We have developed a statistical method named MAP (mutagenesis assistant program) to equip protein engineers with a tool to develop promising directed evolution strategies by comparing 19 mutagenesis methods. Instead of conventional transition/transversion bias indicators as benchmarks for comparison, we propose to use three indicators based on the subset of amino acid substitutions generated on the protein level: (1) protein structure indicator; (2) amino acid diversity indicator with a codon diversity coefficient; and (3) chemical diversity indicator.

A MAP analysis for a single nucleotide substitution was performed for four genes: (1) heme domain of cytochrome P450 BM-3 from *Bacillus megaterium* (EC 1.14.14.1); (2) glucose oxidase from *Aspergillus niger* (EC 1.1.3.4); (3) arylesterase from *Pseudomonas fluorescens* (EC 3.1.1.2); and (4) alcohol dehydrogenase from *Saccharomyces cerevisiae* (EC 1.1.1.1). Based on the MAP analysis of these four genes, 19 mutagenesis methods have been evaluated and criteria for an ideal mutagenesis method have been proposed. The statistical analysis showed that existing gene mutagenesis methods are limited and highly biased. An average amino acid substitution per residue of only 3.15–7.4 can be achieved with current random mutagenesis methods. For the four investigated gene sequences, an average fraction of amino acid substitutions of 0.5–7% results in stop codons and 4.5–23.9% in glycine or proline residues. An average fraction of 16.2–44.2% of the amino acid substitutions are preserved, and 45.6% (epPCR method) are chemically different. The diversity remains low even when applying a non-biased method: an average of seven amino acid substitutions per residue, 2.9–4.7% stop codons, 11.1–16% glycine/proline residues, 21–25.8% preserved amino acids, and 55.5% are amino acids with chemically different side-chains.

Statistical information for each mutagenesis method can further be used to investigate the mutational spectra in protein regions regarded as important for the property of interest.

**Keywords:** protein engineering; directed evolution; random mutagenesis; bias indicator; MAP
purine or pyrimidine to pyrimidine) to transversions (\(T_v\); nucleotide substitutions: purine to pyrimidine or pyrimidine to purine). Since there are four possible transitions and eight possible transversions, a non-biased mutational spectrum (equal percentage of A→N, T→N, G→N, C→N nucleotide substitutions) will have a \(T_v/T_s\) ratio of 0.5 and an AT→GC/GC→AT ratio of 1.

An ideal method would allow substituting every amino acid of a protein sequence by its 19 counterparts in a statistical manner. For a protein engineer developing a directed evolution strategy or performing a directed evolution experiment, it is important to know which amino acid substitutions can be generated on the protein level at positions that have been identified as important by sequence alignments or rational design studies of proteins. The current transition/transversion bias indicator used to describe the mutational spectra is insufficient to provide such information. We illustrate, using four proteins with different functions and origins, that a non-biased mutational spectrum fails to generate highly diverse libraries with equal occurrence of 19 possible amino acid substitutions on the protein level. Reasons lie in the redundancy of the genetic code and its organization to minimize mutational errors. For example, transitions at third positions in codons and its organization to minimize mutational errors.9

Further, generating consecutive nucleotide substitutions has been a rare event in directed protein evolution due to methodological limitations of random mutagenesis methods such as error-prone polymerase chain reaction (epPCR).4

Here, we report a statistical analysis tool named MAP (mutagenesis assistant program), which has been designed to compare directed evolution methods by investigating the consequences of their mutational bias on the level of amino acid substitutions in the protein of interest. MAP investigates the amino acid substitution patterns at each amino acid position of a single gene for 19 random mutagenesis methods under the assumption of single nucleotide substitutions in one codon. MAP allows us: (1) to comprehensively analyse current random mutagenesis methods on the amino acid substitution level; (2) to develop mutational strategies by selecting random mutagenesis methods with different mutational bias; (3) to evaluate the performance of existing and novel random mutagenesis methods and (4) to assist in the design of a novel random mutagenesis method (e.g. a transversion-favoured method or a method introducing mainly charged amino acids).

To our knowledge, this method allows us for the first time to compare random mutagenesis methods on the protein level. As benchmarks for assessing a random mutagenesis method, we proposed three protein diversity indicators depending on the type, the extent and the chemical nature of amino acid substitutions. MAP has been implemented on an automatic server† and provides upon DNA sequence input a statistical analysis for 19 random mutagenesis methods.

