Protein Engineering

DOI: 10.1002/anie.200900953

Taking Fingerprints of DNA Polymerases: Multiplex Enzyme Profiling on DNA Arrays**

*Ramon Kranaster and Andreas Marx**

DNA polymerases are used in a plethora of biotechnical applications, especially in the polymerase chain reaction (PCR), genetic cloning procedures, genome sequencing, and diagnostic methods.[1] Highly processive and accurate DNA polymerases are desired for cloning procedures in order to give shorter extension times as well as more robust and high-yield amplification. A higher DNA polymerase fidelity may increase the reliability of genome sequencing and diagnostic systems.[2] Amplification of ancient DNA samples requires DNA polymerases with an increased substrate spectrum to efficiently overcome typical DNA lesions.[3] To enhance the efficiency of forensic DNA testing, DNA polymerases resistant to inhibitors from blood and soil allow PCR without prior DNA purification.[4] Further improvements of DNA polymerases are required, for example, to meet the requirements of real-time DNA single-molecule sequencing, which relies on the ability of DNA polymerases to efficiently process modified nucleotides.[5] Overall, customized and artificially engineered DNA polymerases that lead to more robust and specific reaction systems are urgently needed.

Directed evolution holds promise for engineering nucleic acid polymerases with altered properties.[6] Alterations are mainly achieved by directed molecular evolution using genetic complementation and/or screening, phage display, or in vitro compartmentalization.[7–10] To our knowledge all reported methods for DNA polymerase evolution are restricted to a single enzyme property, for example, increased selectivity or the ability to efficiently process DNA lesions.[7–10] To overcome these obvious limitations we set out to develop devices that allow the multiplexed screening of several enzyme features in parallel.

Here we report on a chip-based screening format that allows simultaneous and multiplexed profiling of several enzyme features with high throughput. The system is based on the spatial separation of different covalently attached DNA substrates on a glass slide and their selective addressing by oligonucleotide hybridization. Standard microarray equipment is sufficient to conduct and quantify the reactions. The developed method is time and cost efficient, and requires only minimal amounts of reagents.

The principle behind our approach, termed oligonucleotide-addressing enzyme assay (OAEA), is depicted in Figure 1a. It mainly consists of two spotting steps and an incubation step. In the first step, 5′-NH_2-$(CH_2)_6$-modified DNA oligonucleotides are covalently attached in defined rows of spots on phenylenediisothiocyanate-activated glass slides.[11] These oligonucleotides act as primer strands in the DNA polymerase catalyzed reactions. Noteworthy, the slides

