FEBS Letters 507 (2001) 16–20 FEBS 25326

GFP-like chromoproteins as a source of far-red fluorescent proteins[1]

Nadya G. Gurskaya[a,2], Arkady F. Fradkov[a,2], Alexey Terskikh[b], Mikhail V. Matz[a,3], Yulii A. Labas[c], Vladimir I. Martynov[a], Yurii G. Yanushevich[a], Konstantin A. Lukyanov[a], Sergey A. Lukyanov[a,*]

[a] *Shemiakin and Ovchinnikov Institute of Bioorganic Chemistry RAS, Miklukho-Maklaya 16/10, 117997 Moscow, Russia*
[b] *School of Medicine, Stanford University, Stanford, CA 94305, USA*
[c] *Institute of Ecology and Evolution RAS, Leninsky pr. 33, 117071 Moscow, Russia*

Received 17 July 2001; revised 10 August 2001; accepted 10 September 2001

First published online 26 September 2001

Edited by Richard Cogdell

Abstract We have employed a new approach to generate novel fluorescent proteins (FPs) from red absorbing chromoproteins. An identical single amino acid substitution converted novel chromoproteins from the species *Anthozoa* (*Heteractis crispa*, *Condylactis gigantea*, and *Goniopora tenuidens*) into far-red FPs (emission λ_{max} = 615–640 nm). Moreover, coupled site-directed and random mutagenesis of the chromoprotein from *H. crispa* resulted in a unique far-red FP (HcRed) that exhibited bright emission at 645 nm. A clear red shift in fluorescence of HcRed, compared to drFP583 (by more than 60 nm), makes it an ideal additional color for multi-color labeling. Importantly, HcRed is excitable by 600 nm dye laser, thus promoting new detection channels for multi-color flow cytometry applications. In addition, we generated a dimeric mutant with similar maturation and spectral properties to tetrameric HcRed. © 2001 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Green fluorescent protein; Anthozoa; drFP583; DsRed; Multi-color labeling

1. Introduction

Green fluorescent protein (GFP) from *Aequorea victoria* is an excellent marker for gene expression in vivo [1–4]. Despite considerable efforts to generate GFP-based red fluorescence, the most red-shifted GFP variant displayed emission at 529 nm, in the green–yellow region [5]. The real 'red' breakthrough was the discovery of GFP-like fluorescent proteins (FPs) from the species *Anthozoa* [6]. Two FPs showed fluorescence emission in the orange–red region of the visible spectrum at 583 nm and 593 nm [6,7]. Extensive mutagenesis of these proteins generated a red-shifted mutant, with emission maximum at 616 nm [7].

Red-shifted FPs are in high demand for various applications. Firstly, far-red FPs provide an additional color for multi-color labeling. Secondly, animal tissues are almost translucent to far-red light (typically near infra-red window at around 650–900 nm [8]), but effectively absorb shorter-wave light. Thus, efficiency of fluorescence detection greatly depends on the color, i.e. the longer the light wavelength, the higher the sensitivity. At present, the discovery of additional natural far-red FPs seems unlikely. Current data on the fluorescence of live coral (observed emission at 590–620 nm) indicate that maximum red shift is probably attained in the cloned red FPs [9,10].

We previously described a GFP-like non-fluorescent chromoprotein (CP) and it's mutant emitting red fluorescence at 595 nm [11]. Here, we suggest a general strategy to generate far-red FPs, by mutagenesis of novel non-fluorescent GFP-like CPs. Following this procedure, we have successfully produced the most red-shifted FP known to date, with an emission maximum at 645 nm.

