next, we conducted a second round of randomization by employing epPCR, parallel expression, and screening as described above. After screening of approximately 2000 PCR active mutants we identified one variant (termed M2) that was purified and further analyzed. M2 has RT PCR activity and yields the desired PCR product (Figure 1a).

Sequencing of the variant M1 revealed that six mutations have occurred that are distributed throughout the enzyme: L322M, L459M, S515R, I638F, S739G, and E773G (Figure 2a). M2 bears only one additional mutation compared to the mutations already present in M1 (L789F).

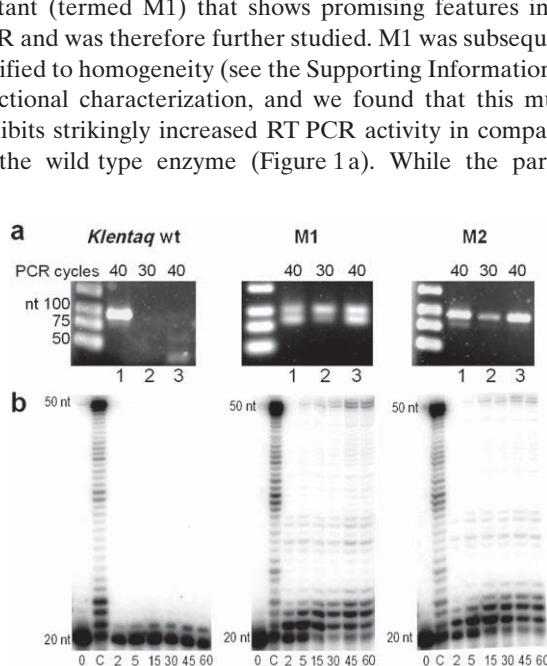


Figure 1. Analysis of the selected variants in comparison to the wild type. a) PCR activity of the respective enzyme using either a DNA target (lanes 1) or RNA target (lanes 2, 3). All reactions were conducted under identical reaction conditions (nt = nucleotides). b) Reverse transcriptase activity conducted with the indicated enzyme for the indicated reaction times [min] under identical reaction conditions. C = control reaction conducted with the corresponding DNA template.

enzyme was unable to form any significant PCR product from RNA after 40 cycles, under identical conditions M1 is clearly competent. All reactions were conducted by employing originally described PCR buffer and reaction conditions that were used for *Klentaq* and enable PCR product formation when DNA amplification is performed using a DNA target.^[6c,9] Primer strands were designed to anneal at the 3' terminus of the RNA target, and their design was not further optimized. However, the enzyme yields two products that are distinguishable through analysis by gel electrophoresis. The major product has the desired length and the expected sequence. However, this feature seems to be sequence dependent, since when using RNA templates from different origin, for example, *E. coli* S16 mRNA, M1 yields

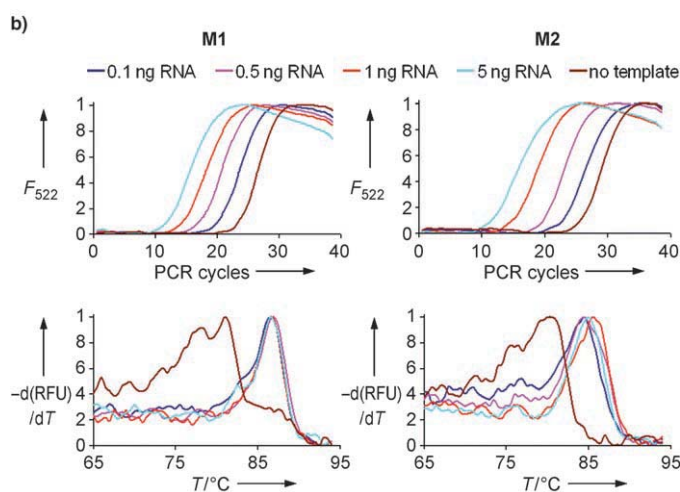
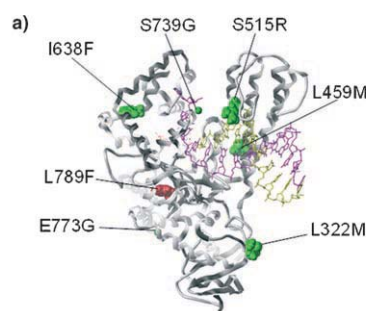


Figure 2. a) Mutations in the evolved RT PCR active DNA polymerases M1 (green) and M2 (green and red) are mapped on a ribbon representation of *Klentaq* (PDB code 1QSS).^[11] The DNA template strand is shown in magenta, the primer strand in yellow. b) Upper: Results of real time RT PCR experiments using M1 and M2 with various amounts of RNA employing conventional real time PCR equipment. Lower: Melting curve analysis of the reaction products employing conventional real time PCR equipment.

To further investigate the properties of M1 and M2 in comparison to the wild type enzyme, we conducted primer extension experiments (Figure 1b). A DNA primer strand was annealed to its complementary site on a 50 nucleotide RNA template strand. When the primer/template complex was incubated with the wild type *Klentaq*, the extension of the DNA primer by only one nucleotide was detected after gel analysis. Interestingly, similar low activity of *Tth* DNA polymerase in the absence of Mn^{2+} is observed when copying

the RNA template (Supporting Information). However, both mutants M1 and M2 reverse transcribe RNA significantly more efficiently. Quantification of the respective activities under steady state conditions indicates that the M1 and M2 mutants are about 40 and 22 fold, respectively, more active on the RNA template than the wild type enzyme (Supporting Information). In marked contrast, the variants M1 and M2 are able to yield significantly longer reaction products that are sufficient for RT PCR under identical conditions. However, quantification of the activity on DNA directed primer extension indicates that both mutants have about twice the activity of the wild type enzyme (Supporting Information). Thus, a clear activity gain for RNA directed DNA synthesis was achieved by two rounds of iterative randomization and screening. Further assessment of the mutants in comparison to the wild type enzyme indicates that this gain was not accompanied by a loss in thermostability, but by a loss of some selectivity (Supporting Information).

