BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Structure of the green fluorescent protein NowGFP with an anionic tryptophan-based chromophore

Vladimir Z. Pletnev,[a]* Nadya V. Pletneva,[a] Karen S. Sarkisyan,[a] Alexander S. Mishin,[a] Konstantin A. Lukyanov,[a] Ekaterina A. Goryacheva,[a] Rustam H. Ziganshin,[a] Zbigniew Dauter[b] and Sergei Pletnev[b,c]*

[a]Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, GSP-7, Moscow, Russian Federation, [b]Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Argonne, IL 60439, USA, and [c]Basic Research Program, Leidos Biomedical Research Inc., Argonne, IL 60439, USA. *Correspondence e-mail: vzpletnev@gmail.com, pletnevs@mail.nih.gov

Received 13 February 2015
Accepted 26 May 2015

Edited by R. McKenna, University of Florida, USA

Keywords: fluorescent proteins; tryptophan-based chromophore; anionic tryptophan; crystal structure; spectral properties.

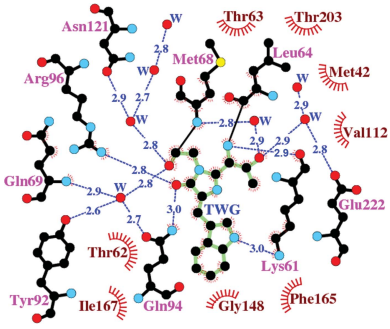
PDB references: NowGFP, pH 4.8, 4rys; pH 9.0, 4rtc; NowGFP_conv, 4ryw

Supporting information: this article has supporting information at journals.iucr.org/d

A green-emitting fluorescent variant, NowGFP, with a tryptophan-based chromophore (Thr65-Trp66-Gly67) was recently developed from the cyan mCerulean by introducing 18 point mutations. NowGFP is characterized by bright green fluorescence at physiological and higher pH and by weak cyan fluorescence at low pH. Illumination with blue light induces irreversible photoconversion of NowGFP from a green-emitting to a cyan-emitting form. Here, the X-ray structures of intact NowGFP at pH 9.0 and pH 4.8 and of its photoconverted variant, NowGFP_conv, are reported at 1.35, 1.18 and 2.5 Å resolution, respectively. The structure of NowGFP at pH 9.0 suggests the anionic state of Trp66 of the chromophore to be the primary cause of its green fluorescence. At both examined pH values Trp66 predominantly adopted a *cis* conformation; only ∼20% of the *trans* conformation was observed at pH 4.8. It was shown that Lys61, which adopts two distinct pH-dependent conformations, is a key residue playing a central role in chromophore ionization. At high pH the side chain of Lys61 forms two hydrogen bonds, one to the indole N atom of Trp66 and the other to the carboxyl group of the catalytic Glu222, enabling an indirect noncovalent connection between them that in turn promotes Trp66 deprotonation. At low pH, the side chain of Lys61 is directed away from Trp66 and forms a hydrogen bond to Gln207. It has been shown that photoconversion of NowGFP is accompanied by decomposition of Lys61, with a predominant cleavage of its side chain at the $C^{\gamma}-C^{\delta}$ bond. Lys61, Glu222, Thr203 and Ser205 form a local hydrogen-bond network connected to the indole ring of the chromophore Trp66; mutation of any of these residues dramatically affects the spectral properties of NowGFP. On the other hand, an Ala150Val replacement in the vicinity of the chromophore indole ring resulted in a new advanced variant with a 2.5-fold improved photostability.

1. Introduction

Green fluorescent protein (GFP) and related fluorescent proteins (FPs) have become established as efficient, non-invasive molecular instruments in cell biology and biomedicine. They are used for the visualization and monitoring of internal processes within cells and whole organisms (Chudakov *et al.*, 2010; Lam *et al.*, 2012; Wu *et al.*, 2011). All wild-type fluorescent proteins known to date have a tyrosine as the central residue of the chromophore-forming triad. Usually, the phenolic hydroxyl of the chromophore Tyr ionizes easily, giving rise to longer wavelength fluorescence (from green to far red) compared with FPs with neutral Tyr. Replacement of Tyr with Trp or His/Phe results in a shift in fluorescence to shorter wavelengths in the cyan/blue region (Chudakov *et al.*, 2010). This blue shift has been attributed to



uncharged state of Trp-, His- or Phe-based chromophores. Until recently, no fluorescent proteins with charged non-Tyr-based chromophores had been described.

