# An Error-Prone T7 RNA Polymerase Mutant Generated by Directed Evolution

Susanne Brakmann* and Sandra Grzeszik[a]

*Viruses replicate their genomes at exceptionally high mutation rates. Their offspring evolve rapidly and therefore, are able to evade common immunological and chemical antiviral agents. In parallel, virus genomes cannot tolerate a further increase in mutation rate: Experimental evidence exists that even few additional mutations are sufficient for the extinction of a viral population. A future antiviral strategy might therefore aim at increasing the error-producing capacity of viral replication enzymes. We employed the principles of directed evolution and developed a scheme for the stringent positive selection of error-prone polymerase activity. A mutant T7 RNA polymerase with a nucleotide substitution error rate at least 20-fold greater than that of the wild-type was selected. This enzyme synthesized highly heterogeneous RNA products in vitro or in vivo and also decreased the replication efficiency of wild-type bacteriophage T7 during infection.*

**KEYWORDS:**

antiviral agents · directed evolution · gene technology · RNA polymerase · transferases

## Introduction

During the past two decades, experiments with a steadily increasing number of viruses have demonstrated that the offspring of these viruses within one infected organism are mutated at such high rates that each virus genome differs from another in one or more nucleotides on average. A majority of viruses have also been shown to deviate from the average or consensus sequence.[1–3] The resulting populations of virus genomes therefore exist as complex distributions of nonidentical, but closely related sequences called "quasispecies".[4, 5] The inhomogeneous nature of the quasispecies gives virus populations an enormous flexibility, which supports their rapid adaptation to less favorable replication conditions, for example in circumventing the influence of antiviral therapeutics, or, in the case of a mammalian host, escaping from the organism's immune response. Obviously, high variability is a tremendous selective advantage for any virus but, from the medical point of view, represents a major disadvantage in disease control and in vaccine development.

The exceptionally high flexibility of virus populations usually has the disadvantage of leading to a large fraction of partially defective or even lethal offspring. In fact, it has been established experimentally that a minor fraction of bacteriophage Q$\beta$ population of only 0.7 – 1.6 % contained infectious particles.[6] With eukaryotic picorna viruses, the relation of infectious to total virus particles even decreases to the level of $10^{-3}$ – $10^{-4}$.[7] These findings suggest that virus genomes cannot tolerate many additional mutations without substantially losing their viability. Even with a small increase in the mutation rate, virus populations reach an error threshold beyond which their replication capacity and infectivity are completely lost. This fact has already been recognized as an opportune way to influence viral infections[8–10] and recently led to the finding of proof of "lethal mutagenesis" of HIV. This was achieved by decreasing the replication fidelity with mutagenic deoxynucleoside analogues during sequential passages in culture.[11]

Considering a future antiviral strategy, a promising and challenging alternative to the application of mutagenic nucleosides might directly aim at increasing the error-producing capacity of viral replication enzymes beyond their limiting threshold. An error-prone polymerase would intrinsically catalyze the transition from "readable" genetic information to "nonsense messages" without any ability to control the necessary replication pathways. As a consequence, the virus population would die out. Although a true "error catastrophe" could only be achieved at the level of genome replication, in certain cases inaccurate transcription and translation may also lead to permanent and inheritable changes within the genome. Indeed, evidence has recently been presented that translational miscoding may give rise to a mutator phenotype in *Escherichia coli* due to the production of mutant DNA polymerases lacking proofreading activity.[12] Similarly, erroneous transcripts may result in dysfunctional proteins after translation that are highly deleterious in maintaining integrity and stability of the genetic information.

We started the search for an error-prone polymerase with the RNA polymerase of bacteriophage T7 (T7 RNAP). This enzyme has several advantages for use in directed evolution of polymerase function. Unlike the multi-subunit DNA-dependent RNA polymerases of eukaryotes and prokaryotes, T7 RNAP consists of a single 98-kDa protein chain that is responsible for the

[a] *Dr. S. Brakmann, S. Grzeszik*
*Evotec Biosystems AG*
*Schnackenburgallee 114, 22525 Hamburg (Germany)*
*Fax: (+49) 551-201-1435*
*E-mail: sbrakma@gwdg.de*

  1439-4227/01/02/03 $ 17.50+.50/0 CHEMBIOCHEM **2001**, 2, 212 – 219

transcription of complete phage DNA. The enzyme initiates transcription after stringent recognition of T7 promoter DNA in vivo and in vitro, and therefore allows for specific high-level transcription and expression of cloned genes.[13] Beyond that, extensive mutational[14–20] and structural studies[21–23] with T7 RNAP provide substantial information for understanding and exploring the differences between wild-type enzyme and its functional variants. These insights, together with the unique catalytic features of T7 RNAP, allowed us to successfully develop a bacterial in vivo system to select RNA polymerases capable of erroneous copying.