2. Materials and methods

2.1. Cloning, expression and mutagenesis

Total RNA from the colored tissues of *Heteractis crispa*, *Condylactis gigantea*, *Condylactis passiflora* and *Goniopora tenuidens* was isolated by guanidine thiocyanate extraction [12]. Synthesis, amplification of fragments of interest using degenerate primers, and generation of full-length cDNA were performed as described in [6]. To produce cDNA of *C. passiflora*, we employed PCR with primers corresponding to *C. gigantea* cDNA, due to high sequence homology between their CP coding regions. For heterological expression of novel CPs and their mutants, full-length coding regions were cloned into the pQE30 vector (Qiagen). Proteins fused to an N-terminal 6×His tag were expressed in *Escherichia coli* (at 37°C for cgCP, cpCP and gtCP, and at 20°C for hcCP), and purified using the TALON metal-affinity resin (Clontech). Site-directed mutagenesis was performed by PCR using the overlap extension method, with primers containing appropriate target substitutions [13]. The Diversity PCR Random Mutagenesis kit (Clontech) was used for random mutagenesis of hcCP-C148S, in conditions optimal for four to five mutations per 1000 bp. *E. coli* expressing mutant proteins were grown at 37°C and visually screened with a fluorescent microscope. The brightest variants were selected and subjected to the next round of random mutagenesis.

2.2. Spectroscopy

Absorption spectra were recorded on a Beckman DU520 UV/vis Spectrophotometer. A Perkin-Elmer LS50B Fluorescence Spectrophotometer was used for measuring excitation–emission spectra. Spectra were corrected for photomultiplier response and monochromator

*Corresponding author. Fax: (7)-95-330 7056.
E-mail address: luk@ibch.ru (S.A. Lukyanov).

[1] The accession numbers of genes described in this article are: cgCP (AF363775), hcCP (AF363776), cpCP (AF383155), gtCP (AF383156).
[2] These authors contributed equally to this work.
[3] Present address: Whitney Laboratory, University of Florida, 9505 Ocean Shore Boulevard, St. Augustine, FL 32086, USA.

Abbreviations: GFP, green fluorescent protein; FP, fluorescent protein; CP, chromoprotein; cgCP, *Condylactis gigantea* chromoprotein; cpCP, *Condylactis passiflora* chromoprotein; gtCP, *Goniopora tenuidens* chromoprotein; hcCP, *Heteractis crispa* chromoprotein; HcRed, *Heteractis crispa* red fluorescent protein

0014-5793/01/$20.00 © 2001 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
PII: S0014-5793(01)02930-1

transmittance. The brightness of hcCP-C148S and HcRed proteins was compared by measurement of relative fluorescence intensity of the purified proteins at the same concentration.

A fluorescent microscope (Nikon Optiphot) equipped with 546/10 band-pass excitation filter and 580 long-pass suppression filter was used for imaging *E. coli* colonies expressing FPs.

2.3. Flow cytometry analysis

pDsRed-N1 (Clontech) and analogous vector containing HcRed instead of DsRed were used. HEK 293T cells were transfected with DsRed (drFP583) or HcRed expressing plasmids and analyzed after 48 h, using highly modified double laser (488 nm and 600 nm) FACS Vantage SE, available at the Stanford shared FACS facility.

2.4. Gel filtration

Purified protein samples (~1 mg/ml) were loaded onto a Sephadex-100 column (0.7×60 cm). Proteins were eluted from the column with 50 mM phosphate buffer (pH 7.0), 100 mM NaCl, and absorbance was monitored at the appropriate wavelength, using a multi-wavelength detector.

3. Results and discussion

We selected four Indo–Pacific *Anthozoa* species, specifically sea anemone *H. crispa*, *C. gigantea* and *C. passiflora* and stony coral *G. tenuidens*, to generate novel GFP-like CPs. Specimens displayed clearly visible coloration on the tentacles. For each organism, cDNA encoding the novel GFP-like protein was cloned, using the procedure described in [6]. Each protein was designated lower-case letters identifying the species, followed by CP. Thus, novel proteins were denoted hcCP, cgCP, cpCP, and gtCP, and the protein earlier identified as FP595 was designated asCP. Sequence comparisons of novel CPs and FPs revealed conserved amino acids at two characteristic positions that distinguish between these groups. An invariant Phe occupies position 64 in all FPs (GFP numbering, see Fig. 1 for details), while CPs contain a small residue (Cys or Ser) at this position. Furthermore, all *Anthozoa* FPs display Ser at position 148 (or His in GFP), while in CPs, this is occupied by Ala, Cys, or Asn. This residue variation at position 148 is probably structurally significant, since in GFP and drFP583, His/Ser-148 serves to stabilize the fluorescent state of the chromophore by hydrogen bonding with phenolic oxygen [5,14,15].

