МОЛЕКУЛЯРНАЯ БИОЛОГИЯ, 2012, том 46, № 1, с. 156–161

СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

UDC 577.112.012

THE CONFORMATIONAL POLYMORPHISM OF THE GREEN FLUORESCENT PROTEIN

© 2012 Haidong Tan[1]*, Yueguang Li[2], Ling Chen[3], Takayuki Kudoh[3], Tomonari Kasai[3], Masaharu Seno[3]*

[1]Dalian Institute of Chemical Physics, CAS, Dalian 116023, China
[2]Department of General Surgery, Tianjin 4th Centre Hospital, Zhongshanlu, Hebeiqu, Tianjin 300140, People's Republic of China
[3]Department of Medical and Bioengineering Science, Graduate School of Natural Science and Technology, Okayama University, Okayama 700-8530, Japan

Received November 09, 2010
Accepted for publication January 13, 2011

Green fluorescent protein (GFPuv) has been widely used as a reporter fused to individual targeting sequences. However, its state in liquid and its effect on other proteins are still unclear. The conformational polymorphisms of glutathione-S-transferase-green fluorescent protein (GST-GFPuv), GFPuv and GST were analyzed by native polyacrylamide gel, indicating that GST was in many different states while GFPuv and GST-GFPuv were only in four and two slightly different states. Four different circular dichroism spectra were obtained from the GFPuv polymorphisms. The single molecular behavior of GST-GFPuv and GFPuv was also characterized by MALDI-TOF MS. Thus, we demonstrated that: (1) there might be four different structural polymorphisms for the native GFPuv; (2) GFPuv could reduce its partner's polymorphism as a fusion protein. Although GFPuv had many merits as a reporter, its unreliability was found in the study.

Keywords: **green fluorescent protein, conformational polymorphism, glutathione-S-transferase, native polyacrylamide gel electrophoresis, MALDI-TOF MS, circular dichroism spectra.**

Although crystal structures of many proteins have been extensively provided [1, 2], the exact structures of the proteins in their native states are still poorly understood [3]. Thus, the structural polymorphisms of native proteins are the basic information for understanding their exact structure.

The green fluorescent protein (GFPuv) is exceptionally widely utilized in cell biology as a natural, brightly fluorescent marker for gene expression, localization of gene products and intracellular protein to protein interactions [4–8]. In order to understand the complex photo physics of GFPuv, a considerable number of studies has been published [9–12]. Thus, GFPuv can be a potential candidate to help us to understand such polymorphisms. The crystal structure of GFPuv is a cylinder, which ensures stable positioning [13] and makes the protein very stable even under strict conditions [14, 15].

In liquid, the secondary structure of proteins can be determined by CD spectroscopy in the far-UV spectral region (190–250 nm). Thus, the method has been widely used for measuring the structure of GFPuv in its native state [16, 17]. However, the structure of GFPuv is a cylinder with two flexible loops at the two ends, so the structure might have more than only one conformation in its native state. The present study is directed towards investigation of the structural polymorphisms of GFPuv and its effects on its' partner in a fusion protein.

On the other hand, GFPuv has been widely used as a reporter fused to individual targeting sequences in the articles, published in most of the leading journals, including *Cell*, *Science* and *Nature*. However, the routine lab method's accuracy may be called into question because GFPuv may affect its partner's function and location in the cell [18, 19]. Also, a strong constant signal from the GFPuv may remain even when its partner is degraded, because GFPuv is highly resistant to protease degradation [20] and can be fused to ubiquitin during the expression [21, 22]. Here, this problem was also found.

EXPERIMENTAL

Materials. Expression vector pET41a was purchased from "Novagen" (USA). Modified pEGFP-1 plasmid (Avoiding restriction digestion site existed in the GFPuv gene, site-directed mutagenesis was used to change the corresponding sites as A178G site and C423G site) was kindly supplied by Zongbao Zhao. *Escherichia coli* BL21 (DE3) and *E. coli* DH5α, plasmid mini-preparation kit, high fidelity polymerase

* E-mail: mseno@cc.okayama-u.ac.jp

and other reagents were purchased from "Dingguo Biotech" (China). Standard EK was purchased from "New England BioLabs Ltd" (China). Primers, restriction enzymes and the T4 DNA ligase were from "TaKaRa Biotechnology" (China).