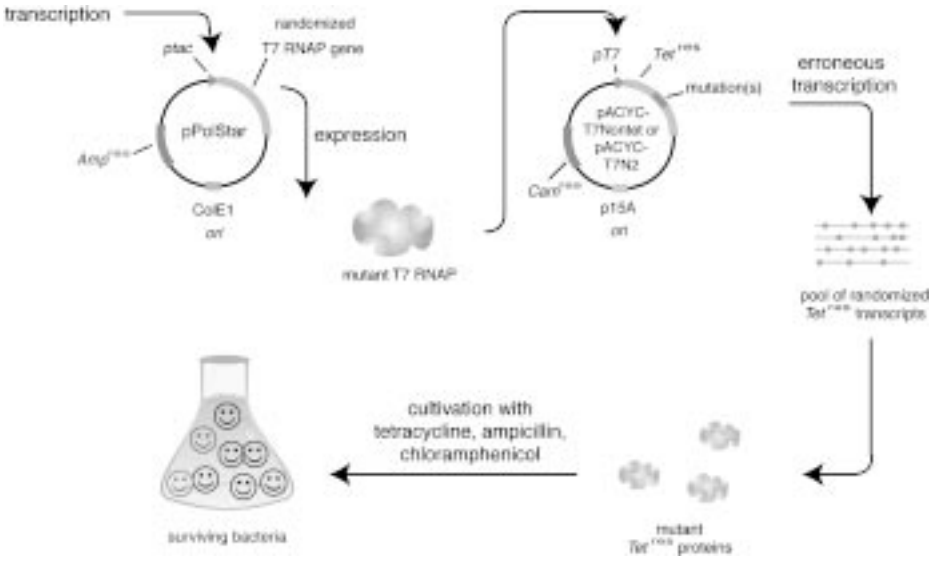


***Figure 1.*** *Positive selection of an error-prone T7 RNA polymerase (T7 RNAP). Plasmid pPolStar contains the ColE1 origin of replication (*ori*), which is compatible with the p15A* ori *of pACYC-T7Nontet. Bacteria co-transformed with both plasmids start expressing a T7 RNAP mutant upon induction with isopropyl-β-thiogalactoside (IPTG). The enzyme subsequently recognizes the T7 promoter on pACYC-T7Nontet and produces a pool of randomly altered transcripts among which one or more might exist that encode restored resistance activity, for example as a consequence of a second-site mutation, and that support the survival of the respective host bacteria.* Amp$^{res}$, Cam$^{res}$, Tet$^{res}$ *= genes for resistance to ampicillin, chloramphenicol, and tetracycline, respectively.*

## Results

### Selection of an error-prone T7 RNA polymerase mutant

The isolation of an error-prone RNA polymerase from a plasmid library containing random sequence substitutions was enabled through a stringent positive selection scheme, that is, by rewarding inaccurate transcription by a T7 RNAP mutant with the survival of bacteria. Therefore, a system of two compatible plasmids was constructed, which couples mutant polymerase genes in a feedback loop to the essential, but inactivated tetracycline resistance gene (Figure 1).

Feedback coupling was achieved by introducing a T7 promoter upstream of the tetracycline resistance gene of plasmid pACYC184 for expressing the resistance under the exclusive control of a T7 promoter. Bacteria transformed with this construct (pACYC-T7Tet) did not survive exposure to tetracycline (10–25 μg mL$^{-1}$) without T7 RNAP being present at the same time. In the presence of wild-type T7 RNAP, however, viable transformants retained resistance to 25 μg mL$^{-1}$ of the antibiotic. Subsequently, the tetracycline resistance of pACYC184, which is mediated by a hydrophobic, membrane-associated efflux pump,[24] was inactivated by the single amino acid replacement Tyr 100 →Pro (Y100P) (Figure 2 B). This mutation is located within a periplasmic region of the membrane-bound protein which previously had been identified to be critical for the performance of resistance.[25] Bacteria transformed with the resulting plasmid pACYC-T7Nontet, and expressing wild-type T7 RNAP, did not survive upon exposure to tetracycline (10 μg mL$^{-1}$). This finding verified that substitution of a single residue abolished the antibiotic resistance. Genetic analysis of the tetracycline resistance also indicated that functional interactions exist between the N- and C-terminal domains of the tetracycline efflux pump.[26] As a consequence of this interaction, a deleterious mutation in one domain of the protein might be suppressed by some

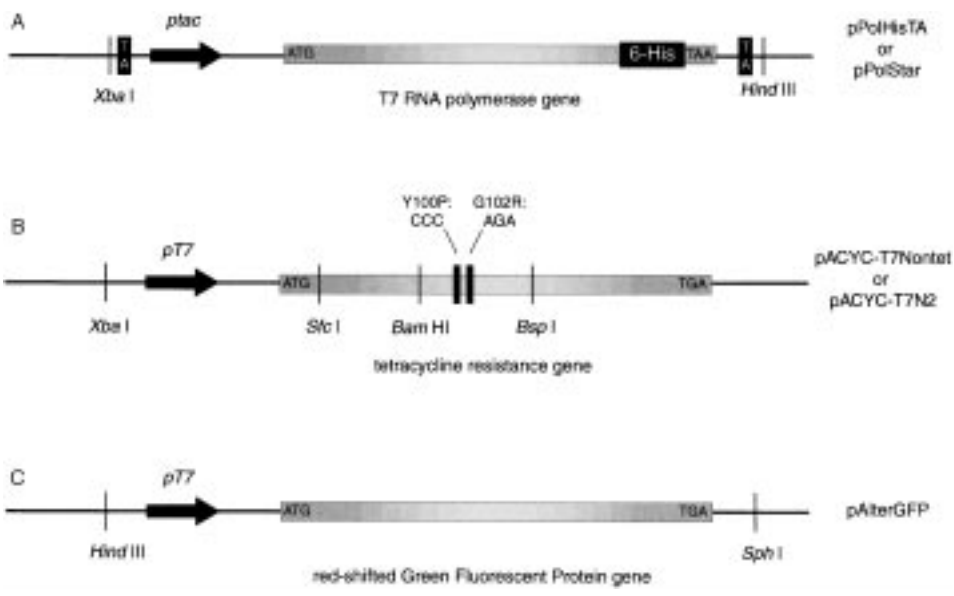


***Figure 2.*** *A: Genetic construction of vectors pPolHisTA and pPolStar. B: Genetic construction of pACYC-T7Tet (wild-type), pACYC-T7Nontet (Y100P), and pACYC-T7N2 (Y100P, G102R). C: Genetic construction of pAlterGFP.*

second-site mutation within the other domain and thus enables the selection scheme described here.

Selection experiments were performed after co-transformation of XL1-Blue cells harboring plasmid pACYC-T7Nontet with a pool of plasmids carrying randomly mutated T7 RNAP genes (pPolStar). These mutant libraries were generated by either PCR mutagenesis or mutator strain passage. In accordance with the error-prone PCR protocol of Joyce and Cadwell,[27] genes were randomized at frequencies of $0.93 \pm 0.06$% per position over the total course of amplification. PCR-generated fragments were submitted to forced cloning (i.e. using two incompatible restriction overhangs and dephosphorylated vector), or were ligated directly into a linear T vector.[45] Ligation efficiencies were limited with both strategies, yielding on average $10^3$ colony-forming units (cfu) per μg of linearized vector DNA. Therefore, mutator strain passage of the target T7 RNAP gene cloned into an expression construct (pPolHisTA) was chosen as an alternative to the PCR mutagenesis – cloning procedure. *E. coli* strain XL1-Red on average exhibits a mutation frequency of $10^{-6}$ per replication cycle and cell.[28] The overall mutation rate was increased to a maximum of ca. $10^{-3}$ by repeating mutator strain passage five times. In parallel, however, recombination events accumulated significantly to reach 5% after the fourth passage, and 20% after the fifth passage.

