

Replacing 32 Proline Residues by a Noncanonical Amino Acid Results in a Highly Active DNA Polymerase

Bastian Holzberger and Andreas Marx*

Department of Chemistry and Konstanz Research School Chemical Biology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany

Received July 26, 2010; E-mail: andreas.marx@uni-konstanz.de

Abstract: Protein engineering may be achieved by rational design, directed evolution-based methods, or computational protein design. Mostly these methods make recourse to the restricted pool of the 20 natural amino acids. With the ability to introduce different new kinds of functionalities into proteins, the use of noncanonical amino acids became a promising new method in protein engineering. Here, we report on the generation of a multifluorinated DNA polymerase. DNA polymerases are highly dynamic enzymes that catalyze DNA synthesis in a template-dependent manner, thereby passing several conformational states during the catalytic cycle. Here, we globally replaced 32 proline residues by the noncanonical imino acid (4*R*)-fluoroproline in a DNA polymerase of 540 amino acids (KlenTaq DNA polymerase). Interestingly, the substitution level of the proline residues was very efficient (92%). Nonetheless, the introduction of (4*R*)-fluoroproline into the DNA polymerase resulted in a highly active fluorinated enzyme, which was investigated in primer extension and PCR assays to analyze activity, selectivity, and stability in comparison to the parental enzyme. The DNA polymerase retained fidelity, activity, and sensitivity as the parental wild-type enzyme accompanied by some loss in thermostability. These results demonstrate that a noncanonical amino acid can be used for substitutions of natural counterparts in a highly dynamic enzyme with high molecular weight without effecting crucial enzyme properties. Furthermore, the employed DNA polymerase represents a promising starting point for directed DNA polymerase evolution with noncanonical amino acids.

Introduction

Protein engineering is mainly achieved by rational design, directed evolution-based methods, or computational protein design.^{1–5} Mostly these engineering strategies make recourse to the restricted pool of the 20 canonical amino acids. With the ability to introduce different new kinds of functionalities into proteins, the use of noncanonical amino acids became a promising new method in protein engineering. Especially in the past decade it has been shown that noncanonical amino acids can be effectively incorporated into proteins to expand the genetic code.^{6–12} In general, it is possible to either introduce noncanonical amino acids site-specifically at single distinct

positions^{6–9} or replace one natural amino acid in complete by a noncanonical amino acid analogue.^{10–12} Thereby, the global replacement of one amino acid all over the protein always runs the risk to disrupt protein stability and function, but it has also the ability to markedly alter properties like protein folding or stability that are often based on synergistic effects at multiple sites.¹³ Furthermore, it is possible to specifically and carefully adjust the characters of amino acid modifications by introducing atomic mutations or new functional groups. On the basis of the use of auxotrophic *E. coli* strains unable to synthesize one specific amino acid on their own, a noncanonical analogue can be incorporated in response to the respective sense codon.¹⁴ Doing so, the creation of novel proteins with chemical diversity not found in nature and noncanonical physicochemical properties seems to be feasible.^{6–12}

In terms of protein folding and stability, the use of fluorinated proline (Pro) residues has shown interesting results. Stereoelectronic effects of the C4-substituent directly influence the trans/cis ratio of the prolyl peptide bond and the pucker of the pyrrolidine ring, favoring, for example, in case of (4*R*)-fluoroproline ((4*R*)-FPro) a C4-exo puckering, whereas in case of (4*S*)-FPro the endo puckering is favored (Figure 1).^{15–18} These effects have been extensively studied to investigate the conformation, folding, and stability of different peptides and

- (1) Jäckel, C.; Kast, P.; Hilvert, D. *Annu. Rev. Biophys.* **2008**, *37*, 153–173.
- (2) Bloom, J. D.; Meyer, M. M.; Meinhold, P.; Otey, C. R.; MacMillan, D.; Arnold, F. H. *Curr. Opin. Struct. Biol.* **2005**, *15*, 447–452.
- (3) Brannigan, J. A.; Wilkinson, A. J. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 964–970.
- (4) Arnold, F. H. *Nature* **2001**, *409*, 253–257.
- (5) Van der Sloot, A. M.; Kiel, C.; Serrano, L.; Stricher, F. *Protein Eng., Des. Sel.* **2009**, *22*, 537–542.
- (6) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441–444.
- (7) Neumann, H.; Slusarczyk, A. L.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132*, 2142–2144.
- (8) Wang, L.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2005**, *44*, 34–66.
- (9) Ryu, Y.; Schultz, P. G. *Nat. Methods* **2006**, *3*, 263–265.
- (10) Budisa, N. *Engineering the Genetic Code*; Wiley-VCH: Weinheim, Germany, 2006.
- (11) Budisa, N. *Angew. Chem., Int. Ed.* **2004**, *43*, 6426–6463.
- (12) Link, A. J.; Mock, M. L.; Tirrell, D. A. *Curr. Opin. Biotechnol.* **2003**, *14*, 603–609.

- (13) Moroder, L.; Budisa, N. *ChemPhysChem* **2010**, *11*, 1181–1187.
- (14) Minks, C.; Alefelder, S.; Moroder, L.; Huber, R.; Budisa, N. *Tetrahedron* **2000**, *56*, 9431–9442.