**Results**

**Enzyme selection to avoid codon bias by origin and enzyme class**

Four enzymes of different origins and of different enzyme classes were selected for MAP analysis in order to avoid a codon bias by the host organisms or enzyme classes. The sense strand sequences encoding the four selected enzymes have the following composition: heme domain of cytochrome P450 BM-3 (A 33%, C 19%, G 22%, T 26%; 1368 bases), glucose oxidase (A 19%, C 36%, G 27%, T 18%; 1812 bases), arylesterase (A 20%, C 32%, G 30%, T 18%; 708 bases) and alcohol dehydrogenase (A 23%, C 23%, G 26%, T 28%; 1044 bases). A MAP analysis for these four proteins was performed by comparing on the protein level the mutational spectra of 19 mutagenesis methods (Table 1).

Results indicate how a random mutagenesis method with a non-biased mutational spectrum will perform on the protein level.

**Protein structure indicator: protein structure/function-disrupting substitutions**

The first important performance criterion of a random mutagenesis method is to limit the fraction of protein structure/function-disrupting (stop codons) and likely destabilizing amino acid substitutions (Gly and Pro residues). A mutation generating a stop codon in the gene sequence results on the protein level in a truncated enzyme variant that often loses its activity and Gly or Pro substitutions destabilize helical structures.

*Escherichia coli* is by far the workhorse of successful directed evolution studies.11,12 A stop codon (TAA, TGA and TAG) is generated with a frequency of 0.5–7% in the case of a single nucleotide substitution in each of the four genes analyzed. The general trend for stop codon frequencies is very similar for all four genes; however, in the case of the glucose oxidase gene, the stop codon frequency is, on average, significantly lower (Figure 1). The Taq (dPTP/8-oxodGTP) method introduces the least number of stop codons (0.5–1.8%) due to low percentages of N→A and N→T nucleotide substitutions. A non-biased random mutagenesis method will have a higher stop codon frequency (2.9–4.7%) than the Taq (dPTP/8-oxodGTP) method.

† http://gis-web.iu-bremen.de/MAP/

---

859
<table>
<thead>
<tr>
<th>Method</th>
<th>Designator</th>
<th>Transitions</th>
<th>Transversions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-biased method</td>
<td>Non-biased</td>
<td>A→G 8.33</td>
<td>T→C 8.33</td>
</tr>
<tr>
<td>Taq-Pol (unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (−, G&gt;A=C=T)</td>
<td>19.34</td>
<td>4.82</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / balanced dNTPs)&lt;sup&gt;33&lt;/sup&gt;</td>
<td>Taq (+, G&gt;A=C=T)</td>
<td>32.69</td>
<td>4.47</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (−, G&gt;A=C=T)</td>
<td>22.52</td>
<td>6.07</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (+, G&gt;A=C=T)</td>
<td>37.30</td>
<td>2.47</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (−, G&gt;T=A=C=T)</td>
<td>14.45</td>
<td>7.12</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (+, G&gt;T=A=C=T)</td>
<td>30.45</td>
<td>9.05</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (−, G&gt;T=A=C=T)</td>
<td>10.19</td>
<td>11.68</td>
</tr>
<tr>
<td>Taq-Pol (Mn&lt;sup&gt;2+&lt;/sup&gt; / unbalanced dNTPs)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Taq (+, G&gt;T=A=C=T)</td>
<td>14.14</td>
<td>13.49</td>
</tr>
<tr>
<td>Taq-Pol (nucleotide analogues dPTP and 8-oxodGTP)&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Taq (dPTP/8-oxodGTP)</td>
<td>39.10</td>
<td>6.60</td>
</tr>
<tr>
<td>Error-prone rolling circle amplification&lt;sup&gt;43&lt;/sup&gt;</td>
<td>epRCA</td>
<td>3.83</td>
<td>33.11</td>
</tr>
<tr>
<td>Pol I method&lt;sup&gt;9&lt;/sup&gt;</td>
<td>Pol I</td>
<td>17.35</td>
<td>22.45</td>
</tr>
<tr>
<td>E. coli expressing mutA allele of glyV gene&lt;sup&gt;16&lt;/sup&gt;</td>
<td>E. coli (mutA)</td>
<td>2.87</td>
<td>13.22</td>
</tr>
<tr>
<td>Chemical mutagen (nitrous acid)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Nitrous acid</td>
<td>27.27</td>
<td>22.27</td>
</tr>
<tr>
<td>Chemical mutagen (formic acid)</td>
<td>Formic acid</td>
<td>2.17</td>
<td>6.52</td>
</tr>
<tr>
<td>Chemical mutagen (hydrazine)</td>
<td>Hydrazine</td>
<td>22.92</td>
<td>10.91</td>
</tr>
<tr>
<td>Chemical mutagen (ethyl methane sulfonate)</td>
<td>EMS&lt;sup&gt;b&lt;/sup&gt;</td>
<td>17.10</td>
<td>21.05</td>
</tr>
</tbody>
</table>

These 19 random mutagenesis methods comprise (from top to down): 14 enzymatic methods, one whole cell method and four chemical methods. In the parentheses of the enzymatic methods a + sign indicates the supplement and a – sign the absence of Mn<sup>2+</sup>. The balance of the nucleotide composition is indicated after the + / − sign.