For selection, a bacterial culture comprising a population of ca. $10^6$ T7 RNAP mutants (originating from mutator strain passage) was grown in liquid media containing 40% of the standard tetracycline dose. The combined selective constraints aimed at rewarding bacteria expressing T7 RNAP variants capable of transcribing both fast as well as with decreased fidelity. Surviving bacteria emerged immediately and amplified for 15 h until a steady state was reached. Purification of both plasmids and the sequencing of the 2.7-kb T7 RNAP gene fragment of pPolStar demonstrated the predominance of a single variant with three amino acid substitutions, F11 →L, C515 →Y, and T613 →A. Sequencing of plasmid pACYC-T7Nontet harbored by the same culture revealed that the correct transcription context, as well as the inactivation of the *tet* gene, had been retained. These findings suggested that indeed an error-prone T7 RNAP mutant—and not an artefact—had been isolated.

### Fidelity of in vitro transcription

Mutant T7 RNAP protein encoded on plasmid pPolStar was overexpressed in *E. coli* strain BL21/pRep4 and purified in accordance with a protocol that was initially developed for the wild-type enzyme.[29] Both wild-type and mutant T7 RNAP were tested in run-off transcription experiments by using linearized plasmid pAlterGFP (rsGFP under T7 promoter control; see below) as a template. Differences in amount and composition of RNA products synthesized by either wild-type or mutant T7 RNAP already became apparent upon electrophoretic analysis in 1.2% (w/v) agarose gels: In contrast to the wild-type enzyme that synthesized few RNA products of distinct length, RNA resulting from transcription by mutant enzyme was homogeneously distributed between DNA marker bands at 500 and 8000 bp, respectively. Isolation, subcloning, and sequencing of the appertaining cDNA revealed a nucleotide substitution error rate of $< 6 \times 10^{-5}$, and an error rate of $6 \times 10^{-5}$ for nucleotide deletions produced by the wild-type enzyme. The implementation of the same procedure with RNA of comparable mobility synthesized by mutant enzyme demonstrated a nucleotide substitution error rate of $1.25 \times 10^{-3}$. The error rate for single-nucleotide deletions was estimated to be lower than $4 \times 10^{-5}$ because frameshift mutations (insertions or deletions) had not been detected so far (ca. 23 000 bp analyzed; Table 1). Thus, the frequency of substitution errors produced by the selected T7 RNAP mutant was increased at least 20-fold as compared to the wild-type enzyme, whereas the frequency of single-nucleotide deletions remained almost unchanged. Among all substitutions 55% transversions and 45% transitions were observed, which showed almost no preference for a certain sequence context (Figure 3).

### In vivo fidelity assays

In an initial test, the influence of mutant T7 RNAP (F11 →L, C515 →Y, T613 →A) on the infection cycle of wild-type phage T7 was monitored. This test was based on the finding that bacteriophage T7 populations that grow on bacterial hosts harboring a certain T7 gene, for example that coding for RNA polymerase, tend to be outgrown by phage mutants defective in the respective part of their genome: These mutants occur by spontaneous recombination (once in $\leq 10^4$ phage replications) and obtain a substantial advantage by complementation, that is, by utilizing the host-expressed enzyme which is supplied by a plasmid. Consequently, the mutants replicate more efficiently than the wild-type.[30]

Differences between wild-type and mutant T7 RNAP became apparent upon growth of wild-type bacteriophage T7 in shaker cultures of host bacteria harboring either of the plasmid-encoded enzymes. In a normal lytic cycle at 30 °C, lysis of the bacterial culture begins about 25 min after infection.[31] In our experiments, the first burst of wild-type phage progeny appeared 35 – 40 min after the addition of phage stock solution and reduced the cell densities to 25 – 35%, and to 10 – 15% after a further burst. The remaining bacteria represent a fraction of

***Table 1.*** *Mutation frequency and mutation types observed in the run-off transcription assay with the rsGFP gene as a template.*

| | | Nucleotide substitutions | | | | Single-nucleotide frameshifts | | | |
|---|---|---|---|---|---|---|---|---|---|
| T7 RNAP | bp sequenced | transitions | transversions | total | mutation frequency | insertions | deletions | total | mutation frequency |
| wild-type | 16750 | – | – | – | $< 6 \times 10^{-5}$ | – | 1 | 1 | $6 \times 10^{-5}$ |
| mutant | 23076 | 13 | 16 | 29 | $1.25 \times 10^{-3}$ | – | – | – | $< 4 \times 10^{-5}$ |

```
TCAGTTGTAC AGTTCATCCA TGCCATGTGT AATCCCAGCA GCTGTTACAA ACTCAAGAAG GACCATGTGG TCTCTCTTTT CGTTGGGATC TTTCGAAAGG
                                       A   C    A           C
                                           A    A
                                                A

GCAGATTGTG TGGACAGGTA ATGGTTGTCT GGTAAAAGGA CAGGGCCATC GCCAATTGGA GTATTTTGTT GATAATGGTC TGCTAGTTGA ACGCTTCCAT
                                                                                 A          T       T
                                                                                 C

CTTCAATGTT GTGGCGGGTC TTGAAGTTCA CTTTGATTCC ATTCTTTTGT TTGTCTGCCA TGATGTATAC ATTGTGTGAG TTATAGTTGT ATTCCAATTT
                                                                                                       C

GTGTCCCAGA ATGTTGCCAT CTTCCTTGAA GTCAATACCT TTTAACTCGA TTCTATTAAC AAGGGTATCA CCTTCAAACT TGACTTCAGC ACGTGTCTTG
                           A                G     C                              C
                           A

TAGTTGCCGT CATCTTTGAA GAAGATGGTC CTTTCCTGTA CATAACCTTC GGGCATGGCA CTCTTGAAAA AGTCATGCCG TTTCATATGA TCCGGGTATC
    G      T                                                                                    A      A

TTGAAAAGCA TTGAACACCA TAGCACAGAG TAGTGACTAG TGTTGGCCAT GGAACAGGCA GTTTGCCAGT AGTGCAGATG AACTTCAGGG TAAGTTTTCC
   GT                                                    G

GTATGTTGCA TCACCTTCAC CCTCTCCACT GACAGAGAAC TTGTGGCCGT TAACATCACC ATCTAATTCA ACAAGAATTG GGACAACTCC AGTGAAGAGT
        TG                                                                 GG   T

TCTTCTCCTT TGCTAGCCAT
```