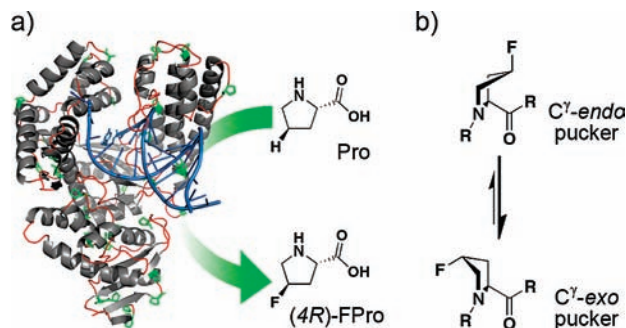


Figure 1. (a) KlenTaq DNA polymerase (PDB code: 3KTQ). Loops are shown in red, while Pro residues are marked with green sticks. The DNA primer-template complex is blue. (b) Ring conformations of (4R)-FPro: The C^γ-exo conformation is favored strongly by stereoelectronic effects.

proteins^{19–25} and gave rise to the assumption that the bias of 4-FPro to adopt either an exo or an endo puckering can have antagonistic effects on peptide stability depending on the puckering and the structural context in the natural state. As a breaker of β -sheet and α -helical structures, natural Pro residues are mainly found in turns and kinks and are rarely involved in substrate binding or active sites.²⁶ Thus, we considered to apply the global replacement of Pro by 4-FPro for enzyme engineering, as highly dynamic enzymes presumably possess proline residues that alter their puckering during catalysis. So far, 4-FPro has been used only in few studies in protein engineering.^{27–31} Along these lines, Budisa and co-workers recently introduced (4S)-FPro together with two other fluorinated amino acid analogues in one single experiment into a small enzyme.²⁷ Aside from the substitution of 16 Phe and two Trp residues by fluorinated analogues, six Pro residues were simultaneously replaced by (4S)-FPro in a lipase of 267 amino acids (30 kDa). The fluorinated lipase retained structural integrity and showed about 60% of the enzymatic activity of the parental enzyme. Nevertheless, further investigations along these lines are indispensable

to understand how to use noncanonical building blocks for highly sophisticated purposes like enzyme engineering. Detailed studies of large and highly dynamic naturally evolved enzymes are necessary to examine, if a high number of noncanonical amino acids spread all over the entire protein can be accommodated by the framework of these proteins. Furthermore, the influence of noncanonical amino acids on enzyme properties like stability, sensitivity, substrate selectivity, or activity has still to be investigated in more detail.

Herein, we present results derived from replacing all Pro residues of a DNA polymerase. DNA polymerases catalyze DNA synthesis in a template-dependent manner. During this process, they have to move along the template strand and to pass several conformational states during the catalytic cycle of DNA polymerization.^{32–34} Thus, they can be regarded as highly dynamic molecular machines. We were able to replace the 32 Pro residues of the thermophilic KlenTaq DNA polymerase, the large fragment of DNA polymerase I from *Thermus aquaticus*, almost quantitatively with a substitution level of approximately 92% with (4R)-FPro. With 32 Pro residues approximately 6% of the amino acids of the KlenTaq DNA polymerase (540 amino acids, 63 kDa) were substituted by the noncanonical imino acid (4R)-FPro. The natural Pro residues are mainly located in loops and at positions where α -helices change direction or pass into loop structures (Figure 1). Remarkably, despite the decrease of stability, the multi (4R)-FPro-labeled KlenTaq DNA polymerase was still fully enzymatically active. Aside from the activity, the enzyme displayed also similar selectivity and sensitivity as compared to the parental wild-type enzyme.

Materials and Methods

General. *H-trans*-4-Fluoroproline ((4R)-FPro) was from Bachem. DNA oligonucleotides were synthesized by Metabion or Purimex. Primers were further purified by preparative PAGE. dNTPs were purchased from Fermentas. VivaSpins were from Sartorius Stedim Biotech, Roti-Blue Colloidal Coomassie was from Carl Roth, Ni-IDA chelating sepharose fast flow was from GE Healthcare, QIAquick Gel Extraction Kit and MinElute Reaction Cleanup Kit were from Qiagen, peqGOLD Gel Extraction Kit was from Peqlab, restriction enzymes were from New England Biolabs, and Rapid DNA Ligation Kit was from Fermentas. Sequencing of isolated plasmids was performed by GATC. LB medium was from Roth (Lennox, LB-broth). Denaturing PAGE was analyzed with Molecular Imager FX (BioRad) and QuantityOne (BioRad). Real-time PCR was performed in Chromo4 instrument from BioRad. SYBRgreen I was from Fluka. ESI-IT-MS and HPLC-ESI-MS/MS were performed on a Bruker Daltonics esquire 3000+. HPLC was performed using a Vydac MS C18 column (Grace).