E. coli-purified CPs displayed similar absorption spectra, with peaks at 570–580 nm (Fig. 2A). However, minor differences in the position of the absorption maxima evidently allow the eye to distinguish between protein colors. The hue changes from purple for asCP, through to mauve for cgCP and cpCP and lilac for hcCP and gtCP.

In an earlier study on non-fluorescent asCP, we showed that a single substitution at position 148 to Ser (A148S) makes the protein fluorescent [11]. To verify the effect of this modification on spectral properties of novel CPs, we placed a serine residue at position 148 in hcCP, cgCP, and gtCP (we did not mutate cpCP because cgCP and cpCP are very similar to each other and we do not expect any significant differences

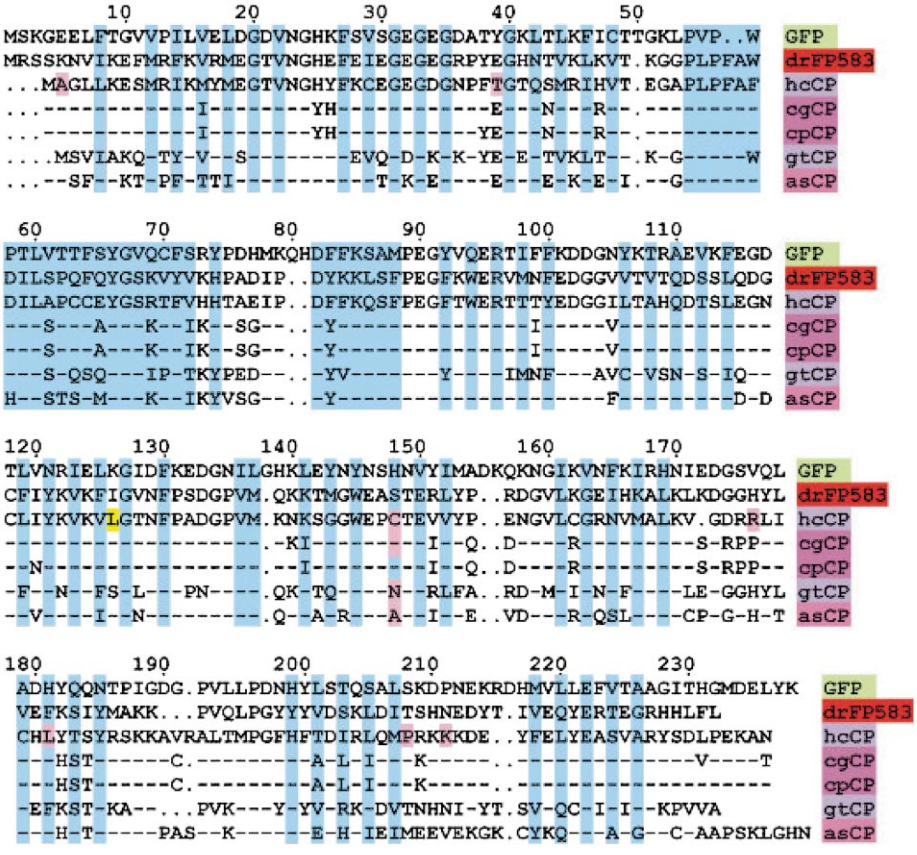

Fig. 1. Multiple alignment of GFP, drFP583, and GFP-like CPs. Numbering is based on *A. victoria* GFP. Protein names are shaded in colors corresponding to their normal coloration. All CP sequences are compared with hcCP. Residues identical to the corresponding amino acids in hcCP are represented by dashes. Gaps introduced in the sequence are represented by dots. Residues with side-chains that form the interior of the b-can are shaded in blue. Mutated residues are shaded in pink. In the hcCP sequence, position 126, differentiating between dimeric HcRed-2 and tetrameric HcRed, is shaded in yellow.