***Figure 3.*** *Distribution of point mutations generated by T7 RNAP using the rsGFP template.*

naturally occurring mutants with an altered surface that does not permit adsorption and penetration of T7 phage. In contrast to these findings, lysis of the bacteria harboring mutant T7 RNAP proceeded more slowly and to a lesser extent (Figure 4): About 60% of the cell density remained after the first infection cycle, and eventually reached 25 – 30%. This observation suggested that the efficiency of viral replication was significantly influenced by the mutant T7 RNAP.

In a further test, the transcriptional activity and accuracy of the isolated T7 RNAP mutant was assayed in vivo by using another system of two functionally linked plasmids which resembles the selection system described above: Plasmid pAlterGFP was constructed for expressing Green Fluorescent Protein (GFP) with a red-shifted (rs) excitation maximum under the control of a T7 promoter carried on a p15A vector compatible with ColE1 plasmids pPolStar (mutant T7 RNAP) or pAR1219 (wild-type T7 RNAP), respectively (Figure 2C). Mutational studies with GFP variants have recently demonstrated that mutant GFPs usually retain the overall structural organization of the wild-type, but often display significant wavelength shifts or alterations of fluorescence intensity.[32–34] Error-prone transcription by a T7 RNAP mutant and consequent expression of GFP variants in *E. coli* therefore should be detectable by screening bacterial cultures for changes in fluorescence behavior. Experiments performed in parallel with cells expressing GFP under control of either wild-type or mutant T7 RNAP clearly revealed mutator activity for the selected variant F11 →L/C515 →Y/T613 →A (Figure 5): Bacterial cultures expressing rsGFP after transcription by wild-type polymerase showed a uniform (Gaussian) distribution of fluorescence with an average relative intensity of 7525 ± 448. In contrast to these, cultures expressing rsGFP following transcription by the selected polymerase mutant exhibited a wide, non-Gaussian distribution of fluorescence intensities with an average relative intensity of 8163 ± 173. The observed deviation from standard wild-type behavior was attributed to contributions of GFP mutants exhibiting wavelength shifts of the normal excitation wavelength of 485 nm.

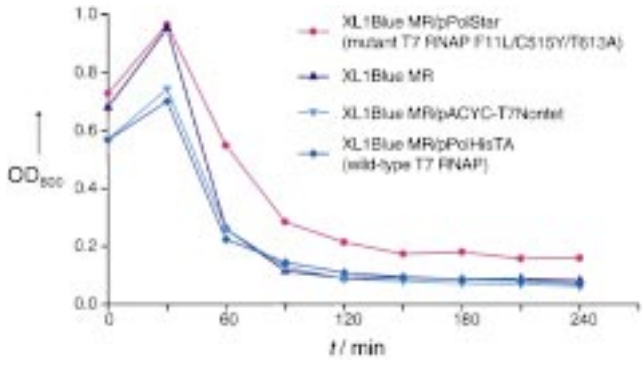


***Figure 4.*** *Dependence of bacterial lysis on the presence of plasmid-encoded T7 RNAP mutant F11L/C515Y/T613A. Wild-type bacteriophage T7 was grown in 25-mL shaker cultures of* E. coli *XL1-Blue MR cells harboring no plasmid, plasmid pACYC-T7Nontet with no T7 polymerase encoded, plasmid pAR1219 with wild-type T7 RNAP, or plasmid pPolStar with mutant T7 RNAP. Following infection the lysis of cells was monitored by determining the absorbance ($OD_{600}$) in 30-min intervals.*

## Discussion

We have used random mutagenesis together with a genetic selection scheme to discover active T7 RNA polymerase mutants with decreased transcriptional fidelity. The wild-type enzyme is the most commonly used DNA-dependent RNA polymerase for synthesizing RNA in vitro. Studies of its error rate revealed