The first fluorescent protein with a presumably anionic Trp-based chromophore (Thr65-Trp66-Gly67) was developed in 2012 (Sarkisyan *et al.*, 2012). This protein, named WasCFP, was created from the cyan-fluorescent protein mCerulean (Rizzo *et al.*, 2004) by introducing a key Val61Lys substitution and four other random mutations (Fig. 1). As expected, ionization of the chromophore resulted in a strong red shift of the spectrum, with fluorescence in the green region ($\lambda_{ex}/\lambda_{em}$ = ~493/502 nm). WasCFP was found to be very unstable with respect to pH and temperature, demonstrating a reversible conversion to the common cyan state at low pH (<7) or at elevated temperatures (>25°C). Its improved variant, NowGFP, has an additional 13 mutations, possesses a dominant green fluorescence under physiological conditions and demonstrates high pH and temperature stability (Sarkisyan *et al.*, 2015).

Under blue-light illumination (450–480 nm) both WasCFP and NowGFP undergo irreversible photoconversion to their cyan form. Such green-to-cyan photoconversion has not been observed previously. The irreversible character of the spectral changes implies the occurrence of some chemical reaction(s) in the protein. However, the molecular mechanism of photoconversion in this type of FP remains unknown.

In terms of their practical application, the main advantages of WasCFP and NowGFP are the extremely long fluorescence lifetimes and high quantum yields of the fluorescence of their green form. These makes them potentially useful tags for fluorescence lifetime imaging microscopy (FLIM), both *in vivo* and *in vitro*, as well as excellent donors for fluorescence resonance energy transfer (FRET) applications (Sarkisyan *et al.*, 2012, 2015). The improved variant NowGFP is 30% brighter than EGFP and exhibits the longest fluorescent lifetime of 5.1 ns, compared with 2–3 ns for other green FPs, making it suitable for simultaneous high-contrast FLIM imaging with regular green FP.

Here, we present the results of a crystallographic study of NowGFP in the green-emitting and cyan-emitting forms at pH 9.0 and pH 4.8, respectively, and that of the photoconverted form, NowGFP_conv, at pH 7.0. The structures of these three forms were determined at 1.18, 1.35 and 2.50 Å resolution, respectively. The conclusions of the structural studies were validated by structure-guided site-directed mutagenesis.

2. Materials and methods

2.1. Expression, purification, crystallization and mutagenesis

For protein expression, the fragment encoding NowGFP with an N-terminal His tag was cloned into pQE-30 vector (Qiagen, USA) and transformed into *Escherichia coli* XL1 Blue strain (Invitrogen, USA). Bacterial cultures were grown overnight at 37°C. No IPTG induction was necessary since promoter leakage was sufficient for effective expression. The cells were pelleted by centrifugation, resuspended in phosphate-buffered saline and lysed by sonication. NowGFP was purified by immobilized metal-affinity chromatography using TALON resin (Clontech Laboratories, USA) followed by size-exclusion chromatography using a Superdex 200 (16/60) column (GE Healthcare, USA).

To obtain the photoconverted form, NowGFP_conv, a 1 mg ml^{-1} solution of NowGFP (at pH 7.0) was irradiated with 472 nm light (a LED array assembly of seven Luxeon Star LXMLPB010040 Rebel LEDs; Lumileds; light power 1 W cm^{-2}, λ_{max} = 470 nm) for 30 min. The photoconversion was monitored from the absorption spectra (a decrease in the green band at 493 nm and an increase in the cyan band at 429 nm). The spectra showed that ~75% of the NowGFP underwent photoconversion. The photoconverted protein was repurified by gel filtration using a Superdex 200 (16/60) column (GE Healthcare, USA).

For crystallization, NowGFP and NowGFP_conv were transferred to 20 m*M* Tris pH 8.0,

```
                10        20        30        40        50        60
                 |         |         |         |         |         |
NowGFP      VSKGEKLFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKMSLKFICTTGKLPVPWPTL
WasCFP      VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL
Cerulean    VSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTL
            *****:**********************************::******************

                70        80        90       100       110       120
                 |         |         |         |         |         |
NowGFP      KTTLTWGMQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV
WasCFP      KTTLTWGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV
Cerulean    VTTLTWGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV
             ******:****************************************************

               130       140       150       160       170       180
                 |         |         |         |         |         |
NowGFP      NRIELKGVDFKEDGNILGHKLEYNAISGNANITADKQKNGIKAYFTIRHDVEDGSVLLAD
WasCFP      NRIELKGIDFKEDGNILGHKLEYNAISGNVNITADKQKNGIKANFKIRHNIEDGSVQLAD
Cerulean    NRIELKGIDFKEDGNILGHKLEYNAISDNVYITADKQKNGIKANFKIRHNIEDGSVQLAD
            *******:******************.*. ************ *.***::***** ***

               190       200       210       220       230
                 |         |         |         |         |
NowGFP      HYQQNTPIGDGPVLLPDNHYLSTQSKQSKDPNEKRDHMVLLEFVTAAGIPLGADELYK
WasCFP      HYQQNTPIGDGPVLLPDNHYLSTQSKQSKDPNEKRDHMVLLEFVTAAGITLGMDELYK
Cerulean    HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK
            ************************  **********************.** *****
```

