GFPs of insertion mutation generated by molecular size-altering block shuffling

Koichiro Kitamura[a], Chuya Yoshida[a], Koichi Nishigaki[a,b,*]

[a] *Department of Functional Materials Science, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama 338-8570, Japan*
[b] *Rational Evolutionary Design of Advanced Biomolecules, Saitama Prefecture Collaboration of Regional Entities for the Advancement of Technological Excellence, JST, Saitama, Japan*

Received 9 September 2003; revised 20 October 2003; accepted 29 October 2003

First published online 18 November 2003

Edited by Lev Kisselev

Abstract Insertion and deletion analyses of a protein have been less common than point mutation analyses, partly due to the lack in effective methods. This is the case with the green fluorescent protein (GFP), which is so widely applied in molecular biology and other fields. In this paper we first introduce a systematic approach for generating insertion/deletion mutants of GFP. A new technology of Y-ligation-based block shuffling (YLBS) was successfully applied to produce size-altered GFPs, providing insertion-containing GFPs of fluorescence, though no deletion type of fluorescence was obtained so far as examined. The analysis of these proteins suggested that size alteration (deletion/insertion) is acceptable so far as some type of rearrangement in a local structure can accommodate it. This paper demonstrates that YLBS can generate insertion and deletion mutant libraries systematically, which are beneficial in the study of structure–function relationship.

Key words: Green fluorescent protein; Y-ligation-based block shuffling; Block shuffling; Molecular diversity; Insertion mutation; Deletion

1. Introduction

Recent advance in genome science has clearly shown that proteins have a hierarchical structure composed of domains and modules which widely appear among different proteins [1–3]. In other words, rearrangement of genes by insertion and deletion has played a pivotal role in generating proteins of a novel function, e.g. bacterial replication protein PriB evolved from a single-stranded (ss) DNA-binding protein gene (*Ssb*) by duplication and rearrangement [4] and two different *Caenorhabditis elegans* FGF receptors of different physiological functions generated by alternative splicing [5]. A computer simulation also supported a rapid molecular evolution with mediation of non-homologous DNA recombination and rearrangement [6]. From these facts it is of no doubt that the shuffling mechanism must be useful in creation of proteins of novel functions. Thus, it is natural to seek a mimic approach in protein engineering of making the molecular evolution as a model.

There have already appeared several technologies in this context: 'micro gene' methods which random-assembles DNA fragments [7], systemic circular permutation of proteins [8], random multirecombinant polymerase chain reaction (RM-PCR) [9] and incremental truncation for creating hybrid enzymes [10]. Recently, we developed the Y-ligation-based block shuffling (YLBS) method for the purpose of general recombination [11]. The most beneficial of this method is the readiness in integrating blocks of a sense as starting elements (which can be of any size: amino acid monomer to domain or protein itself). In this paper, we applied this YLBS method to generation of insertion/deletion-containing green fluorescent protein (GFP) libraries. GFP has been used for a variety of purposes and is mostly made fused at either an amino or carboxyl terminus with a relevant protein. Moreover, GFP is already known to allow insertions at several sites, mainly in loop regions, of peptides (6–20 amino acids) [12] or of Xenopus CaM plus zinc finger domains [13]. As a very limited number of insertion species have yet been examined, we systemically looked into the insertion effect on the fluorescence function of GFP by generating half a million mutant proteins. We report here novel species of three amino acids insertion-containing GFPs together with an intriguing fact that there was no finding of three amino acids deletion-containing GFPs, suggesting the existence of a certain kind of structural constraints. This kind of systemic approach using molecular size-altering block shuffling is shown to be useful for structural analysis of proteins in general.

2. Materials and methods

2.1. YLBS

YLBS is termed after its core step, ligation of two ssDNA branches extruded from a double-stranded stem by T4 RNA ligase. This construct makes a ligation reaction of two ssDNAs to be an *intramolecular* reaction of high efficiency. As compactly shown in Fig. 1, the whole steps of YLBS are rather simple and composed of hybridization, ligation, PCR, restriction cleavage and ss separation (for more detail, see the original paper [11]). The starting blocks and primers used for this experiment are listed in Table 1. By repeating the steps, the number of ligated blocks increases exponentially (i.e. first cycle (Y^1) = 2 blocks; second, 4; third, 8; and so on). As the block size to be ligated is arbitrary, we can ligate m blocks with n blocks of a different size ($m \neq n$). Therefore, YLBS has a very general nature as a synthetic method of proteins.