<sup>a</sup> Values adapted from the user manual of the Diversify PCR Random Mutagenesis Kit (Clontech).

<sup>b</sup> Taq-Pol and [dATP]=[dCTP]=[dTTP]=0.2 mM, [dGTP]=0.24 mM.

<sup>c</sup> Taq-Pol in the presence of [Mn<sup>2+</sup>]=0.25 mM and four standard nucleotides (each 0.02 mM).

<sup>d</sup> Taq-Pol in the presence of [Mn<sup>2+</sup>]=0.64 mM and [dATP]=[dCTP]=[dTTP]=0.2 mM, [dGTP]=0.24 mM.

<sup>e</sup> Taq-Pol in the presence of [Mn<sup>2+</sup>]=0.64 mM and [dATP]=[dCTP]=[dTTP]=0.2 mM, [dGTP]=0.40 mM.

<sup>f</sup> Taq-Pol in the presence of Mn<sup>2+</sup> and [dATP]=[dGTP]=0.2 mM, [dCTP]=[dTTP]=1 mM.

<sup>g</sup> Taq-Pol in the presence of Mn<sup>2+</sup>=0.5 mM and [dCTP]=[dATP]=0.03 mM, [dGTP]=[dTTP]=1 mM.

<sup>h</sup> EMS stands for ethyl methane sulfonate.
The percentage of variants resulting in Gly or Pro substitutions varies between 4.5% and 23.9% (Figure 2). The absolute values differ significantly among the mutagenesis methods and might influence the random mutagenesis strategy. The relative distribution of substitutions is again similar for all four genes. Interestingly, there is an opposite trend between random mutagenesis methods generating stop codons and Gly/Pro substitutions. The Taq (dPTP/8-oxoGTP) method introduces the highest number of Gly or Pro residues due to the high percentages of N/G and N/C nucleotide substitutions. Using a non-biased random mutagenesis method, we expect 11.1–16.0% of Gly/Pro substitutions.

**Amino acid diversity indicator: substitutions leading to amino acid changes and the average number of amino acid substitutions per residue**

The second important performance criterion for adjusting the mutation frequency in a random mutagenesis method is measured by the fraction of variants with preserved amino acids and the average number of amino acid substitutions per residue. Fraction of preserved amino acids is shown in Figure 3 for all 19 random mutagenesis methods and the four investigated genes. It ranges from 16.2–44.2% and clusters around 25–30% preserved amino acids. The relative values of preserved amino acids differ much less (25–34.8%) for arylesterase compared to the other three proteins. The overall trends are, except for arylesterase, similar with differences in absolute values. Fractions of preserved amino acids generated by various epPCR methods differ significantly from gene to gene and are lower for glucose oxidase. It is important to note that a method with non-biased mutational spectrum that hits every nucleotide with an equal frequency results in 21–25.8% of preserved amino acids.

The average number of amino acid substitutions per residue ranges for all four genes from 3.15 to 7.4
amino acids per residue (Figure 4). The absolute and relative values are similar for all four genes. Most commonly used epPCR methods give an average of 6.5 amino acids per residue. Despite the bias of epPCR, the average number of amino acid substitutions per residue is comparable to a method with non-biased mutational spectrum, which results in an average of seven amino acid substitutions per residue. All chemical methods perform poorer than enzymatic epPCR methods, with averages around 4.7 amino acid substitutions per residue.

Natural diversity offers the possibility of exchanging each amino acid with 19 other amino acids. Current random mutagenesis methods such as epPCR access only ~34% of that diversity. A method with a non-biased mutational spectrum which is able to target every nucleotide with equal frequency, performs only slightly better than current epPCR methods (37%).

