Analytical Biochemistry 388 (2009) 71–80

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment

Tiina S. Rasila [a], Maria I. Pajunen [a,b], Harri Savilahti [a,b,*]

[a] *Program in Cellular Biotechnology, Institute of Biotechnology, Viikki Biocenter, University of Helsinki, FIN-00014 Helsinki, Finland*
[b] *Division of Genetics and Physiology, Department of Biology, University of Turku, FIN-20014 Turku, Finland*

ARTICLE INFO

Article history:
Received 28 November 2008
Available online 10 February 2009

Keywords:
Protein engineering
Mutation frequency
Mutation spectrum
Transition
Transversion
Amino acid substitution pattern
XL1–Red
Taq DNA polymerase
Mutazyme II DNA polymerase
8-oxo-dGTP
dPTP
NH_2OH–HCl

ABSTRACT

Random mutagenesis methods constitute a valuable protein modification toolbox with applications ranging from protein engineering to directed protein evolution studies. Although a variety of techniques are currently available, the field is lacking studies that would directly compare the performance parameters and operational range of different methods. In this study, we have scrutinized several of the most commonly used random mutagenesis techniques by critically evaluating popular error-prone polymerase chain reaction (PCR) protocols as well as hydroxylamine and a mutator *Escherichia coli* strain mutagenesis methods. Relative mutation frequencies were analyzed using a reporter plasmid that allowed direct comparison of the methods. Error-prone PCR methods yielded the highest mutation rates and the widest operational ranges, whereas the chemical and biological methods generated a low level of mutations and exhibited a narrow range of operation. The repertoire of transitions versus transversions varied among the methods, suggesting the use of a combination of methods for high-diversity full-scale mutagenesis. Using the parameters defined in this study, the evaluated mutagenesis methods can be used for controlled mutagenesis, where the intended average frequency of induced mutations can be adjusted to a desirable level.

Random mutagenesis is commonly used to generate mutant libraries for protein engineering applications with the ultimate goal of generating protein variants with desired properties [1]. Related disciplines, such as directed protein evolution studies and functional analyses of uncharacterized open reading frames, also take advantage of random mutagenesis [2–6]. Decisive for the success in such endeavors is the quality of the mutant libraries produced, with the aim most often intending to change each amino acid into all possible alternatives or to generate multiple mutations within a single polypeptide. The most important performance criteria for a random mutagenesis method include an unbiased mutation spectrum and a controllable mutation frequency [7].

Random mutagenesis strategies involve methods that create diversity throughout a given DNA sequence or introduce randomization at more specific locations along the sequence [2]. In general, the strategies can be classified into the groups of enzymatic and chemistry- or cell-based approaches as well as their combinations. Enzymatic strategies exploit the properties of DNA-modifying enzymes, and particularly DNA polymerase-based error-prone polymerase chain reaction (PCR)[1] protocols have been popular [1]. Error-prone methods are based on intrinsic high error rates of DNA polymerases, either natural or engineered [8–10], and the fidelity characteristics can be further modulated by altering the reaction conditions, for example by unbalancing nucleotide concentrations [11,12], including nucleotide analogues [13,14], or changing the divalent cation(s) in the reaction [11,12,15].

Purely chemical methods induce mutations directly with DNA-modifying reagents, whereas other chemistry-based methods, such as oligonucleotide synthesis, are more indirect and often used in combination with other types of methods [1,2]. Traditionally, chemical modification has been used for in vivo mutagenesis of plasmids as well as genomes of bacteria and eukaryotes [16–18], and a variety of reagents have been applied for this purpose, including ethyl methane sulfonate (EMS), methylnitronitrosoguanidine (MNNG), and ethylnitrosourea (ENU). In contrast, only a

* Corresponding author. Fax: +358 2 333 6680.
E-mail address: harri.savilahti@utu.fi (H. Savilahti).

[1] *Abbreviations used:* PCR, polymerase chain reaction; EMS, ethyl methane sulfonate; MNNG, methylnitronitrosoguanidine; ENU, ethylnitrosourea; NH_2OH–HCl, hydroxylamine hydrochloride; UV, ultraviolet; PAC, P1-derived artificial chromosome; BAC, bacterial artificial chromosome; IPTG, isopropyl-β-D-thiogalactopyranoside; X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; 8-oxo-dGTP, 8-oxo-2′-deoxyguanosine-5′-triphosphate; dPTP, 2′-deoxy-P-nucleoside-5′-triphosphate; LB, Luria–Bertani; Ap, ampicillin; BSA, bovine serum albumin; EDTA, ethylenediaminetetraacetic acid; MAP, mutagenesis assistant program.

0003-2697/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2009.02.008

few reagents have successfully been used for in vitro mutagenesis [19], including EMS [20] and hydroxylamine hydrochloride (NH_2OH–HCl) [1,5,21].

Cell-based mutagenesis methods are based on the increased rate of DNA replication errors. The methods take advantage of environmental/physiological stress, such as that caused by ionizing radiation or ultraviolet (UV) light, or they may exploit defective replication/repair functions of cells [1], such as bacterial mutator strains harboring critical mutations in DNA repair pathway genes. With cell-based methods, mutations are induced simply by in vivo propagation of the DNA of interest cloned in an appropriate replicative vector such as plasmid, phage, cosmid, P1-derived artificial chromosome (PAC), or bacterial artificial chromosome (BAC) [16,22,23]. Commercial *Escherichia coli* XL1–Red strain is deficient in three of the primary DNA repair pathways—*mutS*, *mutT*, and *mutD* [22]—and has successfully been used for effective mutagenesis in a number of studies [24–28].

Random mutagenesis methods are continuously evolving for a better performance, generating the need for studies that would directly compare different protocols. Several critical reviews and computer programs are available for the general evaluation of a variety of methods [7,29,30], but these comparisons are not direct given that they exploit performance parameters reported previously from disparate studies. Here we report a comparative evaluation of several commonly used random mutagenesis methods. These included error-prone PCR protocols with *Taq* and Mutazyme II DNA polymerases as well as *E. coli* mutator strain and NH_2OH–HCl mutagenesis procedures. The data defined the operational range of each protocol and revealed the corresponding mutation spectra. The parameters determined are useful for controlling the efficiency of mutagenesis, and the choice of the mutagenesis method(s) can be used to adjust the intended mutation pattern(s).

Materials and methods

Enzymes and reagents

Restriction endonucleases, Vent DNA polymerase, and T4 DNA ligase were obtained from New England Biolabs. *Taq* DNA polymerase was obtained from Promega. Deoxynucleotide mix and $MgCl_2$ were purchased from Finnzymes. Isopropyl-β-D-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), dCTP, and dTTP were purchased from Fermentas. 8-oxo-2′-Deoxyguanosine-5′-triphosphate (8-oxo-dGTP) and 2′-deoxy-P-nucleoside-5′-triphosphate (dPTP) were obtained from TriLink. $MnCl_2$ was obtained from Merck, NH_2OH–HCl was obtained from Sigma–Aldrich, and sodium hydroxide was obtained from J.T. Baker.

Escherichia coli strains, culture conditions, plasmids, and DNA techniques

DH10B [31] was used as a standard cloning host and for routine plasmid DNA isolation. DH5α (Invitrogen) was used for the determination of relative mutation frequencies. XL1–Red (Stratagene) was used for in vivo mutagenesis. For standard use, bacteria were grown in Luria–Bertani (LB) medium as described previously [32]. Electrocompetent cells were prepared as described previously [33]. Plasmid pUC19 was obtained from New England Biolabs. The mutagenesis reporter plasmid was constructed as follows. Initially, the *MuA* gene (GenBank Accession No. P07636) from pMK591 [34] was cloned as an *Nco*I–*Bam*HI fragment into pBADHisA (Invitrogen) cleaved with *Nco*I and *Bam*HI to yield pALH6. Next, a PCR fragment was amplified with Vent DNA polymerase from *E. coli* genomic DNA (Cat. No. 14380, USB) using primers HSP489 (5′-GCGCGCAAGCTTGCGCAACGCAATTAATGTG-3′) and HSP490 (5′-GCGCCAAGCTTCTCGAGCGCCATTCGCCATTCAGG-3′). This DNA segment encodes the α-complementing protein fragment of β-galactosidase (amino acids M1–R60, Swiss–Prot Accession No. P00722) and also includes the *lac* promoter as well two primer-derived *Hin*dIII sites for cloning. Finally, the amplified fragment was cleaved with *Hin*dIII and cloned into pALH6 that had been cleaved with *Hin*dIII to yield pTLH2. Plasmid maintenance was selected using 100 μg/ml ampicillin (Ap, Sigma) in the growth medium.

