Directed Evolution

DOI: 10.1002/ange.200601217

A Population of Thermostable Reverse Transcriptases Evolved from *Thermus aquaticus* DNA Polymerase I by Phage Display**

*Sophie Vichier-Guerre, Stéphane Ferris, Nicolas Auberger, Karim Mahiddine, and Jean-Luc Jestin**

Nucleic acid replicases are unique catalysts as they do not only catalyze conversion of substrates into products, but also amplify the substrate. For DNA polymerase to be an active enzyme that catalyses phosphodiester bond formation, magnesium dications are necessary. Substitution by manganese

[*] Dr. S. Vichier-Guerre, S. Ferris, N. Auberger, K. Mahiddine, Dr. J.-L. Jestin
Unité de Chimie Organique URA 2128 CNRS
Département de Biologie Structurale et Chimie
Institut Pasteur
28 rue du Dr. Roux, 75724 Paris 15 (France)
Fax: (+33) 145-688-404
E-mail: jjestin@pasteur.fr

[**] We thank F. Bahrami, M. Delarue, P. A. Kaminski, and S. Wain-Hobson for critical comments on the manuscript, T. Huynh-Dinh and P. E. Bost for advice, and C. Gouyette, O. Helynck, and V. Huteau for technical help. This work was supported by the Ministère de la Recherche (ACI blanche), the CNRS (Programme Physique et Chimie du Vivant), and the Pasteur–Weizmann Foundation.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

dications or addition of manganese ions to magnesium cations is known to broaden the enzyme substrate specificity: DNA polymerases then incorporate ribonucleotides or dideoxynucleotides.[1] Enlarged enzyme specificity has also been reported for template DNA analogues, such as RNA, threose nucleic acids (TNA), and hexitol nucleic acids (HNA).[2] Reverse transcription by *Thermus aquaticus* DNA polymerase I depends on the presence of manganese ions,[3] which are known to alter fidelity. Furthermore, site-directed mutagenesis has been used for the study of polymerase mutants and for the design of DNA polymerase variants with improved incorporation of deoxynucleotide analogues.[4] So far, directed DNA polymerase evolution experiments focused largely on the alteration of nucleotide triphosphate specificity.[5] To gain an insight into the amino acid side chains that can be mutated to change the template specificity of DNA polymerases, a directed-evolution experiment of the Stoffel fragment of *Thermus aquaticus* DNA polymerase I towards thermostable reverse transcriptases (RTs) was carried out. The process started from a large library of polymerase variants displayed on the surface of filamentous bacteriophages (*Inovirus*).

For the construction of library I (6×10^6 clones), mutations within the polymerase domain including the conserved motifs A, B, and C were introduced by polymerase chain reaction (PCR) mutagenesis in the presence of manganese ions.[6,7] To construct library II, the oligonucleotides assembled by PCR were degenerate at positions corresponding to eleven amino acids (483, 486, 539, 545, 547, 548, 570, 575, 578, 736, and 739) in *T. aquaticus* DNA polymerase I. The amino acids with homologous positions in T7 DNA polymerase are known to interact with the template strand.[8] A bias towards the hydrophilic and neutral amino acids, namely glutamine, asparagine, serine, and threonine, was introduced. A library of about 1.1×10^7 mutant polymerases displayed on phage resulted from the combination of libraries I and II.

For the isolation of rare thermostable RTs that are active in the presence of magnesium ions only as dications, we adapted an in vitro selection that we had previously established for the isolation of DNA-dependent DNA polymerases and their genes according to their catalytic activity.[9] Briefly, the selection consists of substrate coupling to phage proteins and the isolation of labeled reaction products by affinity chromatography (Figure 1). High enrichment factors of more than 10^3 were found and were similar to those measured for selections in vitro of proteins according to their affinity for an antigen.[10] It was further shown that the selection for DNA-dependent DNA polymerase activity allows the isolation of very rare polymerases on-phage even among a 10^8-fold excess of an other phage.[11] This method to enrich populations of proteins into catalysts was then considered to be efficient enough for use in a directed-enzyme-evolution experiment. DNA primers modified at their 5′ end by a maleimidyl group were coupled to the phage polymerase mutant library, hybridised to a template RNA, and reacted with deoxynucleotide triphosphates whose uridine was labeled by a biotin group. Selection of active polymerases and their genes was then performed by affinity chromatography for biotin. Cycles of phage–polymerase selection for RT activity, of amplification of the corresponding genes, and of phage preparation were iterated (Figure 2). The increasing activity of the phage population with the cycles showed that the in vitro selection of polymerases is also a selection of proteins according to their catalytic turnover: the higher the catalytic turnover, the higher the number of products bound to the phage, and the higher the avidity of the phage particle for the solid phase used for affinity chromatography. Between cycles five and six, the RT activity of the population of phage polymerase variants did not improve significantly (Figure 2); we therefore decided to stop the directed-enzyme-evolution experiment after the sixth cycle.