The specific detection of RNA through reverse transcription and subsequent cDNA amplification is fundamental for many applications in diagnosis, for example, pathogen detection or gene expression analysis.^[4] To amplify RNA by RT PCR, usually two enzymes have to be employed: a reverse transcriptase that synthesizes the complementary DNA strand and a thermostable DNA polymerase for PCR amplification. However, the first step is prone to failure owing to formation of stable RNA secondary structures and the diminished thermostability of the employed reverse transcriptases.^[3] Thermostable M1 and M2 might be particularly interesting for a one step RT PCR approach. Thus, we investigated whether the mutants are applicable in quantitative detection of RNA by real time one step RT PCR.^[12] This method is extensively used for pathogen and gene expression diagnosis.^[4] We followed the reactions by employing varied RNA concentrations in a conventional real time PCR cycler with SYBRgreenI for detection of double stranded DNA. Indeed, M1 and M2 are able to amplify DNA starting from RNA targets at various concentrations under standard conditions (Figure 2b). The observed RNA detection limit using the evolved enzyme is comparable to the performance of commercially available kits. When the wild type enzyme was used, no such amplification curves were detected.

In summary, through screening of DNA polymerase libraries that contain arbitrary randomized mutants generated by epPCR, we were able to identify enzymes that exhibit RT PCR function, a function that is imperceptible in the wild type enzyme. As demonstrated, the identified mutants might find immediate applications and provide the basis for the development of new means for single step RT PCR technologies like pathogen RNA detection or gene expression analysis in real time. Functional and structural studies to elucidate the origin of the new function are underway.

- [1] A. Kornberg, T. A. Baker, *DNA Replication*, 2nd ed., Freeman, New York, **1991**.
- [2] J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2001**.
- [3] a) T. W. Myers, D. H. Gelfand, *Biochemistry* **1991**, *30*, 182–192; b) M. D. Jones, *Methods Enzymol.* **1993**, *218*, 413–419; c) V. I. Grabko, L. G. Chistyakova, V. N. Lyapustin, V. G. Korobko, A. I. Miroshnikov, *FEBS Lett.* **1996**, *387*, 189–192.
- [4] a) S. A. Bustin, R. Mueller, *Clin. Sci.* **2005**, *109*, 365–379; b) A. K. Sandvik, B. K. Alsberg, K. G. Nørsett, F. Yadetie, H. L. Waldum, A. Lægreid, *Clin. Chim. Acta* **2006**, *363*, 157–164; c) S. Mocellin, C. R. Rossi, P. Pilati, D. Nitti, F. M. Marincola, *Trends Mol. Med.* **2003**, *9*, 189–195.
- [5] a) A. A. Henry, F. E. Romesberg, *Curr. Opin. Biotechnol.* **2005**, *16*, 370–377; b) S. Brakmann, *Cell. Mol. Life Sci.* **2005**, *62*, 2634–2646; c) E. Loh, L. A. Loeb, *DNA Repair* **2005**, *4*, 1390–1398.
- [6] a) P. H. Patel, L. A. Loeb, *J. Biol. Chem.* **2000**, *275*, 40266–40272; b) P. H. Patel, H. Kawate, E. Adman, M. Ashbach, L. A. Loeb, *J. Biol. Chem.* **2001**, *276*, 5044–5051; c) S. Brakmann, S. Grzeszik, *ChemBioChem* **2001**, *2*, 212–219; d) M. B. Kermekchiev, A. Tzekov, W. M. Barnes, *Nucleic Acids Res.* **2003**, *31*, 6139–6147; e) D. Summerer, N. Z. Rudinger, I. Detmer, A. Marx, *Angew. Chem.* **2005**, *117*, 4791–4794; *Angew. Chem. Int. Ed.* **2005**, *44*, 4712–4715.
- [7] a) M. Fa, A. Radeghieri, A. A. Henry, F. E. Romesberg, *J. Am. Chem. Soc.* **2004**, *126*, 1748–1754; b) A. M. Leconte, L. Chen, F. E. Romesberg, *J. Am. Chem. Soc.* **2005**, *127*, 12470–12471; c) G. Xia, L. Chen, T. Sera, M. Fa, P. G. Schultz, F. E. Romesberg, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6597–6602.
- [8] a) F. J. Ghadessy, J. L. Ong, P. Holliger, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4552–4557; b) F. J. Ghadessy, N. Ramsay, F. Boudsocq, D. Loakes, A. Brown, S. Iwai, A. Vaisman, R. Woodgate, P. Holliger, *Nat. Biotechnol.* **2004**, *22*, 755–759.
- [9] W. M. Barnes, *Gene* **1992**, *112*, 29–35.
- [10] R. C. Cadwell, G. F. Joyce, *PCR Methods Appl.* **1992**, *2*, 28–33.
- [11] Y. Li, V. Mitaxov, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9491–9496.
- [12] J. Wilhelm, A. Pingoud, *ChemBioChem* **2003**, *4*, 1120–1128.