Figure 1
Alignment of the amino-acid sequences: residues in the nearest environment of the Thr65-Trp66-Gly67 chromophore (highlighted in grey) in WasCFP and NowGFP are shown in bold red and sequence differences are highlighted in yellow.

200 m*M* NaCl buffer and concentrated to 25 and 6.3 mg ml^{-1}, respectively. Initial screening for crystallization conditions was performed with a Mosquito Robotic Crystallization System (TTP Labtech, UK). Crystals suitable for data collection were obtained by the hanging-drop vapour-diffusion method. Typically, 2 µl protein solution was mixed with an equal amount of reservoir solution and incubated at 20°C for two weeks. The best crystals of NowGFP and NowGFP_conv were obtained from 14 m*M* KH_2PO_4 pH 4.8, 14% PEG 3350 (NowGFP_pH4.8) and from 0.1 *M* sodium acetate pH 4.5, 30%(*v*/*v*) PEG 5000 monomethyl ether, respectively. To obtain structures of NowGFP at pH 9.0 (NowGFP_pH9.0) and NowGFP_conv at pH 7.0 (NowGFP_conv_pH7.0) the pH of the cryoprotectant was adjusted with 4 *M* Tris base to a pH of ∼9.0 and a pH of ∼7.0, respectively.

Preparation of mutant variants by site-directed mutagenesis was performed by self-assembly cloning with primers containing appropriate target substitutions (Matsumoto & Itoh, 2011). Synthetic DNA oligonucleotides for mutagenesis were purchased from Evrogen. PCR reactions were carried out using a PTC-100 thermal cycler (MJ Research, Canada). Purification of PCR products and the products of digestion was performed by gel electrophoresis and extraction using a Cleanup Standard Kit (Evrogen, Russia). Small-scale expression was performed in *E. coli* XL1 Blue strain (Invitrogen, USA) grown at 37°C in Petri dishes with LB agar (100 mg ml^{-1} ampicillin) with no IPTG induction. After cell sonication, the mutant proteins were purified using TALON metal-affinity resin (Clontech, USA).

Absorption and excitation–emission spectra of the purified proteins were recorded with a Varian Cary 100 UV–Vis spectrophotometer and a Varian Cary Eclipse fluorescence spectrophotometer, respectively. Experiments under low-oxygen conditions were performed in an argon-purged solution in a closed cuvette.

2.2. X-ray data collection, structure solution and crystallographic refinement

X-ray diffraction data were collected on SER-CAT beamlines 22-ID and 22-BM at the Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois, USA. Prior to data collection, the crystals were briefly soaked in a cryoprotectant solution containing 20% glycerol and 80% reservoir solution and were flash-cooled in a 100 K nitrogen stream.

Table 1
Crystallographic data and refinement statistics for NowGFP.