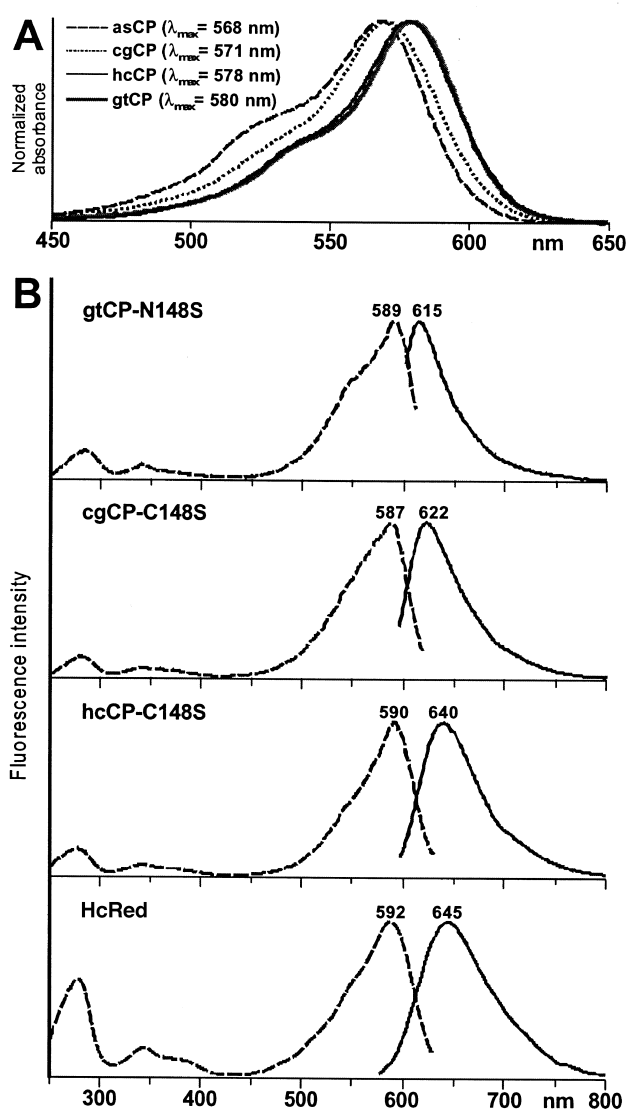

Fig. 2. Spectral characteristics of the CPs and fluorescent mutants. A: Absorption spectra for wild-type CPs (note that the spectrum for cpCP is not shown, since cgCP and cpCP display very similar absorption curves). B: Excitation (dashed lines) and emission (solid lines) spectra for the fluorescent mutants of the CPs.

in their properties). The mutation invariably induced the appearance of red fluorescence in all three CPs. Moreover, the novel FPs were considerably red-shifted compared to all known natural FPs to date. Emission spectra showed peaks at 615, 622, and 640 nm for gtCP, cgCP, and hcCP mutants, respectively (Fig. 2B).

Random mutagenesis was employed to progressively enhance the most red-shifted FP, hcCP-C148S. The main aim of the first round of mutagenesis was to generate a protein that maturates at 37°C, since both wild-type hcCP and hcCP-C148S are expressed in *E. coli* at low temperatures only (20°C). A mutant (hc-4) containing two additional substitutions, A5S and P208L, was selected. This protein possessed similar spectral characteristics to parental hcCP-C148S, but was not temperature-sensitive, and attained maximum brightness after overnight growth at 37°C in *E. coli*. A second round of random mutagenesis on hc-4 yielded a clone (hc-44) containing additional substitutions, T39A and L181H. This mutant protein was 3-fold brighter than hc-4, and displayed slightly more red-shifted fluorescence that peaked at 645 nm. Subsequently, hc-44 was subjected to a third round of mutagenesis that resulted in a 2-fold brighter variant, containing one additional substitution, K211E. The protein, denoted HcRed (from *Heteractis crispa Red* FP), contained six substitutions, specifically, A5S, T39A, C148S, L181H, P208L, and K211E, was selected as the final mutant, since no further improvement in far-red fluorescence was observed with further rounds of mutagenesis.