Codon diversity coefficient

Codon diversity coefficient is a coefficient that measures how random mutations are distributed among codons of a gene (Figure 5). In a method with non-biased mutational spectrum, the codon diversity coefficient has a value of 0 considering each nucleotide and each codon substitution to occur with equal frequency. Biased methods show preferences toward certain nucleotide exchanges (e.g., Taq (+, G > A = C = T) (biased to AT → GC, 74.6%), Taq (dPTP/8-oxodGTP) (biased to AT → GC, 78.2%), epRCA (biased to GC → AT, 66.22%) and nitrous acid (biased to GC → AT, 65.46%)) and mutate certain nucleotides in codons preferentially. In other words, biased mutagenesis methods generate “hot spots” of mutagenesis that compromise genetic diversity. It is therefore important to add as performance criterion to the amino acid diversity, a codon diversity coefficient that indicates
the extent of hot spots generated by a random mutagenesis method. Apart from the bias of the mutagenesis method, the codon diversity coefficient is influenced by the codon usage of the gene's origin. Most epPCR and chemical methods have high codon diversity coefficient values ranging from 14.3 to 56 (Figure 5). A more balanced distribution of mutations among codons can be achieved using Pol I, Mutazyme II, and E. coli (mutA).

Chemical diversity indicator: substitutions leading to amino acids with changed chemical properties (hydrophobic, charged, neutral)

The third important performance criterion for a random mutagenesis method is to analyze how chemically different the substituted amino acids are. Such a chemical diversity indicator might be helpful in the future to develop property-specific mutagenesis strategies. Non-conservative substitutions might be more promising for evolving protein towards non-natural conditions and properties such as organic or thermal resistance. Conservative mutations might be more promising for altering enantioselectivity, which requires often only slight energetic changes. For thermal resistance it might be beneficial to use a random mutagenesis method that introduces mainly charged amino acids, which form additional salt-bridges in loop areas.

For the chemical diversity indicator, we grouped the amino acids into four categories depending on chemical properties of amino acids. As mentioned, an ideal method allows us to substitute each residue equally with 19 other amino acids at each amino acid position. After neglecting stop codons, we expect an ideal “chemical distribution” of: 25% aliphatic amino acids (G, A, V, L, I; 5/20), 15%...
aromatic amino acids (F, Y, W; 3/20), 35% neutral amino acids (C, M, P, S, T, N, Q; 7/20) and 25% charged amino acids (D, E, H, K, R; 5/20).

In Table 2, we performed an analysis of 19 random mutagenesis methods using cytochrome P450 BM-3 heme domain as example. The results of the analysis are reported as the deviation of each random mutagenesis method from the ideal "chemical distribution" as described above. The analysis for a single nucleotide substitution reveals that: (1) neutral and aromatic amino acid substitutions are under-represented; (2) aliphatic amino acids are significantly over-represented; (3) charged amino acid substitutions are over or under-represented depending on the employed mutagenesis methods; and (4) stop codons occur frequently (6.3% using E. coli (mutA); 5.8% epRCA).

A method with a non-biased mutational spectrum gives a chemical distribution of 35.6% aliphatic, 5.9% aromatic, 28.9% neutral, 24.9% charged amino acids and 4.7% stop codons.

The reasons for the bias lie in the organization of the genetic code and in the different number of codons encoding for the same amino acid. The average number of codons per amino acid is: (1) 4.2 for aliphatic amino acids (21 codons encoding five amino acids); (2) 1.67 for aromatic amino acids (five codons encoding three amino acids); (3) three for neutral amino acids (21 codons encoding seven amino acids); (4) 2.8 for charged amino acids (14 codons encoding five amino acids); and (5) three for stop codons (three codons for translation termination).

It is important for the chemical diversity indicator to analyze to what extent each amino acid species is generated. Figure 6 shows a detailed analysis of the substitution pattern for each of the 20 amino acids of P450 BM-3 heme domain by comparing a method with a non-biased mutational spectrum.
with non-biased mutational spectrum (left; equal occurrence of each nucleotide substitution) to an epPCR method (right; Taq (+, G=A=C=T)). The values on the Y-axis of Figure 6 represent the fraction of amino acid substitutions with different amino acid side-chains. For a non-biased method only 55.5% of the substitutions are amino acids with chemically different side-chains. A biased epPCR method results in even less chemical diversity (45.6%). The redundancy of the genetic code preserves strongly the chemical nature of amino acids. The white spots in Figure 6 show the low diversity of amino acid substitution patterns that can be generated with state of the art mutagenesis methods such as epPCR.

Figure 7 shows the percentage of aliphatic, aromatic, neutral and charged amino acid substitutions generated by all 19 random mutagenesis methods for the four genes of interest. There are significant variations for each method when applied to different genes, especially in the percentage values of aromatic amino acids. The differences are larger for chemical methods compared to epPCR methods or a whole cell method.

