Protein evolution by hypermutation and selection in the B cell line DT40

Hiroshi Arakawa¹, Hiroaki Kudo¹, Vera Batrak¹, Randolph B. Caldwell¹, Michael A. Rieger², Joachim W. Ellwart³ and Jean-Marie Buerstedde¹*

¹Institute for Molecular Radiobiology, ²Institute of Stem Cell Research, GSF-National Research Center for Environment and Health, Ingolstaedter Landstrasse 1, D-85764 Neuherberg-Munich and ³Institute of Molecular Immunology, GSF National Research Center for Environment and Health, Marchioninistrasse 25, 81377 Munich, Germany

Received May 24, 2007; Revised July 3, 2007; Accepted July 29, 2007

ABSTRACT

Genome-wide mutations and selection within a population are the basis of natural evolution. A similar process occurs during antibody affinity maturation when immunoglobulin genes are hypermutated and only those B cells which express antibodies of improved antigen-binding specificity are expanded. Protein evolution might be simulated in cell culture, if transgene-specific hypermutation can be combined with the selection of cells carrying beneficial mutations. Here, we describe the optimization of a GFP transgene in the B cell line DT40 by hypermutation and iterative fluorescence activated cell sorting. Artificial evolution in DT40 offers unique advantages and may be easily adapted to other transgenes, if the selection for desirable mutations is feasible.

INTRODUCTION

Natural evolution, based on the selection of beneficial mutations within a population of genetic variants, has created the amazing diversity of life on our planet. While natural selection works on whole genomes, the evolution of individual proteins can be tracked by the analysis of intra-species polymorphisms and inter-species divergence. A fascinating example for the evolution of a single protein by mutation and selection is the affinity maturation of antibodies in B cells (1). Developing B cells activate and diversify their immunoglobulin (Ig) genes by recombination, but only cells encoding antibodies with antigen-binding specificity are stimulated to expand after antigen challenge. These B cells hypermutate their Ig genes and cycles of hypermutation and selection continue until antibodies of sufficiently high antigen-binding affinity emerge.

Recombinant DNA technologies are able to recapitulate the evolution of proteins by mutagenesis and selection in vitro. Most approaches combine the expression of diverse gene libraries in phage, bacteria or translation systems with selection for the encoded proteins (2,3). If the structure-function relationship of a protein is sufficiently understood, site-directed mutagenesis of critical amino acids can be used instead or in addition to random mutagenesis (4). The potential of in vitro protein evolution is illustrated by the example of the green fluorescent protein (GFP) (5) which could be changed into yellow, cyan or blue variants by random and site-directed mutagenesis and selection in vitro (6–8).

Affinity maturation of antibodies can be simulated ex vivo in cultures of hypermutating B cell lines by the enrichment of cells expressing antigen-specific antibodies (9,10). B cell lines could also be used for the optimization of non-immunoglobulin proteins, if the hypermutating activity were directed toward transfected transgenes and cells carrying beneficial mutations were selected. Advantages of this approach are the possibilities to generate an enormous amount of genetic diversity and to select for improved protein variants within a living cell culture. It was recently reported that a gene encoding a red fluorescent protein (FP) was transferred into the hypermutating human B cell line RAMOS, and the selection of cells emitting red-shifted fluorescence yielded the most far red-shifted FP protein known to date (11).

Despite this impressive result, hypermutation is difficult to control in RAMOS, because transgenes usually integrate at random chromosomal sites outside the hypermutating Ig loci. In contrast, transgenes can be easily inserted into the Ig loci of the chicken B cell line DT40 (12). DT40 diversifies its rearranged Ig light chain gene by pseudo V (ψV) gene-templated gene conversion.