An equivalent comparison of *E. coli* colonies expressing drFP583 and HcRed under the fluorescence microscope reveals a marked difference between the orange color of drFP583 and the truly red color of HcRed (Fig. 3). Additionally, HcRed maturates more rapidly than drFP583, as confirmed by the brighter fluorescence displayed by colonies containing the former protein after overnight expression in *E. coli*. However, it must be noted that the ultimate fluorescence intensity of mature drFP583 is considerably higher (about 5-fold) than that of HcRed.

Spectral properties of HcRed (excitation maximum at 592 nm, emission maximum at 645 nm, Fig. 2B) make it suitable for FACS analysis with the 600 nm dye laser. HEK 293 cells transfected with drFP583 were readily detected with a 488 nm argon ion laser (PE channel), but not with a 600 nm dye laser. Conversely, HcRed-transfected cells displayed no fluorescence when excited at 488 nm, but were clearly visible with the 600 nm laser (TxRed and APC channels) (Fig. 4). Our data indicate that the mutant HcRed may be employed for additional far-red color in multi-color flow cytometry applications.

Recently, Baird and coworkers showed that drFP583 is an obligate tetramer [16]. We tested the oligomeric state of the novel proteins by gel filtration analysis, using EGFP (enhanced GFP) and drFP583 as monomer and tetramer standards, respectively. All proteins behaved chromatographically in a similar manner to drFP583. Results showed that all wild-type CPs and nearly all the fluorescent mutants (including HcRed) were tetrameric. A single dimeric mutant of hcCP (denoted hc-41) was detected, which migrated between EGFP and drFP583. This protein contained eight amino acid substitutions, specifically, A5S, S43G, E65A, L126H, K141E, C148S, R186K and P208L. Within these substitutions, only one mutation, L126H, may be responsible for modifying the oligomeric state of the protein, since the Ile residue at this position participates in formation of the tetrameric interface in drFP583 [14,15]. As hc-41 possessed rela-

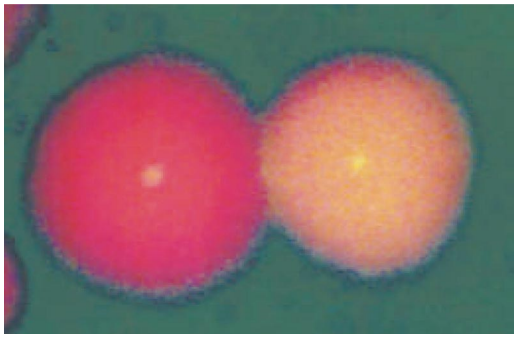

Fig. 3. *E. coli* colonies expressing HcRed (left) or drFP583 (right) after overnight growth, as viewed by fluorescence microscopy. The colors of these colonies are plainly distinguishable.