Figure 8 shows the crystal structures of P450 BM-3 heme domain, colored according to probabilities of (a) amino acid substitutions, (b) stop codons and (c) Gly/Pro substitutions using an epPCR method (Taq (+, G=A=C=T), left) and a chemical method (EMS, right) for mutagenesis. The graphical representation visualizes mutagenic hot spots and positions of structurally/functionally less favorable substitutions (stop codons and Gly/Pro). A closer look at catalytically important and well characterized residues (R47, Y51, F87 and L188) of P450 BM-3 heme domain reveal interesting differences between an epPCR method (Taq (+, G=A=C=T)) and a chemical method (EMS). For the chemical mutagenesis method an almost equal substitution probability was calculated for the catalytically
important residues. For the epPCR method, the probability of F87 substitution is twice compared to that of R47 and L188. An average amino acid substitution per residue of 5.75 (epPCR method) and 2.5 (chemical method) can be achieved for these four residues. Position R47 plays an important role in substrate binding through an ionic interaction between the positively charged R residue and the negatively charged carboxylate of the natural fatty acid substrate. For position R47, we obtain the following substitution pattern: (a) epPCR method 7.7% H, 3.2% S, 1.2% P, 9.0% C, 2.6% L and 1.2% G; and (b) chemical method 33.3% S, 33.3% C, 14.3% G. The epPCR method generates substitutions to S and L that are not generated by the chemical method. The epPCR method has also a 19.6-fold higher probability of Gly/Pro substitutions. None of the methods generates the chemically similar lysine residue. Position Y51 is the second amino acid proposed to be involved in substrate orientation by forming a H-bond between the carbonyl group of the fatty acid and hydroxyl group of Y residue. For position Y51, we obtain the following substitution pattern: (a) epPCR method 2.8% stop codons, 34.1% H, 3.2% D, 9.7% N, 1.2% S, 35.4% C, 8.0% F; and (b) chemical method 17.7% stop codons, 25.5% H, 25.5% C. The epPCR method generates again a higher diversity on the protein level (six substitutions compared to two substitutions) with a low probability of generating a chemically similar serine. For the chemical method, the stop codon probability is very high in absolute and relative numbers (5.5-fold higher than epPCR method). Both methods fail to generate the Y51F substitution, which increases the hydroxylation efficiency towards aromatic compounds such as phenoxotoluene. Position 87 is in P450 BM-3 important for the regio- and stereoselectivity of the fatty acid hydroxylation. For position F87, we obtain the following substitution pattern: (a) epPCR method 24.2% S, 2.3% C, 6.9% Y, 6.9% L, 2.3% V; and (b) chemical method 33.3% S, 33.3% L. For position 87, the same trends as for residues R47 and Y51 can be observed in terms of substitution pattern and diversity. The epPCR method generates, in contrast to the chemical method, the important F87V mutation, which converts P450 BM-3 into a regio- and stereoselective epoxygenase. For position L188 in P450 BM-3 has been proven to influence substrate specificities in several studies. For position L188, we obtain the following substitution pattern: (a) epPCR method 5.1% R, 15.5% Q, 54.5% P, 3.2% M, 1.2% V; and (b) chemical method 20.6% P, 14.3% V. Again the epPCR method proves to be more diverse than the chemical method and generates important substitutions such as L188Q. In summary, the epPCR method is shown, for the four investigated residues of P450 BM-3, to be more diverse than the chemical method. This might, however, be different for other codons and chemical methods. The generated amino acid substitution patterns are, however, far below natural diversity and many important substitutions would have been missed by both methods.

Discussion

We have developed a statistical method (MAP) to provide protein engineers with a tool to develop directed evolution strategies by investigating the consequences of mutational bias on the protein level. The analyses of amino acid substitution
patterns for four genes of different origins and different functions revealed that current benchmarks based on transition/transversion bias indicators are insufficient to describe and to compare the performance of random mutagenesis methods. We therefore propose three indicators and one coefficient to describe the performance of random mutagenesis methods: (1) protein structure indicator; (2) amino acid diversity indicator with a codon diversity coefficient; and (3) chemical diversity indicator.

The protein structure indicator describes the influence of structure/function-disrupting amino acid substitutions and comprises two parts, percentage of stop codons and percentage of Gly/Pro substitutions. MAP analysis of four genes revealed that a random mutagenesis with non-biased mutational spectrum ($T_s/T_v = 0.5$, $AT \rightarrow GC/\text{GC} \rightarrow AT = 1$) generates 2.9–4.7% of stop codons and 11.1–16% of Gly/Pro substitutions. These high absolute values of structurally less favorable substitutions are hurdles for creating functional mutant libraries with high mutation frequencies (15–20 mutations per 1000 bp). In the case of 20 mutations per gene, 44.5–61.8% of variants would contain at least one stop codon and 90.5–96.9% of variants would contain at least one substitution to a Gly or Pro residue.