Expression and Purification of KlenTaq Wild-Type and (4R)-FPro-KlenTaq. The pGDR11 vector^{35,36} harboring the KlenTaq wild-type gene (amino acids 293–832 of *Taq* gene³⁷ cloned into pGDR11 using the restriction sites *Sph*I and *Hind*III) was transformed into *E. coli* JM83. KlenTaq wild-type was expressed

- (15) Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009**, *78*, 929–958.
- (16) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497–2505.
- (17) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem., Int. Ed.* **2001**, *40*, 923–925.
- (18) Eberhardt, E. S.; Panasiak, N.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261–12266.
- (19) Zheng, T. Y.; Lin, Y. J.; Horng, J. C. *Biochemistry* **2010**, *49*, 4255–4263.
- (20) Chiang, Y. C.; Lin, Y. J.; Horng, J. C. *Protein Sci.* **2009**, *18*, 1967–1977.
- (21) Boulègue, C.; Milbradt, A. G.; Renner, C.; Moroder, L. *J. Mol. Biol.* **2006**, *358*, 846–856.
- (22) Horng, J. C.; Raines, R. T. *Protein Sci.* **2006**, *15*, 74–83.
- (23) Kim, W.; McMillan, R. A.; Snyder, J. P.; Conticello, V. P. *J. Am. Chem. Soc.* **2005**, *127*, 18121–18132.
- (24) Shoulders, M. D.; Satyshur, K. A.; Forest, K. T.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 559–564.
- (25) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777–778.
- (26) Betts, M. J.; Russell, R. B. In *Bioinformatics for Geneticists*, 2nd ed.; Barnes, M. R., Ed.; Wiley-VCH: Weinheim, Germany, 2007.
- (27) Merkel, L.; Schauer, M.; Antranikian, G.; Budisa, N. *ChemBioChem* **2010**, *11*, 1505–1507.
- (28) Lepthien, S.; Merkel, L.; Budisa, N. *Angew. Chem., Int. Ed.* **2010**, *49*, 5446–5450.
- (29) Steiner, T.; Hess, P.; Bae, J. H.; Wiltschi, B.; Moroder, L.; Budisa, N. *PLoS One* **2008**, *3*, e1680.
- (30) Naduthambi, D.; Zondlo, N. J. *J. Am. Chem. Soc.* **2006**, *128*, 12430–12431.
- (31) Golbik, R.; Yu, C.; Weyher-Stingl, E.; Huber, R.; Moroder, L.; Budisa, N.; Schiene-Fischer, C. *Biochemistry* **2005**, *44*, 16026–16034.

- (32) Santoso, Y.; Joyce, C. M.; Potapova, O.; Le Reste, L.; Hohlbein, J.; Torella, J. P.; Grindley, N. D. F.; Kapanidis, A. N. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 715–720.
- (33) Allen, W. J.; Rothwell, P. J.; Waksman, G. *Protein Sci.* **2008**, *17*, 401–408.
- (34) Rothwell, P. J.; Waksman, G. *Adv. Protein Chem.* **2005**, *71*, 401–440.
- (35) Holzberger, B.; Rubini, M.; Möller, H. M.; Marx, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 1324–1327.
- (36) Peist, R.; Koch, A.; Bolek, P.; Sewitz, S.; Kolbus, T.; Boos, W. *J. Bacteriol.* **1997**, *179*, 7679–7686.
- (37) Barnes, W. M. *Gene* **1992**, *112*, 29–35.

as previously described.³⁵ (4*R*)-FPro-KlenTaq: A 400 mL culture was grown in M9 minimal medium (NH₄Cl 1 g/L, NaCl 0.5 g/L, KH₂PO₄ 3 g/L, Na₂HPO₄·2H₂O 8.5 g/L, MgSO₄ 2 mM, CaCl₂ 0.1 mM, glucose 22.2 mM, thiamine 10 mg/L, biotin 10 mg/L, carbenicillin 100 mg/L, all amino acids except Pro 50 mg/L) with 0.035 mM Pro until stationary phase was reached (OD₆₀₀ 0.5–0.7). The cells were additionally incubated for up to 6.5 h at 37 °C. Cells were harvested (3800 rpm, 12 min, 4 °C), washed with 200 mL of PBS buffer and resuspended in 400 mL of M9 minimal media without Pro. (4*R*)-FPro was added to a final concentration of 1 mM. After incubation for 20 min at 37 °C, protein synthesis was induced (1 mM IPTG) and carried out overnight. After harvesting (4500g, 30 min, 4 °C), cells were lysed (15 min, 37 °C) in 25 mL of lysis buffer (10 mM Tris·HCl pH 9.2, 300 mM NaCl, 2.5 mM MgCl₂, 0.1% Triton X 100, 1 mg/mL lysozyme) followed by heat denaturation of host proteins at 75 °C for 45 min and centrifugation at 20 000g for 30 min at 4 °C. Supernatants were incubated with Ni-IDA sepharose slurry and 5 mM imidazole. After washing with buffer (10 mM Tris·HCl pH 9.2, 300 mM NaCl, 2.5 mM MgCl₂, 0.1% Triton X 100, 20 mM imidazole), elution was carried out using elution buffer (100 mM Tris·HCl pH 9.2, 5 mM MgCl₂, 200 mM imidazole). The elution buffer was exchanged by storage buffer (elution buffer without imidazole) using VivaSpins. For storage, glycerol was added to 50%, (NH₄)₂SO₄ to a final concentration of 16 mM, and Tween 20 to 0.1%. Enzyme purity was controlled by SDS-PAGE. Protein concentrations were determined using Bradford assay. Purified enzymes were stored at –20 °C.

ESI Mass Spectrometry. For electrospray ionization mass spectrometry (ESI), the protein sample was desalted (VivaSpin) to 5 mM Tris·HCl pH 9.2 and 0.25 mM MgCl₂. After 1:1 dilution with 2% acetic acid in water/CH₃CN (1:1), the samples were directly injected, and data were analyzed with DataAnalysis from Bruker. To calculate the replacement level of Pro by (4*R*)-FPro, we used the molecular weights of (4*R*)-FPro-KlenTaq and KlenTaq wild-type, determined by the deconvoluted ESI-MS spectra (Figure 2). The difference of the molecular weights divided by the respective difference of a single Pro→(4*R*)-FPro substitution results in 30.4 (92%) Pro residues replaced by (4*R*)-FPro.