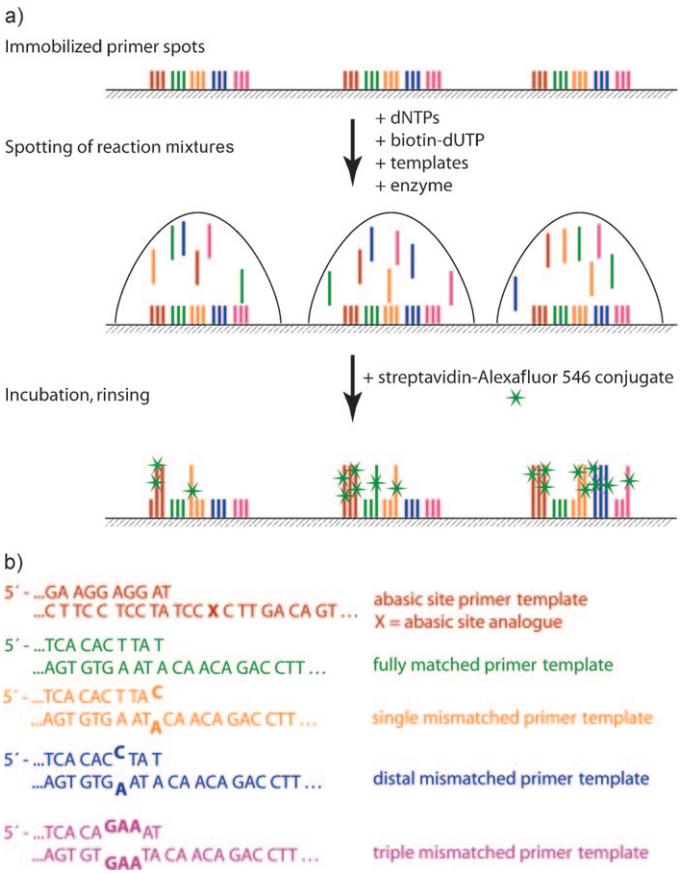

Figure 1. Principle of oligonucleotide-addressing enzyme assay (OAEA). a) Short colored strips represent immobilized primer strands; long colored strips represent templates, which are selectively addressed by hybridization with the complementary immobilized primer strands. Green stars represent streptavidin–Alexafluor546 conjugates which bind to incorporated biotin on the extended primer strands. b) Partial sequences used in this study.

[*] Dipl.-Chem. R. Kranaster, Prof. Dr. A. Marx
Department of Chemistry and
Konstanz Research School Chemical Biology
University of Konstanz
Universitätsstrasse 10, 78457 Konstanz (Germany)
Fax: (+49) 7531-88-5140
E-mail: andreas.marx@uni-konstanz.de

[**] We gratefully acknowledge funding by the DFG (SPP1170) and BMBF (BioChancePlus).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200900953.

can be stored at 4 °C for several weeks without compromising their usability. Then, in a second spotting step, enzyme mutants suspended in a buffer, which contains the respective DNA templates, dNTPs, and a biotin-dUTP derivative, are applied. The mixtures are distributed in nanoliter quantities on the same positions as the previously spotted primers. During this procedure the spots dry out. Subsequently the glass slides are incubated in a humidity chamber at ≈50 °C, which causes the spots to rehydrate. We found that drying and rehydration had little effect on the activity of the employed DNA polymerase (see below) after independent studies in solution (see the Supporting Information).

Rehydration of the spots creates separate reaction entities in which the respective DNA polymerase mutant is expected to process the primer–template duplexes covered by the respective spot. All reactions on the glass slide are stopped by repeated rinsing with an aqueous detergent solution. In case of primer extension, biotin-dUTP will be incorporated in the respective extended primer strand. A fluorescence signal is generated by incubation with a streptavidin–Alexafluor546 conjugate, which binds to the incorporated biotin (Figure 1 a).

To first evaluate our approach we investigated whether the depicted microarray system is able to distinguish between active and non-active DNA polymerase mutants. Thus, we first spotted primer strands on the slide that bind to a template (120 nt long) in a fully matched fashion. For screening purposes we used a library of an N-terminal-shortened form of the DNA polymerase from *Thermus aquaticus* (KlenTaq),[12] obtained by error-prone PCR.[8b,d] The polymerase library was first expressed in *E. coli* cells distributed in 96-well plates. After expression, cell lysis, and heat denaturation of the host proteins, the crude lysates were mixed with a buffer containing DNA templates, dNTPs, and a biotin-dUTP derivative. The polymerase-containing solutions were spotted such that each enzyme variant covered two of each set of primer spots. Thus, duplicated results under identical conditions were obtained.

We estimated the amount of immobilized DNA to be about 100 amol and of each enzyme to be 200–300 amol (for details see the Supporting Information). After incubation, reaction termination, and rinsing, the slides were treated with a solution containing a streptavidin–Alexafluor546 conjugate, rinsed again, and quantified using a standard microarray reader.

We also tested whether fluorescently labeled 2′-deoxynucleoside-5′triphosphates can be used instead of the biotin streptavidine-based approach. Interestingly, we obtained high fluorescence background values resulting from unspecific binding of the modified triphosphate even after extensive washing conditions; thus this approach is not suitable for our purpose.