The numbers of stop codons and of Gly/Pro substitutions differ significantly among the 19 random mutagenesis methods (factors up to 10.8 for stop codon, Figure 1; and factor up to 4.4 for Gly/Pro, Figure 2). Interestingly, the random mutagenesis method with the highest percentage of stop codons generates for all four investigated enzymes the least Gly/Pro substitutions and vice versa. The protein structure indicator can therefore assist in adjusting mutation frequencies in random mutagenesis method for obtaining functional libraries as a compromise between stop codons and Gly/Pro residues.

The amino acid diversity indicator can be regarded as a valuable benchmark for the power of a random mutagenesis method to generate diversity on the
protein level. It is a measure for the redundancy of the genetic code and the bias of random mutagenesis methods. Biased epPCR methods generate, on average, only 34% of the possible diversity on the protein level and a large fraction of the substitutions result in unchanged amino acids (16.3–34.6%). The performance of epPCR methods is comparable to that of a random mutagenesis method with non-biased mutational spectrum (37% of possible diversity on the protein level, Figure 4). The latter result proves that current bias indicators (transition and transversion bias) are insufficient for describing the performance of random mutagenesis methods.

The amino acid diversity indicator should be complemented by a codon diversity coefficient that measures the distribution of amino acid substitutions over a whole protein sequence (Figure 5). Mutations occur at preferred codons (mutagenic hot spots), since mutational spectra of most random mutagenesis methods are biased. MAP analysis allows us to compare mutagenic preferences of different mutagenesis methods and to combine random mutagenesis methods with different preferences in order to maximize diversity. The codon diversity coefficient values differ significantly among the random mutagenesis methods (Figure 5).

Chemical diversity indicator is an important measure of the chemical diversity of amino acid substitutions generated by a random mutagenesis method. We hope that the chemical diversity indicator will assist in the future to develop property-specific mutagenesis strategies in which a bias to charged, hydrophobic or sterically demanding amino acid substitutions is beneficial for the property of interest (for examples, see Results). A detailed analysis for the P450 BM-3 heme domain revealed that the organization of the genetic code favors strongly conservative amino acid substitutions, resulting in chemically similar amino acids (Table 2 and Figure 6). These findings are in good agreement with the physicochemical and ambiguity reduction theory, which proposes that the genetic code structure was organized to reduce the deleterious effect of physicochemical properties between amino acids differing in one base of a codon. The current bias indicators (ratios of $T_v$/$T_c$ and AT→GC/GC→AT) do not consider the chemical diversity of amino acid substitutions.

### Conclusions and Propositions for Novel Random Mutagenesis Methods

MAP analyses shed light on the performance of current random mutagenesis methods. The main conclusions that can be drawn on the performance of current genetic engineering methods are: (a) lack of mutagenesis methods bypassing stop codons and minimizing structurally destabilizing amino acid substitutions (protein structure indicator); (b) low average number of amino acid substitutions per residue (amino acid diversity indicator); and (c) a non-biased mutagenesis method with single nucleotide substitution in one codon performs poorly due to the organization of the genetic code and is far from being chemically diverse (chemical diversity indicator).

The analysis of a non-biased mutagenesis method has proven that the current bias indicators (transition and transversion bias) are insufficient to evaluate the performance of a random mutagenesis method and cannot be used as a quality measure or benchmark for a mutagenesis method. The example of the Taq (dPTP/8-oxodGTP) method shows that a biased method can even lead to a lower number of stop codons. The organization of the genetic code

### Table 2. Amino acid substitution patterns generated by a non-biased random mutagenesis method and 19 random mutagenesis methods for P450 BM-3 heme domain