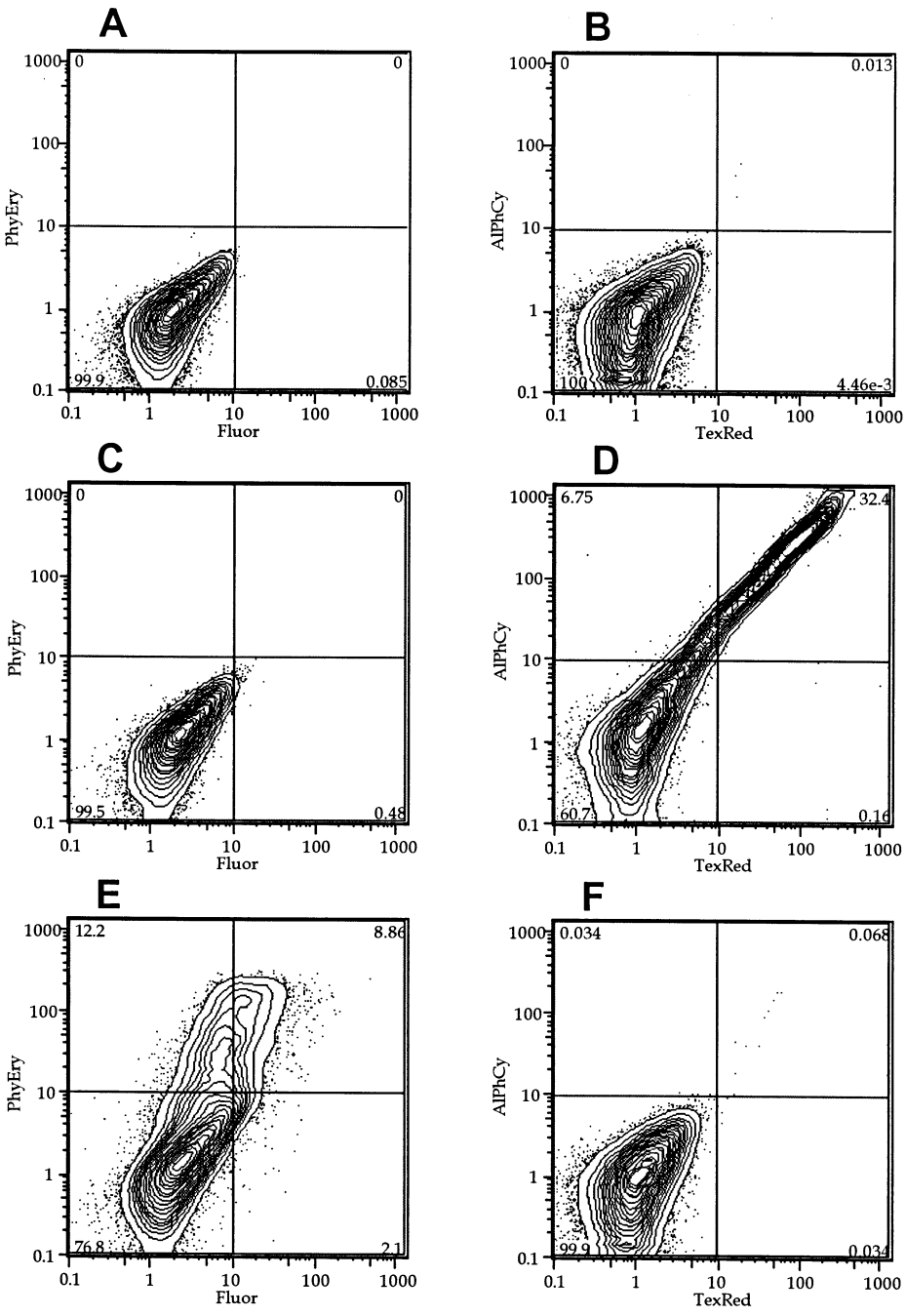

Fig. 4. Flow cytometry analysis of HEK 293 cells transfected with HcRed (C, D), DsRed (E, F) or mock-transfected (A, B) 48 h post-transfection. The 488 nm laser was used in A, C, E and 600 nm laser was used in B, D, F. Contour plots (5% probability with outliners) are shown. Dead cells were excluded from analysis. Standard compensations and filter sets for Vanford FACS machine were used.

tively blue-shifted fluorescence (emission $\lambda_{max} = 625$ nm), we introduced the L126H mutation into HcRed. The resultant mutant was designated HcRed-2. Gel filtration analyses revealed the dimeric nature of this protein. Following random mutagenesis of HcRed-2 resulted in brighter dimeric variant HcRed-2A containing two additional amino acid substitutions, specifically, R176H and E211K. Spectral and maturation properties of HcRed-2A were similar to those of HcRed. *E. coli*-expressed HcRed-2A maturated nearly completely during overnight growth and displayed excitation and emission maxima at 590 and 640 nm, respectively. However, further site-directed and random mutagenesis failed to generate a bright monomeric mutant of HcRed.

Oligomerization does not significantly limit the use of FPs simply as reporters of gene expression, but is a very serious problem in many potential applications where FP is fused to a host protein. Despite this, several examples of successfully expressed drFP583-fused proteins were published last year [17–20]. Recently, Lauf and coworkers suggested an approach to rescue the function of DsRed-tagged proteins [21]. We believe that dimeric HcRed-2A is more suitable than tetrameric drFP583 for functional fusion protein construction.