Next, we randomly chose ten mutants, identified as non-active, and 10 mutants, identified as active by the formation of fluorescent spots on the microarray, for further characterization (Figure 2 a). Indeed, in solution the OAEA non-active mutants showed only little primer extension in contrast to the OAEA active mutants, which yielded full-length product (38 nucleotides long, Figure 2 b). We used a shorter template for primer extension reactions in solution which was in the same sequence context as the one used in screening in order to obtain better resolution in product analysis by gel electrophoresis. The enzymes add an additional nucleotide in non-templated manner as has been observed for 3′-5′-exonuclease-deficient DNA polymerases before.[13] Thus, the findings in the solution phase are in excellent agreement with the results obtained on the solid phase.

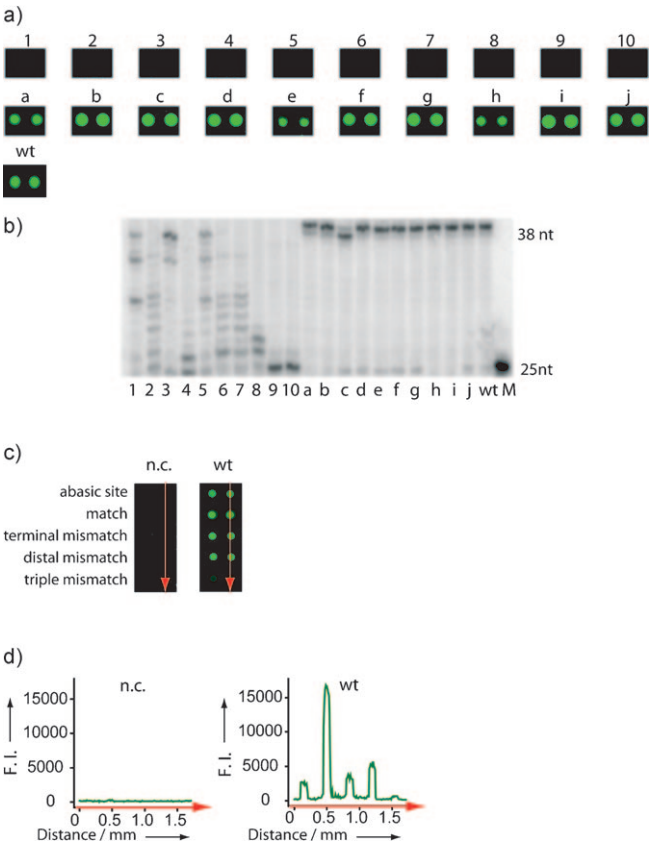

Figure 2. OAEA evaluation. a) Ten randomly chosen non-active mutants (1–10) and ten randomly chosen active mutants (a–j). A complete depiction of the data can be found in the Supporting Information. b) Denaturing polyacrylamide gel electrophoresis analysis of primer extensions performed in solution by the mutants depicted in (a). M: marker, reaction without cell lysate. c, d) OAEA results derived from cell lysates expressing the *KlenTaq* wild-type (wt) gene and a negative control (n.c.) from cell lysates harboring a plasmid without the *KlenTaq* gene. d) Fluorescence intensity profiles (F.I.) along the red arrows as indicated in (c). For experimental details see the Supporting Information.

Next, we investigated the simultaneous processing of five different primer–template duplexes by a library of DNA polymerase mutants. We employed a template harboring an abasic site analogue, as abasic sites are known to hinder numerous DNA polymerases.[14] We also tested several different substrates that are mismatched at the 3′-end of the primer including a single terminal mismatch, a single distal mismatch, and one triple mismatch duplex (for detailed DNA sequences see Figure 1 b and the Supporting Information). As a reference, a non-modified fully matched primer–template complex was used. Without further optimization, we screened a small library of 736 *KlenTaq* mutants with these five primer–

template duplexes. Each enzyme-containing entity was spotted with all five primer–template duplexes in duplicate. Generally, for the fully matched case the highest fluorescence intensity was found, as expected. The signal-to-noise ratio of wild-type (wt) *KlenTaq* with the non-modified primer–template substrate was consistantly greater than 35:1, whereas for a negative control reaction (n.c.) using a bacterial extract without the *KlenTaq* gene, no fluorescence was detected (see Figure 2c,d). This signal-to-noise ratio exceeds values reported previously in fluorescent-based screening approaches.[8,15]