<table>
<thead>
<tr>
<th></th>
<th>Stop (%)</th>
<th>Charged (%)</th>
<th>Neutral (%)</th>
<th>Aromatic (%)</th>
<th>Aliphatic (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideal chemical distribution</td>
<td>0</td>
<td>25</td>
<td>35</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Non-biased</td>
<td>4.65</td>
<td>-0.07</td>
<td>-6.08</td>
<td>-9.10</td>
<td>+10.60</td>
</tr>
<tr>
<td>Taq (−, G&gt;A=C=T)</td>
<td>+3.23</td>
<td>-0.35</td>
<td>-5.58</td>
<td>-9.79</td>
<td>+12.49</td>
</tr>
<tr>
<td>Taq (+, G=A=C=T)</td>
<td>+2.33</td>
<td>+3.24</td>
<td>-9.28</td>
<td>-9.89</td>
<td>+13.61</td>
</tr>
<tr>
<td>Taq (+, G&gt;A=C=T)</td>
<td>+3.75</td>
<td>+0.30</td>
<td>-8.48</td>
<td>-8.50</td>
<td>+12.92</td>
</tr>
<tr>
<td>Taq (+, G=A=C=T)</td>
<td>+5.06</td>
<td>-1.88</td>
<td>-9.15</td>
<td>-7.05</td>
<td>+13.03</td>
</tr>
<tr>
<td>Taq (+, G=T, A=C)</td>
<td>+2.16</td>
<td>+2.94</td>
<td>-8.05</td>
<td>-10.03</td>
<td>+12.97</td>
</tr>
<tr>
<td>Taq (I614K)</td>
<td>+3.74</td>
<td>+0.43</td>
<td>-8.20</td>
<td>-8.46</td>
<td>+12.49</td>
</tr>
<tr>
<td>Mutazyme I</td>
<td>+5.76</td>
<td>-1.20</td>
<td>-4.68</td>
<td>-8.00</td>
<td>+8.12</td>
</tr>
<tr>
<td>Mutazyme II</td>
<td>+5.67</td>
<td>-1.85</td>
<td>-7.57</td>
<td>-7.13</td>
<td>+10.88</td>
</tr>
<tr>
<td>Pfu (exo-, D473G)</td>
<td>+5.20</td>
<td>-1.40</td>
<td>-7.56</td>
<td>-7.63</td>
<td>+11.39</td>
</tr>
<tr>
<td>Transcriptase</td>
<td>+3.51</td>
<td>-1.40</td>
<td>-4.78</td>
<td>-8.33</td>
<td>+11.00</td>
</tr>
<tr>
<td>Taq (dPTP/8-oxodGTP)</td>
<td>+0.92</td>
<td>+4.49</td>
<td>-7.42</td>
<td>-11.30</td>
<td>+13.30</td>
</tr>
<tr>
<td>EpRCA</td>
<td>+5.79</td>
<td>-2.12</td>
<td>-2.83</td>
<td>-7.85</td>
<td>+7.01</td>
</tr>
<tr>
<td>Pol I</td>
<td>+4.07</td>
<td>-0.19</td>
<td>-5.89</td>
<td>-8.49</td>
<td>+10.49</td>
</tr>
<tr>
<td>E. coli (mutA)</td>
<td>+6.32</td>
<td>-2.65</td>
<td>-5.67</td>
<td>-7.39</td>
<td>+9.40</td>
</tr>
<tr>
<td>Nitrous acid</td>
<td>+4.77</td>
<td>-1.04</td>
<td>-10.82</td>
<td>-7.77</td>
<td>+14.86</td>
</tr>
<tr>
<td>Formic acid</td>
<td>+5.73</td>
<td>-7.30</td>
<td>-2.32</td>
<td>-5.90</td>
<td>+9.80</td>
</tr>
<tr>
<td>Hydrazine</td>
<td>+2.73</td>
<td>-0.31</td>
<td>-6.83</td>
<td>-8.96</td>
<td>+13.36</td>
</tr>
<tr>
<td>EMS</td>
<td>+2.54</td>
<td>+3.13</td>
<td>-3.48</td>
<td>-10.69</td>
<td>+10.50</td>
</tr>
</tbody>
</table>

Deviations in composition of categories (stop, charged, neutral, aromatic and aliphatic) from an ideal chemical distribution (amino acids: 25% charged, 35% neutral, 15% aromatic, and 25% aliphatic) are given.
with stop codons and structurally destabilizing amino acids prevents current random mutagenesis methods from generating functional mutant libraries with high mutation frequencies (15–20 mutations per 1000 bp). The analysis of the transition bias on the chemical properties revealed that amino acids are preferentially substituted by chemically similar amino acids (e.g. hydrophobic amino acids substituted by hydrophobic ones).

Instead of developing methods without transition and transversion bias, we have identified, using MAP analysis, the need for mutagenesis methods that generate more diversity on the protein level or generate subsets of amino acid substitutions. One desirable example would be to replace each amino acid of a protein by charged or hydrophobic amino acids. In this regard, a transversion biased mutagenesis method would be beneficial, since such a method would increase the number of amino acid substitutions with different chemical properties. The diversity of random mutagenesis methods can also be increased if methods with two or three consecutive nucleotide substitutions could be developed. Avoiding, for instance, CC and GG pairs would further reduce Gly and Pro substitutions. A notable step forward has been achieved by the random insertion and deletion (RID) mutagenesis method, which, however, lacks technical simplicity (nine ligation steps) and success stories are yet to be reported.