Error-prone PCR with Taq DNA polymerase in mutagenic buffer

The reaction conditions for error-prone PCR with *Taq* DNA polymerase in mutagenic buffer were modified from those described previously [12,35]. Each standard PCR (50 μl) contained 50 ng (12 fmol) of plasmid pTLH2 as a template, 0.3 μM each of the primers HSP492 (5′-ATCAGACCGCTTCTGCGTTC-3′) and HSP493 (5′-GATTAGCGGATCCTACCTGAC-3′), 200 μM each of the dNTPs, and 5 U of *Taq* DNA polymerase in *Taq* DNA polymerase reaction buffer (10 mM Tris–HCl [pH 9.0 at 25 °C], 50 mM KCl, 1.5 mM $MgCl_2$, and 0.1% [v/v] Triton X-100). When indicated, 1, 2, 4, 6, 8, or 10 μl of mutagenic buffer (4 mM dTTP, 4 mM dCTP, 27.5 mM $MgCl_2$, and 2.5 mM $MnCl_2$) was included in the standard 50-μl reaction. PCR employed initial 2 min at 94 °C, 25 cycles of amplification (1 min at 94 °C, 1 min at 59 °C, and 2.5 min at 72 °C), and final 10 min at 72 °C. The PCR products were purified by using the QIAquick PCR Purification Kit (Qiagen), after which the products were digested with *Nco*I and *Xho*I and subjected to preparative electrophoresis on a 1% SeaPlaque GTG agarose gel (Lonza) in TAE buffer [32]. The approximately 2.3-kb PCR fragment was isolated by using the QIAquick MinElute Gel Extraction Kit (Qiagen). The gel-purified PCR products were cloned into pBADHisA cleaved with *Nco*I and *Xho*I.

Error-prone PCR with Taq DNA polymerase and nucleotide analogues

The reaction conditions for error-prone PCR with *Taq* DNA polymerase and nucleotide analogues (8-oxo-dGTP and dPTP) were modified from those described previously [14]. PCRs were set up essentially as described above for the standard reaction with mutagenic buffer. However, instead of mutagenic buffer, 0.2, 1, 2, 10, 20, 100, or 200 μM both 8-oxo-dGTP and dPTP was included in the standard 50-μl reaction when indicated. Plasmid pTLH2 (260 ng) was used as a template (corresponding to 100 ng of amplicon) in PCR amplification that employed initial 2 min at 94 °C, 10 cycles of amplification (1 min at 94 °C, 1 min at 59 °C, and 2.5 min at 72 °C), and final 10 min at 72 °C. PCR products were extracted from a preparative agarose gel as described above and further amplified using Vent DNA polymerase and nonmutagenic reaction conditions. Each amplification reaction (50 μl) contained approximately 40 ng of gel-purified PCR product as template, 0.5 μM each of the primers HSP492 and HSP493, 200 μM each of the dNTPs, and 1 U of Vent DNA polymerase in ThermoPol reaction buffer (20 mM Tris–HCl [pH 8.8 at 25 °C], 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 4 mM $MgSO_4$, and 0.1% [v/v] Triton X-100). PCR employed initial 5 min at 95 °C, 25 cycles of amplification (45 s at 95 °C, 1 min at 59 °C, and 2.5 min at 72 °C), and final 5 min at 72 °C. PCR products were purified, digested, gel-isolated, and cloned into pBADHisA as described above.

Error-prone PCR with Mutazyme II DNA polymerase

Error-prone PCR with Mutazyme II DNA polymerase was performed using the GeneMorph II Random Mutagenesis Kit (Stratagene) as recommended by the supplier. Each amplification reaction (50 μl) contained 0.5 μM each of the primers HSP492

and HSP493, 200 μM each of the dNTPs, and 2.5 U of Mutazyme II DNA polymerase in Mutazyme II reaction buffer (Stratagene). PCR employed initial 2 min at 95 °C, 30 cycles of amplification (1 min at 95 °C, 1 min at 54 °C, and 2.5 min at 72 °C), and final 10 min at 72 °C. Two separate sets of mutagenic amplification reactions were conducted. The amount of template amplicon varied in the first set of experiments (100, 10, 1, 0.1, or 0.001 ng were used). In the second set of experiments, 100 ng of template amplicon was used, but the number of amplification cycles varied (1, 3, 5, 7, 10, 15, or 20 cycles were used). PCR products were treated and cloned into pBADHisA as above.

Mutagenesis with XL1–Red mutator strain

Selecting for Ap resistance, plasmid pTLH2 (50 ng) was transformed into competent cells of XL1–Red mutator strain (Cat. No. 200129, Stratagene) according to the manufacturer's instructions. Approximately 1000 Ap-resistant colonies were pooled and grown in 10 ml of LB–Ap medium at 26 °C for 24 h (first cycle). The bacterial culture was diluted 1:1000 [36], and culturing was continued at 26 °C for an additional 24 h in 10 ml of LB–Ap medium (second cycle). Culturing was repeated, and altogether 10 successive 24-h cycles were employed. Plasmid DNA was recovered and analyzed for mutations after each passage.

NH_2OH–HCl mutagenesis

NH_2OH–HCl mutagenesis was performed as described previously [37]. Briefly, plasmid pTLH2 (10 μg) was mixed with 500 μl of freshly made NH_2OH–HCl solution (1 M NH_2OH–HCl in 0.45 M NaOH, pH 7.0), and the mixture was incubated at 37 °C for different time periods (8, 10, 15, 17, 20, 24, and 40 h were tested). Each reaction was stopped by the addition of 10 μl of 5 M NaCl, 50 μl of bovine serum albumin (BSA, 1 mg/ml), and 1 ml of ethanol. Precipitated DNA was collected by centrifugation and resuspended in 100 μl of TE buffer (10 mM Tris [pH 7.5] and 0.5 mM ethylenediaminetetraacetic acid [EDTA]). DNA was reprecipitated by the addition of NaCl and ethanol, and following a rinse with ethanol, DNA was resuspended in 100 μl of TE buffer.

Determination of relative mutation frequencies

Relative mutation frequencies were determined using a lactose utilization phenotype change on indicator plates. To ensure that mutagenized plasmid clones represented homoduplexes with segregated wild-type and mutant alleles, following primary culturing on selection plates, approximately 1×10^4 colonies were pooled from each mutagenesis experiment and grown in LB–Ap medium at 37 °C for 2.5 h. DNA was isolated, and 1 ng of purified homoduplex plasmid DNA was electroporated into *E. coli* DH5α. The cells were then plated on color indicator plates containing Ap (100 μg/ml), X-gal (80 μg/ml), and IPTG (0.1 mM). The total colony number ranged from 4.0×10^3 to 1.2×10^4 among experiments. Colonies were evaluated according to their color, and relative mutation frequencies were determined as the proportion of white and pale blue colonies from the total number of all colonies. Note that this visual screen does not detect mutations not affecting lactose utilization phenotype (e.g., those not changing the respective amino acid). In cases where a cloning step was included in the protocol, a fraction of white colonies was generated by plasmid clones containing no inserts, thereby causing a small systematic bias, typically representing 1 to 2% of total colonies. The bias was eliminated using control vector ligations and subtracting the control colony numbers from the numbers of non-blue colonies in the experimental samples.

Sequence analysis

DNA sequence determination of selected plasmids was performed at the DNA sequencing facility of the Institute of Biotechnology (University of Helsinki) using the primer HSP492. In general, two blue colonies, one pale blue colony, and one white colony were chosen for the analysis per time point or PCR condition. The sequencing reads covered the entire *lacZα*-encoding region, extending into the *MuA*-encoding segment. Total read lengths ranged from 539 to 800 nt. The total number of sequenced plasmid clones per mutagenesis method ranged from 15 to 27. A common 500-bp region of mutagenized plasmids (obtained using the primer HSP492) was used to directly compare the number of mutations induced by different methods. For the analysis of mutations on protein level, conservative and nonconservative amino acid substitutions were defined according to the BLOSUM62 matrix [38].