*To whom correspondence should be addressed. Tel: +49 89 3187 2871; Fax: +49 89 3187 4093; Email: buersted@gsf.de

Present address:
Hiroaki Kudo, Graduate School of Medicine, Kyoto University, Yoshida Konoe, Sakyo-ku, Kyoto 606-8501, Japan

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
(13), but if gene conversion is blocked due to the inactivation of RAD51 paralogues or the deletion of gene conversion donors, hypermutation occurs (14,15). Ig gene diversification by conversion or hypermutation requires expression of the AID gene (15,16). Based on these results, we reasoned that transgenes inserted into the Ig loci of DT40 without nearby gene conversion donors would be diversified by hypermutation in AID positive cells. To test this hypothesis, we inserted the enhanced GFP (eGFP) gene (17) into the rearranged Ig light chain locus and searched for cells displaying increased fluorescence. Only three rounds of fluorescence activated cell sorting (FACS) were sufficient to isolate cells expressing eGFP variants whose fluorescent intensity appears to be superior to the best GFPs currently available for vertebrate cell labeling.

MATERIALS AND METHODS

Cell lines

Cells were cultured in chicken medium (RPMI-1640 or DMEM/F-12 with 10% fetal bovine serum, 1% chicken serum, 2 mM l-glutamine, 0.1 mM β-mercaptoethanol and penicillin/streptomycin) at 41°C with 5% CO2. The AID expressing DT40 clones, AID R1CL1 and AID R2CL2, used in the study for hypermutation were derived from the AID knockout cell clone DT40 Cre1AID2/2 (16) by stable transfection of a floxed AID—IRES (internal ribosome entry site)—gpt (guanine phosphoribosyl transferase) bicistronic cassette. The IRES sequence was derived from Encephalomyocarditis virus. Because the bi-cistronic cassette is driven by the strong β-actin promoter and since AID is the first gene downstream of the promoter, AID protein expression in AID R1CL1 and CL2 is higher than in wild-type DT40 cells (unpublished data). The cell clones express the Cre recombinase as a MerCreMer fusion protein which is inactive due to its retention in the endoplasmic reticulum in the absence of estrogen derivatives (18). However, since background activity of MerCreMer can lead to undesired excision of the floxed AID-gpt expression cassette during prolonged culture, we selected for cells retaining AID by culturing in media containing 0.5 μg/ml of mycophenolic acid for 3 days following each preparative FACS sort.

An AID negative subclone of AID R1CL1 was generated by culturing the cells in chicken medium containing 1 μM 4-hydroxitamoxifen (Sigma) for 2 days and subsequent subcloning.

Targeted integration of transgenes into the rearranged Ig locus

Cells were transfected by electroporation using the Gene Pulser Xcell (BIO-RAD) at 25μF and 700 V (16) and stable transfectants were selected for using 1 μg/ml of puromycin. Transfectants having integrated the pHypermut1- or pHypermut2-derived constructs by targeted integration were identified by PCR using primer pairs P1/P2 and P1/P3, respectively. The frequency of targeted integration after transfection of pHypermut constructs was consistently more than 70% (data not shown). Since the constructs can target either the rearranged or the unrearranged Ig locus of DT40, integration into the rearranged Ig light chain locus was verified by PCR amplification of the VJ intervening sequence of the unrearranged locus using primer pairs P4/P5.

Flow cytometry

To quantify the appearance of GFP negative cells within proliferating cultures, the AID R1IgLeGFP1/AID−/−IgL eGFP and ψVAID R1IgLeGFP clones were subcloned by limited dilution, and 24 subclones of each were analyzed 2 weeks after subcloning by FACS. The preparative FACS sorts to select cells of increased fluorescence were performed using the MoFlo high-speed cell sorter (Cytomation).

PCR, cloning and sequencing

The screenings for targeted integration and the amplification of the eGFP transgenes were performed with the expand long template PCR system (Roche) under the following conditions: 2 min of initial incubation at 93°C; 35 cycles consisting of 10 s at 93°C, 30 s at 65°C and 5 min (plus an added 20 s per cycle) at 68°C and a final elongation step of 7 min at 68°C. The primer pair for the amplification of the eGFP gene from sorted cells was P6/P7. The PCR products were digested with HindIII and XhoI, cloned into the pUC119 plasmid vector and sequenced using primers P8 and P9. The mutated eGFP genes were then amplified using Pfu Ultra hotstart polymerase (Stratagene) and primer pair P10/P11, digested by AvrII and cloned into the NheI site of pHypermut2. The orientation and the sequence of the mutant eGFPs in pHypermut2 was verified by sequencing using primers P12 and P13.