Protein	NowGFP_pH9.0 (PDB entry 4rtc)	NowGFP_pH4.8 (PDB entry 4rys)	NowGFP_conv (PDB entry 4ryw)
Crystallographic data			
Space group	*C*2	*C*2	$P2_1$
Unit-cell parameters (Å, °)	*a* = 111.1, *b* = 51.2, *c* = 55.7, *β* = 100.0	*a* = 111.6, *b* = 51.3, *c* = 61.0, *β* = 108.3	*a* = 52.1, *b* = 98.0, *c* = 71.7, *β* = 111.6
Z/*Z*′	4/1	4/1	6/3
Estimated solvent content (%)	59	61	44
Temperature (K)	100	100	100
Wavelength (Å)	1.00	1.00	1.00
Resolution range† (Å)	27.42–1.35 (1.40–1.35)	26.19–1.18 (1.22–1.18)	30.00–2.50 (2.59–2.50)
Total observations	250298	388492	86273
Unique reflections observed	67619	107191	23175
Multiplicity†	3.7 (3.4)	3.6 (3.0)	3.7 (3.6)
$\langle I/\sigma(I)\rangle$†	21.5 (2.0)	19.7 (2.2)	10.4 (2.4)
R_{merge}†	0.056 (0.547)	0.057 (0.425)	0.117 (0.502)
Completeness† (%)	99.9 (99.9)	99.5 (96.8)	100.0 (99.9)
Refinement statistics			
Non-H atoms in model			
Protein	1911 (residues 2–231)	1974 (residues 0–231)	5441
Water	271	326	170
Glycerol	12	36	54
R_{work}‡	0.161 (98.0%)	0.132 (98.5%)	0.205 (94.9%)
R_{free}‡	0.193 (2.0%)	0.142 (1.5%)	0.295 (5.1%)
R.m.s.d. from ideal values			
Bond lengths (Å)	0.027	0.029	0.012
Bond angles (°)	2.1	3.1	1.5
Torsion angles (period 3; °)	14	14	17
Chirality (Å^3)	0.15	0.25	0.08
General planes (Å)	0.014	0.013	0.006
Ramachandran statistics (%)			
Preferred	97.1	97.9	95.4
Allowed	2.9	2.1	4.3
Outliers	0.0	0.0	0.3

† Values in parentheses are for the highest resolution shell. ‡ Values in parentheses are the percentages of the data reserved for the working and free sets.

For NowGFP_pH9.0 and NowGFP_conv_pH7.0 the pH of the cryoprotectant was adjusted with 4 *M* Tris base to a pH of ∼9.0 and a pH of ∼7.0, respectively. All diffraction images were processed with *HKL*-2000 (Otwinowski & Minor, 1997). Crystal structures were solved by the molecular-replacement method with *MOLREP* (Vagin & Teplyakov, 2010; Winn *et al.*, 2011) using the coordinates of Cerulean (92% sequence identity; PDB entry 2wso; Lelimousin *et al.*, 2009) as a search model. Crystallographic refinement was performed with *REFMAC*5 (Murshudov *et al.*, 2011), alternating with manual revision of the model with *Coot* (Emsley *et al.*, 2010). The location of water molecules and structure validation were performed with *Coot*. Crystallographic data and refinement statistics are given in Table 1. The coordinates and structure factors of NowGFP_pH9.0, NowGFP_pH4.8 and NowGFP_conv_pH7.0 were deposited in the Protein Data Bank under accession codes 4rtc, 4rys and 4ryw, respectively.

2.3. Mass-spectrometric experiment

Lyophilized protein samples (100 µg each) of blue-light illuminated and non-illuminated NowGFP were dissolved in 10 µl buffer consisting of 100 m*M* ammonium bicarbonate (ABC), 5% sodium deoxycholate, 10 m*M* dithiothreitol (DTT) and incubated for 40 min at 60°C. Reduced cysteine

residues were alkylated by adding 5 µl aqueous 50 m*M* iodoacetamide (IAA) and a 30 min incubation at room temperature in the dark. The unreacted IAA was neutralized with 5 µl 50 m*M* aqueous DTT. Protein samples were diluted with 50 µl water and α-chymotrypsin from bovine pancreas (Sigma–Aldrich, USA) was added in a 1:100 (enzyme:protein) ratio to a final volume of 100 µl. Samples were incubated for 5 h at 37°C. Chymotrypsin activity was inhibited by adding