Results

Experimental outline

We were interested in directly comparing the operational range and mutation spectrum of several of the most commonly used random mutagenesis methods (Fig. 1). Initially, we constructed a reporter plasmid that includes a screening system based on the enzymatic activity of *E. coli* β-galactosidase encoded by *lacZ*. The plasmid includes a gene fragment (*lacZα*) that encodes an N-terminal peptide capable of complementing an enzyme defect caused by a deletion mutation in the *E. coli* chromosome. Mutations in the plasmid that affect the reporter system can be screened by plating plasmid-containing bacterial clones on X-gal/IPTG indicator plates. With regard to the reporter system, critically mutated plasmid clones generate white or light blue colonies, whereas original (and noncritically mutated) plasmids yield deep blue colonies, forming a visual assay to determine relative mutation frequencies. We also included the *MuA* gene in our reporter plasmid because the encoded MuA transposase protein will be mutated in our currently ongoing projects; it also allowed mutation spectrum analysis along a longer region of protein-encoding DNA. Initially, we subjected the reporter plasmid to different mutagenic treatments and then propagated the plasmids in nonmutagenic conditions for a short time period to ensure segregation of mutated DNA strands, yielding homoduplexed plasmid DNA for the analyses. For mutagenesis, we used *Taq* DNA polymerase with mutagenic buffer or with a combination of nucleotide analogues (8-oxo-dGTP and dPTP), Mutazyme II DNA polymerase with variable amplicon amounts or PCR cycles, mutator strain XL1–Red, and chemical mutagen NH_2OH–HCl. We varied critical parameters and determined relative mutation frequencies for each of these methods. In addition, we analyzed the number and spectrum of nucleotide changes induced, and in cases where a relevant number of mutations were produced, we also analyzed the mutation diversity at the protein level.

Error-prone PCR with Taq DNA polymerase and mutagenic buffer

Taq DNA polymerase PCR amplification is error prone as such, and the frequency of mutations generated can be modulated using mutagenic buffer that unbalances the nucleotide concentrations and includes $MnCl_2$ in the reaction [12]. The error rate of *Taq* DNA polymerase was examined by increasing the proportion of mutagenic buffer in the PCR (see Materials and methods and Fig. 2A). Linear correlation was observed between the added volume of the buffer and the percentage of discernible mutant clones

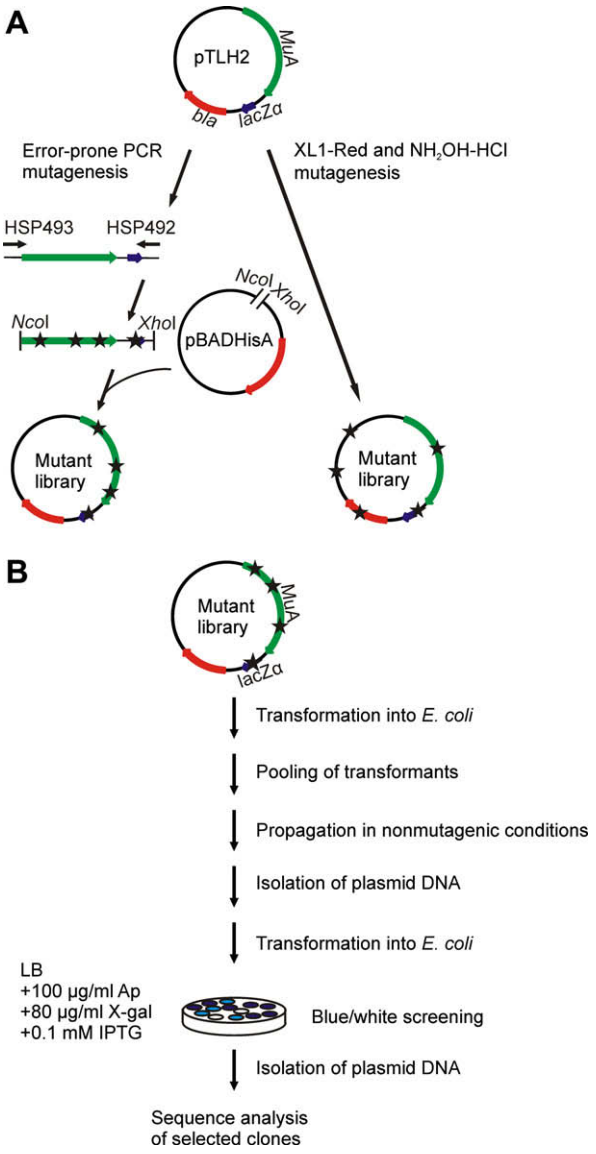

Fig. 1. Experimental outline. (A) Construction of mutant plasmid pools. In error-prone PCR mutagenesis protocols, the target gene fragment in pTLH2, containing the *lacZα* marker gene and the *MuA* gene, was amplified in mutagenic conditions using the primer pair HSP492/HSP493. The PCR product was digested with *Nco*I and *Xho*I and then cloned into pBADHisA cleaved with *Nco*I and *Xho*I to construct a mutant plasmid pool. In XL1–Red and NH_2OH–HCl mutagenesis protocols, the entire pTLH2 plasmid was subjected to mutagenic conditions to generate a pool of mutant plasmids. Thus, in error-prone PCR mutagenesis, only the cloned target gene fragments gain random nucleotide changes. In XL1–Red and NH_2OH–HCl mutagenesis, random nucleotide changes are generated throughout the entire plasmid. (B) Probing the mutation generation capacity of random mutagenesis methods. Mutant plasmid libraries were transformed into *E. coli*, and 1 to 3×10^4 colonies were pooled and propagated in nonmutagenic conditions to ensure segregation of mutated DNA strands. Isolated plasmid DNA was transformed into *E. coli*, and the cells were plated on color indicator plates to determine the proportion of discernible mutant colonies (pale blue and white colonies) from the total number of colonies. Sequencing with primer HSP492 was used to analyze the number of mutations along the target sequence in selected plasmid clones.

generated. The proportion of observed mutants ranged from 7 to 56%, and sequence analysis revealed up to 12 mutations per 500 nt analyzed (extrapolated for 1000 bp in Fig. 2A), indicating strong mutagenic activity and a wide operational range of this mutagenesis system, particularly desirable features in cases where simultaneous generation of many mutations is needed.

Error-prone PCR with nucleotide analogues

A PCR method that employs the nucleotide analogues 8-oxo-dGTP and dPTP is one of the most efficient means to introduce point mutations in DNA [14,39]. The fidelity of *Taq* DNA polymerase in the presence of these analogues was examined by increasing their concentration in the PCR (see Materials and methods and Fig. 2B). Within the range of 2 to 100 μM, the relative mutation frequency observed was proportional to the concentration of nucleotide analogues and ranged from 30% to more than 90%. Sequencing revealed 2 to 36 mutations within the 500 nt analyzed (extrapolated for 1000 bp in Fig. 2B). Thus, the method has a very high mutational capacity as well as a wide operational range that manifests particularly at higher mutation frequencies. The data suggest that nearly one-tenth of the nucleotides can be readily mutated within a given DNA molecule.

Error-prone PCR with Mutazyme II DNA polymerase

Mutazyme II DNA polymerase is a commercial error-prone PCR enzyme mixture formulated to provide useful mutation rates with minimal mutational bias. The mutation rate can be controlled by varying the initial amount of target DNA in the reaction or the number of amplification cycles performed [8,40,41]. We first tested the effect of target DNA amount in the reaction and used an invariable 30 cycles of amplification (see Materials and methods and Fig. 2C). The proportion of discernible mutant clones increased from 10% to nearly 40% when the target amplicon amount was reduced logarithmically from 100 to 0.001 ng. Sequencing revealed 1 to 9 mutations within the segment of the 500 nt analyzed (extrapolated for 1000 bp in Fig. 2C). Next, we tested the effect of amplification cycles and used an invariable amount (100 ng) of the template amplicon (see Materials and methods and Fig. 2D). The proportion of discernible mutant clones increased from 2% to more than 11% when the number of PCR cycles was increased from 1 to 20. Sequence analysis revealed 0 to 5 mutations within the 500-nt segment analyzed (extrapolated for 1000 bp in Fig. 2D). The results imply that Mutazyme II DNA polymerase is a powerful tool for the construction of low- to medium-rate mutation libraries.