2.2. Cloning, sequencing and expression

DNAs of insertion/deletion libraries were cloned with a TA cloning kit (Invitrogen, USA). Plasmid DNAs were purified with a plasmid extraction kit (Wizard Plus SV Minipreps DNA Purification System, Promega, USA). For sequencing DNA, DNA sequencing kit (Amer-

*Corresponding author. Fax: (81)-48-858 3533.
E-mail address: koichi@fms.saitama-u.ac.jp (K. Nishigaki).

0014-5793/03/$22.00 © 2003 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)01308-5

sham Bioscience, USA) was used and sequencing was performed with DNA sequencer DSQ2000 (Shimadzu, Japan). Plasmid pGFPuv4 [14] was used for preparation of GFPuv4 gene. The gene was excised with *Mlu*I and *Eco*RI and purified using a QIA quick column (Qiagen, Germany).

2.3. Selection of fluorescent GFPs from insertion/deletion libraries

GFP proteins containing insertion/deletion were constructed following the procedures depicted in Fig. 2. *Areas* 1 to 3 of GFP were selected for block shuffling regions (Fig. 2a), excised, divided into eight blocks of nonanucleotides (equivalent to tripeptides), and shuffled, respectively. The library type of insertion/deletion could be determined by the number of blocks which were returned back to the remaining portion of GFP (see Fig. 2b). The primers used for each PCR are given in Fig. 2b and Table 1. To re-integrate the excised DNA fragment, sealing at restriction sites was adopted using different pairs of restriction enzymes ('*Res*X' and *Eco*RI; *Res*X = *Bsr*GI for *Area* 1 and *Mlu*I for *Areas* 2 and 3, respectively). DNAs of 400 fmol were treated with the restriction enzymes *Mlu*I and *Eco*RI, and incubated in a solution containing 4 fmol of expression vector DNA and 4 Weiss units of T4 DNA ligase at 16°C for 18 h, and subjected to electroporation with 80 μl of *Escherichia coli* competent cells (strain DH5α) using Electropulser (BioRad, USA) under the conditions of 1.8 kV of applied voltage and 4 ms of pulse duration. After addition of 2 ml of SOB (20 g/l Bacto tryptone, 5 g/l Bacto yeast extract, 5 M NaCl, 2 M KCl and 2 M $MgCl_2$), the mixture was incubated at 37°C for 1 h, then plated, by 1 ml for selection, onto an LB agar plate (210×280 mm^2) containing 100 μg/ml ampicillin, and incubated at 37°C for 18 h. Fluorescence of colonies was detected by a fluoroimager, Typhoon 9400 (Amersham Biosciences, USA) adopting the wavelength of 488 nm (excitation) and 511 nm (emission).

2.4. Computer works

Peptide sequences were drawn using planar representation of sequences (PRS) [15], which makes a sequence of amino acids a kind of zigzag line, stitching the points defined by the type of amino acid (*x*-axis direction) and the sequence site (*y*-axis direction). The merit of this representation is that it can make rather monotonous sequence information illustrative and has easy to grasp features. Random sequences of *Area* 3 were generated by a Monte Carlo simulation approach by making the probability of occurrence of each amino acid equal to that in *Area* 3. To estimate the substitution number, the amino acid sequences of a wild-type and an insertion mutant were aligned so as to give a maximum match. The resulting mismatches and insertions/deletions were counted to be 2 and 3 units in distance, respectively, and a summation of these units was taken over the while sequence.