Plasmid construction. The GFPuv gene was amplified using a modified pEGFP-1 plasmid as template by PCR with a forward primer 5'-GCGGAATTCAGTAAAGGAGAAGAAC-3' and, a reverse primer 5'-CGGATCCTCGAGTTATTTGTAGAGCTC-3'. The PCR was performed according to the following thermocycle: an initial denaturing step at 95°C for 2 min, 30 cycles (95°C for 30 s, 58°C for 30 s, 72°C for 1 min), and additional extension at 72°C for 10 min. The PCR product was cleaved by EcoRI and XhoI and cloned into the corresponding sites in pET41a to obtain the expression plasmid pET41a-GFPuv and then transformed into *E. coli* BL21 (DE3).

Fusion protein expression and purification. One colony of *E. coli* BL21(DE3) containing pET41a-GFPuv plasmids was picked and transferred into 25 ml LB medium containing kanamycin (50 μg/ml) and cultured at 200 rpm shaking at 37°C overnight. 10 ml culture was added to 1 l of LB broth containing kanamycin (50 μg/ml) and 0.1 mM IPTG and incubated at 37°C for 10 h.

All of the cells were harvested by centrifugation at 5000 g for 15 min and washed twice with 500 mM NaCl, 20 mM Tris-HCl, pH 8.0 buffer. Finally, ultrasonic methods were used to disrupt the host strain. The disrupted cell liquid was centrifuged at 12000g for 15 min, and the supernatants were collected, added 70 mM imidazole and loaded onto a Ni-NTA column, pre-equilibrated with the same buffer and the same concentration of imidazole, for purification. The proteins were eluted from the column by 250 mM imidazole, 20 mM sodium phosphate and 500 mM NaCl. The fusion protein glutathione-S-transferase-green fluorescent protein (GST-GFPuv), harboring an enterokinase site between GST and GFPuv, was purified.

Conformational polymorphism analysis by native polyacrylamide gel electrophoresis (PAGE). Initially the purified proteins were dialyzed against the buffer (2 mM $CaCl_2$, 100 mM NaCl and 20 mM Tris-HCl, pH 8.0) at 4°C for 24 h. 50 ug GST-GFPuv were cleaved by 2 u enterokinase at 25°C for 24 h. GFPuv with a His-tag at the C terminus was further purified using the above mentioned Ni-NTA purification system. All of the samples were analyzed by native PAGE [23]. Using UV lamps (300 nm) as an illumination source, the conformational polymorphisms of GST-GFPuv and GFPuv were determined by native PAGE since the proteins were separated according to the net charge, size and shape of their native structures. For direct visualization of the proteins, the gel was stained with Coomassie Brilliant Blue R-250.

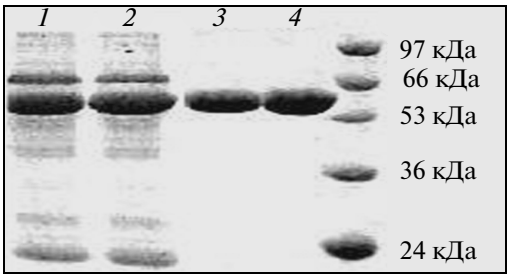

Fig. 1. SDS-PAGE analysis of GST-GFPuv purification using Ni-NTA purification system. Lanes *1* and *2* – Unpurified GST-GFPuv. Lanes *3* and *4* – GST-GFPuv purified through two steps (loading buffer with 70 mM imidazole and eluting liquid with 250 mM imidazole).

Circular Dichroism Measurements. The four bands of the cleaved GFPuv (fig. 1) on the SDS-PAGE gel were excised respectively and washed in a solution containing 100 mM potassium phosphate (pH 8.0), 50 mM NaCl, and 1 mM EDTA. CD wavelength scans were recorded with a Jasco J-715 spectropolarimeter ("Jasco", Japan). The protein samples were prepared with a concentration of 0.1 mg ml^{-1}. All spectra were recorded at 25°C with 10 scans. The data was normalized to a molar ellipticity with a path length of 0.1 cm.