Proteolytic Digest and LC-ESI-MS/MS. For the MS/MS analysis of the peptide fragments, tryptic in-gel digest and in-solution digest with pepsin were performed. After resolution of the resulting peptide fragments in 0.2% formic acid in water, HPLC-ESI-MS/MS was performed. Data were analyzed using DataAnalysis from Bruker and Mascot Search.

Primer Extension Experiments. The primer was 5'-³²P-labeled using [γ -³²P]-ATP according to standard techniques. The reaction mixture (20 μ L) contained DNA polymerase, 200 nM template (5'-d(AAA TCA ACC TAT CCT CCT TCA GGA CCA ACG TAC)-3'), 150 nM radioactive labeled primer (5'-d(CGT TGG TCC TGA AGG AGG AT)-3'), and 200 μ M dNTPs in KlenTaq 1x reaction buffer (50 mM Tris·HCl pH 9.2, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% Tween 20). The reaction mixtures without dNTPs and DNA polymerase were annealed by heating for 4 min at 95 °C and stepwise cooling to 0 °C. Used DNA polymerase concentrations were (i) for reactions depicted in Figure 3a,b 300 pM and (ii) for reactions performed for activity assays (Figure 3c,d) 0.015, 0.025, 0.05, 0.15, 0.25, 0.3, and 0.5 nM. The reactions were incubated (i) for reactions depicted in Figure 3a,b for 10 min at 37.0, 37.8, 40.1, 43.7, 47.6, 51.5, 55.4, 59.4, 63.3, 66.9, 69.2, and 70.0 °C, respectively, and (ii) for reactions performed for activity assays (Figure 3c,d) for 10 min at 72 °C in a thermocycler and were stopped by addition of 40 μ L of stop solution (80% [v/v] formamide, 20 mM EDTA, 0.025% [w/v] bromophenol blue, 0.025% [w/v] xylene cyanol). After denaturing at 95 °C for 5 min, the reaction mixtures were separated using a 12% denaturing PAGE gel. Visualization was performed by phosphoimaging.

DNA Polymerase Activity Determination. For activity determination,³⁸ primer extension reactions were carried out with varying concentrations of DNA polymerase at 72 °C (see primer extension experiments). The obtained intensities for each band were transformed into dNTP conversion per minute. The amount of enzyme was plotted against dNTP conversion, and the resulting slopes represent the specific activities of KlenTaq DNA polymerases [dNTP conversion pol⁻¹ min⁻¹] (Figure 3c,d).

Thermostability at 95 °C. For thermostability determination, KlenTaq DNA polymerases (10 nM) were incubated in KlenTaq 1x reaction buffer (50 mM Tris·HCl pH 9.2, 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% Tween 20) at 95 °C. After different time periods, 10 μ L aliquots were removed and stored on ice. Primer extension reactions were performed as described above for 5 min at 72 °C. Seven microliters of the polymerase sample was combined with 7 μ L of Mastermix (containing primer (5'-d(CGT TGG TCC TGA AGG AGG ATA GG)-3'), template (5'-d(AAA TCA ACC TAT CCT CCT TCA GGA CCA ACG TAC)-3'), dNTPs in KlenTaq 1x reaction buffer) and stopped by addition of 28 μ L of stop solution. After denaturing at 95 °C for 5 min, the reaction mixtures were separated using a 12% denaturing PAGE gel. Visualization was performed by phosphoimaging.

Error Spectrum. The error spectrum of (4*R*)-FPro-KlenTaq was determined using a PCR based assay described by Patel et al.³⁹ In brief, the PCR reaction (100 μ L) contained 30 fmol plasmid harboring a gene encoding for KlenTaq wild-type, 500 nM each primer, 250 μ M dNTPs, and 80 nM DNA polymerase in 1x KlenTaq reaction buffer. After initial denaturation for 30 s at 95 °C, 35 cycles were performed with 20 s at 95 °C, 60 s at 65 °C, and 120 s at 72 °C. Products were purified using agarose gel electrophoresis and cloned into the pGDR11 vector.³⁶ The cloned vector was transformed into *E. coli* BL21(DE3), single colonies were randomly picked, and a stretch of 600 bp was sequenced for each isolated plasmid. The mutation frequency equals number of mutations per clone and per bp sequenced (600 bp). The error rate was calculated as mutation frequency per number of replication cycles. The number of replication cycles was calculated from the PCR yield determined by quantification after agarose gel electrophoresis. For (4*R*)-FPro-KlenTaq the number of replication cycles was 10.1, reflecting a DNA amplification of ~1100-fold.

PCR Experiments. PCR reactions (20 μ L) contained 12 pM of a 120-nt template (5'-d(AAA GCT CCT TTC TGA ATA TTG AGC TCA TCA GTG AGA AAA CGG CTG CAT ATT GGT GTC AAA GTG TCA CTG AAC TAA AGG CTG ACT TTC CAG ACA ACA TAA GTG TGA TTT AAC ATC TAA AAC)-3'), 500 nM each primer (5'-d(GTT TTA GAT GTT AAA TCA CAC TTA T)-3'; 5'-d(AAA GCT CCT TTC TGA ATA TTG AG)-3'), 250 μ M dNTPs, and 5, 3, 2, and 1 nM, respectively, DNA polymerase in 1x KlenTaq reaction buffer. After initial denaturation for 30 s at 95 °C, PCR cycles were performed with 20 s at 95 °C, 60 s at 55 °C, and 120 s at 72 °C. PCR products were analyzed by agarose gel electrophoresis in 0.5x TBE buffer.