Table 1
Oligonucleotides used for starting blocks and PCR primers

Sequence (length)	Name
Definition	
GGCTCGCGAATACTTTGAAGATCTCTTTT (29)	= $\mathbf{D}_5$
GGCTCGCGAATACTTTGAACGGGATCCCCTA (31)	= $\mathbf{D}_3$
CTTCAAGAATTCATTAGTGATGGTGATGGTGATGTTTGTAGAGCTCATCC (50)	= **Cc-r**
Starting 5′/3′ halves for *Area* 1 (38/40 nt)	
$\mathbf{D}_5$-ATATCCTTC/ATATCCTTC-$\mathbf{D}_3$	*A15′h*/*A13′h*-A
$\mathbf{D}_5$-AAAGACGAC/AAAGACGAC-$\mathbf{D}_3$	*A15′h*/*A13′h*-B
$\mathbf{D}_5$-GGCAACTAT/GGCAACTAT-$\mathbf{D}_3$	*A15′h*/*A13′h*-C
$\mathbf{D}_5$-AAAACCCGT/AAAACCCGT-$\mathbf{D}_3$	*A15′h*/*A13′h*-D
$\mathbf{D}_5$-GCTGAAGTG/GCTGAAGTG-$\mathbf{D}_3$	*A15′h*/*A13′h*-E
$\mathbf{D}_5$-AAGTTTGAG/AAGTTTGAG-$\mathbf{D}_3$	*A15′h*/*A13′h*-F
$\mathbf{D}_5$-GGTGACACA/GGTGACACA-$\mathbf{D}_3$	*A15′h*/*A13′h*-G
$\mathbf{D}_5$-CTTGTGAAT/CTTGTGAAT-$\mathbf{D}_3$	*A15′h*/*A13′h*-H
Starting 5′/3′ halves for *Area* 2 (38/40 nt)	
$\mathbf{D}_5$-CGCATCGAG/CGCATCGAG-$\mathbf{D}_3$	*A25′h*/*A23′h*-A
$\mathbf{D}_5$-CTGAAAGGT/CTGAAAGGT-$\mathbf{D}_3$	*A25′h*/*A23′h*-B
$\mathbf{D}_5$-ATCGATTTT/ATCGATTTT-$\mathbf{D}_3$	*A25′h*/*A23′h*-C
$\mathbf{D}_5$-AAAGAGGAC/AAAGAGGAC-$\mathbf{D}_3$	*A25′h*/*A23′h*-D
$\mathbf{D}_5$-GGCAACATT/GGCAACATT-$\mathbf{D}_3$	*A25′h*/*A23′h*-E
$\mathbf{D}_5$-CTTGGACAC/CTTGGACAC-$\mathbf{D}_3$	*A25′h*/*A23′h*-F
$\mathbf{D}_5$-AAACTGGAA/AAACTGGAA-$\mathbf{D}_3$	*A25′h*/*A23′h*-G
$\mathbf{D}_5$-TACAACTAC/TACAACTAC-$\mathbf{D}_3$	*A25′h*/*A23′h*-H
Starting 5′/3′ halves for *Area* 3 (38/40 nt)	
$\mathbf{D}_5$-ATCAAAGCT/ATCAAAGCT-$\mathbf{D}_3$	*A35′h*/*A33′h*-A
$\mathbf{D}_5$-AATTTCAAG/AATTTCAAG-$\mathbf{D}_3$	*A35′h*/*A33′h*-B
$\mathbf{D}_5$-ATTCGTCAC/ATTCGTCAC-$\mathbf{D}_3$	*A35′h*/*A33′h*-C
$\mathbf{D}_5$-AACATCGAG/AACATCGAG-$\mathbf{D}_3$	*A35′h*/*A33′h*-D
$\mathbf{D}_5$-GATGGGAGC/GATGGGAGC-$\mathbf{D}_3$	*A35′h*/*A33′h*-E
$\mathbf{D}_5$-GTCCAACTG/GTCCAACTG-$\mathbf{D}_3$	*A35′h*/*A33′h*-F
$\mathbf{D}_5$-GCAGACCAC/GCAGACCAC-$\mathbf{D}_3$	*A35′h*/*A33′h*-G
$\mathbf{D}_5$-TACCAACAA/TACCAACAA-$\mathbf{D}_3$	*A35′h*/*A33′h*-H
PCR primers for amplification of ligated products	
GGCTCGCGAATACTGCGAAGGAACGCACT (29)	PS_5
GGCTCGCGAATACTGCGAACAGCTCGAGATC (31)	PS_3
CTGCGAAGAAACGCACT (17)	PD_5
GCGAACAGCTCGAGATC (17)	PD_3
PCR primers for GFP expression	
AGGTTATGTACAGGAACGCACTATATCCTTC/AACTCGATACGATTCACAAG	*A1*-Xp/Xp-r
CCCTTGTTAATCGCATCGAG/TTGTGTGAGTTGTAGTTGTA	*A2*-Xp/Xp-r
AAAGAATGGAATCAAAGCT/ATTGGAGTATTTTGTTGGTA	*A3*-Xp/Xp-r
AACTACAAGACGCGTGCTGA/AGCTTTGATTCCATTCTT	*A1*-Nc/Nc-r
AACTACAAGACGCGTGCTGA/GAGCTACGCTAATTGTTCCC	*A2*-Nc/Nc-r
AACTACAAGACGCGTGCTGA/AGCTTTGATTCCATTCTT	*A3*-Nc/Nc-r
CTTGTGAATCGTATCGAGTT/**Cc-r**	*A1*-Cc/Cc-r
TACAACTACAACTCACAC/**Cc-r**	*A2*-Cc/Cc-r
TACCAACAAAATACTCCAAT/**Cc-r**	*A3*-Cc/Cc-r