Site-directed mutagenesis

To code the codon changes found in different eGFP variants, we designed primers (P14–P26) which included the intended mutations in the center of the primer sequence. In the first step, parts of the eGFP gene were amplified by PCR using at least one mutation-containing primer. In the second step, PCR fragments containing mutations were mixed with other PCR fragments to cover the full eGFP coding sequence. This mixture served as a template for chimeric PCR using the primer pair of P10/P11. The full length mutant eGFP PCR fragments were digested with AvrII and cloned into the NheI site of pHypermut2. The orientation and the sequence of the mutant eGFPs in pHypermut2 were confirmed by sequencing.

Primers

P1 GGGACTAGTAAAAATGATGCATAACCTTTTG CACA
P2 CGATTGAAGAACTCTATCCACTCAAATATA CCC
P3 CCCACCGACTCTAGAGGATCATATAATCGCC CCC
P4 TACAAAAACCTCCTGGCAGTGCAGAGGCA GCTGATGGTTTTTACTGTCT
The AIDR1IgLeGFP1 and AID−/−IgLeGFP1 clones were subcloned by limited dilution and 24 subclones of each were analyzed by FACS for GFP brightness 14 days after subcloning. Subclones of the AIDR1IgLeGFP1 clone contained cell populations showing lower (3.8%) and higher (0.5%) fluorescence than the dominant GFP positive cell population, whereas subclones of the AID−/−IgLeGFP1 clone consisted only of a homogeneous GFP positive population (Figure 1B). This suggests that the eGFP gene inserted into the rearranged Ig light chain locus is diversified in AIDR1IgLeGFP1 cells giving rise to variants of increased and decreased fluorescent intensity.

**Mutation activity of artificial evolution system**

To analyze the mutation spectrum of the eGFP transgenes, the coding sequence was amplified by PCR from AIDR1IgLeGFP1 and AID−/−IgLeGFP1 cells 6 weeks after subcloning, cloned into a plasmid vector and sequenced. A total of 13 nt changes were found within the 0.7 kb eGFP coding region of 39 sequences from the AIDR1IgLeGFP1 clone (Table 1). Assuming a DT40 doubling time of 10 h, the mutation rate of the eGFP coding region in the rearranged Ig light chain locus of AIDR1IgLeGFP1 cells is approximately $4.7 \times 10^{-9}$ mutations per base pair and division. This mutation frequency is approximately three times lower than the mutation rate of the rearranged VJ segment in the pseudogene deleted $\psi^V$ AIDR cell line (15) perhaps reflecting the lower incidence of hypermutation hotspots in the eGFP sequence. In both cases, the majority of mutations occurred at C/G base pairs (Table 1). Only 1 nt change which possibly represents a PCR artifact was found in 26 sequences of the AIDIgLeGFP1 clone confirming that the mutation activity of the eGFP transgene is dependent on AID.

**Sorting and sequence analysis of cells displaying increased fluorescence**

The FACS analysis suggested that a few cells within the AIDR1IgLeGFP1 culture had accumulated mutations which increased either the abundance or the fluorescent intensity of the eGFP protein. To enrich these cells, 10 million AIDR1IgLeGFP1 cells were sorted by preparative FACS using a gate which included only the 0.05% brightest cells (Figure 1C). The cells collected by the FACS sort were analyzed by FACS for GFP brightness 14 days after subcloning. Subcloning by limited dilution (16). In this way, AID positive (AIDR1IgLeGFP1) and negative (AID−/−IgLeGFP1) clones could be isolated which had deleted the bsr marker gene from the light chain locus.
Figure 1. Strategy for artificial evolution of eGFP gene. (A) A physical map of the chicken rearranged Ig light chain locus, the pHypermut1-eGFP targeting construct and the rearranged Ig light chain locus after targeted integration and marker excision is shown. The positions of primers used for the identification of targeted integration events are shown by arrows. (B) FACS profiles of AID R1IgLeGFP1 and AID−/−IgLeGFP1 clones. The average percentages of events falling into the GFP high and GFP low gates based on the measurement of 24 subclones are shown. (C) Sorting strategy for cells of increased fluorescence activity.