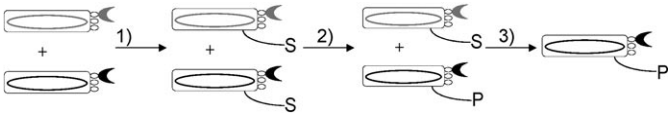

Figure 1. Principle for the isolation of active enzymes and their genes by phage display. Active and inactive enzymes are represented as black and gray, respectively. Each variant enzyme is expressed on the surface of a phage particle that contains the corresponding gene that can be amplified and sequenced to characterize the displayed enzyme. The in vitro selection of catalysts consists of three steps: 1) the cross-linking of substrate (S) to the phage particle; 2) the intramolecular conversion of the substrate into product (P); and 3) the affinity chromatography for the product. Selection cycles are iterated to reduce the background of the selection and thereby to enrich the population of the mutant enzymes into catalysts.

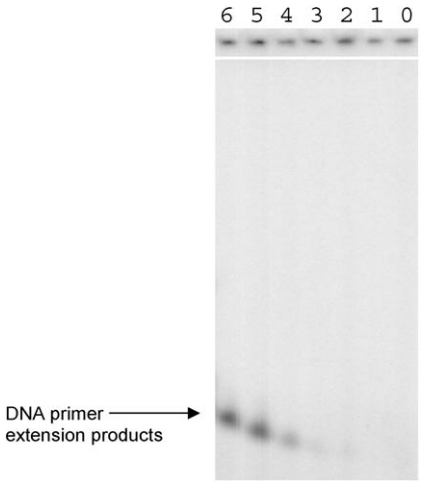

Figure 2. RT activity of the selected phage-variant populations. Column 0 highlights the RT activity of the initial library of more than 10^7 variants (not detectable). Column 6 shows the RT activity of the polyclonal phage–polymerase population after the sixth selection cycle. RT activity was detected by using a DNA-primer extension assay: DNA primer 13 hybridized to RNA template 12 was extended with radioactively labeled α^{32}P-deoxythymidine triphosphate (dTTP) by 10^8 phage–polymerases; electrophoresis of the products on a 20% polyacrylamide gel was followed by detection of the extended primers by using a Phosphorimager screen (see the Supporting Information).

To characterize the selected population of catalysts after the sixth selection cycle, 116 polymerase variants were isolated and sequenced. Twenty eight variants were reported

Table 1: Variants and their mutations associated with the catalytic activity shift from DNA- to RNA-dependent DNA polymerization.

Mutations[a]	Variant	Abundance[b]	Phage RT activity[c]
L552H, T509S, M779K	**1**	1	0.3
L552P, A661T, A800V, **E820K**	**2**	2	0.2
A608V	**3**	1	0.1
A608V, I553F, **F749S**	**4**	2	0.4
A608T, E520G, **W827R**	**5**	1	1.4
S612N, **P816S**	**6**	6	0.5
S612N, N483Q, S486Q, T539N, Y545Q, D547T, P548Q, A570Q, D578Q, A597T W604R, V730L, R736Q, S739N, **M747R**	**7**	3	1.0
Y696N	**8**	1	0.2
Y696C	**9**	2	0.3
Y696C, K767E	**10**	1	0.3
Y696C, E530G, A803V	**11**	1	0.4
M747R	**12**	3	1.0
M747K	**13**	1	1.3
M747K, E742K	**14**	1	5.5
M747K, Q698L, **P816L**	**15**	1	0.7
m747K, V586I, I638F, G776E	**16**	11	1.8
F749Y, A568V	**17**	4	0.4
F749Y, P550Q, R556S, V740E, V819A	**18**	3	0.7
F749Y, F482L, T664A, P770T, M775T	**19**	1	0.3
M761V	**20**	4	0.5
M761T, D547G, I584V	**21**	34	1
E820K, I599N, L780P	**22**	1	0.2
W827R	**23**	7	0.7
W827R, H480R	**24**	1	0.4
W827R, A517V, T664S, F769S	**25**	2	0.7
D551G, V783I	**26**	1	0.3
Q534R, A764T	**27**	1	0.5
R651Q	**28**	1	0.2