MALDI-TOF MS analysis of GST-GFPuv and GFPuv. To avoid the mutants of GST-GFPuv and GFPuv causing the polymorphism of the molecules, MALDI-TOF MS analysis of GST-GFPuv and purified GFPuv were performed. The experimental protocol for sample preparation and MALDI analysis was described elsewhere [24]. The samples were analyzed on a MALDI-TOF mass spectrometer ("Bruker Autoflex, Bruker-Daltonics", Germany). Mass calibration was performed with aqueous solutions of lysine, trehalose and alanine. The accuracy was ±0.1 apparent mass units.

Unreliability for GFPuv as a reporter. GST-GFPuv was incubated at 37°C for 0, 24, 36, 42, 45, 47 and 48 h, respectively. The samples were detected by SDS-PAGE using UV lamps (300 nm) as an illumination source. For direct visualization of the proteins, the gel was stained with Coomassie Brilliant Blue R-250. The molecular weight of the samples was determined by SDS-PAGE.

RESULTS AND DISCUSSION

Amplification of the GFPuv gene and plasmid construction

The GFPuv gene was amplified by PCR using a modified pEGFP-1 plasmid as a template. A DNA fragment of about 0.7 kb was obtained, what coincided with the expected fragment length. The PCR product was purified and digested with EcoR I/Xho I, and cloned into plasmid pET-41a between the EcoR I/Xho I restric-

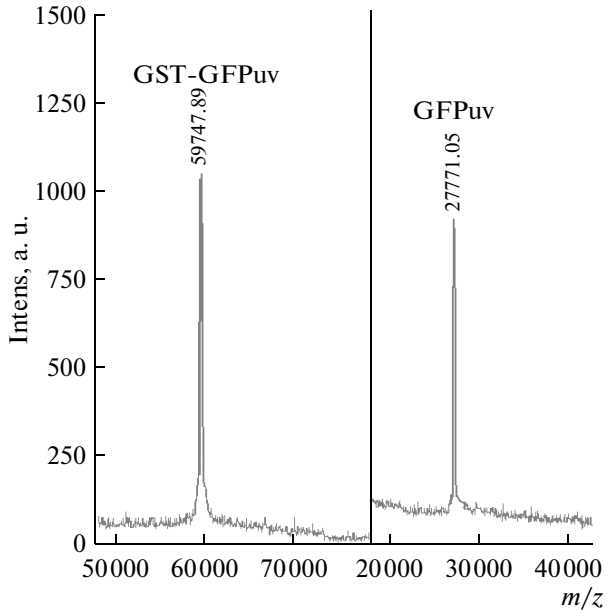

Fig. 4. MALDI-TOF MS analysis of purified GST-GFPuv and GFPuv. The purified GST-GFPuv appeared as two states on the native PAGE while the purified GFPuv appeared as four states on the PAGE (fig. 3). One microliter of the GST-GFPuv or GFPuv was desalted through a C18 ZipTip (Millipore). The ZipTip was activated, equilibrated, and loaded according to the manufacturer's instructions. The bound material was then eluted with α-cyano-4-hydroxycinnamic acid (10 mg/ml) in 70% acetonitrile containing 0.1% TFA before being spotted and co-crystallized.

tion sites downstream of the fusion partner GST to give pET41a-GFPuv.

Fusion protein expression and purification

Upon induction with IPTG, the fusion protein could be highly expressed (500 mg GST-GFPuv l^{-1}). Illuminating the Ni-NTA affinity column with a long UV lamp ($\lambda = 365$ nm), protein purification could be clearly tracked and optimized. Thus highly purified GST-GFPuv was obtained without a washing step (fig. 2).

The conformational polymorphism of GFPuv

The structural polymorphisms of GST-GFPuv and GFPuv (the latter was obtained by digesting the fusion protein with an enterokinase) were observed by native PAGE using UV lamps (300 nm) as an illumination source. As fig. 2 showed, GST was in a smear, while GFPuv and GST-GFPuv were in four and two slightly different states, respectively. The slight difference of GFPuv or GST-GFPuv should be caused by the structural polymorphisms since the proteins were separated according to the net charge, size and shape of their native structures.

Unlike the mess of polymorphisms of GST observed in its free state in the fusion protein only two states of GST were present (fig. 2), suggesting that GFPuv could reduce the number of its partner's conformational polymorphisms. Thus, GFPuv was very useful to avoid hydrophobic interactions and improve the refolding of its partner during recombinant protein expression. GST was also fused to other proteins but was often presented as a mess on the native gel [23].