Table 1. Mutation profile

<table>
<thead>
<tr>
<th>Cell source</th>
<th>Gene Source</th>
<th>Mutations</th>
<th>Number of sequences</th>
<th>Mutations/sequence</th>
<th>Mutations at C</th>
<th>Mutations at G</th>
<th>Mutations at A</th>
<th>Mutations at T</th>
<th>Duplication</th>
<th>Deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>AID−/−IgLeGFP1</td>
<td>eGFP (723 bp)</td>
<td>1</td>
<td>26</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AID R1IgLeGFP1</td>
<td>eGFP (723 bp)</td>
<td>13</td>
<td>39</td>
<td>0.33</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>AID R1IgLeGFP1 Sort I</td>
<td>eGFP (723 bp)</td>
<td>15</td>
<td>95</td>
<td>0.16</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AID R1IgLeGFP1 Sort I</td>
<td>polyA-L (300 bp)</td>
<td>88</td>
<td>95</td>
<td>0.93</td>
<td>1</td>
<td>87</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AID R1IgLeGFP1 Sort III</td>
<td>eGFP (723 bp)</td>
<td>221</td>
<td>85</td>
<td>2.60</td>
<td>130</td>
<td>93</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Their sequences were derived from the different region of the same plasmids.
of the eGFP coding sequence by primer walking. This analysis revealed frequent mutations at the splice donor site of the Ig light chain leader intron (88 mutations/95 seq) (Figure 2A, Table 1). Although the role of these mutations downstream of the eGFP polyA signal remains speculative, it is possible that they enhance fluorescence by increasing eGFP mRNA stability.

In contrast, most of the sequences after the third sort showed mutations in the eGFP coding sequence (221 mutations/85 seq) (Figure 2B, Table 1). Comparison of the sequences of each sorted subclone allowed us to reconstruct evolutionary trees showing the stepwise accumulation of eGFP codon changes (Figure 2C). Many of the codon changes were not shared by different subclones and their functional significance remains uncertain. However, a number of codon changes occurred independently in different subclones suggesting that they are related to the increased fluorescence of the sorted cells. Mutated eGFP sequences which were found more than once were named GFP variants (GFPv1–9).

**Development of a new vector for transgene evolution in DT40**

Although we were able to use pHypermut1 for the enhancement of eGFP, this vector is not ideally suited for transgene evolution in DT40, as mutations outside the transgene coding region had apparently been selected for after the first sort. We therefore developed a new targeting vector named pHypermut2 in which the expression of the rearranged Ig light chain gene, because transgene transcription is driven by the Rous sarcoma virus (RSV) promoter in the opposite direction of the light chain gene (Figure 3A), pHypermut2 allows transgenes to be cloned into multiple cloning sites (NheI, EcoRV and BglII). The vector’s puromycin resistance (puromycin resistance) cassette, located downstream of the transgene can be excised by Cre recombinase (19), but it is unlikely to interfere with transgene transcription or mutagenesis. A convenient feature of pHypermut2 is the presence of a SpeI site upstream of the RSV promoter which can be used for the insertion of gene conversion donor sequences (20).

We inserted the eGFP transgene into pHypermut2, and transfected the construct into the V discovery cell line (15) (Figure 3B). To estimate the hypermutation rate within the eGFP transgene, one of the stable transfected clones 14 days of culture revealed that on average 10.9% of the cells showed decreased or lost fluorescence (Figure 3C). This result indicates that transgenes inserted by pHypermut2 into the rearranged Ig light chain locus are efficiently diversified by hypermutation.