Real-Time PCR Experiments. Real-time PCR reactions (20 μ L) contained a dilution series of a 120-nt DNA template (5'-d(AAA GCT CCT TTC TGA ATA TTG AGC TCA TCA GTG AGA AAA CGG CTG CAT ATT GGT GTC AAA GTG TCA CTG AAC TAA AGG CTG ACT TTC CAG ACA ACA TAA GTG TGA TTT AAC ATC TAA AAC)-3') with 10 nM, 2 nM, 400 pM, 80 pM, and 16 pM, respectively, 200 nM each primer (5'-d(GTT TTA GAT GTT AAA TCA CAC TTA T)-3'; 5'-d(AAA GCT CCT TTC TGA ATA TTG AG)-3'), 250 μ M dNTPs, 0.6x SYBRgreen I, and 15 nM DNA

(38) Rudinger, N. Z.; Kranaster, R.; Marx, A. *Chem. Biol.* **2007**, *14*, 185–194.

(39) Patel, P. H.; Kawate, H.; Adman, E.; Ashbach, M.; Loeb, L. A. *J. Biol. Chem.* **2001**, *276*, 5044–5051.

polymerase in 1x KlenTaq reaction buffer. Thirty cycles were performed with 5 s at 95 °C, 20 s at 55 °C, and 30 s at 72 °C.

Results

Incorporation of (4R)-FPro into the KlenTaq DNA Polymerase. To incorporate (4R)-FPro in KlenTaq DNA polymerase, we used the Pro-auxotrophic *E. coli* strain JM83. Initially, cells were grown in a defined minimal medium under a limited concentration of natural Pro to circumvent any toxic effects of the noncanonical analogue. After depletion of natural Pro, cells were incubated for several hours in stationary phase and afterward harvested, washed, and resuspended in minimal medium without proline. Induction of protein expression by simultaneously adding (4R)-FPro led to the incorporation of the noncanonical analogue. Interestingly, attempts to incorporate (4S)-FPro failed, as the level of protein expression decreased dramatically (Figure S1). The yield (0.2–0.5 mg L⁻¹) of (4R)-FPro-KlenTaq was lower than that of the wild-type enzyme (8 mg L⁻¹), which was expressed as previously reported in standard LB medium.³⁵ Both N-terminally His-tagged enzymes were afterward purified by affinity chromatography (Figure S1). To analyze if the Pro residues were quantitatively substituted by (4R)-FPro, we performed electrospray-ionization (ESI) mass spectrometry of the modified and the wild-type enzyme (Figures 2, S2, and S3). By comparison of the molecular weights, it turned out that on average approximately 92% of the 33 Pro imino acids (32 from KlenTaq gene, 1 Pro from the N-terminal vector sequence) were replaced by (4R)-FPro. To further verify the sites and fidelity of (4R)-FPro incorporation, we performed proteolytic digests using trypsin and pepsin, respectively. LC-ESI-MS analysis of the resulting peptide fragments allowed us to identify 21 of the 33 Pro positions by MS/MS analysis (Table S1). Apart from position P701, which was only detected as Pro, for the remaining 20 residues we found always the (4R)-FPro modified peptide whereas in 11 of them natural Pro was detectable, too.

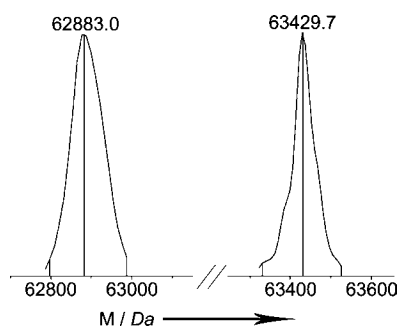


Figure 2. Deconvoluted ESI mass spectra profiles of KlenTaq wild-type (62 883.0 Da, calcd: 62 829.8 Da) and (4R)-FPro-KlenTaq (63 429.7 Da). *M*: molecular mass.

Activity, Fidelity, and Thermostability. Next, we performed primer extension experiments at various temperatures (Figure 3a). It turned out that the KlenTaq DNA polymerase containing the noncanonical imino acid (4R)-FPro ((4R)-FPro-KlenTaq) was enzymatically active. Similar to the wild-type enzyme (4R)-FPro-KlenTaq was able to extend the primer strand to full-length (30 nt) at the optimal temperatures between 50 and 65 °C, showing a temperature optimum similar to that of the wild-type enzyme (Figure 3b). To analyze the activity of the DNA polymerase in more detail, we performed primer extension reactions with various DNA polymerase concentrations at 72 °C (Figure 3c). Doing so, it was possible to determine the