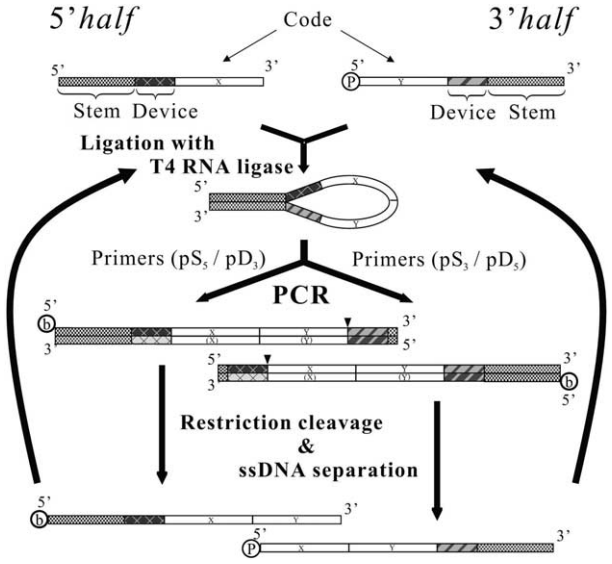

Fig. 1. Schematic diagram of YLBS. Following the hybridization of 5′ half and 3′ half at the stem region, ligation occurs at the tips of Y branches by the action of T4 RNA ligase. The succeeding steps of PCR, restriction cleavage and ss separation constitute the cycle of YLBS. Stem, Device and Code regions are assigned as a complementary sequence for hybridization, a restriction enzyme recognition/PCR priming site and an arbitrary block sequence, respectively. A filled arrowhead, encircled B or P indicates a restriction cleavage site, biotin, or phosphate attached at 5′-terminal, respectively.

3. Results and discussion

In this study, a particular portion of GFP protein was dealt for size-altering block shuffling. As mentioned in Section 2, we generated a DNA library which contains extra three amino acids in the region of 164–181 in the GFP amino acid sequence (*Area* 3). Each resultant protein has a modified *Area* 3, which is well-shuffled constituted with seven blocks arbitrarily chosen (in a repetition-allowing mode) out of the original six blocks, resulting in insertion into GFP by one block (= three amino acids). The DNA library thus generated was subjected to sequencing for the sake of confirmation of quality and then the DNA library was cloned with an expression vector in *E. coli* to recover proteins. The expression of GFP was monitored by fluorescence. As shown in Fig. 3, 10 clones were found fluorescent after an intensive screening of around 40 000 clones. These fluorescent proteins were purified and characterized: (i) all of these have the same excitation and emission maximum as those of the wild-type GFP (494 and 511 nm, respectively), (ii) the quantum yield of the mutants was in the range of 22–54% lower than that of the wild-type (0.53) and (iii) the extinction coefficient was also 46–82% lower than that of the wild-type (67 000 M^{-1} cm^{-1}). The colony fluorescence of these mutant clones are shown in Fig. 3. These clones together with non-fluorescent ones (10 out of 20 clones) were sequenced and are shown in Table 2.

The fluorescent proteins of insertion mutation have a common feature that they seem to hold most of an antiparallel β-sheet structure in *Area* 3, as shown in Fig. 4 while the non-fluorescent proteins have a diversity of structures. This difference is evidently shown by PRS as a pattern difference between both sets of 10 sequences. Though the number of fluorescent mutants is limited, we can see two clusters for the currently obtained proteins as indicated in Fig. 4a (two clusters (reddish and bluish in combined regions) are evident along s2 to s3 and s7 to s8 in this figure). This clustering in fluorescent mutants is also confirmed by the fact that the lines in Fig. 4a (fluorescent mutants) look sparser than those of Fig. 4b (non-fluorescent ones) though both contain the same number of lines (10 entries). Another prominent feature common to these fluorescent mutant proteins is that they contain a high ratio of charged amino acids in the coil region between two β-sheet strands (designated as the most permissive site for insertion by Kamb et al. [12]; see Fig. 2), which is obviously different from that of non-fluorescent mutant proteins. This