Mutagenesis with XL1–Red mutator strain

Escherichia coli XL1–Red is a mutator bacterial strain that lacks three of the primary DNA repair pathways and induces point mutations into a resident target plasmid. The number of generated mutations can be increased by successive growth cycles [36]. We tested the mutational capacity of this strain by propagating the reporter plasmid for several 24-h passages (see Materials and methods and Fig. 2E). XL1–Red strain induced 0.6 to 2.0% discernible mutants when the number of successive passages grew from 1 to 10. Sequencing revealed no mutations or a single mutation along the sequenced 500-nt segment in selected clones (extrapolated for 1000 bp in Fig. 2E). All of the detected mutations were in plasmid clones generating colonies that grew white or pale blue on indicator plates. The data imply that whenever a low mutation rate is needed (e.g., for the generation of single amino acid changes in a protein encoded by a relative long target DNA segment), the XL1–Red mutator strain represents an attractive mutagenesis alternative.

NH_2OH–HCl mutagenesis

NH_2OH–HCl is a chemical mutagen that can be used for in vitro mutagenesis [16,21]. The number of generated mutations can be

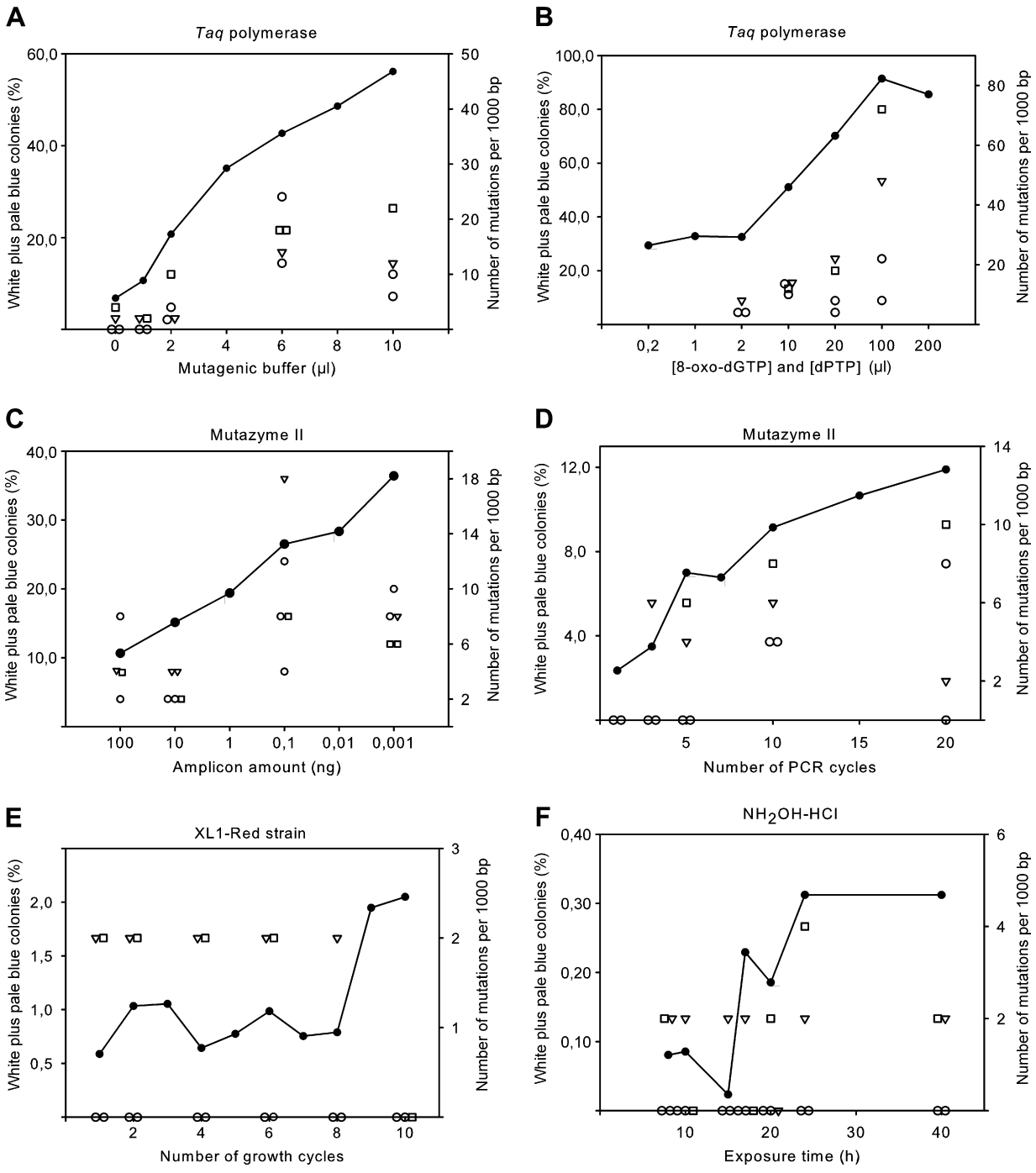

Fig. 2. Relative mutation frequencies and numbers of observed mutations. Relative mutation frequency is defined as the percentage of white and pale blue colonies from the total number of colonies. In each panel, the relative mutation frequency is shown with filled circles connected by black lines with the scale on the left. The scale on the right in each panel indicates the numbers of mutations per 1000 bp (extrapolated from the 500-bp region analyzed). The number of mutations from blue colonies are shown with open circles, those from pale blue colonies are shown with open triangles, and those from white colonies are shown with open squares. For clarity, the symbols for the clones having the same number of mutations are depicted adjacent to each other. (A) Error-prone PCR using *Taq* polymerase with mutagenic buffer. (B) Error-prone PCR using *Taq* polymerase and 8-oxo-dGTP and dPTP nucleotide analogues. (C) Error-prone PCR using Mutazyme II DNA polymerase and variable amounts of the amplicon. (D) Error-prone PCR using Mutazyme II DNA polymerase and variable numbers of PCR cycles. (E) Mutator strain XL1–Red. (F) NH_2OH–HCl mutagenesis.

increased by lengthening the exposure time [37]. We incubated the reporter plasmid for different time periods (see Materials and methods and Fig. 2F) and observed a relative mutation frequency that ranged from 0.02 to 0.31% when the exposure time was varied from 8 to 40 h. Sequencing revealed 0 to 2 mutations along the 500 nt analyzed (extrapolated for 1000 bp in Fig. 2F), and all of the mutations were detected in plasmid clones originating from colonies that grew white or pale blue on indicator plates. Thus, NH_2OH–HCl mutagenesis generates mutations with a low rate, and so it is most useful in protein modification application where single amino acid substitutions are needed.

Mutation types at the DNA level

The mutation spectrum generated is an important performance criterion for a random mutagenesis method, and nonbiased spectra

are generally preferred [1]. We examined, by DNA sequence analysis, the mutation spectrum of the six evaluated mutagenesis methods (Table 1 and Fig. 3). The number of clones analyzed per mutagenesis method ranged from 15 to 27, and the total number of sequenced nucleotides varied between 12,000 and 21,000 per method.

Taq DNA polymerase with mutagenic buffer induced nearly all possible point mutation types (Table 1 and Fig. 3). Transition mutations dominated with nearly 60% prevalence; in particular, A:T → G:C transitions were common (~38%). The vast majority (85%) of mutations were of the type A:T → G:C, G:C → A:T, or A:T → T:A. Also, 1-bp deletions were found, but no insertions or G:C → C:G transversions were detected.

The inclusion of nucleotide analogues biased the mutational spectrum of *Taq* DNA polymerase substantially (Table 1 and Fig. 3). Amplification with nucleotide analogues 8-oxo-dGTP and dPTP primarily induced transitions (93%), and A:T → G:C transitions dominated with a 65% prevalence. A small number of A:T → T:A and A:T → G:C changes, as well as a few 1-bp deletions, were also found.