In native PAGE proteins are separated according to the net charge, size and shape of their native structure. According to the protocol mentioned in the previous report [25], the relation between electrophoresis velocity and protein compact degree was established as: $V = \frac{P\sqrt{-Q}}{CDM}$, where V stands for electrophoresis velocity (cm h^{-1}), P stands for power (watt), Q – protein charge, C – protein compact degree, D stands for the percentage of the gel and M – for the molecular weight of the protein (Dalton). The structural polymorphism of protein is mainly caused by the parameter C. In our results, the C values were almost the same for GST-GFPuv and GFPuv, but quite different from the value of GST. Thus, we suggested that GFPuv has a mooring function and reduces the number of its partner's polymorphisms.

Circular Dichroism Measurements

The four single protein bands on the SDS-PAGE gel were analyzed by Jasco J-715 spectropolarimeter. They show four different circular dichroism spectra, in which the α-helix, β-sheet and β-turn were almost the same, while the remainder was different (fig. 3). We guess that the remainder might be the side loops on the both ends, which are flexible and may cause the different structural polymorphisms in GFPuv (fig. 2). Other structures, such as the α-helix, β-sheet and β-turn, were very stable, forming a cylinder with a rigid structure. Our results were different from previous reports [16, 17], in which only one native structure was reported.

MALDI-TOF MS analysis of GST-GFPuv and GFPuv

To avoid the polymorphism caused by the mutants of GST-GFPuv and GFPuv, MALDI-TOF MS analysis was performed. The single molecular behavior of GST-GFPuv and GFPuv was characterized by MALDI-TOF MS although they appeared as two states and four states in native PAGE (fig. 2 and fig. 4). The molecular weights of GST-GFPuv and GFPuv were consistent with their theoretical values (fig. 4), so there were no mutants among the molecules. Thus, these results further proved that the different polymorphisms might be caused by different conformations, but not by the mutants. For GST, there were many polymorphisms, which were hard to analyze by CD spectra. However, the

Previous reports for GFPuv's partners and possible mechanisms of the fusion protein degradation

GFPuv's partner used in the publications of *Cell*, *Nature*, and *Science*	Possible degradation mechanism
P450 (*1*, *2*)	Ubiquitin-dependent proteasomal degradation (*3–22*)
Arginine vasopressin (*23*)	Vasopressinase (*24-29*)
Alkaline phosphatase (*30*)	Ubiquitin-dependent proteasomal degradation (*31–33*)
Leptin (*34*)	Ubiquitin-dependent proteasomal degradation (*35–37*)
Actin (*38*, *39*)	–"– (*40–52*)
Myc (*53*)	–"– (*54–78*)
c-kit (*79*)	–"– (*80–83*)
Raf (*84*)	–"– (*85–90*)
CLIP170 (*91*)	–"– (*92*)
GluR-A (*93*)	–"– (*94–96*)
Nras/Kras4B (*97*)	–"– (*98–100*)
VSVG (*101*)	VSV acid-activation (*102–105*)
Calmodulin (*106*)	20S proteasome (*107–111*)
Integrin (*112*)	Ubiquitin-dependent proteasomal degradation (*113–116*)
Glucocorticoid receptor (*117*)	Ubiquitin-dependent proteasomal degradation (*118–125*)
ErbB1 (*126*)	c-Cbl-regulation (*127–130*)
Histone H1 (*131*)	Ubiquitin-dependent proteasomal degradation (*132–135*)
Dynamin (*136*)	Ubiquitin-dependent proteasomal degradation (*137–140*)
Myosin II (*141*)	–"– (*142–144*)
VEGF (*145*)	–"– (*146–148*)

Note: GFPuv is highly resistant to protease degradation (*149*) and can be fused to ubiquitin during expression (*150*, *151*). The number stands for the corresponding references, all of which can be viewed in supplementary material (please contact with corresponding author).

polymorphisms could be reduced by its partner in the fusion protein (fig. 2).