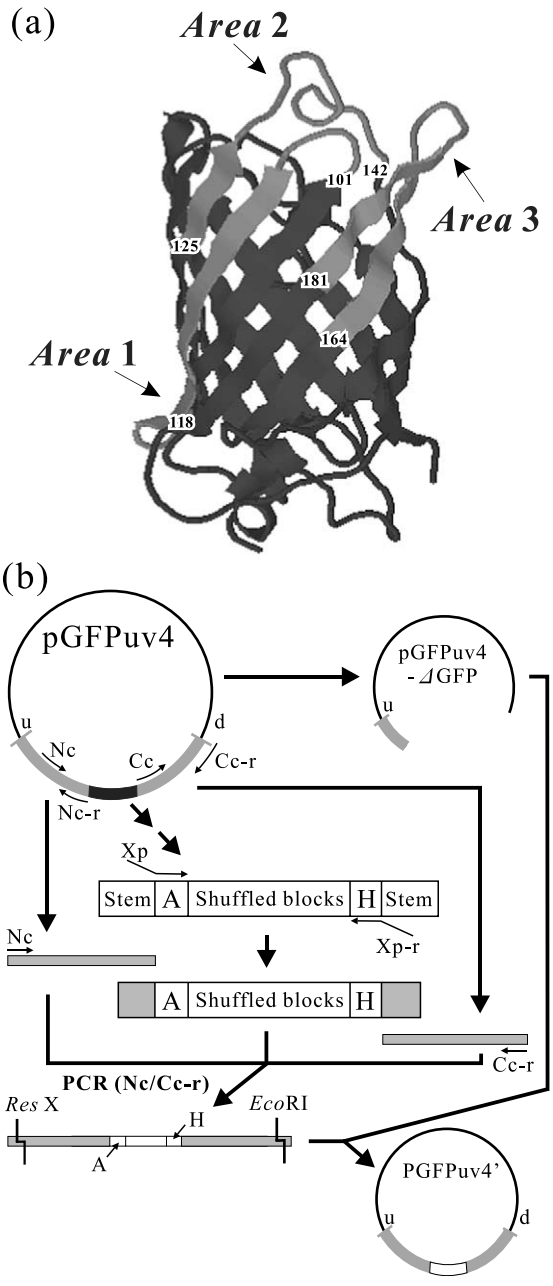

Fig. 2. Areas in GFP used for block shuffling (a) and its excision and integration (b). a: *Area* 3 (residues 164–181) was selected for an insertion library, which has a β-loop-β motif. For deletion libraries, *Areas* 1–3 (residues 101–118, 125–142, and 164–181) were selected. A ribbon model of GFP is drawn using RasMol. b: PCR primers are drawn for each purpose with their names listed in Table 1.

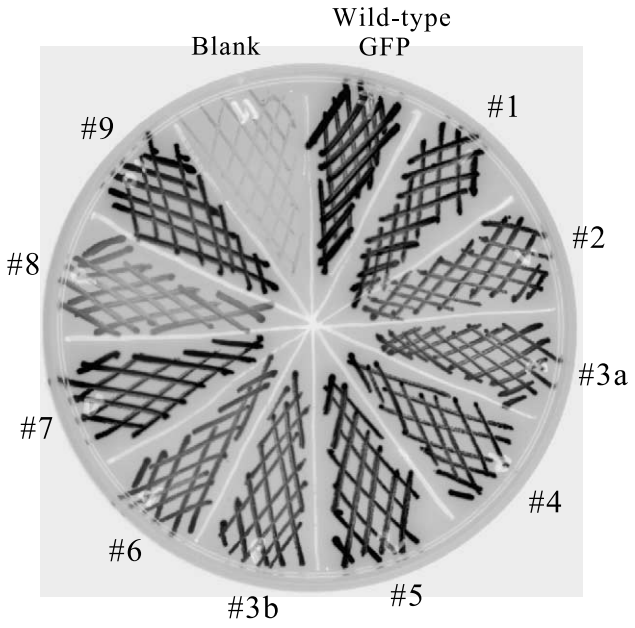

Fig. 3. Colonies of fluorescent GFPs. Fluorescence-expected clones (10 samples) were streaked on an LB agar plate and monitored by fluorescence at 488 nm (excitation) and 511 nm (emission). The clone numbers correspond to those in Table 2 (a and b are different clones but have the same sequence). The wild-type is a GFPuv4-harboring clone, while the blank is a colony of no GFP.