[a] Mutations are indicated by their amino acid position within *T. aquaticus* DNA-polymerase I and are noted in bold if the mutated position occurred at least twice in distinct variants. [b] The abundance of each variant within a sample of the selected population after the sixth selection cycle is indicated. [c] The RT activity of the phage–polymerase variants is given as quantities of products formed relative to those of the most abundant variant that is, **21**. This activity assay made use of RNA template 31 (X = A) hybridized to DNA primer 32 that was extended by radioactively labeled α^{32}P-dTTP. Reaction products were separated by electrophoresis on a 20% polyacrylamide gel for detection of extended primers on a Phosphorimager screen prior to quantification using the Imagequant software.

with their population-sample abundances and relative RT activity (Table 1). The most abundant clone after the sixth cycle (variant **21**) represented about one third of the clones in the population sample. Among the 27 other distinct variants, 15 were represented by a single clone, thereby highlighting the diversity of the selected population. Variant **7** was the only variant derived from library II, which was obtained by PCR assembly of degenerate oligonucleotides: 10 non-neutral mutations occurred at targeted amino acid positions and five mutations were introduced during the PCR assembly and amplification. The other 27 variants with RT activity had, on average, 2.4 mutations. The mutations E742K, M747K, and E820K could be structurally interpreted in terms of salt-bridge formation with negatively charged template phosphodiester groups located close to the amino acids 742 and 747 or with the nucleotide's triphosphate group close to amino acid 820 (Figure 3). Interestingly, amino acids 608 and 612, which were mutated in several variants, surround in a beta strand the aspartate 610 known to bind a magnesium ion at the polymerase active site (compare 2KTQ and 1LV5 structures in the protein databank (PDB))[12] The Stoffel fragment, the most-abundant polymerase **21**, and two of the most-active variants **5** and **14** were then expressed in *E. coli* and purified by two affinity-chromatography steps for kinetic characterization (Table 2). For the reverse-transcription reaction, the most-abundant polymerase **21** had an almost tenfold improved catalytic efficiency as measured by the k_{cat}/K_m ratio when compared with the Stoffel fragment. Variant **5** had rate constants k_{cat} of about two orders of magnitude larger than those of the Stoffel fragment. Remarkably, variants **5** and **14** had catalytic efficiencies for DNA-dependent DNA polymerization that were similar or higher than those of the wild-type Stoffel fragment even though no selection for the wild-type activity was applied during the experimental evolution. Our directed-polymerase-evolution experiment yielded variants endowed with two catalytic activities, namely RNA-dependent DNA polymerization (EC 2.7.7.49), and DNA-dependent DNA polymerization (EC 2.7.7.7). To further estimate the fidelity during reverse transcription, the variants **5**, **14**, and **21** were then used to copy a part of messenger RNA (mRNA) into complementary DNA (cDNA) prior to subcloning for sequencing of reverse transcription products. The sequences showed that selected variants could polymerize more than 300 nucleotides and were not limited to the addition of single nucleotides at the 3′ end of DNA primers. The substitution rates per base for RNA-dependent DNA polymerization of the most active variants, **5** (1.5×10^{-3}) and **14** (2.2×10^{-3}), were found to be similar or lower than that of avian myeloblastosis virus (AMV) RT (2.2×10^{-3}), which was used as a standard. Interestingly, the most abundant variant **21** (2.3×10^{-4}) had a fidelity which was about 10-times higher than that of AMV RT.

Enzymes generally evolved to have finely tuned substrate specificities. This directed-enzyme-evolution experiment allowed the identification of mutations associated with a broadened template specificity for DNA polymerase. The catalytic efficiency of a selected variant was improved by about two orders of magnitude for reverse transcription when compared with the Stoffel fragment of Taq DNA polymerase. Some of these engineered and thermostable DNA polymerases, which can copy RNA as well as DNA with high

Table 2: Kinetic parameters of the DNA- and RNA-dependent DNA polymerization reactions.[a]