Unreliability of GFPuv as a reporter

GFPuv has been widely used as a reporter fused to individual targeting sequences (table). However, the accuracy of routine lab methods may be called into question because GFPuv is highly stable and resistant to protease degradation, heat and chemical denaturation, while many of its partners may be unstable and easily proteolyzed (table). Here, GFPuv's unreliability was also found. Its partner was degraded after 42 h of incubation at 37°C, but a strong fluorescent signal remained constant (fig. 5). However, this seemed very unreasonable, because it would be expected that the degradation would have occured during the entire period. We guess that the degradation is too low to be detected at the beginning, and the conformation will change slightly preparing for degradation. When the changes accumulate to a certain degree, degradation can be promoted abruptly (fig. 5). Meanwhile, mass spectrum analysis revealed that there were still eight amino acid residues located upstream of the GFPuv after fusion protein degradation.

Frankly, the stability of GFPuv is not absolute. GFPuv was stable, but could be destabilized by the addition of putative proteolytic signal sequences [26]. Additionally, a destabilized GFP could be created by fusing amino acids 422–461 of the degradation domain of mouse ornithine decarboxylase (MODC) to the C-terminal end of GFPuv. The fusion protein was unstable in the presence of cycloheximide and had a fluorescence half-life of 2 h [27]. In the same way, the fusion protein GFPuv-IkB can be degraded as well as

IkB upon TNFα treatment [28]. Likely, short-lived green fluorescent proteins could also be created by fusing with ubiquitin [29]. Therefore, on the contrary, if some proteins were more stable than GFPuv and GFPuv would be degraded in a short time, the target protein would still exist although the signal would be gone.

Thus, to improve standardization, the intactness of its partner should be checked if GFPuv was used as a reporter. Sometimes western blotting may be essential.

CONCLUSION

There might be four slightly different structural polymorphisms for the native GFPuv. Thus, it is a potential candidate to explore the exact structure of the native protein. Because of its polymorphism, GFPuv could reduce its partner's polymorphisms in a fusion protein. Although GFPuv has many merits as a reporter, its unreliability should be considered in many fields of research.

We appreciate Ms. Mami Asakura for her excellent technical assistance. We also thank the anonymous reviewer for his or her insightful suggestions.

The financial support of the National Basic Research Program of China (973 Program) (No. 2007CB707802) is gratefully acknowledged.