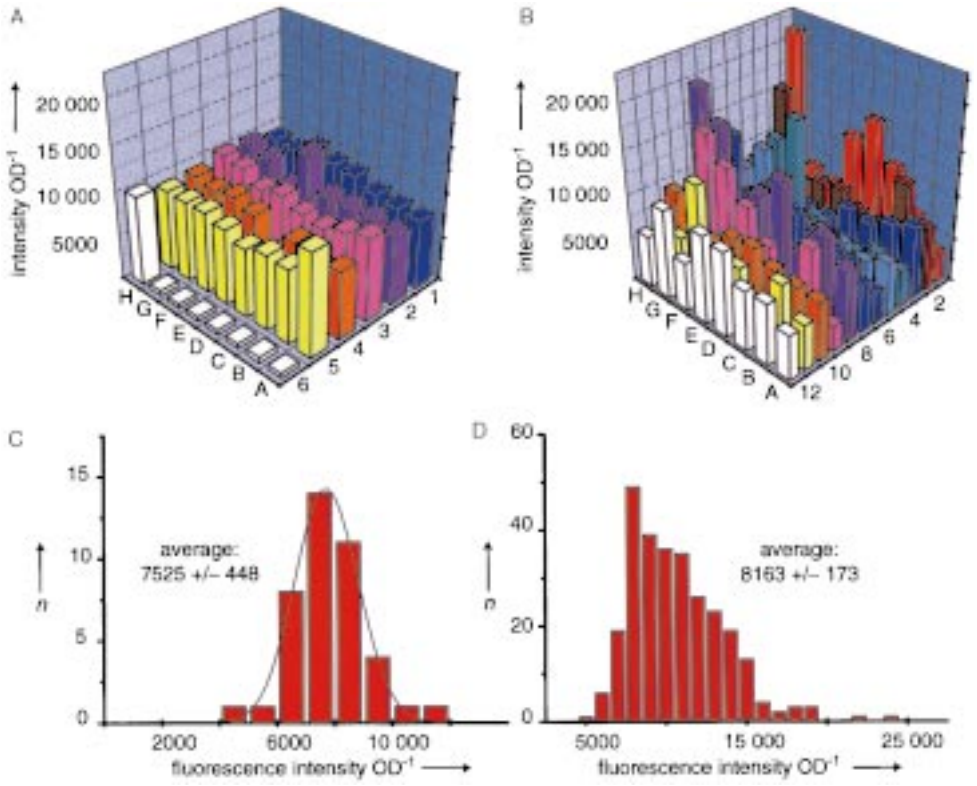


***Figure 5.*** *Relation between rsGFP fluorescence intensity and transcriptional fidelity of T7 RNA polymerase. Cultures of bacteria grown overnight at 37 °C in a microtiter format were measured for fluorescence intensity at 535 nm (excitation at 485 nm), and for optical density at 590 nm ($OD_{590}$). To account for varying amounts of bacteria, all fluorescence intensity data were normalized by the respective cell densities. A: Fluorescence intensity of rsGFP in dependence of wild-type T7 RNAP (bars A – G in row 6 show the background fluorescence intensity of the LB medium). B: Fluorescence intensity of rsGFP in cells harboring mutant T7 RNAP. C: Gaussian distribution of fluorescence intensities exhibited in cells expressing wild-type T7 RNAP (n = 32). D: Distribution of fluorescence intensities resulting from mutant T7 RNAP (n = 280).* n = *number of clones screened.*

estimates of ca. $0.7 \times 10^{-5}$ by using a mutation assay in vivo,[35] and of ca. $3 \times 10^{-5}$ by using an in vitro system for quantitative assessment of RNA synthesis.[36] Both findings confirm that the fidelity of transcription lacks proofreading mechanisms and therefore is limited by base pairing of base tautomers, leading to mutation frequencies of $10^{-4} - 10^{-5}$.[37]

The selective constraints generated with the plasmid-based selection system described here required erroneous transcription for compensating one mutation within the *tet* gene. The idea was based on the finding that inactive tetracycline mutants may be reverted by random second-site mutagenesis.[26] By compensating one missense mutation, the error rate of potential T7 RNAP mutants could theoretically reach ca. $0.7 \times 10^{-3}$. Indeed, bacteria surviving the applied selective constraint expressed a mutant T7 RNAP that exhibited a substitution error rate of $1.25 \times 10^{-3}$. By yielding a variant polymerase with a mutation frequency of the expected order of magnitude ($10^{-3}$), the applied selection system completely fulfilled the expectations. Attempts to obtain a further increase in the transcriptional error rate by compensating two missense mutations failed: Surviving bacteria represented artefacts harboring either no T7 RNAP gene or nonfunctional polymerase fragments.

The selected mutant F11 →L/C515 →Y/T613 →A introduced base substitutions without significant preference for a certain type of mutation (Table 1) or a certain sequence context (Figure 3). Due to a mutation frequency that is in the range of error-prone PCR,[38] RNA synthesized by mutant T7 RNAP in vivo or in vitro is highly heterogeneous. This fact had already been observed by analyzing the crude RNA products with gel electrophoresis under native conditions and, most strikingly, was demonstrated by detecting the variance of rsGFP fluorescence after transcription in vivo: Assuming that every cell contained only about 1000 copies of rsGFP mRNA, at least one mRNA would have been translated into a mutant protein with spectral properties different from those of the wild-type. However, the copy number of rsGFP mRNA is certainly much higher than 1000: RNA transcription by the host RNA polymerase is outrivaled, and almost all transcription is effected by T7 RNAP.[13] Indeed, mRNA concentrations in cells containing pACYC-T7Nontet, or pAlterGFP accumulated to levels observable by ethidium bromide staining of electropherograms (ca. 5 μg RNA per μL of culture; data not shown) and most likely were capable of saturating the protein-synthesizing machinery of *E. coli*. Consequently, large amounts of GFP mutants with spectral properties different from wild-type rsGFP were synthesized.

To determine the nucleotide substitution errors of comparable RNA products, transcripts generated in vitro by either wild-type or mutant T7 RNAP were size-selected by using electrophoretic separation. Thereby, the broad size distribution of in vitro transcripts generated by the mutant enzyme became obvious. This observation might indicate an altered processivity or an improper termination behavior of the mutant T7 RNAP, in addition to the increased error rate. Despite this indication, a major contribution of immature transcripts to the observed in vivo effects remains doubtful because 3′-end heterogeneity would result in dysfunctional protein fragments. By studying the transcription of rsGFP by mutant T7 RNAP in vivo, however, we observed an increase in the average fluorescence intensity.