		DNA.DNAP				RNA.DNAP			
		Stoffel fragment	Variant **21**	Variant **5**	Variant **14**	Stoffel fragment	Variant **21**	Variant **5**	Variant **14**
k_{cat} [s^{-1}]	dATP	$9.6\ (0.6)\times10^{-2}$	$2.8\ (0.2)\times10^{-2}$	$6.5\ (0.3)\times10^{-1}$	$2.0\ (0.1)\times10^{-1}$	$1.7\ (0.1)\times10^{-2}$	$7.5\ (0.7)\times10^{-2}$	3.3 (0.2)	1.4 (0.1)
	dCTP	$1.3\ (0.6)\times10^{-1}$	$1.7\ (0.2)\times10^{-2}$	$3.1\ (0.3)\times10^{-1}$	$2.8\ (0.2)\times10^{-1}$	$6.0\ (0.4)\times10^{-2}$	$9.3\ (1.0)\times10^{-2}$	2.2 (0.2)	$8.5\ (1.0)\times10^{-1}$
	dTTP	$1.2\ (0.1)\times10^{-1}$	$1.2\ (0.1)\times10^{-2}$	$2.8\ (0.1)\times10^{-1}$	$1.8\ (0.1)\times10^{-1}$	$1.0\ (0.1)\times10^{-2}$	$8.5\ (0.6)\times10^{-2}$	2.6 (0.6)	$9.7\ (1.6)\times10^{-1}$
K_m [μM]	dATP	3.0 (1.0)	45 (16)	5.7 (1.1)	2.7 (0.4)	60 (15)	33 (7)	292 (48)	86 (20)
	dCTP	6.1 (1.2)	33 (22)	8.5 (2.7)	3.4 (1.0)	177 (31)	31 (6)	344 (61)	72 (37)
	dTTP	5.2 (1.9)	22 (10)	7.9 (1.6)	10 (2)	77 (34)	63 (7)	240 (83)	250 (112)
k_{cat}/K_m [$L\,mol^{-1}\,s^{-1}$]	dATP	3.1×10^4	6.1×10^2	1.1×10^5	7.4×10^4	2.9×10^2	2.2×10^3	1.1×10^4	1.6×10^4
	dCTP	2.3×10^4	5.3×10^2	3.6×10^4	8.2×10^4	3.4×10^2	3.0×10^3	6.4×10^3	1.2×10^4
	dTTP	2.3×10^4	5.6×10^2	3.6×10^4	1.8×10^4	1.5×10^2	1.3×10^3	1.1×10^4	3.9×10^3

[a] Reactions performed at 37°C for the three polymerase variants **5**, **14** and **21** and for the Stoffel fragment of *T. aquaticus* DNA-polymerase I. The polymerases were overexpressed in *E. coli* and purified by two-step affinity chromatography. The single-turnover kinetic experiments at the steady state made use of 5′ ^{32}P labeled DNA primer 32 and of the template RNA 31 (X = A, U or G) or of the template DNA 30 (X = A, T or G). The labeled substrates and products were separated, detected, and quantified as indicated in Table 1. DNAP = DNA polymerization.

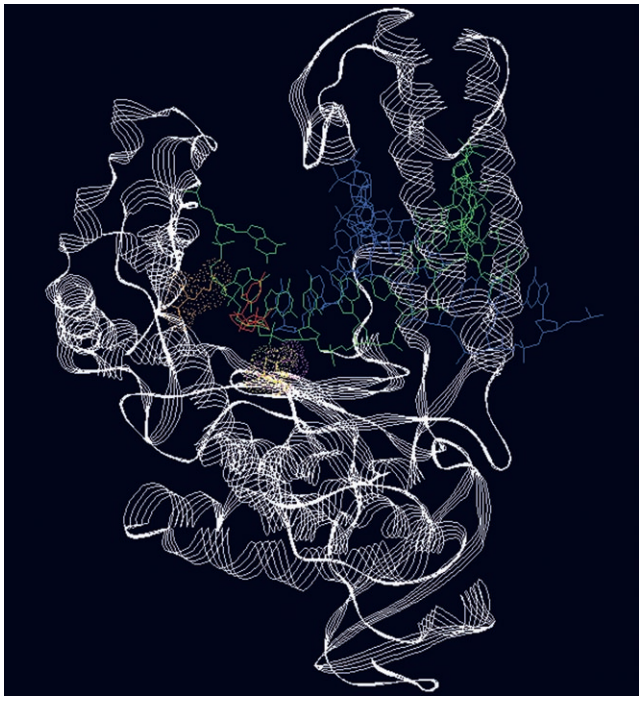

Figure 3. Structure of the Stoffel fragment of *Thermus aquaticus* DNA polymerase I (2KTQ) plotted by using Swiss PdbViewer. The polymerase (white) is in a complex with dideoxycytidine triphosphate (red) and a DNA template (green) that is hybridized to a DNA primer (blue). The amino acids M747 (orange), E820 (yellow) and D610 (purple) are highlighted.

fidelity, may find applications in the field of diagnostics for the detection of rare RNAs by reverse transcription-PCR.

Received: March 28, 2006
Revised: May 29, 2006
Published online: July 12, 2006

Keywords: directed evolution · DNA structures · enzyme catalysis · enzymes · selection

[1] a) J. H. Van de Sande, P. C. Loewen, H. G. Khorana, *J. Biol. Chem.* **1972**, *247*, 6140–6148; b) S. Tabor, C. C. Richardson, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 4076–4080; c) S. Tabor, C. C. Richardson, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6339–6343.