Mutazyme II induced all possible point mutation types (Table 1 and Fig. 3). Transitions and transversions were detected in equal numbers, and they totaled 90% of the mutations. Approximately 10% of the induced mutations were 1-bp insertions or deletions. Very similar results were obtained for both experimental series, either with a variable amount of amplicon DNA or with a variable number of amplification cycles.

Both XL1–Red strain and NH_2OH–HCl treatment induced very low levels of mutations (Table 1 and Fig. 3), and most of them were G:C → A:T transitions. A few other types of mutations were also observed, including 1-bp deletions with XL1–Red strain.

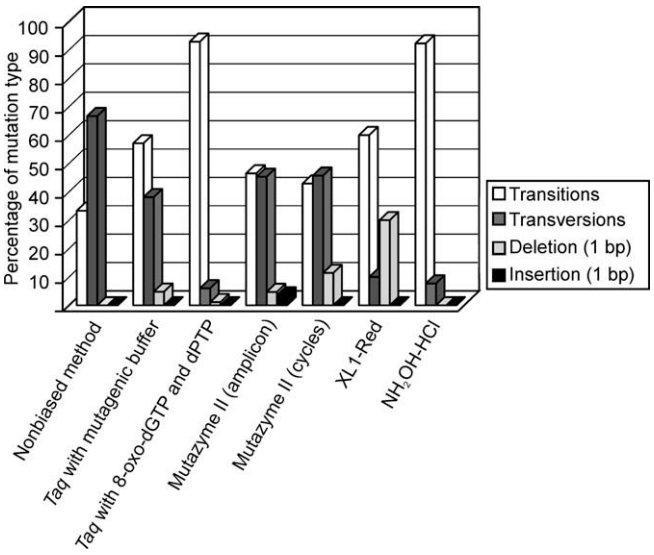

Fig. 3. Mutation types induced by the evaluated random mutagenesis methods. Percentages are shown for transitions with white columns, for transversions with dark gray columns, for deletions with light gray columns, and for insertions with black columns. An ideal nonbiased method would produce a pattern shown on the left.

Amino acid substitution patterns

The mutation spectrum of a given mutagenesis method influences the repertoire of changed amino acids at the protein level. We were able to evaluate the amino acid substitution patterns generated by the four error-prone mutagenesis protocols using mutation data from the protein-encoding regions in the reporter plasmid (Table 2). However, the chemical and mutator strain mutagenesis protocols generated too few mutations, thereby preventing a meaningful evaluation of these methods. Amino acid substitutions were analyzed from the LacZα sequence (60 amino acids) and from the C terminus of the MuA sequence (143 amino acids). The total number of amino acids analyzed ranged from 2998 to 4235 among the different methods. Mutations were distributed throughout the LacZα and MuA sequences without evident bias toward any particular region (data not shown). Because the spectrum of amino acid changes was similar for both proteins (data not shown), the final analysis took advantage of the additive data from both sequences (Table 2). Below, we highlight the most significant differences among the methods.

Initially, we classified mutations into those that preserved the corresponding amino acid, changed it, or generated a stop codon. A clear difference was seen between the polymerases used in mutagenesis, *Taq* DNA polymerase versus Mutazyme II DNA polymerase. Of the total mutations, *Taq* DNA polymerase preserved the amino acid in more than 30% of the cases with both mutagenesis protocols (mutagenic buffer and nucleotide analogues). With Mutazyme II DNA polymerase, the amino acid preservation percentage was much lower (~13 and 26% with the two separate protocols). The rate of stop codon generation was also very different for these enzymes. Whereas *Taq* DNA polymerase produced less than 4% of stop codons, Mutazyme II DNA polymerase generated nearly 15% of them.

Amino acid changes were classified into conservative and nonconservative substitutions. Conservative substitutions change the

Table 1
Mutation spectrum of random mutagenesis methods.

	Taq DNA polymerase				Mutazyme II polymerase				XL1–Red		NH_2OH–HCl	
	Mutagenic buffer		8-oxo-dGTP and dPTP		Amount of amplicon		Number of cycles					
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Transitions	72	57.1	169	92.9	40	46.5	15	42.9	6	60.0	12	92.3
A:T → G:C	48	38.1	119	65.4	12	14.0	6	17.1	0	0	2	15.4
G:C → A:T	24	19.0	50	27.5	28	32.5	9	25.7	6	60.0	10	76.9
Transversions	48	38.1	11	6.0	39	45.3	16	45.7	1	10.0	1	7.7
A:T → T:A	35	27.8	3	1.6	13	15.1	10	28.6	1	10.0	1	7.7
A:T → C:G	5	4.0	8	4.4	2	2.3	1	2.9	0	0	0	0
G:C → T:A	8	6.3	0	0	19	22.1	5	14.3	0	0	0	0
G:C → C:G	0	0	0	0	5	5.8	0	0	0	0	0	0
Deletion (1 bp)	6	4.8	2	1.1	4	4.7	4	11.4	3	30.0	0	0
Insertion (1 bp)	0	0	0	0	3	3.5	0	0	0	0	0	0
Total mutations	126	100	182	100	86	100	35	100	10	100	13	100
Clones sequenced	21		15		19		17		22		27	
Nucleotides sequenced	16,510		11,790		14,639		13,512		15,810		21,485	

Table 2
Mutations at protein level.

	Taq DNA polymerase				Mutazyme II DNA polymerase				Nonbiased method[a]	Ideal method
	Mutagenic buffer		8-oxo-dGTP and dPTP		Amount of amplicon		Number of cycles			
	n	%	*n*	%	*n*	%	*n*	%	%	%
Total mutations	98	100	143	100	56	100	23	100		
Preserved amino acid	36	36.7	46	32.2	7	12.5	6	26.1	21.0–25.8	
Amino acid change	60	61.2	92	64.3	41	73.2	14	60.9	74.3–69.5	
Stop	2	2.1	5	3.5	8	14.3	3	13	4.7	
Impact of amino acid change	60	100	92	100	41	100	14	100		
Conservative[b]	31	51.7	48	52.2	19	46.3	5	35.7		
Nonconservative[b]	29	48.3	44	47.8	22	53.7	9	64.3		
Substitution	62	100	97	100	49	100	17	100		
Charged[c]	19	30.7	21	21.7	8	16.3	3	17.7		
Neutral[d]	18	29	31	32	16	32.7	9	52.9		
Aromatic[e]	4	6.5	3	3	2	4.1	0	0		
Aliphatic[f]	19	30.6	37	38.1	15	30.6	2	11.8		
Stop	2	3.2	5	5.2	8	16.3	3	17.6		
Stop and Gly/Pro codons	13	20.9	28	28.9	13	26.5	4	23.5	15.8–20.7	
Stop	2	3.2	5	5.2	8	16.3	3	17.6	4.7	
Gly/Pro	11	17.7	23	23.7	5	10.2	1	5.9	11.1–16.0	
Change of chemical group	46	100	66	100	34	100	9	100	100	100
Charged	11	24	11	16.7	5	14.7	2	22.2	24.9	25
Neutral	14	30.4	19	28.8	11	32.4	3	33.4	28.9	35
Aromatic	4	8.7	3	4.5	2	5.9	0	0	5.9	15
Aliphatic	15	32.6	28	42.4	8	23.5	1	11.1	35.6	25
Stop	2	4.3	5	7.6	8	23.5	3	33.3	4.7	0
Analyzed amino acids	4235		2998		3712		3451			

[a] Wong and coworkers [7].
[B] Conservative and nonconservative amino acid substitutions were defined according to the BLOSUM62 matrix [38]. For comparison, using the matrix for a nonbiased method, the impact of conservative and nonconservative amino acid changes would give the percentage values of 27.8 and 72.2, respectively.
[c] Charged amino acids (D, E, H, K, R).
[d] Neutral amino acids (C, M, P, S, T, N, Q).
[e] Aromatic amino acids (F, Y, W).
[f] Aliphatic amino acids (G, A, V, L, I).

encoded amino acid to a similar amino acid according to certain criteria. In our analysis, we applied the commonly used BLOSUM62 matrix [38]. Of amino acid changes, *Taq* DNA polymerase produced somewhat more conservative substitutions than did Mutazyme II DNA polymerase. With nonconservative substitutions, the trend was the opposite, Mutazyme II DNA polymerase yielding more changes.