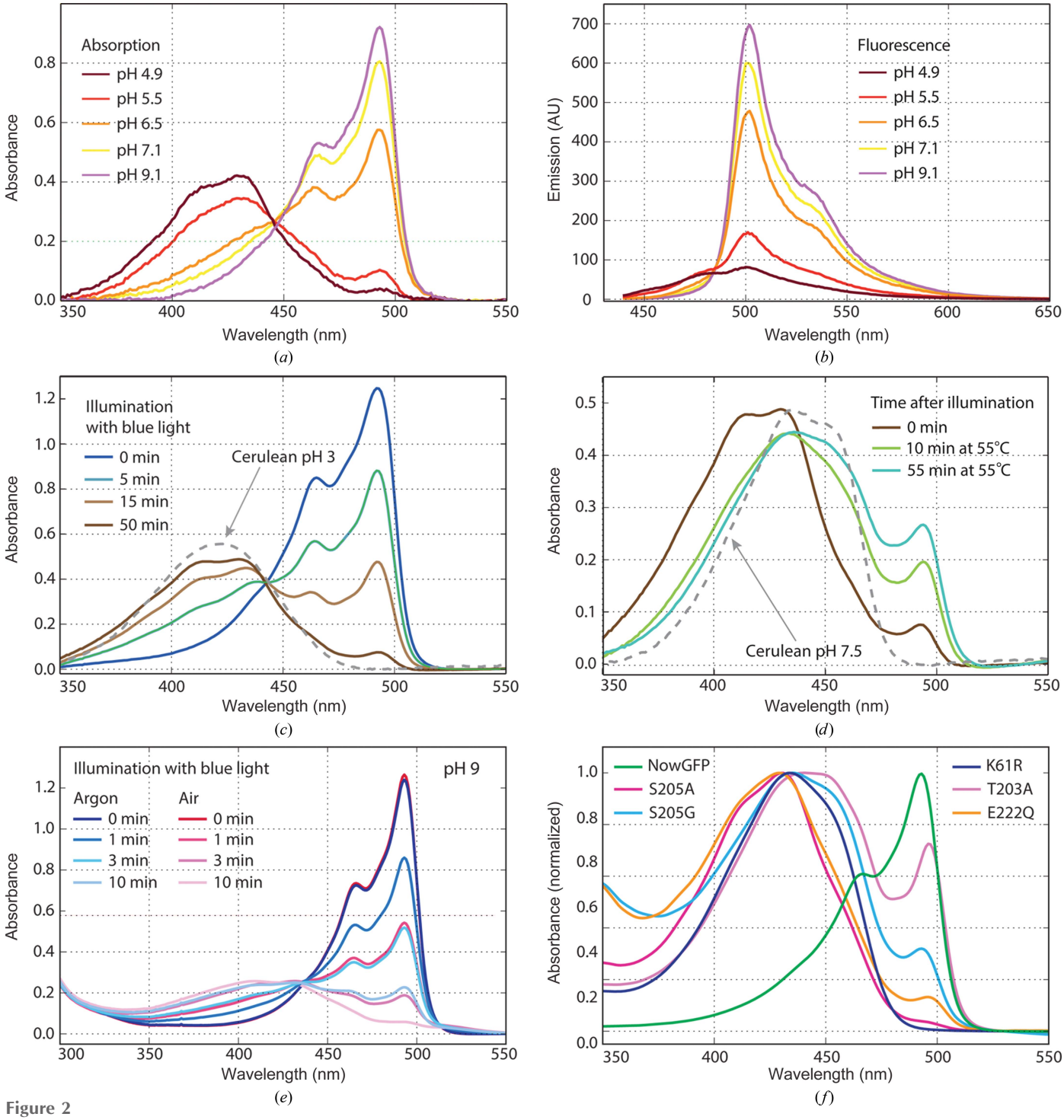


Figure 2
Spectral properties of NowGFP. (*a*) Absorption and (*b*) emission spectra at various pH values. (*c*, *d*) Changes in the absorption spectra accompanying green-to-cyan photoconversion under blue-light illumination (*c*) and a further relaxation period in the dark (*d*). Spectra of Cerulean at pH 3 (with the chromophore in the *trans* conformation) and pH 7.5 (with the chromophore in the *cis* conformation) are shown as grey dashed lines for comparison; similarity to these spectra implies formation of the *trans* chromophore conformation during photoconversion of NowGFP (*c*) followed by limited *trans*–*cis* isomerization during dark relaxation (*d*). (*e*) Influence of oxygen on green-to-cyan photoconversion with limited (under argon) and natural (atmospheric) access of oxygen. (*f*) Absorption spectra of engineered variants with single mutations: Thr203Ala, Ser205Ala/Gly, Glu222Gln and Lys61Arg. The mutated residues form a local hydrogen-bond network with the Trp66 indole ring.

Table 2
Dependence of NowGFP green fluorescence ($\lambda_{abs}^{max}/\lambda_{em}^{max}$ = 493/502 nm) on pH.

pH	EC† (M^{-1} cm^{-1})	QY‡	B§
4.9	2700	0.71	1.9
5.5	6900	0.90	7.1
6.5	38300	0.73	28.0
7.1	53600	0.68	36.4
9.1	61300	0.68	41.7

† EC, molar extinction coefficient at 493 nm. ‡ QY, fluorescence quantum yield. § B, brightness: (EC × QY)/1000.

5 µl 10% TFA and the resulting peptides were desalted using Discovery DSC-18 reverse-phase solid extraction cartridges (100 mg; Supelco). The eluted peptides were vacuum-dried and stored at −80°C prior to MS analysis.