**Confirmation of the variant GFP phenotypes**

The phenotype of mutations is best confirmed by retransfection into the genetically stable AID−/− clone and the analysis of transfectants which have integrated the gene as a single copy into the rearranged locus. The pHypermut2 vector is well suited for this task and all isolated GFP variants, as well as the eGFP and the Emerald (8) control sequences, were cloned into pHypermut2. These constructs were then transfected into AID−/−. More than 50% of all stable transfectants had integrated the constructs into the rearranged light chain locus and these clones were named according to the transgene (AID−/−/ILGFPv5, AID−/−/ILEmerald, AID−/−/ILGFPv1, etc.).

The fluorescent intensity of the different transfectants was compared by FACS (Figure 4A). Two independent transfectants of each variant GFP gene were included in the experiment to account for variation among transfectants of the same gene. This analysis revealed that the Emerald transfectant was 1.2 times brighter than the eGFP transfectant, four of the nine GFP variant transfectants (AID−/−/ILGFPv2, AID−/−/ILGFPv5, AID−/−/ILGFPv6 and AID−/−/ILGFPv8) were brighter than the eGFP transfectant. Three variant GFP transfectants (AID−/−/ILGFPv5, AID−/−/ILGFPv6 and AID−/−/ILGFPv8) had even higher fluorescent intensity than the Emerald transfectant (Table 2). Of all variant GFP transfectants, AID−/−/ILGFPv6 had the highest fluorescent intensity (2.5-fold more than the AID−/−/ILGFPv5) and AID−/−/ILGFPv5 had the second highest intensity (2.0-fold brighter than the AID−/−/ILGFP) (Figure 4A, Table 2). Little fluorescent variation was observed among transfectants of the same transgene.

To compare the fluorescent intensity of eGFP and GFPv6 in human cells, the pHypermut2 constructs of GFPv6 and eGFP were transiently transfected into the human embryonic kidney cell line HEK293T. By FACS analysis, cells transfected with the GFPv6 construct showed on average more fluorescence than cells transfected with the eGFP construct (data not shown).

**Increased fluorescence by the combination of mutations**

As certain combinations of codon changes were not found among the GFP variants, we wondered whether the accumulation of mutations in a single sequence may lead to further enhancement of fluorescent intensity (Figure 4B). At first, we combined Y145F and S202T of GFPv6 with V163A of GFPv5 to produce GFPv10. Because GFPv1 harboring the single Q80E mutation showed less fluorescent intensity than eGFP, we subtracted Q80E from GFPv10 to generate GFPv11. We also added L221V of GFPv8 to GFPv11, thereby generating GFPv12. Finally, we added Q184E and A206T of GFPv2 as well as L141V of GFPv9 to GFPv12 to generate GFPv13. These new variants were cloned into pHypermut2 and inserted into the rearranged Ig light chain locus of the clone AID−/−.

Transfectants of all new GFP variants showed higher fluorescent intensity than AID−/−/ILGFPv6 by FACS analysis (Figure 4C, Table 2). AID−/−/ILGFPv10 had 3.0-fold higher fluorescent intensity than AID−/−/ILGFPv6, but AID−/−/ILGFPv11 and AID−/−/ILGFPv12 were even brighter, displaying 3.2- and 3.3-fold more fluorescence respectively than AID−/−/ILGFPv6. AID−/−/ILGFPv13 did not surpass the fluorescent intensity of the other variant.
Figure 2. Mutations downstream and within the eGFP transgene. (A) Mutations identified at the exon/intron border of the Ig light chain leader sequence. The number of times a mutation was found is shown by superscript. (B) Mutations within the eGFP coding sequence. Mutations were mapped below the reference eGFP sequence together with the corresponding amino acid codons. The eGFP sequence which we used had one codon (Val) inserted next to the start codon to yield an optimal translation-initiation sequence (Kozak motif). This Val was numbered as codon 1a according to Crameri (26) for the easier comparison to wild-type GFP and previously reported GFP mutants. (C) Pedigrees for the evolution of the eGFP transgenes in culture. The number of times each sequence was found within each subclone is indicated at the right side of the circles. The amino acid changes of each step are shown beside the arrows. Sequences identified in more than one subclone are named v1–v9.
GFP transfectants. The variants GFPv11 and GFPv12 which offer the most potential for the labeling of vertebrate cells were named GFPnovo1 and GFPnovo2, respectively.