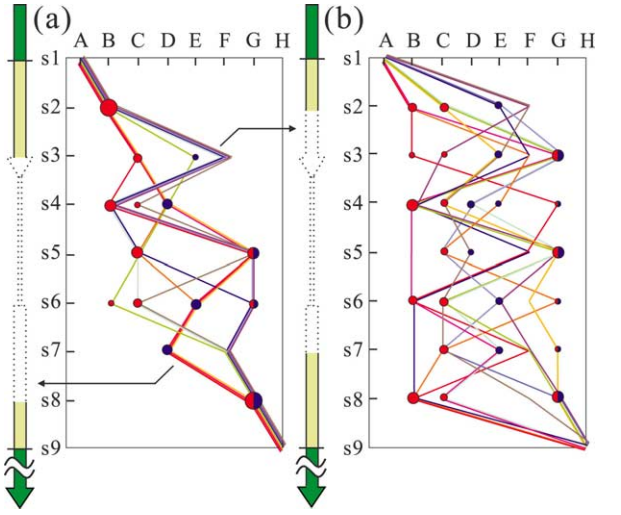

Fig. 4. Insertion mutants of fluorescence (a) and no fluorescence (b). PRS is employed for the sake of clarity where the abscissa represents the block type (A–H) and the ordinate represents the site along the sequence (s1–s9). Each mutant (10 species) is plotted in a line of different color on PRS. Those blocks which contain positively, negatively or both charged amino acids are shown in red, blue, and half-red-half-blue nodes, respectively. Note that both ends are fixed as A and H for all. The original secondary structures of GFP are shown by filled or empty arrows of β-sheets along the corresponding sequence ('filled with dark green' means 'structured' in the relevant cluster of proteins while 'filled in light green' or 'empty' means 'supposedly structured' or 'not structured', respectively). (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

situation is clearly shown in Fig. 5, where the concept of sequence space (which presents the fitness of each protein against its sequence [16]) is employed by showing the activity (fitness) of fluorescence as the size of a green circle at each point in the 2-D plane (where proteins are placed closer depending on the similarity in sequence (1st dimension)) (i.e. the closer the sequence, the nearer they are placed) and the total charge (2nd dimension). Evidently, the fluorescent proteins (green circles) are positioned closely to the wild-type (in the inner skirt with the wild-type centered) and segregated from the non-fluorescent ones (which are located in the outer skirt), suggesting a Mt. Fuji-shaped (not rugged) landscape which allows a progressive evolution [16]. Mutant proteins were tentatively generated by computer simulation in a mode of point mutation and plotted on the same plane as shown by crosses in Fig. 5. Although the substitution ratio was tuned to be roughly equivalent to that of block shuffling performed here (i.e. 0.51), the resultant products were more widely distributed in the 2D-plane than in the case of block shuffling (reflecting the independence of each amino acid in the sequence in the case of point mutation while, in the case of block shuffling, the amino acid sequence within a block is fixed). Since it is difficult to compare the two methods that are quite different by nature, the above discussion may have the less rationale. Nevertheless, a methodological difference is evident between point mutation-based and shuffling-based methods: the latter is well controlled within a range of sequences due to the conservation of partial sequences as blocks. A block, a stretch of sequence, can represent some meaning per se as in the cases of domains and modules. Therefore, the technology that can generate a mutation library with integrating 'meanings' must be of high potential in the field of protein engineering. This is especially the case when insertion mutation is involved since the insertion (block) may bring a new function into proteins.

What is interesting with insertion mutation is that none of the proteins selected here has the construct of single block insertion (at the loop region) without disturbing the remaining portion. This fact can be interpreted by two possibilities, either that (i) the number of samples collected here is too small to find such a type of mutant proteins or that (ii) any combi-

Table 2
Block sequence of insertion mutants

No.	Sequence
Wild-type	A-B-C-D-E-F-G-H
Fluorescent mutants	
1	A-B-F-B-G-G-F-G-H
2	A-B-E-D-C-B-F-G-H
3	A-B-C-B-G-E-D-G-H
4	A-B-F-B-C-C-F-G-H
5	A-B-C-D-G-E-D-G-H
6	A-B-F-D-G-C-F-G-H
7	A-B-F-B[a]-G-G-F-G-H
8	A-B-F-B-C-G-F-G-H
9	A-B-C-D-C-E-D-G-H
Non-fluorescent mutants	
11	A-F-C-B-G-E-F-C-H
12	A-E-G-D-C-E-C-G-H
13	A-B-B-G-F-B-F-B-H
14	A-C-G-B-G-C-F-G-H
15	A-E-F-B-E-B-B-B-H
16	A-B-G-B-B-B-E-C-H
17	A-F-E-C-D-C-C-F-H
18	A-C-E-C-G-E-G-G-H
19	A-C-G-D-G-C-E-C-H
20	A-B-F-E-C-G-C-B-H