[2] a) J. C. Chaput, J. K. Ichida, J. W. Szostak, *J. Am. Chem. Soc.* **2003**, *125*, 856–857; b) S. Pochet, P. A. Kaminski, A. Van Aerschot, P. Herdewijn, P. Marliere, *C. R. Biol.* **2003**, *326*, 1175–1184; c) H. Shandilya, K. Griffiths, E. K. Flynn, M. Astatke, P. J. Shih, J. E. Lee, G. F. Gerard, M. D. Gibbs, P. L. Bergquist, *Extremophiles* **2004**, *8*, 243–251.

[3] V. I. Grabko, L. G. Chistyakova, V. N. Lyapustin, V. G. Korobko, A. I. Miroshnikov, *FEBS Lett.* **1996**, *387*, 189–192.

[4] a) R. Sousa, R. Padilla, *EMBO J.* **1995**, *14*, 4609–4621; b) D. T. Minnick, M. Astatke, C. M. Joyce, T. A. Kunkel, *J. Biol. Chem.* **1996**, *271*, 24954–24961; c) A. Tosaka, M. Ogawa, S. Yoshida, M. Suzuki, *J. Biol. Chem.* **2001**, *276*, 27562–27567; d) M. Astatke, K. Ng, N. D. Grindley, C. M. Joyce, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3402–3407; e) M. Ogawa, A. Tosaka, Y. Ito, S. Yoshida, M. Suzuki, *Mut. Res. DNA Repair* **2001**, *485*, 197–207; f) Y. Li, V. Mitaxov, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9491–9496.

[5] a) S. Brakmann, *ChemBioChem* **2001**, *2*, 212–219; b) A. Shinkai, P. H. Patel, L. A. Loeb, *J. Biol. Chem.* **2001**, *276*, 18836–18842; c) G. Xia, L. Chen, T. Sera, M. Fa, P. G. Schultz, F. E. Romesberg, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6597–6602; d) M. Fa, A. Radeghieri, A. A. Henry, F. E. Romesberg, *J. Am. Chem. Soc.* **2004**, *126*, 1748–1754; e) 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; f) J. Chelliserrykattil, A. D. Ellington, *Nat. Biotechnol.* **2004**, *22*, 1155–1160; g) S. Brakmann, *Cell. Mol. Life Sci.* **2005**, *62*, 2634–2646; h) R. C. Holmberg, A. A. Henry, F. E. Romesberg, *Biomol. Eng.* **2005**, *22*, 39–49; i) D. Summerer, N. Z. Rudinger, I. Detmer, A. Marx, *Angew. Chem.* **2005**, *117*, 4791–4794; *Angew. Chem. Int. Ed.* **2005**, *44*, 4712–4715.

[6] M. Fromant, S. Blanquet, P. Plateau, *Anal. Biochem.* **1995**, *224*, 347–353.

[7] R. C. Cadwell, G. F. Joyce, *PCR Methods Appl.* **1994**, *3*, 136–140.

[8] S. Doublié, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **1998**, *391*, 251–258.

[9] J. L. Jestin, P. Kristensen, G. Winter, *Angew. Chem.* **1999**, *111*, 1196–1200; *Angew. Chem. Int. Ed.* **1999**, *38*, 1124–1127.

[10] a) E. Orsi, J. L. Jestin, *C. R. Chim.* **2003**, *6*, 501 – 506; b) J. D. Marks, H. R. Hoogenboom, A. D. Griffiths, G. Winter, *J. Biol. Chem.* **1992**, *267*, 16007 – 16010.
[11] S. Vichier-Guerre, J. L. Jestin, *Biocatal. Biotransform.* **2003**, *21*, 75 – 78.
[12] a) Y. Li, S. Korolev, G. Waksman, *EMBO J.* **1998**, *17*, 7514 – 7525; b) S. J. Johnson, J. S. Taylor, L. S. Beese, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3895 – 3900.
[13] F. C. Lawyer, S. Stoffel, R. K. Saiki, K. Myambo, R. Drummond, D. H. Gelfand, *J. Biol. Chem.* **1989**, *264*, 6427 – 6437.
[14] H. Strobel, D. Ladant, J. L. Jestin, *Mol. Biotechnol.* **2003**, *24*, 1 – 9.
[15] P. Kristensen, G. Winter, *Folding Des.* **1998**, *3*, 321 – 328.
[16] S. Creighton, L. B. Bloom, M. F. Goodman, *Methods Enzymol.* **1995**, *262*, 232 – 256.