Extension activities for the other duplex systems like the templates with the basic site, the single terminal mismatch, and the single distal mismatch, were characteristically decreased to values between 15–30% of the matched case, while for the triple mismatch case values of 3–5% were observed. We found a broad spectrum of different activities within the screened 736 enzyme mutants. This makes it possible to profile and classify DNA polymerases by specific activity "fingerprints". The most interesting three mutants (m1–m3) were expressed, purified, and further characterized in primer extension reactions in solution (see the Supporting Information).

Mutant m1 showed extraordinary discrimination against mismatches within the primer template construct and showed lower bypass activity for the substrate with the abasic site than the wild-type enzyme did (Figure 3a). Sequencing of the mutated gene revealed only a single mutation for m1, namely serine 460 (S460) to proline (P). Interestingly, S460 is located in helix H of the thumb domain of *KlenTaq* (Figure 3b). As a result of the known helix-breaking ability of proline,[17] it can be assumed that the H helix has lost its conformation in the S460P mutant. Interestingly, mutants m2 and m3 showed the opposite behavior and were more efficient at bypassing the abasic site (located at nucleotide position 24) than the wild-type enzyme. Mutants m2 and m3 also showed greater activity in extension reactions with the mismatched than that displayed by either the wild-type enzyme or m1 is observed.

Sequencing these mutants revealed two mutations for m2 (Y455N, V766A) and four mutations for m3 (L359P, R457G, E537G, V586I). Currently it is not known which amino acid exchange contributes most to the observed effects. Nevertheless, interestingly in both the m2 and m3 mutants, as well as the m1 mutant, mutations in the H helix are observed (Figure 3b). The H helix does not have direct van der Waals contact with the DNA substrate, but our findings show that mutations at this helix are able to influence the processing of the various substrates.

Interestingly, with our screening approach we can take "fingerprints" of DNA polymerases by direct comparison of their properties in processing different substrates. The findings indicate that the ability of lesion bypass in m2 and m3 is linked to their lower discrimination against mismatches. This is in accord with previously reported results.[8d,10b] For m1 it appears that this enzyme is in general more discriminatory against aberrant DNA structures such as mismatched primer ends and DNA lesions.

In summary, we have described a new microarray-based approach for DNA polymerase evolution which we term oligonucleotide-addressing enzyme assay (OAEA). First studies have proven the practicability of our approach and identified new DNA polymerase mutants with altered properties. In comparison with other known directed evolution approaches for DNA polymerases, OAEA offers several significant advantages. First, this approach allows the multiplex detection of various DNA polymerase activities in parallel under identical conditions. In addition, in OAEA each reaction can be duplicated readily. These features render OAEA reliable and less prone to false positives and negatives. Furthermore, all steps can be performed by automated pipetting devices, allowing high-throughput analysis requiring only minuscule amounts of reagents. Given the recent advances in microarray fabrication with more than 6000000 possible discrete features[18] on one chip, the depicted assay can be extended for the simultaneous ultrahigh-throughput multiplexed screening of extensive libraries with thousands of mutants. Furthermore, other possible applications can be foreseen. As all reactions are separately addressable by single enzyme entities, the method allows for parallel profiling of DNA polymerases from different origins. Additionally, other DNA-modifying enzymes like ligases and endonucleases can