[a]This block contains a point mutation (Phe→Lys). Otherwise, sequence 7 is the same with sequence 1.

nation of blocks cannot be such a construct so far as it remains fluorescent enough. Since we have found 10 fluorescent mutants out of 40 000 clones and the number of the entire combinations (diversities) for taking seven out of six blocks (blocks B–G) by allowing repetitive use of blocks is 6^7 (= 279 936), we have already observed one seventh (= 40 000/279 936) of the whole diversity. Thus, it is less probable that any one-block insertion at the loop region (i.e. three insertion points, 169–170, 172–173 and 175–176) can be accepted as its probability is only 0.03. This will be confirmed directly by producing such GFPs of insertion mutation or collecting more fluorescent clones (as possible future work), which is of value from the viewpoint of structural study of proteins.

On the other hand, from the experiments on the deletion mutants, none of fluorescent mutants were found for all libraries of *Areas* 1 to 3. When their sequences were analyzed (partly shown in Fig. 6), great diversities were observed (only a trivial partial clustering was observed at s6 to s7 in *Area* 1 (Fig. 6a)). This may indicate that deletion by three amino acids in the relevant areas is structurally more difficult than insertion by three amino acids (the insertion amino acids allowed were rather charged and hydrophilic and thus may be easily extruded outside from the core of a folding protein, resulting in being less inhibitory to the structure formation of a functional GFP). If it is true that the deletion of only three amino acids leads to a severe structural stress and eventually a loss of function, it could be a useful finding for protein engineering (though a larger amount of data are necessary for the final conclusion).

Once the methodology and its effectiveness have been established, other possible future work will be to change the size and number of blocks to be deleted or inserted in addition to changing the location where such mutations are introduced. This will be equivalent to a high-throughput version of the previous efforts performed on nested deletion [17,18] or shuffling of proteins such as hemoglobin [19]. These works are useful not only from a molecular engineering viewpoint (such as creating minimum-sized GFP) but also from a protein science viewpoint (such as exploring the protein sequence space). Furthermore, this methodology can utilize any type of blocks of both native and artificial origin; shuffling of undesigned blocks is highly promising in generating proteins of novel functions through exhaustive panning. On the other hand, the most beneficial aspect of this methodology must reside in the ability of doping designed blocks in proteins, which have intrinsic subfunctions like domains and modules. Hence, self-shuffling adopted here is a small step toward the world of block-shuffling-based protein engineering.

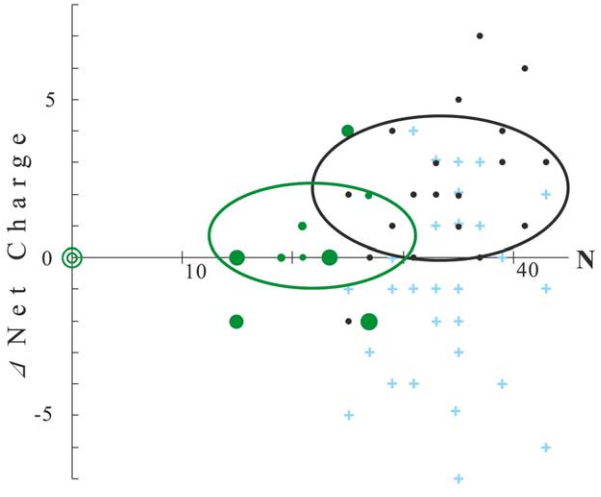

Fig. 5. Fluorescent and non-fluorescent mutant proteins distributed on a kind of sequence space. Each insertion mutant of GFP is plotted on a plane of substitution number (abscissa) and Δ net charge (ordinate). The substitution number is defined in Section 2 and represents how far from the wild-type sequence the mutant one is. On the other hand, Δ net charge is a difference in net charge between a mutant and the wild-type. The green circles represent fluorescent insertion mutants of GFP with the size of circle indicating the intensity of fluorescence. Non-fluorescent GFPs of insertion mutant are shown as a black circle. For the sake of reference, random polypeptide sequences generated by computer simulation are also plotted as crosses (+) to underscore the deviation in the distribution of block-shuffled ones. The ellipses show the range of within-a-standard deviation of mutant proteins (one ellipse radius for one parameter). (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