REFERENCES

1. May P., Kreuchwig A., Steinke T., Koch I. 2010. PTGL: a database for secondary structure-based protein topologies. *Nucl. Acids Res.* **38**, D326–330.
2. Lobanov M., Bogatyreva N.S., Ivankov D.N., Finkel'shtein A.V. 2009. Protein structure prediction from analogy. I. New database of spatially similar and dissimilar structures of protein domains for testing and optimization of prediction methods. *Mol. Biol. (Mosk.)*. **43**, 722–732.
3. Eyal E., Gerzon S., Potapov V., Edelman M., Sobolev V. 2005. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *J. Mol. Biol.* **351**, 431–442.
4. Leclercq J., Lardet L., Martin F., Chapuset T., Oliver G., Montoro P. 2010. The green fluorescent protein as an efficient selection marker for *Agrobacterium tumefaciens*-mediated transformation in *Hevea brasiliensis* (Mull. Arg). *Plant Cell Rep.* **29**, 513–522.
5. Zhang Z., Dharmakumar R., Mascheri N., Fan Z., Wu S., Li D. 2009. Comparison of superparamagnetic and ultrasmall superparamagnetic iron oxide cell labeling for tracking green fluorescent protein gene marker with negative and positive contrast magnetic resonance imaging. *Mol. Imaging.* **8**, 148–155.
6. Li W., Zhang N., Liang X., Li J., Gong P., Yu X., Ma G., Ryan U.M., Zhang X. 2009. Transient transfection of *Cryptosporidium parvum* using green fluorescent protein (GFP) as a marker. *Mol. Biochem. Parasitol.* **168**, 143–148.
7. Kim H.S., Cho H.R., Choi S.H., Woo J.S., Moon W.K. 2010. *In vivo* imaging of tumor transduced with bimodal lentiviral vector encoding human ferritin and green fluorescent protein on a 1.5T clinical magnetic resonance scanner. *Cancer Res.* **70**, 7315–7324.
8. Viczian A., Kircher S. 2010. Luciferase and green fluorescent protein reporter genes as tools to determine protein abundance and intracellular dynamics. *Meth. Mol. Biol.* **655**, 293–312.
9. Garcia-Parajo M.F., Segers-Nolten G.M., Veerman J.A., Greve J., van Hulst N.F. 2000. Real-time light-driven dynamics of the fluorescence emission in single green fluorescent protein molecules. *Proc. Natl. Acad. Sci. USA.* **97**, 7237–7242.
10. Kubitscheck U., Kuckmann O., Kues T., Peters R. 2000. Imaging and tracking of single GFP molecules in solution. *Biophys. J.* **78**, 2170–2179.
11. Haupts U., Maiti S., Schwille P., Webb W.W. 1998. Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *Proc. Natl. Acad. Sci. USA.* **95**, 13573–13578.
12. Creemers T.M., Lock A.J., Subramaniam V., Jovin T.M., Volker S. 2000. Photophysics and optical switching in green fluorescent protein mutants. *Proc. Natl. Acad. Sci. USA.* **97**, 2974–2978.
13. *Y*ang F., Moss L.G., Phillips G.N., Jr. 1996. The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246–1251.
14. Ishii M., Kunimura J.S., Jeng H.T., Penna T.C., Cholewa O. 2007. Evaluation of the pH- and thermal stability of the recombinant green fluorescent protein (GFP) in the presence of sodium chloride. *Appl. Biochem. Biotechnol.* **137–140**, 555–571.
15. Penna T.C., Ishii M., Junior A.P., Cholewa O. 2004. Thermal stability of recombinant green fluorescent protein (GFPuv) at various pH values. *Appl. Biochem. Biotechnol.* **113–116**, 469–483.
16. Iwai H., Lingel A., Pluckthun A. 2001. Cyclic green fluorescent protein produced *in vivo* using an artificially split PI-PfuI intein from *Pyrococcus furiosus*. *J. Biol. Chem.* **276**, 16548–16554.
17. Visser N.V., Hink M.A., Borst J.W., van der Krogt G.N., Visser A.J. 2002. Circular dichroism spectroscopy of fluorescent proteins. *FEBS Lett.* **521**, 31–35.
18. Hanson D.A., Ziegler S.F. 2004. Fusion of green fluorescent protein to the C-terminus of granulysin alters its intracellular localization in comparison to the native molecule. *J. Negat. Results Biomed.* **3**, 2.
19. Prescott M., Nowakowski S., Nagley P., Devenish R.J. 1999. The length of polypeptide linker affects the stability of green fluorescent protein fusion proteins. *Anal. Biochem.* **273**, 305–307.
20. Roucou X., Prescott M., Devenish R.J., Nagley P. 2000. A cytochrome c-GFP fusion is not released from mitochondria into the cytoplasm upon expression of Bax in yeast cells. *FEBS Lett.* **471**, 235–239.
21. Perroy J., Pontier S., Charest P.G., Aubry M., Bouvier M. 2004. Real-time monitoring of ubiquitination in living cells by BRET. *Nat. Methods.* **1**, 203–208.

22. Dantuma N.P., Groothuis T.A., Salomons F.A., Neefjes J. 2006. A dynamic ubiquitin equilibrium couples proteasomal activity to chromatin remodeling. *J. Cell Biol.* **173**, 19–26.
23. Wang C., Meier U.T. 2004. Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J.* **23**, 1857–1867.
24. Liao C.C., Mehta A., Ward N.J., Marsh S., Arulampalam T., Norton J.D. 2010. Analysis of post-operative changes in serum protein expression profiles from colorectal cancer patients by MALDI-TOF mass spectrometry: a pilot methodological study. *World J. Surg. Oncol.* **8**, 33.
25. Gallagher S. 2001. One-dimensional electrophoresis using nondenaturing conditions. *Curr. Protoc. Protein Sci.* **Chapter 10**, Unit 10 13.
26. Corish P., Tyler-Smith C. 1999. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng.* **12**, 1035–1040.
27. Li X., Zhao X., Fang Y., Jiang X., Duong T., Fan C., Huang C.C., Kain S.R. 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* **273**, 34970–34975.
28. Li X., Fang Y., Zhao X., Jiang X., Duong T., Kain S.R. 1999. Characterization of NFkappaB activation by detection of green fluorescent protein-tagged IkappaB degradation in living cells. *J. Biol. Chem.* **274**, 21244–21250.
29. Dantuma N.P., Lindsten K., Glas R., Jellne M., Masucci M.G. 2000. Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells. *Nat. Biotechnol.* **18**, 538–543.