Until those novel random mutagenesis methods are developed, it is advisable to change the mutational bias by changing the random mutagenesis method used for mutant library generation. In this respect, MAP can be a valuable tool for developing mutational strategies by comparing the bias of random mutagenesis methods. As benchmarks for a random mutagenesis method, we propose to use in the future three protein

Figure 8. P450 BM-3 heme domain colored according to (a) amino acid substitution probability, (b) stop codon probability and (c) Gly/Pro substitution probability using an epPCR method (Taq (+, G=A=C=T), left) and a chemical method (EMS, right). The probability values increase from red (lowest probability) to blue (highest probability).
diversity indicators describing the extent and the chemical nature of amino acid substitutions. Currently, we are extending the MAP analysis to take secondary structure elements and steric constraints in proteins with known 3D structures into account.

MAP has been made publicly available† and it allows, upon a DNA sequence input, a statistical analysis for 19 random mutagenesis methods.

**Methods**

**Enzymes selected for MAP analysis**

MAP analysis has been performed on four genes encoding: (1) heme domain of cytochrome P450 BM-3 from *B. megaterium* (NCBI GenBank accession number J04832);29 (2) glucose oxidase from *A. niger* (AJ294936);30 (3) arylesterase from *P. fluorescens* (D12484);31 and (4) alcohol dehydrogenase from *S. cerevisiae* (V01292).32

**Random mutagenesis methods included in MAP**

Nineteen random mutagenesis methods were implemented into MAP. These methods comprise 14 enzymatic methods (six error-prone polymerase chain reactions employing Mn+2 and/or unbalanced nucleotide concentrations,33–38 Taq polymerase mutant J614K,36 Mutzyme I,37 Mutzyme II,37 Pfu polymerase (exo-) mutant D437C,38 reverse transcriptase,39 Taq polymerase with nucleotide analogues dPTP/8-oxodGTP,13 error-prone rolling circle amplification,40 Polymerase I method14), one whole cells-based method (*E. coli* expressing mutA allele of glyV gene16) and four chemical methods (employing nitrous acid,41 formic acid,41 hydrazine42 and ethyl methane sulfonate 42).

**Mutational spectra included in MAP**

The mutational spectra were changed in two ways compared to the reported data.4 Firstly, insertion and deletion with an occurrence frequency between 0.8% and 13.9% were neglected and occurrence of remaining nucleotide substitutions was scaled proportionally to 100%. Secondly, mutations in upper and lower DNA strand resulting in equivalent substitutions (e.g. A→C (lower strand), T→G (upper strand)) were considered to occur with equal frequency. The final mutational spectra used in a statistical MAP analysis are summarized in Table 1.

A Perl script was written to analyze all the possible amino acid substitutions resulting from a single nucleotide exchange in one codon of the gene of interest. Amino acid substitutions caused by two or more consecutive nucleotide exchanges were excluded due to limitations of current random mutagenesis methods.3 Four subsequent nucleotides (one codon) encode one amino acid in a protein. For a protein with *N* amino acids and considering a single nucleotide exchange in a codon, a DNA sequence space of $N \times [3+3+3] \times 2$ variants was sampled. Every nucleotide position of a codon was investigated independently by replacing every nucleotide position of a codon by three other nucleotides, resulting in $(3+3+3)$ exchanges per codon. Since substitutions occur at the sense and the anti-sense strand, a factor of two has been added.

**Statistics and definitions**

Amino acids were grouped into categories depending on the chemical properties of amino acid side-chains: category 1, aliphatic side-chains (G, A, V, L, I); category 2, aromatic side-chains (F, Y, W); category 3, neutral side-chains (C, M, P, S, N, Q); category 4, charged side-chains (D, E, H, K, R), and stop codons. The following technical terms were used: (1) fraction of variants with stop codons: fraction of single nucleotide substitutions resulting in a stop codon (TAA, TGA or TAG); (2) fraction of variants with Gly or Pro: fraction of single nucleotide substitutions resulting in one of the following codons (GGA, GGT, GGC, CCA, CCT, CCG or CCC); (3) fraction of variants with preserved amino acids: fraction of single nucleotide substitutions that do not change the encoded amino acid; (4) average amino acid substitution per residue: average number of amino acid substitutions after single nucleotide exchange of a codon; (5) codon diversity coefficient: coefficient of variance calculated on the probabilities of generating single nucleotide substitution in each codon encoding an amino acid in the protein of interest.

**Acknowledgements**

We thank German Academic Exchange Service (DAAD, Bonn, Germany) and Deutsche Forschungsgemeinschaft (DFG, award SCHW-790/2-1) for financial support.

**References**


† http://gis-web.iu-bremen.de/MAP/