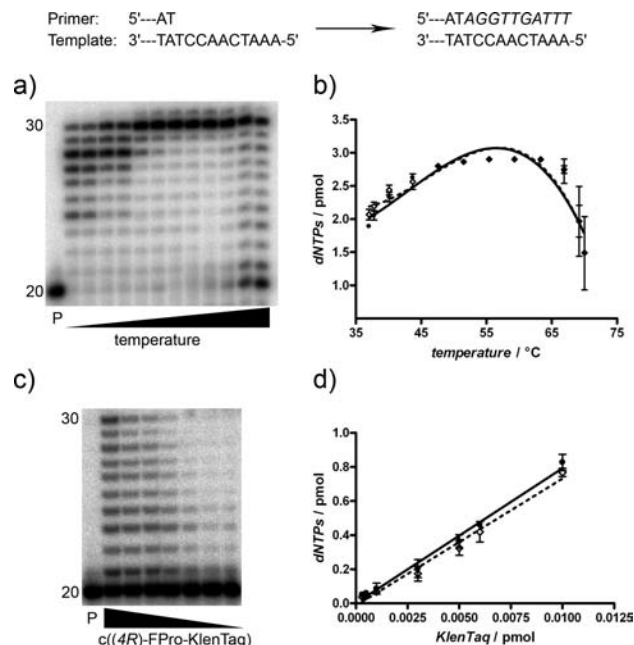


Figure 3. Primer extension activity of (4R)-FPro-KlenTaq. Partial primer template sequences employed are depicted on top. (a,b) Primer extension reactions at various temperatures catalyzed by (4R)-FPro-KlenTaq. (a) P, primer only (20 nt); temperature, primer extension at 37.0, 37.8, 40.1, 43.7, 47.6, 51.5, 55.4, 59.4, 63.3, 66.9, 69.2, and 70.0 °C. (b) quantified dNTP conversion (mean \pm s.d.) at various temperatures: (4R)-FPro-KlenTaq (\diamond ; dashed line) and KlenTaq wild-type (\blacklozenge ; solid line). (c,d) Primer extension reactions in the presence of various KlenTaq DNA polymerase concentrations. (c) P; primer only (20 nt); c((4R)-FPro-KlenTaq), primer extension in the presence of 0.5, 0.3, 0.25, 0.15, 0.05, 0.025, and 0.015 nM (4R)-FPro-KlenTaq. (d) Quantified dNTP conversion per minute (mean \pm s.d.) of (4R)-FPro-KlenTaq and KlenTaq wild-type using various KlenTaq DNA polymerase concentrations. Calculated DNA polymerase activities at 72 °C: (4R)-FPro-KlenTaq (\diamond ; dashed line), 74.5 ± 3.2 ; KlenTaq wild-type (\blacklozenge ; solid line), 79.1 ± 3.1 ³⁵ [dNTPs $\text{pol}^{-1} \text{min}^{-1}$].

specific activity of (4R)-FPro-KlenTaq, represented by the linear slope of dNTP conversion per min in presence of various amounts of DNA polymerase (Figure 3d). With 74.5 dNTPs $\text{pol}^{-1} \text{min}^{-1}$, the specific activity of (4R)-FPro-KlenTaq was in the same order of magnitude as the wild-type activity.³⁵

The error spectrum of (4R)-FPro-KlenTaq was investigated employing a reported PCR-based assay.³⁹ In short, a 1657 bp fragment encoding for KlenTaq wild-type was amplified followed by subsequent cloning of the PCR product and sequencing of several clones. In comparison, the error rate of (4R)-FPro-KlenTaq was quite similar to that found for KlenTaq wild-type⁴⁰ and TFM-KlenTaq, a previously reported KlenTaq variant with methionine (Met) substituted by trifluoromethionine (TFM)³⁵ (Table 1). Thus, the KlenTaq variants are at least as selective in PCR amplification as the wild-type enzyme. Transitions are thereby the predominant errors in every case. Interestingly, in case of KlenTaq wild-type, deletions are the second most occurring errors, whereas (4R)-FPro-KlenTaq and TFM-KlenTaq do not show this kind of error in the analyzed clones.

As (4R)-FPro is known to affect peptide and protein stability, we also studied the thermostability of (4R)-FPro-KlenTaq. Therefore, the enzyme was incubated for different time periods at 95 °C. Afterward, we examined the activity in primer extension reactions. We observed that (4R)-FPro-KlenTaq lost

(40) Gloeckner, C.; Sauter, K. B. M.; Marx, A. *Angew. Chem., Int. Ed.* **2007**, *46*, 3115–3117.

Table 1. Error Rate and Error Spectrum of (4*R*)-FPro-KlenTaq, TFM-KlenTaq, and KlenTaq Wild-Type

KlenTaq	no. of analyzed clones	no. of mutations per clone ^a	error rate ^b [$\times 10^{-5}$]	no. of individual substitutions		no. of deletions	no. of insertions
				transitions A \leftrightarrow G/C \leftrightarrow T	transversions T \leftrightarrow G/C \leftrightarrow A T \leftrightarrow A/C \leftrightarrow G		
(4 <i>R</i>)-FPro	28	0.2	3.5	6	0	0	0
TFM ^c	27	0.2	3.2	3	1	0	1
wild-type ^c	32	0.4	8.8	9	1	3	0

^a Average number of mutations per 600 bases sequenced per clone. ^b Error rate equals number of mutations per base per division. ^c Data were taken from ref 35.