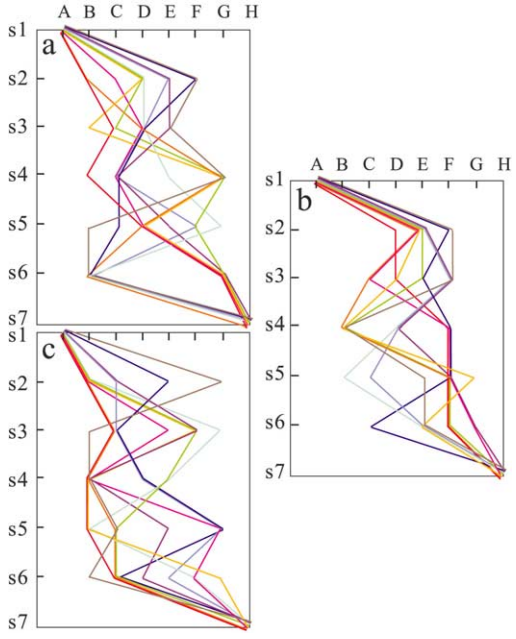

Fig. 6. PRS applied to deletion-containing GFPs. Sequences of deletion-containing mutants for *Area* 1 (a), *Area* 2 (b) and *Area* 3 (c). Note that both ends of s1 and s7 were fixed to be A and H, respectively. Each panel contains 10 sequences.

Acknowledgements: This work was partly supported by a fund from the Rational Evolutionary Design of Advanced Biomolecules (REDS) Project, Saitama Prefecture Collaboration of Regional Entities for the Advancement of Technological Excellence supported by JST.

References

[1] Doolittle, R.F. (1995) Annu. Rev. Biochem. 64, 287–314.
[2] Yura, K., Shionyu, M., Kawatani, K. and Go, M. (1999) Cell. Mol. Life Sci. 55, 472–486.
[3] Hegyi, H. and Bork, P. (1997) J. Protein Chem. 16, 545–551.
[4] Ponomarev, V.A., Makarova, K.S., Aravind, L. and Koonin, E.V. (2003) J. Mol. Microbiol. Biotechnol. 5, 225–229.
[5] Goodman, S.J., Branda, C.S., Robinson, M.K., Burdine, R.D. and Stern, M.J. (2003) Development 130, 3757–3766.
[6] Bogarad, L.D. and Deem, M.W. (1999) Proc. Natl. Acad. Sci. USA 96, 2591–2595.
[7] Shiba, K., Takahashi, Y. and Noda, T. (1999) Proc. Natl. Acad. Sci. USA 94, 3805–3810.

[8] Iwakura, M., Takenawa, T. and Nakamura, T. (1998) J. Biochem. 124, 769–777.
[9] Tsuji, T., Onimaru, M. and Yanagawa, H. (2001) Nucleic Acids Res. 29, E97.
[10] Ostermeier, M., Shim, J.H. and Benkovic, S.J. (1999) Nat. Biotechnol. 17, 1205–1209.
[11] Kitamura, K., Kinoshita, Y., Narasaki, S., Nemoto, N., Husimi, Y. and Nishigaki, K. (2002) Protein Eng. 15, 843–853.
[12] Abedi, M.R., Caponigro, G. and Kamb, A. (1998) Nucleic Acids Res. 26, 623–630.
[13] Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (1999) Proc. Natl. Acad. Sci. USA 96, 11241–11246.
[14] Ito, Y., Suzuki, M. and Husimi, Y. (1999) Biochem. Biophys. Res. Commun. 264, 556–560.
[15] Nishigaki, K. and Yokoyama, S. (1996) in: A Pilot Experiment Designed for Evaluation of a Strategy ROSOR for Human Genome Project (Nishigaki, K., Ed.), pp. 97–114, Scientific Report for Grant-in-Aid from the Ministry of Education, Science, Sports, and Culture of Japan, Saitama.
[16] Aita, T. and Husimi, Y. (1996) J. Theor. Biol. 182, 469–485.
[17] Barnes, W.M., Bevan, M. and Son, P.H. (1983) Methods Enzymol. 101, 98–122.
[18] Zhu, L. and Holtz, A.E. (1996) Methods Mol. Biol. 57, 119–137.
[19] Inaba, K., Ishimori, K., Imai, K. and Morishima, I. (2000) J. Biol. Chem. 275, 12438–12445.