MALDI-TOF MS analysis of the resulting protein fragments was performed using an ultraflex II TOF/TOF mass spectrometer (Bruker Daltonics, Germany). A matrix solution was prepared by dissolving 20 mg 2,5-dihydroxybenzoic acid (Bruker Daltonics, Germany) in 1 ml of a 1:1 water: acetonitrile solution of 0.1% TFA. Mass spectra were acquired in the positive-ion reflection mode. The MS data were processed with the Bruker Daltonics *flexAnalysis* 3.3 software. Correlation of the MS data with the protein sequence was carried out with the Bruker Daltonics *BioTools* 3.1 software.

3. Results

3.1. Spectral characteristics

At physiological and higher pH the spectrum of NowGFP exhibits a strong green fluorescence ($\lambda_{ex}/\lambda_{em}$ = 493/502 nm) and it exhibits a weak cyan fluorescence at acidic pH ($\lambda_{ex}/\lambda_{em}$ = 429/475 nm) (Figs. 2*a* and 2*b*). The ratio of the green:cyan fluorescence-band intensities strongly depends on the pH. Alkalization of the solution from pH 5.0 to 9.0 resulted in a nearly 20-fold increase in the green emission and in a nearly complete elimination of the cyan band (Table 2). Irradiation of NowGFP with blue light (450–480 nm) in basic and neutral solutions results in almost irreversible photoconversion to the cyan form (NowGFP_conv). This photoconversion is accompanied by a dramatic decrease in green fluorescence, which undergoes a partial recovery when illumination stops (Figs. 2*c* and 2*d*). Photoconversion shows a noticeable deceleration when the access of oxygen is limited (Fig. 2*e*). Point mutation of any of the Lys61, Glu222, Thr203 and Ser205 residues, which are involved in the local hydrogen-bond network with the chromophore Trp66, significantly reduces the fraction of protein absorbing at 493 nm, hence greatly decreasing its green fluorescence (Fig. 2*f*).

3.2. Overall structure

NowGFP is monomeric both in solution and in the crystal; its principal structural fold, which is shared with all other members of the GFP family, is an 11-stranded β-barrel with the chromophore (Thr65-Trp66-Gly67) embedded in the middle of an internal axial helix. The peptide bond preceding Pro89 adopts a rare *cis* conformation. Multiple conformations of the side chains were observed for a number of residues. Superposition of the atomic structures of NowGFP_pH9, NowGFP_pH4.8 and Cerulean at pH 7 and pH 5 (Lelimousin *et al.*, 2009; Malo *et al.*, 2007) revealed a significant conformational difference between the β-strand fragments 203–208 manifested as a ∼2 Å shift of several C$^{\alpha}$ positions (Fig. 3).

3.3. Structural features of the chromophore environment

3.3.1. NowGFP_pH9.0 and NowGFP_pH4.8. The structure of the chromophore and the stereochemistry of its immediate environment are the key factors determining the spectral properties of GFP-like proteins. Post-translational modification of the chromophore-forming sequence Thr65-Trp66-Gly67 in NowGFP results in a coplanar chromophore consisting of the five-membered imidazolinone heterocycle and the aromatic nine-membered indole ring. The C$^{\alpha}$ atom of Thr65 has sp^3 hybridization with the standard *trans* conformation of the preceding peptide bond. The nearest shell of the chromophore (within 3.9 Å) is composed of 17 residues, most of which are involved in an extensive network of hydrogen bonds, including a number of mediating water molecules (Fig. 4). This network is functionally important for protein maturation and fluorescence. In the NowGFP structures at pH 9.0 and pH 4.8, the chromophore and the nearby residues form four and three direct hydrogen bonds and four and four water-mediated hydrogen bonds, respectively.

At pH 9.0, the TWG chromophore adopts the same entirely *cis* conformation as in Cerulean at pH 7.0, with χ_1 and χ_2 torsion angles of the Trp66 side chain (rotation around the C$^{\alpha}$—C$^{\beta}$ and C$^{\beta}$—C$^{\gamma}$ bonds) of −4 and 5°, respectively (Table 3, Fig. 5). The *cis* conformation remains dominant even at pH 4.8 (∼80%) along with a minor *trans* conformation (∼20%), with χ_1/χ_2 angles of 2 and −167°, respectively. The side chain of the chromophore Thr65, at both high and low pH, assumes two alternative conformations with χ_1 angles of