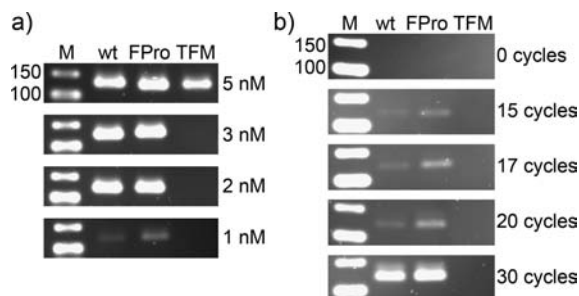


Figure 4. PCR amplification of a 120 bp DNA fragment using KlenTaq wild-type (wt), (4*R*)-FPro-KlenTaq (FPro), and TFM-KlenTaq (TFM) (a) in the presence of 5, 3, 2, and 1 nM DNA polymerase (30 cycles) and (b) in the presence of 2 nM DNA polymerase after performing either 0, 15, 17, 20, or 30 PCR cycles. M: Marker [bp].

stability in comparison to the wild-type. Nevertheless, the (4*R*)-FPro-KlenTaq was still stable enough to maintain more than 50% primer extension activity after incubation for 60 min at 95 °C (Figure S4).

PCR Activity. To investigate the PCR activity of (4*R*)-FPro-KlenTaq in comparison to TFM-KlenTaq and the wild-type enzyme, we performed PCR experiments on a 120-nt DNA template under different DNA polymerase concentrations and by employing different numbers of PCR cycles, respectively (Figure 4). Contrary to TFM-KlenTaq, which was only able to amplify the desired DNA fragment at DNA polymerase concentrations of at least 5 nM, the (4*R*)-FPro-KlenTaq was at least as active as the wild-type enzyme even at lower concentrations (Figure 4a). By performing limited numbers of PCR cycles in the presence of 2 nM DNA polymerase, it turned out that both (4*R*)-FPro-KlenTaq and the wild-type enzyme do not show any saturation of product formation before 30 cycles are completed (Figure 4b). Interestingly, in the presence of 1 nM DNA polymerase and under limited number of PCR cycles, the (4*R*)-FPro-KlenTaq seems to be slightly more active than the wild-type, generating apparently more DNA product (Figure 4).

To further analyze the PCR activity of (4*R*)-FPro-KlenTaq, we performed real-time PCR experiments using SYBRgreen I and a 120-nt DNA template, which was 5-fold diluted stepwise from 10 nM to 16 pM concentration of template (Figure 5). Analyzing the resulting real-time PCR amplification curves, the (4*R*)-FPro-KlenTaq shows a behavior similar to that of the wild-type enzyme. Both of them are able to amplify the DNA fragment even at comparable low template concentrations (Figure 5a and b), whereas the TFM-KlenTaq only shows an amplification in the presence of a high DNA template concentration (Figure 5c). By comparing the threshold-crossing points (Ct) between (4*R*)-FPro-KlenTaq and the wild-type enzyme, we

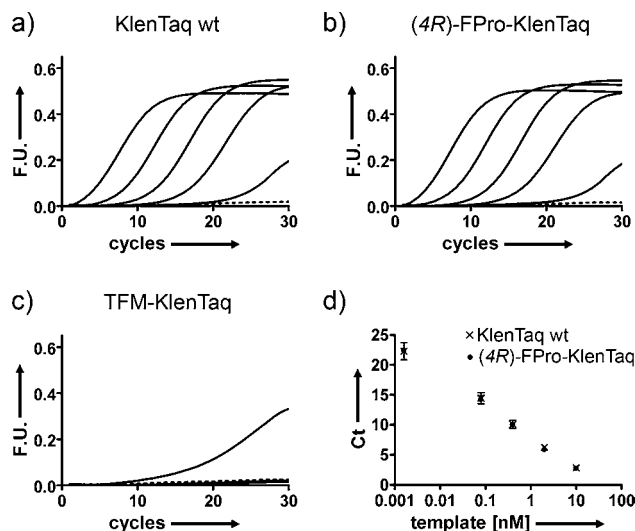


Figure 5. Real-time PCR activity of (4*R*)-FPro-KlenTaq compared to KlenTaq wild-type (wt) and TFM-KlenTaq. (a–c) Real-time PCR curves employing a 5-fold template dilutions series using (a) KlenTaq wild-type, (b) (4*R*)-FPro-KlenTaq, and (c) TFM-KlenTaq including a negative control without template (dashed line). Used DNA template concentrations were 10 nM, 2 nM, 400 pM, 80 pM, and 16 pM. (d) Ct values versus DNA template start concentration (mean \pm s.d.).

found very similar Ct values, demonstrating the same PCR sensitivity of (4*R*)-FPro-KlenTaq and KlenTaq wild-type (Figure 5d).

Discussion and Conclusions

By using the selective pressure incorporation method,¹⁴ we were able to replace the 32 Pro residues of the KlenTaq DNA polymerase by the noncanonical imino acid (4*R*)-FPro almost quantitatively. In contrast, the respective diastereomer (4*S*)-FPro yields no significant amounts of DNA polymerase (Figure S1). This effect appears to be protein specific because it has been shown for another protein (EGFP) that (4*S*)-FPro may be used to complement Pro in selective pressure incorporation.²⁹ However, EGFP is considerably smaller as the herein studied KlenTaq DNA polymerase bearing only 10 Pro residues. The authors report that 9 out of 10 (4*S*)-FPro residues adopt the endo conformation in the crystalline state. This preference might interfere with proper protein folding of nascent KlenTaq DNA polymerase, leading to the observed low yield of protein expression.