Amino acid substitutions can be divided according to their properties into five groups: charged amino acids (D, E, H, K, R), neutral amino acids (C, M, P, S, T, N, Q), aromatic amino acids (F, Y, W), aliphatic amino acids (G, A, V, L, I), and stop codons [1]. Overall, the different mutagenesis methods induced a very similar substitution pattern, although a degree of variation could be observed in charged amino acid and stop codon substitutions.

For the comparison of random mutagenesis methods, Wong and coworkers [1] proposed a useful structure indicator that takes into account Gly and Pro substitutions as well as stop codons. Whereas *Taq* DNA polymerase produced proportionally more Gly/Pro substitutions than stop codons, the trend was the opposite with Mutazyme II DNA polymerase. Together, these protein structure-disturbing amino acid substitutions accounted for approximately 25% of total substitutions independent of the method.

Another useful difference indicator among mutagenesis methods takes into account only those substitutions that result in the change of the chemically categorized group, including stop codons [7]. Overall, the patterns were very similar to those obtained when the total repertoire of substitutions was analyzed except that the variation between polymerases was not detectable with the charged amino acid substitution category.

Discussion

Random mutagenesis methods constitute an important protein modification tool set, with applications ranging from protein engineering to directed evolution studies. Although a vast array of such methods has been described [1–5,42], larger studies that would critically evaluate and directly compare the performance parameters of various methods are lacking. We evaluated several of the most commonly used methods, including error-prone PCR protocols as well as one chemical and one biological mutagenesis method, and determined their operational ranges. We were particularly interested in the most important performance criteria for random mutagenesis—mutation frequency and its controllability. Accordingly, we determined the relative mutation frequencies under different mutagenic conditions for each of the evaluated methods. We also analyzed the induced mutation spectra at the nucleotide level and then assessed the diversity of mutation types at the amino acid level.

The blue/white *E. coli* colony assay was a convenient means to directly compare the operational ranges of the methods. This assay detects those mutations that critically alter the expression or activity of β-galactosidase, and it yields a quantitative measure for relative mutation frequencies. The PCR protocols generated mutations that were targeted to the amplified fragment only. Although the chemical and biological mutagenesis methods yielded mutations that were distributed throughout the reporter plasmid, this difference did not interfere with the interpretation of the data because of the comparatively much lower mutation frequencies. Sequencing was used to analyze the number and type of

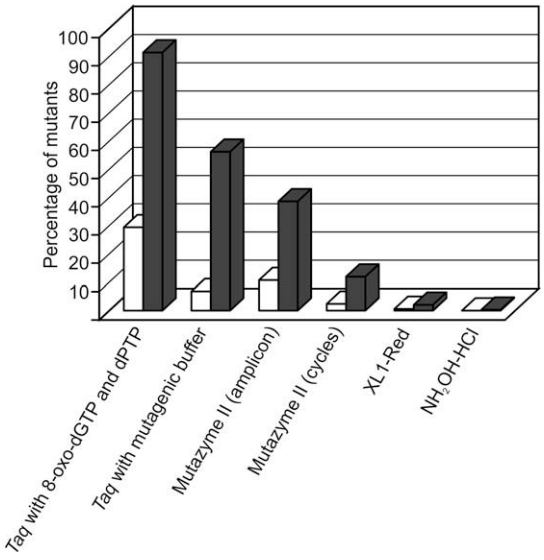

Fig. 4. Controllability of the evaluated random mutagenesis methods. For each method, the minimum and maximum percentages of discernible mutants obtained are shown with white and dark gray columns, respectively.

mutations for selected clones, allowing us to correlate the relative mutation frequencies to the actual number of mutated base pairs. It also produced data for the comparison of mutation spectra and amino acid substitution patterns.

Error-prone PCR methods proved to be highly efficient mutagenesis techniques with controllable mutation frequencies. *Taq* DNA polymerase, particularly with nucleotide analogues 8-oxo-dGTP and dPTP, was a highly potent mutagenesis inducer and provided a good control over a wide range of mutation frequencies (Figs. 2A, 2B, and 4). In previous studies the mutation frequency induced by *Taq* DNA polymerase has been controlled by the number of PCR cycles under a constant 400- or 500-μM concentration of nucleotide analogues, yielding a relatively high initial mutation frequency [14,39]. Our aim was not only to verify a high-frequency mutagenesis protocol but also to establish conditions where an excessive mutation load could be avoided; lowering the concentration of the analogues proved to be a straightforward means to achieve the latter goal.

Error-prone PCR with Mutazyme II DNA polymerase appeared to be a convenient mutagenesis method for the generation of mutant libraries of relatively low to moderate mutation frequency (Figs. 2C, 2D, and 4). The frequency could be regulated by adjusting the amount of the amplicon or the number of amplification cycles in the PCR. There do not appear to be any drastic differences between the mutation rates of different commercial Mutazyme DNA polymerase preparations given that our results with Mutazyme II were very similar to those obtained previously with the earlier Mutazyme I version [43].

With *E. coli* XL1–Red strain, the mutation frequency should be directly proportional to the number of generations that the DNA of interest has been propagated in the mutator strain, resulting in an approximately 5000-fold higher mutation rate than obtained with a wild-type *E. coli* strain [36]. Accordingly, the mutation rate of a colE1 origin-containing plasmid, such as pBluescript, is approximately 1 mutation/2 kb of cloned DNA after 30 generations of growth in XL1–Red mutator strain [36]. In our study, we attained a relative mutation frequency of 2% after approximately 100 generations (Figs. 2E and 4). In addition, the sequence analysis revealed a maximum of 1 mutation/500 bp in the selected (white or pale blue) clones. Although a direct comparison of mutation rates is not possible between the studies due to different analysis strategies, the results appear to be very similar. It is clear that to obtain substantially higher mutation rates with the XL1–Red mutagenesis protocol, a significant increase in the number of growth cycles should be implemented.

With the NH_2OH–HCl protocol, the maximum yield of discernible mutants (0.3%) was reached following approximately 25 h of exposure, and no increase was observed with longer incubation times (Figs. 2F and 4). Previously, a mutant proportion of up to 1.8% was reported in a study involving an experimental setup similar to that used in our study [16]. However, this result was achieved using a somewhat differing protocol (e.g., including pH 6.0 vs. 7.0 in the reaction), possibly explaining the difference. Overall, the NH_2OH–HCl protocol provided a convenient and cheap mutagenesis method for the generation of low mutation frequency libraries.

In general, all random mutagenesis methods developed to date are biased toward transition mutations, although efforts have been made to overcome the bias [8,44–46]. Transition bias was also observed in the current study (Tables 1 and 3 and Fig. 3). In this regard, Mutazyme II DNA polymerase was the least biased, yielding transitions and transversions with roughly equal proportions. (Note that an ideal nonbiased method would generate two-thirds of transversions and one-third of transitions [Fig. 3].) All of the methods generated some insertions and/or deletions, and in most of the cases the frequency was very low. Somewhat higher frequencies were observed with Mutazyme II protocols and XL1-Red. With regard to mutation bias between two transition mutations, A:T → G:C versus G:C → A:T, the *Taq* DNA polymerase protocols were biased toward A:T → G:C mutations and other methods were biased toward G:C → A:T mutations (Table 2 and Fig. 5A). A similar trend among methods was seen in the bias with A and T versus G and C nucleotide mutations (Table 2 and Fig. 5B). In general, the generated types of mutations and biases with all protocols were similar to those observed in previous studies [14–16,40,41,43,47], although in some cases apparent differences could be discerned. For example, Mutazyme II DNA polymerase, with two separate protocols, showed a difference in mutating A and T versus G and C nucleotides (Table 2 and Fig. 5B), and XL1–Red produced a large fraction of deletions (Fig. 3). However, a change, possibly associated with amplification bias and/or a relatively small num-

Table 3
Bias indicators of random mutagenesis methods.