**Spectral properties of the new GFP variants**

The excitation and the emission maximum of eGFP are at 488 and 509 nm wavelength, respectively (7,8). To determine the spectral profiles of the new variant GFP proteins, the control and variant gene transfectants (AID−/−IgL<sup>GFp5</sup>, AID−/−IgL<sup>Emerald</sup>, AID−/−IgL<sup>GFp6</sup>, AID−/−IgL<sup>GFPnovo1</sup> and the AID−/−IgL<sup>GFPnovo2</sup>) were analyzed using a luminescence spectrometer (Figure 4D). This analysis showed that the fluorescent intensity, but not the excitation and emission spectra had been changed in the variant GFP transfectants. The lack of mutants showing altered excitation or emission properties is most likely due to our selection strategy which only enriched cells for increased fluorescent intensity.

We also examined cells of the eGFP, Emerald and variant GFP transfectants by fluorescence microscopy (Figure 4E). Consistent with the FACS and spectrometer analysis, cells of the GFPv5, GFPv6, GFPnovo1 and GFPnovo2 transfectants produced brighter images than cells of the eGFP and Emerald transfectants, but there was no noticeable change of color.

**DISCUSSION**

We describe here an artificial evolution system in the B cell line DT40 which offers advantages over alternative approaches and could be of generic value for the optimization of many types of proteins in cell culture. The main advantages are the high transgene-specific mutation rates, the easy confirmation of mutant
Figure 4. Analysis of variant GFP transfectants. (A) FACS analysis of control and variant GFP transfectants. The average fluorescence of AID London/IgLeGFP is indicated by a green line for easier comparison. (B) Flow chart of site-directed mutagenesis of GFP variants. (C) Relative fluorescence of the transfectants normalized to the fluorescence of AID/C0/IgLeGFP. (D) Excitation and emission spectra. (E) Image of single cells by fluorescence microscopy. The image of the same single cells is shown with fluorescence activation (upper row) and without fluorescence activation (lower row).
Amino acid changes and brightness of GFP variants

<table>
<thead>
<tr>
<th>GFP variant</th>
<th>Amino acid changes</th>
<th>Number of sequences</th>
<th>Subclone</th>
<th>Relative brightnessa</th>
</tr>
</thead>
<tbody>
<tr>
<td>eGFP</td>
<td>Change from wild-type GFP</td>
<td>–</td>
<td>–</td>
<td>1.0 ± 0.1</td>
</tr>
<tr>
<td>GFPv1</td>
<td>Q80E</td>
<td>28</td>
<td>1, 2, 3</td>
<td>0.7 ± 0.0</td>
</tr>
<tr>
<td>GFPv2</td>
<td>Q80E, Q145F, Y163A, L221V</td>
<td>19</td>
<td>1</td>
<td>1.2 ± 0.0</td>
</tr>
<tr>
<td>GFPv3</td>
<td>Q80E, E124D, L141V, Y237stop</td>
<td>11</td>
<td>3</td>
<td>0.3 ± 0.0</td>
</tr>
<tr>
<td>GFPv4</td>
<td>Q80E, S202T</td>
<td>2</td>
<td>3</td>
<td>0.9 ± 0.1</td>
</tr>
<tr>
<td>GFPv5</td>
<td>Q80E, V163A</td>
<td>4</td>
<td>1</td>
<td>2.0 ± 0.0</td>
</tr>
<tr>
<td>GFPv6</td>
<td>Q80E, Y145F, S202T</td>
<td>3</td>
<td>2</td>
<td>2.5 ± 0.0</td>
</tr>
<tr>
<td>GFPv7</td>
<td>Q80E, S208T</td>
<td>2</td>
<td>3</td>
<td>0.7 ± 0.1</td>
</tr>
<tr>
<td>GFPv8</td>
<td>Q80E, L221V</td>
<td>2</td>
<td>3</td>
<td>1.5 ± 0.3</td>
</tr>
<tr>
<td>GFPv9</td>
<td>Q80E, L141V, S202T</td>
<td>5</td>
<td>3</td>
<td>1.0 ± 0.1</td>
</tr>
<tr>
<td>GFPv10</td>
<td>Q80E, Y145F, V163A, S202T</td>
<td>–</td>
<td>–</td>
<td>3.0 ± 0.1</td>
</tr>
<tr>
<td>GFPv11</td>
<td>Y145F, V163A, S202T</td>
<td>–</td>
<td>–</td>
<td>3.2 ± 0.0</td>
</tr>
<tr>
<td>GFPv12</td>
<td>Y145F, V163A, S202T, L221V</td>
<td>–</td>
<td>–</td>
<td>3.3 ± 0.0</td>
</tr>
<tr>
<td>GFPv13</td>
<td>L141V, Y145F, V163A, Q184E, S202T, A206T, L221V</td>
<td>–</td>
<td>–</td>
<td>2.9 ± 0.1</td>
</tr>
<tr>
<td>Emerald</td>
<td>Change from wild-type GFP</td>
<td>–</td>
<td>–</td>
<td>1.2</td>
</tr>
</tbody>
</table>