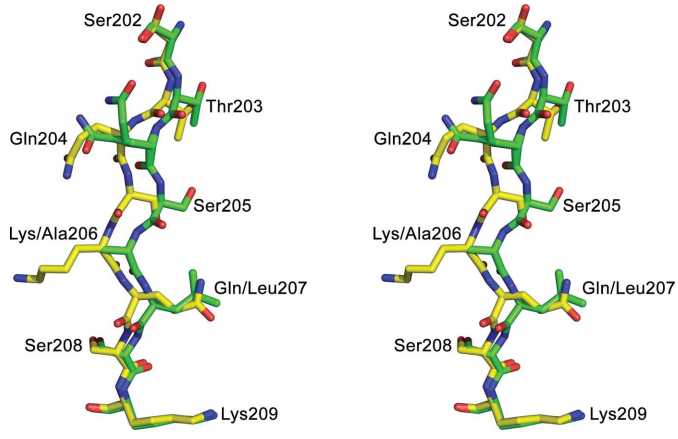


Figure 3
Stereoview of the conformational difference between the β-strand fragment 203–208 in NowGFP (shown in yellow) and the precursor Cerulean (shown in green; sequence identity ∼92%). This figure was created with *PyMOL* (DeLano, 2002).

Table 3
Alternative conformations of side chains of the residues in the chromophore environment.

The alternative values for the χ_1 or χ_1/χ_2 torsion angles are separated by sign, and the percentages of the alternative conformational states are given in parentheses. Cerulean has Val at position 61.

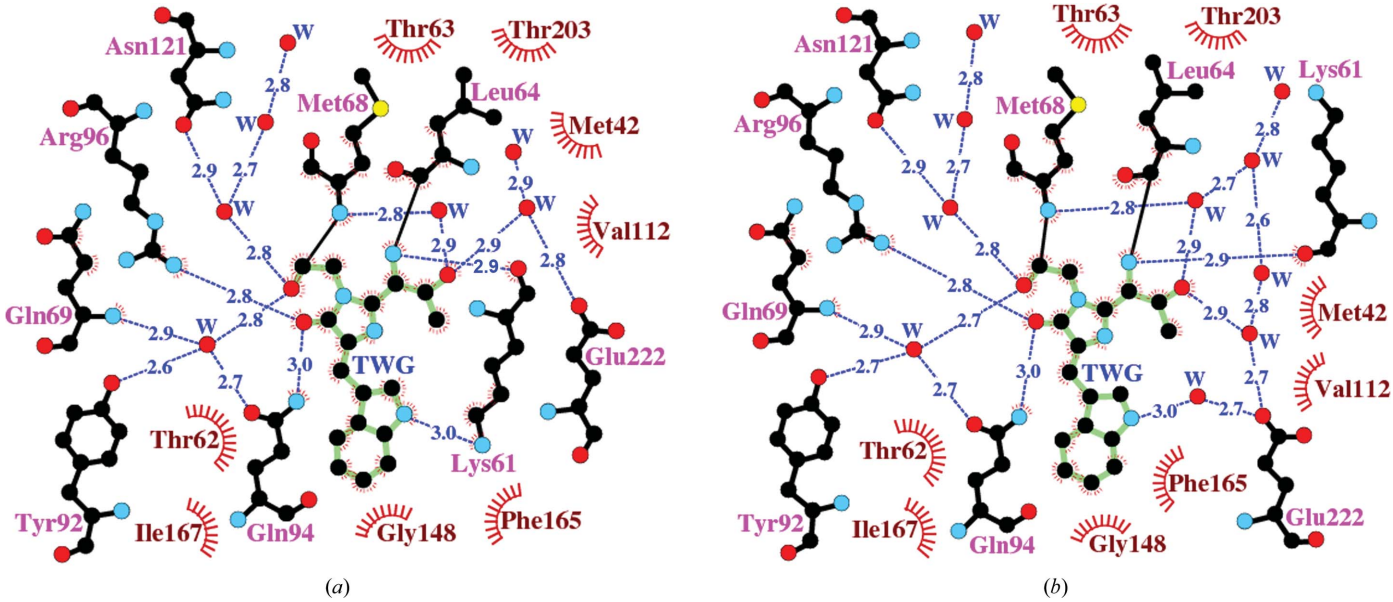
	Torsion angles (°): χ_1 ($C^\alpha-C^\beta$) or χ_1/χ_2 ($C^\alpha-C^\beta/C^\beta-C^\gamma$)			
	NowGFP		Cerulean	
Residue	pH ∼9.0	pH ∼4.8	pH ∼7†	pH ∼5‡
Trp66 (chromophore)	−4/5 (*cis*)	−1/7 (∼80% *cis*) and 2/−167 (∼20% *trans*)	5/−6 (*cis*)	−1/180 (*trans*)
Thr65 (chromophore)	180 (∼70%) and 62 (∼30%)	−179 (∼50%) and 59 (∼50%)	176	58
Lys61	−78/172 (∼20%) and −166/−109 (∼80%)	−81/177 (∼80%) and 166/156 (∼20%)	—	—
Ser205	173	172	−54	−105
Leu220	−70/179	−61/175	−69/−174 and 169/154	−166/58

† Lelimousin *et al.* (2009). ‡ Malo *et al.* (2007).