### Structural considerations

The three mutations that give rise to a T7 RNAP with mutator phenotype map to three different subdomains of the protein (Figure 6): Mutation F11 →L is located within the N-terminal

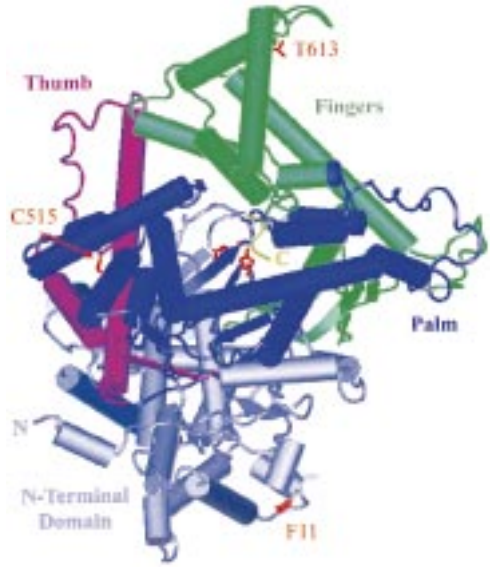


***Figure 6.*** *Modified secondary-structure representation of the crystal structure of T7 RNAP.[23] The coordinates (PDB file 1CEZ) were displayed using the program VMD[44] on a Silicon Graphics workstation. Residues phenylalanine 11, cysteine 515, and threonine 613, which were identified to be mutated, are highlighted in orange, residues aspartate 537 and 812 indicating the active site are highlighted in red, and the C-terminal residues 880 – 883 are colored yellow.*

region, which has been associated with the establishment of processivity,[22] as well as with binding upstream regions of promoter DNA[20, 39] and binding of nascent, single-stranded RNA.[40, 41] The second mutation, C515 →Y, can be attributed to the palm subdomain of T7 RNAP that is central to the catalytic activity of the protein and involved in the distiction between rNTP and dNTP substrates.[22] The third mutation, T613 →A, is located in the fingers subdomain, which enables a contact to the triphosphate moiety of the incoming ribonucleotide substrate.[23]

The T7 RNAP in pPolStar contained a C-terminal six-histidine tag still present in the selected mutant. Although the C-terminal four-residue stretch of wild-type T7 RNAP is known to be flexible in solution,[42] and was only poorly ordered in a crystal of T7 RNAP complexed with its inhibitor T7 lysozyme,[22] it is not clear whether the artificial tag affects the increased mutation rate: On the one hand, mutation of the penultimate residue of the unmodified enzyme, F882, suggests a role in binding the elongating rNTP.[14] Also, F882 and A883 of the wild-type protein were observed to be bound beneath three central β strands in the recently solved structure of a transcribing complex.[23] On the other hand, experiments performed by us to control the selection scheme showed that bacteria harboring the histidine-tagged T7 RNAP did not survive upon exposure to tetracycline. Therefore, an exclusive dependence between error rate and histidine tag seems rather unlikely.

## Conclusion and perspectives

Erroneous replication is a necessary prerequisite for the antiviral concept described above. Our approach represents a promising model for selecting error-prone polymerase activity that could, in principle, be adopted to the positive selection of DNA-dependent DNA polymerases and also, of reverse transcriptases. Beyond that, the selected T7 RNAP variant may be a valuable tool for the generation of RNA or DNA mutant libraries, for example, for in vitro selection by using isothermal amplification reactions (self-sustained sequence replication, 3SR; nucleic acid sequence-based amplification, NASBA) or in vitro transcription. Because of its relaxed interaction with mispaired bases, the mutant polymerase will probably also facilitate the incorporation of a series of modified nucleotides into RNA and thus open up the horizon for generating RNAs with modified properties that are of interest in the fields of antisense RNA, RNA aptamers, and ribozymes.

## Experimental Section

**Generation of functional T7 RNAP mutant libraries:** A 2.8-kb fragment coding for T7 RNAP was PCR-amplified starting from plasmid pAR1219[43] using the primer pair 5′-AGC TAG CTA GC-T AAG CTT CGA TCA TGG CGA CCA CAC CC-3′ and 5′-AGC TAG CTA GCT-GAA TTC GAT CCC GGA TTT ACT AAC TGG AAG AGG-3′ (restriction sites are underlined). Mutagenizing PCR conditions were achieved according to Cadwell and Joyce.[27] The PCR-generated wild-type fragment was ligated into pKK223-3 (Pharmacia) yielding plasmid pPol4-1. For TA cloning and expression of wild-type and mutant T7 RNAP genes, the upper primer 5′-GGC GTT AGT GAT GGT GAT GGT-GAT GCG CGA ACG CGA AGT CCG ACT CTA AG-3′ introduced the sequence coding for a six-histidine tag (underlined) upstream of the stop codon, and the lower primer 5′-TTG ACA ATT AAT-CAT CGG CTC GTA TAA TGT GTG GAA TTG TGA GCG GAT AAC AAT TTC A-CA CAG GAA ACA GAA TTC GAT CCC G-3′ inserted the hybrid *tac* promoter sequence (underlined) flanking the opposite end of the gene. Amplification with this primer pair was achieved starting from template pPol4-1. PCR-generated fragments, either wild-type or randomized, were directly ligated into linear vector pCR2.1 (Invitrogen) yielding plasmid pPolHisTA (wild-type T7 RNAP), or mutant library pPolStar (randomized gene). Alternatively, the mutant library pPolStar was produced by submitting the expression construct pPolHisTA to five successive passages in mutator strain XL1-Red (Stratagene) by using the supplier's protocol.

**Purification of T7 RNA polymerases:** Purification of wild-type or mutant polymerase was achieved by using a standard protocol,[29] starting from strain BL21 (Stratagene) co-transformed with pREP4 (Qiagen) and a compatible T7 RNAP expression plasmid.

**Plasmid-based selection system:** A synthetic 208-bp fragment, 5′-CTA GAT TTC AGT GCA ATT TAT CTC TTC AAA TGT AGC ACC TGA AGT-CAG CCC CAT ACG ATA TAA GTT GCG AAC TTC TGA TAG ACT TCG AAA-TTA ATA CGA CTC ACT ATA GGG AGA CCT TAT CAC AGT TAA ATT GCT-AAC GCA GTC AGG CAC CGT GTA TGA AAT CTA ACA ATG CGC TCA TCG-TCA TCC TCG GCA CCG TCA CCC TGG ATG C-3′, with 49 bp substituting the original plasmid DNA with the T7 Φ10 promoter sequence (−41 to +8; underlined) was inserted into pACYC184 (New England Biolabs) upstream of *tet* to yield plasmid pACYC-T7Tet. Co-transformation with pPolHisTA (wild-type T7 RNAP) in XL1-Blue MR (Stratagene) confirmed the correct transcriptional context for the expression of tetracycline resistance and also the full activity of cloned wild-type T7 RNAP. For inactivation of the tetracycline resistance, a synthetic 114-bp fragment, 5′-G ATC CTC CCC GCC-GGA CGC ATC GTG GCC GGC ATC ACC GGC GCC ACA GGT GCG GTT-GCT GGC GCC TAT ATC GCC GAC ATC ACC GAT GGG GAA GAT CGG-