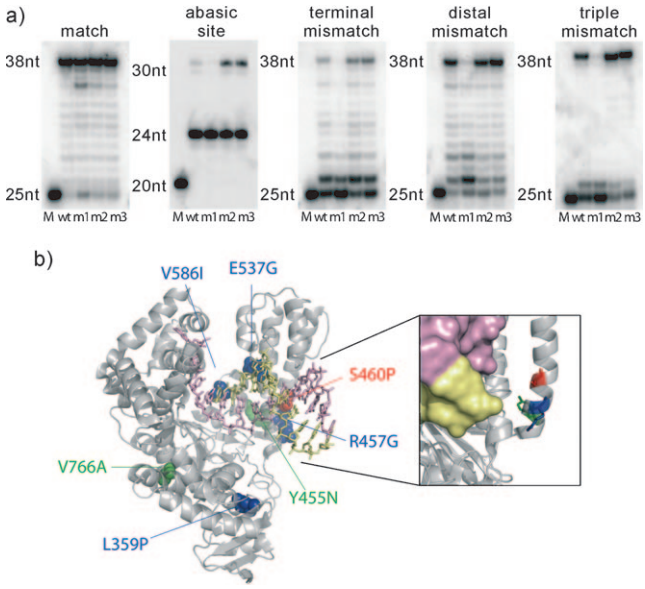

Figure 3. Evaluation of mutants with altered properties identified by OAEA. a) Processing of various DNA substrates in solution by evolved *KlenTaq* mutants (m1–m3) in comparison to the wild-type enzyme. All reactions were performed with identical reaction buffer and dNTP concentrations (100 μM each). Enzyme concentrations and incubation times: for matched, terminal, and distal mismatched: 1 nM, 15 min; for triple mismatched and abasic site template: 50 nM, 60 min. In the latter two cases higher enzyme concentrations and prolonged incubation times were required to promote extension of the more aberrant DNA complexes. For more experimental details see the Supporting Information. M: marker, reaction without enzyme. b) Mutations in the evolved *KlenTaq* DNA polymerases m1 (red), m2 (green), and m3 (blue) are mapped on a ribbon representation of *KlenTaq* (PDB code: 1QSS).[16] The inset highlights the H helix with the observed mutation sites from m1–m3.

be included in multiplex directed evolution approaches using OAEA.

Received: February 18, 2009
Revised: March 24, 2009
Published online: May 14, 2009

Keywords: directed evolution · DNA damage · DNA polymerases · enzymes · microarrays