	Taq DNA polymerase		Mutazyme II polymerase		XL1–Red	NH_2OH–HCl
	Mutagenic buffer	8-oxo-dGTP and dPTP	Amount of amplicon	Number of cycles		
Ts/Tv[a]	1.5	15.5	1.0	0.9	6.0	12.0
A:T → G:C/G:C → A:T[b]	2.0	2.4	0.4	0.7	0.0	0.2
A → N, T → N/G → N, C → N[c]	2.8	2.6	0.5	1.2	0.2	0.3

[a] Ts/Tv is a transition versus transversion ratio. A nonbiased mutational spectrum would yield the ratio of 0.5.
[b] A:T → G:C versus G:C → A:T mutations. A nonbiased mutational spectrum would yield the ratio of 1.
[c] Any change in A or T nucleotide versus any change in G or C nucleotide. A nonbiased mutational spectrum would yield the ratio of 1.

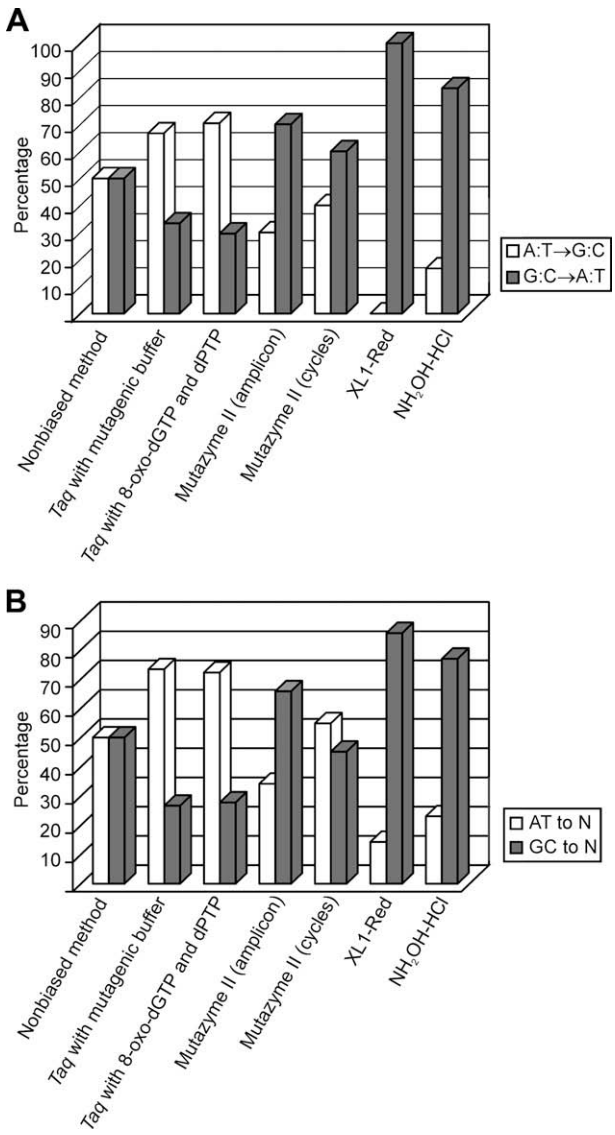

Fig. 5. Biases in the mutation spectra of the evaluated random mutagenesis methods. Insertion and deletion mutations are not taken into account. (A) Biases manifested in A:T → G:C versus G:C → A:T mutations, also called transition bias. (B) Biases manifested when any change from A or T nucleotide was compared with any change from G or C nucleotide.

ber of mutations sequenced, may have caused the apparent disagreements.

At the protein level, the amino acid substitution pattern differed among the error-prone PCR methods applied (Table 2). For example, *Taq* DNA polymerase preserved the amino acid more often than Mutazyme II DNA polymerase. A similar trend was also observed with structure-disturbing amino acid changes (Gly/Pro). However, the most striking difference was observed in the proportion of stop codon changes, with Mutazyme II DNA polymerase generating nearly 15% of them as compared with 2 to 4% generated by *Taq* DNA polymerase protocols. In general, differences were subtle among the rest of the analyzed amino acid change categories. All of the observed biases at the amino acid level conformed with the observed nucleotide changes. The methods that yielded biases toward A and T nucleotides had a higher probability of generating stop codons (TAA, TAG, TGA). Similarly, biases toward G and C nucleotides generated more Gly codons (GGT, GGC, GGA, GGG) and Pro codons (CCT, CCC, CCA, CCG) (Tables 1 and 2).

A publicly available mutagenesis assistant program (MAP) [7] can be used to statistically predict amino acid changes in a given target protein. It takes into account the protein-encoding nucleotide sequence of the target and the reported mutation spectra of several mutagenesis methods, including multiple *Taq* DNA polymerase protocols and a Mutazyme II protocol. This gave us the opportunity to compare the predicted amino acid changes of our target sequence with the actual amino acid changes observed in the current study (Table 4). With most of the amino acid change categories (vs. ideal), the predicted amino acid changes were comparable to those observed. However, charged amino acids were generated somewhat less frequently than predicted with both the *Taq* and Mutazyme II protocols. Significantly larger differences were observed with Mutazyme II DNA polymerase; aliphatic changes were generated less frequently and stop codons were generated excessively more frequently than predicted by the program. The data imply that small differences in the reported mutation spectra at the nucleotide level nevertheless can have significant consequences at the protein level.

Our data demonstrate the usability of different random mutagenesis methods and their operational ranges. In essence, different methods and parameters constitute a continuum from a very low to high mutation rate. Defined parameters can be used for controlled mutagenesis, where the frequency of random mutations can be estimated reliably. To cover the entire sequence space (i.e., to generate the most variable amino acid substitution reper-

Table 4
Amino acid substitution patterns for the analyzed LacZα and MuA protein sequences, including ideal distribution and deviations from the ideal.

	Charged (%)	Neutral (%)	Aromatic (%)	Aliphatic (%)	Stop (%)
Ideal distribution	25	35	15	25	0
Deviations according to this study					
Taq with mutagenic buffer[a]	−1.1	−4.6	−6.3	+7.6	+4.4
Taq with 8-oxo-dGTP and dPTP[b]	−8.3	−6.2	−10.5	+17.4	+7.6
Mutazyme II (amplicon)	−10.3	−2.7	−9.1	−1.5	+23.5
Mutazyme II (cycles)	−2.8	−1.7	−15.0	−13.9	+33.3
Deviations according to MAP[c]					
Nonbiased method	+0.9	−4.1	−10.0	+9.2	+4.0
Taq–Pol (Mn^{2+}/unbalanced dNTPs)[d]	+0.9	−7.1	−8.6	+10.4	+4.4
Taq–Pol with 8-oxo-dGTP and dPTP[e]	+5.8	−5.5	−12.5	+10.6	+1.5
Mutazyme II[f]	−0.2	−5.2	−8.5	+8.8	+5.1

[a] *Taq* with mutagenic buffer, in the presence of [Mg^{2+}] = 1.5 to 7 mM, [Mn^{2+}] = 0 to 5 mM, [dATP] = [dGTP] = 0.2 mM, and [dCTP] = [dTTP] = 0.2 to 1 mM.
[b] *Taq* with 8-oxo-dGTP and dPTP, in the presence of [8-oxo-dGTP] = [dPTP] = 0.2 to 100 μM.
[c] Wong and coworkers [7]. MAP is publicly available at http://map.jacobs-university.de/MAP.html.
[d] Shafikhani and coworkers [15]. *Taq* in the presence of [Mg^{2+}] = 7 mM, [Mn^{2+}] = 0.5 mM, [dATP] = [dGTP] = 0.2 mM, and [dCTP] = [dTTP] = 1 mM.
[e] Zaccolo and coworkers [14], *Taq* in the presence of [8-oxo-dGTP] = [dPTP] = 500 μM.
[f] Stratagene [40].

toire), the use of a combination of differentially biased methods is warranted.

Acknowledgments

We thank Tero Ahola for critical reading of the manuscript. Financial support was obtained from the Finnish National Technology Agency (TEKES, to H.S.) and the Academy of Finland (to M.I.P.).