GCT CGC CAC TTC GGG CT-3′, (substituting Tyr 100 (codon TAC) by Pro (codon CCC; underlined)) was inserted into pACYC-T7Tet substituting the original sequence. Co-transformation of the resulting plasmid pACYC-T7Nontet with pPolHisTA (wild-type T7 RNAP) in XL1-Blue MR, and replica-plating of ampicillin/chloramphenicol-resistant colonies onto plates supplemented with tetracycline (10 μg mL$^{-1}$) did not reveal any surviving bacteria over a period of three days.

**Selection experiments:** Competent cells of XL1-Blue MR/pACYC-T7Nontet (100 μL) were co-transformed with pPolStar (T7 RNAP mutant library; ca. $10^9$ molecules; originating from mutator strain passage). An overnight culture of ampicillin/chloramphenicol-resistant transformants was diluted 100-fold into 2 L Luria – Bertani (LB) medium containing standard ampicillin and chloramphenicol concentrations, as well as tetracycline (10 μg mL$^{-1}$), and allowed to incubate. Tetracycline-resistant cells were grown until the stationary phase after 15 h of incubation. Both plasmids were isolated and analyzed by sequencing of the T7 promoter – tetracycline resistance gene cassette (pACYC-T7Nontet) or the mutant T7 RNAP gene (pPolStar).

**Determination of transcriptional error rates in vitro:** Run-off transcription with either wild-type T7 RNAP or selected mutant polymerase was performed with linear template pAlterGFP (6 μg) and enzyme (2 μg) by using standard conditions.[43] RNA of discrete length was purified by electrophoresis (1.2% (w/v) agarose), recovered by electroelution, and reverse-transcribed with Superscript Reverse Transcriptase (Gibco) and primer 5′-GCT TTG TTA GCA GCC GGA TCC TCA GTT GTA CAG TTC ATC CAT G-3′ (restriction site underlined). Second-strand synthesis was achieved by adding *Vent Exo*$^-$ polymerase (New England Biolabs) as well as primer 5′-CTT TAA GAA GGA GAT CTG CAG ATG GCT AGC AAA GGA GAA G-3′ (restriction site underlined), and followed by PCR amplification of double-stranded cDNA using *Pfu* polymerase (Stratagene). The PCR product was ligated into pUC18, and the ligation reaction was used to transform XL1-Blue MR. Plasmids from individual clones were purified and submitted to sequence analysis of the respective GFP fragment.

**Influence of the transcriptional error rate on phage replication:** A logarithmic-phase culture of selected bacteria XL1-Blue MR/pACYC-T7Nontet/pPolStar was infected with wild-type T7 phage[30] at a multiplicity of infection of 0.1, and incubation was continued. The time course of bacterial lysis was followed by monitoring the optical density at 600 nm ($OD_{600}$) in 30-min intervals. Analogous experiments were performed in parallel with XL1-Blue MR (no plasmid), XL1-Blue MR/pPolHisTA (wild-type T7 RNAP), and XL1-Blue MR/pACYC-T7Nontet (plasmid; no polymerase encoded).

**Assay of transcriptional error rates in vivo:** The T7 promoter – rsGFP gene cassette from plasmid pQBI63 (ColE1 *ori*; Quantum Biotechnologies) was ligated into plasmid pAlter-Ex2 (p15A *ori*; Promega). Co-transformation of the resulting plasmid pAlterGFP with pAR1219 (wild-type T7 RNAP) yielded transformants that exhibited an intense green color upon irradiation with UV light (352 nm; excitation maximum: 473 nm; emission maximum: 509 nm). The distribution of rsGFP fluorescence among large numbers of clones was monitored in a 96-well format by using a fluorescence microplate reader (Tecan Spectra Fluor Plus): Microtiter plates with 250-μL wells and planar bottom (Costar 3595) were filled with LB medium containing appropriate antibiotics, inoculated with individual cultures of XL1-Blue MR/pAlterGFP/pAR1219, and incubated for 15 h. Absorbance was determined at 590 nm, and fluorescence emission was measured at 535 nm (excitation at 485 nm). Analogous experiments were performed by using plasmid pPolStar (T7 RNAP mutant) instead of pAR1219. Fluorescence data were normalized by the respective cell densities.

**Nucleotide sequence analysis:** Dideoxy-sequencing was performed by using the *Taq* dye terminator FS sequencing kit (PE Biosystems) with resolution of the products on an ABI 373A sequencer.

*We are greatly indebted to Manfred Eigen for inspiring these studies on error-prone polymerases and also for providing the necessary intellectual and physical environment. We are grateful to Christof K. Biebricher and Andreas Schwienhorst for many helpful discussions and suggestions. Our work was supported by the Max-Planck-Gesellschaft, BMBF grant 0311003 together with Roche Diagnostics, and DFG grant Ei 411/1-1.*