[1] J. Sambrook, D. W. Russell, *Molecular Cloning:A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2001**.
[2] B. Gilje, R. Heikkila, S. Oltedal, K. Tjensvoll, O. Nordgård, *J. Mol. Diagn.* **2008**, *10*, 325 – 331.
[3] M. d'Abbadie, M. Hofreiter, A. Vaisman, D. Loakes, J. Gasparutto, R. Cadet, S. Woodgate, S. Pääbo, P. Holliger, *Nat. Biotechnol.* **2007**, *25*, 939 – 943.
[4] M. B. Kermekchiev, L. I. Kirilova, E. E. Vail, W. M. Barnes, *Nucleic Acids Res.* **2009**, *37*, e40.
[5] J. Eid, et al., *Science* **2009**, *323*, 133 – 138.
[6] a) A. A. Henry, F. E. Romesberg, *Curr. Opin. Biotechnol.* **2005**, *16*, 370 – 377; b) S. Brakmann, *Cell. Mol. Life Sci.* **2005**, *62*, 2634 – 2646; c) J. F. Davidson, J. Anderson, H. Guo, D. Landis, L. A. Loeb in *Directed Molecular Evolution of Proteins* (Eds.: S. Brakmann, K. Johnsson), Wiley-VCH, Weinheim, **2002**, pp. 281 – 307.
[7] a) P. H. Patel, L. A. Loeb, *J. Biol. Chem.* **2000**, *275*, 40266 – 40272; b) P. H. Patel, H. Kawate, E. Adman, M. Ashbach, L. A. Loeb, *J. Biol. Chem.* **2001**, *276*, 5044 – 5051; c) S. Brakmann, S. Grzeszik, *ChemBioChem* **2001**, *2*, 212 – 219; d) M. B. Kermekchiev, A. Tzekov, W. M. Barnes, *Nucleic Acids Res.* **2003**, *31*, 6139 – 6147; e) E. Loh, J. Choe, L. A. Loeb, *J. Biol. Chem.* **2007**, *282*, 12201 – 12209.
[8] a) D. Summerer, N. Z. Rudinger, I. Detmer, A. Marx, *Angew. Chem.* **2005**, *117*, 4791 – 4794; *Angew. Chem. Int. Ed.* **2005**, *44*, 4712 – 4715; b) K. B. M. Sauter, A. Marx, *Angew. Chem.* **2006**, *118*, 7795 – 7797; *Angew. Chem. Int. Ed.* **2006**, *45*, 7633 – 7635; c) M. Strerath, C. Gloeckner, D. Liu, A. Schnur, A. Marx, *ChemBioChem* **2007**, *8*, 395 – 401; d) C. Gloeckner, K. B. M. Sauter, A. Marx, *Angew. Chem.* **2007**, *119*, 3175 – 3178; *Angew. Chem. Int. Ed.* **2007**, *46*, 3115 – 3117.
[9] a) M. Fa, A. Radeghieri, A. A. Henry, F. E. Romesberg, *J. Am. Chem. Soc.* **2004**, *126*, 1748 – 1754; b) A. M. Leconte, L. Chen, F. E. Romesberg, *J. Am. Chem. Soc.* **2005**, *127*, 12470 – 12471; c) G. Xia, L. Chen, T. Sera, M. Fa, P. G. Schultz, F. E. Romesberg, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6597 – 6602; d) S. Vichier-Guerre, S. Ferris, N. Auberger, K. Mahiddine, J.-L. Jestin, *Angew. Chem.* **2006**, *118*, 6279 – 6283; *Angew. Chem. Int. Ed.* **2006**, *45*, 6133 – 6137.
[10] a) F. J. Ghadessy, J. L. Ong, P. Holliger, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4552 – 4557; b) F. J. Ghadessy, N. Ramsay, F. Boudsocq, D. Loakes, A. Brown, S. Iwai, A. Vaisman, R. Woodgate, P. Holliger, *Nat. Biotechnol.* **2004**, *22*, 755 – 759; c) J. L. Ong, D. Loakes, S. Jaroslawski, K. Too, P. Holliger, *J. Mol. Biol.* **2006**, *361*, 537 – 550.
[11] a) R. Kranaster, P. Ketzer, A. Marx, *ChemBioChem* **2008**, *9*, 694 – 697; b) J. Gaster, G. Rangam, A. Marx, *Chem. Commun.* **2007**, 1692 – 1694.
[12] W. M. Barnes, *Gene* **1992**, *112*, 29 – 35.
[13] J. M. Clark, *Nucleic Acids Res.* **1988**, *16*, 9677 – 9686.
[14] a) L. A. Loeb, B. D. Preston, *Annu. Rev. Genet.* **1986**, *20*, 201 – 230; b) M. F. Goodman, H. Cai, L. B. Bloom, R. Eritja, *Ann. N. Y. Acad. Sci.* **1994**, *726*, 132 – 142.
[15] D. Summerer, A. Marx, *Angew. Chem.* **2002**, *114*, 3778 – 3780; *Angew. Chem. Int. Ed.* **2002**, *41*, 3620 – 3622.
[16] Y. Li, V. Mitaxov, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9491 – 9496.
[17] N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, Wiley-VCH, Weinheim, **2002**.
[18] T. C. Mockler, S. Chan, A. Sundaresan, H. Chen, S. E. Jacobsen, J. R. Ecker, *Genomics* **2005**, *85*, 1 – 15.