The discovery of yellow and red FPs from *Anthozoa* disclosed a novel prospect for developing multi-color in vivo labeling [6,7]. From both scientific and practical standpoints, it would be interesting to ascertain limitations in the fluorescence color of GFP-like proteins towards the far-red region. In 1995, Mazel quantitatively measured the fluorescence of a diverse group of Caribbean live corals [9]. Despite the large sampling population, emission at wavelengths longer than 590 nm was never detected (except the 680 nm fluorescence of chlorophyll from symbiotic algae). Recently, Salih and coworkers [10] described fluorescence of coral tissues that displayed emission maxima between 480 and 620 nm. The large number of diverse specimens analyzed in these experiments suggests that a further quest for natural far-red *Anthozoa* FPs is not likely to be successful. In the present study, we demonstrate a new strategy for generating novel far-red FPs, by mutagenesis of non-fluorescent CPs. Using this method, we have successfully created the most red-shifted GFP-like protein known to date, which significantly exceeds the probable wavelength limit of naturally existing FPs. As different-colored CPs are widespread in coral [22], they are a good natural resource for generating further red-shifted FPs.

Acknowledgements: This work was supported by Clontech Laboratories Inc. and the Russian Foundation for Fundamental Research (Grants 99-04-48873 and 01-04-49037).

References

[1] Kendall, J.M. and Badminton, M.N. (1998) Trends Biotechnol. 16, 216–224.
[2] Pollok, B.A. and Heim, R. (1999) Trends Cell Biol. 9, 57–60.
[3] White, J. and Stelzer, E. (1999) Trends Cell Biol. 9, 61–65.
[4] Margolin, W. (2000) Methods 20, 62–72.
[5] Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y. and Remington, S.J. (1996) Science 273, 1392–1395.
[6] Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L. and Lukyanov, S.A. (1999) Nat. Biotechnol. 17, 969–973.
[7] Fradkov, A.F., Chen, Y., Ding, L., Barsova, E.V., Matz, M.V. and Lukyanov, S.A. (2000) FEBS Lett. 479, 127–130.
[8] Weissleder, R. (2001) Nat. Biotechnol. 19, 316–317.
[9] Mazel, C.H. (1995) Mar. Ecol. Prog. Ser. 120, 185–191.
[10] Salih, A., Larkum, A., Cox, G., Kuhl, M. and Hoegh-Guldberg, O. (2000) Nature 408, 850–853.
[11] Lukyanov, K.A., Fradkov, A.F., Gurskaya, N.G., Matz, M.V., Labas, Y.A., Savitsky, A.P., Markelov, M.L., Zaraisky, A.G., Zhao, X., Fang, Y., Tan, W. and Lukyanov, S.A. (2000) J. Biol. Chem. 275, 25879–25882.
[12] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[13] Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Gene 77, 51–59.
[14] Wall, M.A., Socolich, M. and Ranganathan, R. (2000) Nat. Struct. Biol. 7, 1133–1138.
[15] Yarbrough, D., Wachter, R.M., Kallio, K., Matz, M.V. and Remington, S.J. (2001) Proc. Natl. Acad. Sci. USA 98, 462–467.
[16] Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (2000) Proc. Natl. Acad. Sci. USA 97, 11984–11989.
[17] Mas, P., Devlin, P.F., Panda, S. and Kay, S.A. (2000) Nature 408, 207–211.
[18] Ayoob, J.C., Shaner, N.C., Sanger, J.W. and Sanger, J.M. (2001) Mol. Biotechnol. 17, 65–71.
[19] Laird, D.W., Jordan, K., Thomas, T., Qin, H., Fistouris, P. and Shao, Q. (2001) Microsc. Res. Tech. 52, 263–272.
[20] Charpilienne, A., Nejmeddine, M., Berois, M., Parez, N., Neumann, E., Hewat, E., Trugnan, G. and Cohen, J. (2001) J. Biol. Chem. 276, 29361–29367.
[21] Lauf, U., Lopez, P. and Falk, M.M. (2001) FEBS Lett. 498, 11–15.
[22] Dove, S.G., Hoegh-Guldberg, O. and Ranganathan, S. (2001) Coral Reefs 19, 197–204.