[1] E. Domingo, E. Matínez-Salas, F. Sobrino, J. C. De La Torre, J. Ortín, C. López-Galindez, P. Pérez-Breña, N. Villanueva, R. Nájera, S. Van De Pol, D. Steinhauer, N. De Polo, J. Holland, *Gene* **1985**, *40*, 1 – 8.
[2] J. J. Holland, K. Spindler, F. Horodyski, E. Grabau, S. Nichol, S. Van De Pol, *Science* **1982**, *215*, 1577 – 1585.
[3] E. Domingo, J. J. Holland, *Annu. Rev. Microbiol.* **1997**, *51*, 151 – 178.
[4] M. Eigen, *Naturwissenschaften* **1971**, *58*, 465 – 523.
[5] M. Eigen, *Gene* **1993**, *135*, 37 – 47.
[6] C. K. Biebricher, E.-M. Düker, *J. Gen. Microbiol.* **1984**, *130*, 941 – 949.
[7] *Quasispecies and RNA evolution: Principles and consequences* (Eds.: E. Domingo, C. K. Biebricher, M. Eigen, J. J. Holland), Landes Bioscience, Orlando, in press.
[8] S. Brakmann, *Biophys. Chem.* **1997**, *66*, 133 – 144.
[9] M. Eigen, E. Domingo, J. Holland, A. Schwienhorst, C. Biebricher, B. Lindemann (Diagen GmbH), WO-A 94/01545 (PCT/EP93/01711), **1994**.
[10] C. Escarmís, M. Dávila, N. Charpentier, A. Bracho, A. Moya, E. Domingo, *J. Mol. Biol.* **1996**, *264*, 255 – 267.
[11] L. A. Loeb, J. M. Essigman, F. Kazazi, J. Zhang, K. D. Rose, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1492 – 1497.
[12] M. M. Slupska, C. Baikalov, R. Lloyd, J. H. Miller, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4380 – 4385.
[13] F. W. Studier, B. A. Moffatt, *J. Mol. Biol.* **1986**, *189*, 113 – 130.
[14] D. Patra, E. M. Lafer, R. Sousa, *J. Mol. Biol.* **1992**, *224*, 307 – 318.
[15] P. A. Osumi-Davis, M. C. de Aguilera, R. W. Woody, A.-Y. M. Woody, *J. Mol. Biol.* **1992**, *226*, 37 – 45.
[16] R. Ikeda, L. L. Chang, G. S. Warshamana, *Biochemistry* **1993**, *32*, 9115 – 9124.
[17] V. L. Rechinsky, D. A. Kostyuk, D. L. Lyakhov, B. K. Chernov, S. N. Kochetkov, *Mol. Gen. Genet.* **1993**, *238*, 455 – 458.
[18] D. L. Lyakhov, B. He, X. Zhang, F. W. Studier, J. J. Dunn, W. T. McAllister, *J. Mol. Biol.* **1997**, *269*, 28 – 40.
[19] X. Zhang, F. W. Studier, *J. Mol. Biol.* **1995**, *250*, 156 – 168.
[20] L. Gross, W. J. Chen, W. T. McAllister, *J. Mol. Biol.* **1992**, *228*, 488 – 505.
[21] R. Sousa, Y. J. Chung, J. P. Rose, B. C. Wang, *Nature* **1993**, *364*, 593 – 599.
[22] D. Jeruzalmi, T. A. Steitz, *EMBO J.* **1998**, *17*, 4101 – 4113.
[23] G. M. T. Cheetham, T. A. Steitz, *Science* **1999**, *286*, 2305 – 2309.
[24] D. Schnappinger, W. Hillen, *Arch. Microbiol.* **1996**, *165*, 359 – 369.
[25] P. McNicholas, I. Chopra, D. M. Rothstein, *J. Bacteriol.* **1992**, *174*, 7926 – 7933.
[26] P. McNicholas, M. McGlynn, G. G. Guay, D. M. Rothstein, *J. Bacteriol.* **1995**, *177*, 5355 – 5357.
[27] R. C. Cadwell, G. F. Joyce, *PCR Protoc. Guide Methods Appl.* **1992**, *2*, 28 – 33.
[28] A. Greener, M. Callaghan, B. Jerpseth, *Mol. Biotechnol.* **1997**, *7*, 189 – 195.
[29] J. Grodberg, J. J. Dunn, *J. Bacteriol.* **1988**, *170*, 1245 – 1253.
[30] J. Yin, *J. Bacteriol.* **1993**, *175*, 1272 – 1277.
[31] F. W. Studier, J. Dunn, *Cold Spring Harb. Symp. Quant. Biol.* **1983**, *47*, 999 – 1007.
[32] G. N. Phillips, Jr., *Curr. Opin. Struct. Biol.* **1997**, *7*, 821 – 827.
[33] A. B. Cubitt, L. A. Woollenweber, R. Heim, *Methods Cell Biol.* **1999**, *58*, 19 – 29.
[34] T. Ehrig, D. J. O'Kane, F. G. Prendergast, *FEBS Lett.* **1995**, *367*, 163 – 166.

[35] J. C. Boyer, K. Bebenek, T. A. Kunkel, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6919 – 6923.
[36] K. M. Remington, S. E. Bennett, C. M. Harris, T. M. Harris, K. Bebenek, *J. Biol. Chem.* **1998**, *273*, 13170 – 13176.
[37] U. Englisch, D. Gauss, W. Freist, S. Englisch, H. Sternbach, F. von der Haar, *Angew. Chem.* **1985**, *97*, 1033 – 1043; *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 1015 – 1025.
[38] R. C. Cadwell, G. F. Joyce, *PCR Protoc. Guide Methods Appl.* **1994**, *3*, 136 – 140.
[39] V. L. Tunitskaya, A. K. Akbarov, S. V. Luchin, L. V. Memelova, V. O. Rechinsky, S. N. Kochetkov, *Eur. J. Biochem.* **1990**, *191*, 99 – 103.
[40] D. K. Muller, C. T. Martin, J. E. Coleman, *Biochemistry* **1988**, *27*, 5763 – 5771.
[41] B. He, M. Rong, R. K. Durbin, W. T. McAllister, *J. Mol. Biol.* **1997**, *265*, 275 – 288.
[42] K. A. Mookthiar, P. S. Peluso, D. K. Muller, J. J. Dunn, J. E. Coleman, *Biochemistry* **1991**, *30*, 6305 – 6313.
[43] P. Davanloo, A. H. Rosenberg, J. J. Dunn, F. W. Studier, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 2035 – 3039.
[44] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33 – 38.
[45] T vectors are linear blunt-ended plasmids with a few dT extensions added on by Taq polymerase. Apparently, Taq adds dA overhangs to the ends of PCR products. Thus, most of the PCR products have one or a few dA extensions on the 5′ ends and can be ligated into the T vectors. The whole method is known as TA cloning.

Received: August 23, 2000 [F121]