References

[1] T.S. Wong, D. Zhurina, U. Schwaneberg, The diversity challenge in directed protein evolution, Comb. Chem. High Throughput Screen. 9 (2006) 271–288.
[2] C. Neylon, Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution, Nucleic Acids Res. 32 (2004) 1448–1459.
[3] L. Yuan, I. Kurek, J. English, R. Keenan, Laboratory-directed protein evolution, Microbiol. Mol. Biol. Rev. 69 (2005) 373–392.
[4] S. Lutz, W.M. Patrick, Novel methods for directed evolution of enzymes: quality, not quantity, Curr. Opin. Biotechnol. 15 (2004) 291–297.
[5] J. Kaur, R. Sharma, Directed evolution: an approach to engineer enzymes, Crit. Rev. Biotechnol. 26 (2006) 165–199.
[6] C. Jäckel, P. Kast, D. Hilvert, Protein design by directed evolution, Annu. Rev. Biophys. 37 (2008) 153–173.
[7] T.S. Wong, D. Roccatano, M. Zacharias, U. Schwaneberg, A statistical analysis of random mutagenesis methods used for directed protein evolution, J. Mol. Biol. 355 (2006) 858–871.
[8] J. Cline, H. Hogrefe, Randomize gene sequences with new PCR mutagenesis kit, Strategies 13 (2000) 157–161.
[9] B.D. Biles, B.A. Connolly, Low-fidelity *Pyrococcus furiosus* DNA polymerase mutants useful in error-prone PCR, Nucleic Acids Res. 32 (2004) e176.
[10] S. Emond, P. Mondon, S. Pizzut-Serin, L. Douchy, F. Crozet, K. Bouayadi, H. Kharrat, G. Potocki-Veronese, P. Monsan, M. Remaud-Simeon, A novel random mutagenesis approach using human mutagenic DNA polymerases to generate enzyme variant libraries, Protein Eng. Des. Sel. 21 (2008) 267–274.
[11] D.W. Leung, E. Chen, D.V. Goeddel, A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction, Techniques 1 (1989) 11–15.
[12] R.C. Cadwell, G.F. Joyce, Randomization of genes by PCR mutagenesis, PCR Methods Appl. 2 (1992) 28–33.
[13] J.H. Spee, W.M. de Vos, O.P. Kuipers, Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP, Nucleic Acids Res. 21 (1993) 777–778.
[14] M. Zaccolo, D.M. Williams, D.M. Brown, E. Gherardi, An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues, J. Mol. Biol. 255 (1996) 589–603.
[15] S. Shafikhani, R.A. Siegel, E. Ferrari, V. Schellenberger, Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization, BioTechniques 23 (1997) 304–310.
[16] J.H. Miller, A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992.
[17] P. Anderson, Mutagenesis, in: H.F. Epstein, D.C. Shakes (Eds.), Methods in Cell Biology, Academic Press, San Diego, 1995, pp. 31–58.
[18] B. Singer, J.T. Kusmierek, Chemical mutagenesis, Annu. Rev. Biochem. 51 (1982) 655–693.
[19] M. Smith, In vitro mutagenesis, Annu. Rev. Genet. 19 (1985) 423–462.
[20] Y.P. Lai, J. Huang, L.F. Wang, J. Li, Z.R. Wu, A new approach to random mutagenesis in vitro, Biotechnol. Bioeng. 86 (2004) 622–627.
[21] A.E. Stellwagen, N.L. Craig, Gain-of-function mutations in TnsC, an ATP-dependent transposition protein that activates the bacterial transposon Tn7, Genetics 145 (1997) 573–585.
[22] A. Greener, M. Callahan, B. Jerpseth, An efficient random mutagenesis technique using an *E. coli* mutator strain, Methods Mol. Biol. 57 (1996) 375–385.
[23] M. Camps, J. Naukkarinen, B.P. Johnson, L.A. Loeb, Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I, Proc. Natl. Acad. Sci. USA 100 (2003) 9727–9732.
[24] U.T. Bornscheuer, J. Altenbuchner, H.H. Meyer, Directed evolution of an esterase for the stereoselective resolution of a key intermediate in the synthesis of epothilones, Biotechnol. Bioeng. 58 (1998) 554–559.
[25] E. Henke, U.T. Bornscheuer, Directed evolution of an esterase from *Pseudomonas fluorescens:* random mutagenesis by error-prone PCR or a mutator strain and identification of mutants showing enhanced enantioselectivity by a resorufin-based fluorescence assay, Biol. Chem. 380 (1999) 1029–1033.
[26] X. Lu, H. Hirata, Y. Yamaji, M. Ugaki, S. Namba, Random mutagenesis in a plant viral genome using a DNA repair-deficient mutator *Escherichia coli* strain, J. Virol. Methods 94 (2001) 37–43.
[27] F. Belanger, G. Theberge-Julien, P.R. Cunningham, L. Brakier-Gingras, A functional relationship between helix 1 and the 900 tetraloop of 16S ribosomal RNA within the bacterial ribosome, RNA 11 (2005) 906–913.
[28] M.J. Callanan, W.M. Russell, T.R. Klaenhammer, Modification of *Lactobacillus* β-glucuronidase activity by random mutagenesis, Gene 389 (2007) 122–127.
[29] W.M. Patrick, A.E. Firth, J.M. Blackburn, User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries, Protein Eng. 16 (2003) 451–457.
[30] M.J. Volles, P.T. Lansbury Jr., A computer program for the estimation of protein and nucleic acid sequence diversity in random point mutagenesis libraries, Nucleic Acids Res. 33 (2005) 3667–3677.
[31] S.G. Grant, J. Jessee, F.R. Bloom, D. Hanahan, Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation–restriction mutants, Proc. Natl. Acad. Sci. USA 87 (1990) 4645–4649.
[32] J. Sambrook, D.W. Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
[33] A. Lamberg, S. Nieminen, M. Qiao, H. Savilahti, Efficient insertion mutagenesis strategy for bacterial genomes involving electroporation of in vitro-assembled DNA transposition complexes of bacteriophage Mu, Appl. Environ. Microbiol. 68 (2002) 705–712.
[34] T.A. Baker, M. Mizuuchi, H. Savilahti, K. Mizuuchi, Division of labor among monomers within the Mu transposase tetramer, Cell 74 (1993) 723–733.
[35] P. Saviranta, M. Pajunen, P. Jauria, M. Karp, K. Pettersson, P. Mäntsälä, T. Lövgren, Engineering the steroid-specificity of an anti-17β-estradiol Fab by random mutagenesis and competitive phage panning, Protein Eng. 11 (1998) 143–152.
[36] A. Greener, M. Callahan, B. Jerpseth, An efficient random mutagenesis technique using an *E. coli* mutator strain, Mol. Biotechnol. 7 (1997) 189–195.
[37] D. Burke, D. Dawson, T. Stearns, Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2000.
[38] S. Henikoff, J.G. Henikoff, Amino acid substitution matrices from protein blocks, Proc. Natl. Acad. Sci. USA 89 (1992) 10915–10919.
[39] M. Zaccolo, E. Gherardi, The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 β-lactamase, J. Mol. Biol. 285 (1999) 775–783.
[40] Stratagene, Overcome mutational bias, Strategies 17 (2004) 20–21.
[41] Stratagene, Overcome mutational bias, Strategies 18 (2005) 18–19.
[42] L.G. Otten, W.J. Quax, Directed evolution: selecting today's biocatalysts, Biomol. Eng. 22 (2005) 1–9.
[43] T. Vanhercke, C. Ampe, L. Tirry, P. Denolf, Reducing mutational bias in random protein libraries, Anal. Biochem. 339 (2005) 9–14.
[44] C. Fenton, H. Xu, E.I. Petersen, S.B. Petersen, M.R. el-Gewely, Random mutagenesis for protein breeding, in: J. Braman (Ed.), Methods in Molecular Biology, Humana Press, Totowa, NJ, 2002, pp. 231–241.
[45] T.S. Wong, K.L. Tee, B. Hauer, U. Schwaneberg, Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution, Nucleic Acids Res. 32 (2004) e26.
[46] T.S. Wong, D. Roccatano, D. Loakes, K.L. Tee, A. Schenk, B. Hauer, U. Schwaneberg, Transversion-enriched sequence saturation mutagenesis (SeSaM–Tv+): a random mutagenesis method with consecutive nucleotide exchanges that complements the bias of error-prone PCR, Biotechnol. J. 3 (2008) 74–82.
[47] J.H. Miller, Mutators in *Escherichia coli*, Mutat. Res. 409 (1998) 99–106.