aSDS based on the analysis of two clones.

phenotypes and the selection for beneficial mutations within the environment of live vertebrate cells. We have illustrated the utility of the system by generating new GFP variants which display more than 3-fold higher fluorescence activity than the best GFPs currently available for bio-imaging of vertebrate cells.

We have chosen eGFP as a first example for protein evolution in DT40, because the protein has been extensively optimized by random and site-directed mutagenesis and selection of cells displaying brighter fluorescence is straightforward. Only three rounds of FACS sorting during 2 months of culturing were sufficient to isolate mutants of increased fluorescent intensity. Some of the identified mutations are identical to mutations previously reported by others for GFP variants with an altered excitation and/or emission spectrum. For example, previously reported by others for GFP variants with an altered excitation and/or emission spectrum. For example, previously reported by others for GFP variants with an altered excitation and/or emission spectrum.

Table 2. Amino acid changes and brightness of GFP variants

The high efficiency of targeted gene integration in DT40 allows one to standardize the hypermutation of transgenes and the confirmation of mutant phenotypes (Figure 5). We envision that the transgene of interest is first cloned into the pHypermut2 vector and then transfected into a DT40 clone which conditionally expresses AID. Transfectants, which hypermutate the transgene after targeted integration into the rearranged Ig light chain locus, are subjected to iterative cycles of cell selection and expansion. When a cell population of the desired phenotype is isolated, hypermutation is shut off by excision of the AID expression cassette and the cells are subcloned. Transgene sequences from the most promising subclones are amplified by PCR and cloned into the pHypermut2 vector for sequencing. In the end, the mutated transgenes within the pHypermut2 vector are useful in situations where maximal sensitivity is required for bio-imaging.
transfected back into AID negative cells to verify their phenotype.

In principle, DT40 can be used for the evolution of any coding or non-coding sequence whose transcription is tolerated. However, evolution of transgenes encoding non-fluorescent proteins is only feasible, if rare cells carrying desirable mutations can be enriched from the bulk culture. As FACS sorting is ideal for selection, one might attempt to tie the optimization of non-fluorescent proteins to fluorescent signals. For example, a cell surface receptor expressed on the surface of DT40 might be optimized using a fluorescent coupled ligand. Most likely, the selection strategies will need to be customized for different types of proteins and the outcome of these efforts will determine the success of artificial evolution in DT40.

ACKNOWLEDGEMENTS

We are grateful to Claire Brellinger for excellent technical assistance. This work was supported by the Geninteg grant of the FP6 framework and the DGF program ‘Networks in genome expression and maintenance’. Funding to pay the Open Access publication charges for this article was provided by XXX.

Conflict of interest statement. None declared.

REFERENCES


Figure 5. Scheme of artificial evolution in DT40. The approach only requires a cell line such as AIDR which conditionally expresses AID and the pHypermut2 targeting vector.