Here, (4*R*)-FPro was incorporated site-specifically and apparently equally distributed all over the enzyme with a substitution level of approximately 92%. Thereby, the MS/MS data of proteolytic digests provided at least an indication that every Pro residue is replaced to some extent. Under the basic assumption of an equal distribution of (4*R*)-FPro with every Pro residue replaced by (4*R*)-FPro with a probability of 92%, only less than

0.02% KlenTaq DNA polymerase molecules would exhibit less than 24 (4R)-FPro residues according to simple theory of probabilities. As expected, (4R)-FPro affected the stability of the enzyme and caused a loss of thermostability at 95 °C. Nevertheless, the decrease in stability did not influence any other properties of the enzyme. Even the activity of the modified DNA polymerase at 72 °C was unaltered as compared to the activity of the wild-type enzyme. Thus, it supersedes the activity of the previously reported multifluorinated KlenTaq DNA polymerase with all 13 methionines globally replaced by TFM that showed only about one-third of the activity of the wild-type enzyme.³⁵ Moreover, the error rate of the DNA polymerase, the mutation frequency per number of replications, was similar to the wild-type fidelity. As a result, sufficient stability and high activity were also reflected in PCR experiments. The (4R)-FPro-KlenTaq was similar to the wild-type enzyme able to amplify DNA fragments at limited DNA polymerase concentrations and showed also an unaltered high PCR sensitivity in real-time PCR reactions using a template dilution series. Thus, the (4R)-FPro-KlenTaq clearly differs from the previously reported TFM-KlenTaq, displaying only minor activity in PCR experiments. The multifluorinated (4R)-FPro-KlenTaq is a highly active and thermophilic DNA polymerase that would be still adequate for biotechnological applications as the crucial properties of DNA polymerases like specific activity, fidelity, and sensitivity are not affected.

Including our previous findings with TFM-KlenTaq,³⁵ the KlenTaq DNA polymerase is apparently able to tolerate the incorporation of different size-demanding or destabilizing noncanonical amino acids, even when the natural counterpart is globally replaced all over the enzyme. Aside from the facts that none of the noncanonical amino acids are involved in the active site and that in this case most of the replaced Pro residues are located in loops, we assume that especially the remarkable stability of the wild-type enzyme accounts for the acceptance of the noncanonical amino acids. As the natural KlenTaq DNA polymerase is able to overcome several hours of incubation at 95 °C without compromising enzyme activity, local disturbances might be tolerated without affecting adjacent main-chain conformations and consequently without altering the activity of the enzyme. In summary, this clearly demonstrates that the global replacement of natural amino acids by noncanonical analogues is not limited to small proteins or enzymes. Even a larger and highly dynamic enzyme can tolerate noncanonical amino acids at multiple positions without losing enzymatic activity or fidelity.

DNA polymerases are the crucial workhorses in numerous molecular biological core technologies.⁴¹ Yet for several applications the characteristics of natural evolved DNA polymerases are not satisfying. Thus, several modified DNA polymerases with new characteristics have been developed by directed evolution, a promising method to create nucleic acid polymerases with modified properties.^{42–47} Nevertheless, as noncanonical amino acids in general could be also very promising tools for DNA polymerase engineering,⁴⁸ our studies may open up new opportunities for investigations like the directed evolution of DNA polymerases with new characteristics due to the expanded amino acid repertoire.^{49–51}

Acknowledgment. We gratefully acknowledge Dr. M. Rubini for her contributions to this work, Dr. A. Marquardt (Proteomics Facility, University of Konstanz) for the support with the MS analysis, and the Konstanz Research School Chemical Biology for partial financial support. B.H. thanks the Carl-Zeiss-Stiftung for granting a stipend.

Supporting Information Available: SDS-PAGE gel of purified KlenTaq DNA polymerases, ESI mass spectra profiles of KlenTaq wild-type and (4R)-FPro-KlenTaq, DNA polymerase thermostabilities at 95 °C, and ESI-MS/MS data from proteolytic digests. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA106525Y

- (41) *Molecular Cloning: A Laboratory Manual*; Sambrook, J., Russell, D. W., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.
- (42) Loakes, D.; Holliger, P. *Chem. Commun.* **2009**, *31*, 4619–4631.
- (43) Kranaster, R.; Marx, A. *Angew. Chem., Int. Ed.* **2009**, *48*, 4625–4628.
- (44) Sauter, K. B. M.; Marx, A. *Angew. Chem., Int. Ed.* **2006**, *45*, 7633–7635.
- (45) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Biotechnol.* **2005**, *16*, 370–377.
- (46) Brakmann, S. *Cell. Mol. Life Sci.* **2005**, *62*, 2634–2646.
- (47) Davidson, J. F.; Anderson, J.; Guo, H.; Landis, D.; Loeb, L. A. In *Directed Molecular Evolution of Proteins*; Brakmann, S., Johnsson, K., Eds.; Wiley-VCH: Weinheim, Germany, 2002.
- (48) Chou, C.; Young, D. D.; Deiters, A. *Angew. Chem., Int. Ed.* **2009**, *48*, 5950–5953.
- (49) Liu, C. C.; Mack, A. V.; Tsao, M.-L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17688–17693.
- (50) Yoo, T. H.; Link, A. J.; Tirrell, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 13887–13890.
- (51) Montclare, J. K.; Tirrell, D. A. *Angew. Chem., Int. Ed.* **2006**, *45*, 4518–4521.