∼180° and ∼60°. In Cerulean, these conformations were observed individually at pH 7 and pH 5, respectively (Lelimousin *et al.*, 2009; Malo *et al.*, 2007). A number of other residues in the nearest environment of the NowGFP chromophore also demonstrate a certain conformational mobility (Table 3).

The side chain of the chromophore-proximal Lys61 (Val61 in Cerulean) exhibits a clearly pH-dependent conformational mobility (Table 3, Fig. 5). At pH 9.0, Lys61 (pK_a 10.5) adopts the dominant side-chain conformation *k*1 (∼80%), in which its side chain is directed towards the chromophore. It forms two hydrogen bonds connecting the indole N atom of Trp66 and the carboxyl group of the catalytic Glu222 (pK_a 4.3). Glu222, in turn, extends the hydrogen bonding further, forming a direct hydrogen bond to the adjacent Ser205 and a water-mediated hydrogen bond to the nearby Thr203. In the minor conformation *k*2 (∼20%), the side chain of Lys61 is located within a hydrogen-bonding distance from the Gln207 side chain and the Asn144 carbonyl. At pH 4.8 conformation *k*2 becomes dominant (∼80%), sharing the total occupancy with a minor *k*3 conformation (∼20%; Fig. 5*b*). In acidic media, the position occupied by the ε-amino group of Lys61 in the *k*1 conformation (observed at pH ∼9.0) is occupied by a water molecule recreating a similar hydrogen-bond network, Trp66–(Lys61/W)–Glu222, as found at higher pH. The key mutation Val61Lys, which converts Cerulean to green fluorescent variants, affected the conformational states of some other residues in the chromophore area of NowGFP: the Ser205 and Leu220 residues, proximal to Lys61, adopt conformations different from those in Cerulean (Table 3).

3.3.2. Photoconverted NowGFP_conv. At physiological and higher pH, illumination of NowGFP with blue light (450–480 nm) induces an irreversible photoconversion of its green form to the cyan-emitting form with a low fluorescence quantum yield (Fig. 2*c*). This process is accompanied by *cis*–*trans* isomerization of the Trp66 side chain followed by a gradual decrease in the green emission intensity. The general structural features of the chromophore environment in the



(*a*) (*b*)

Figure 4
The nearest (amino-acid) environment of the chromophore in the structures of NowGFP at pH 9.0 (*a*) and pH 4.8 (*b*). Hydrogen bonds (≤3.3 Å) are shown as blue dashed lines, water molecules (W) as red spheres and van der Waals contacts (≤3.9 Å) as orange 'eyelashes' (figure prepared with *LIGPLOT*/*HBPLUS*; McDonald & Thornton, 1994; Wallace *et al.*, 1995)

photoconverted NowGFP_conv (pH 7) remain nearly identical to those in the intact NowGFP (pH 9). The only exception is the dramatic decomposition of the Lys61 side chain, which is located at a hydrogen-bonding distance from both the Trp66 indole N atom and the Glu222 carboxyl O atom (Fig. 5*c*). The electron density in the direction of the ε-amino group is strongly degraded beyond the C^{β} atom, implying possible cleavage of the Lys side chain at the C—C bond. Mass-spectrometric experiments carried out with illuminated and non-illuminated NowGFP revealed a 43 Da (NH_2CH_2CH–) difference in the mass of the respective Lys61-containing peptides. Both of these findings are in good agreement with the hypothesis of the decomposition of the Lys61 side chain in NowGFP_conv.

Illumination carried out in an argon atmosphere revealed the importance of atmospheric oxygen for the photoconversion: a limited oxygen supply significantly decreases the rate of green-to-cyan photoconversion (Fig. 2*e*). Surprisingly, illumination of NowGFP with blue light at low pH does not cause any irreversible supression of the green fluorescence: the illuminated acidic solution restores the green fluorescence on an increase in pH. Thus, in the *k*2 conformation the side chain of Lys61, which is located far away from Trp66 and Glu222, shows no signs of degradation (Fig. 5*b*).

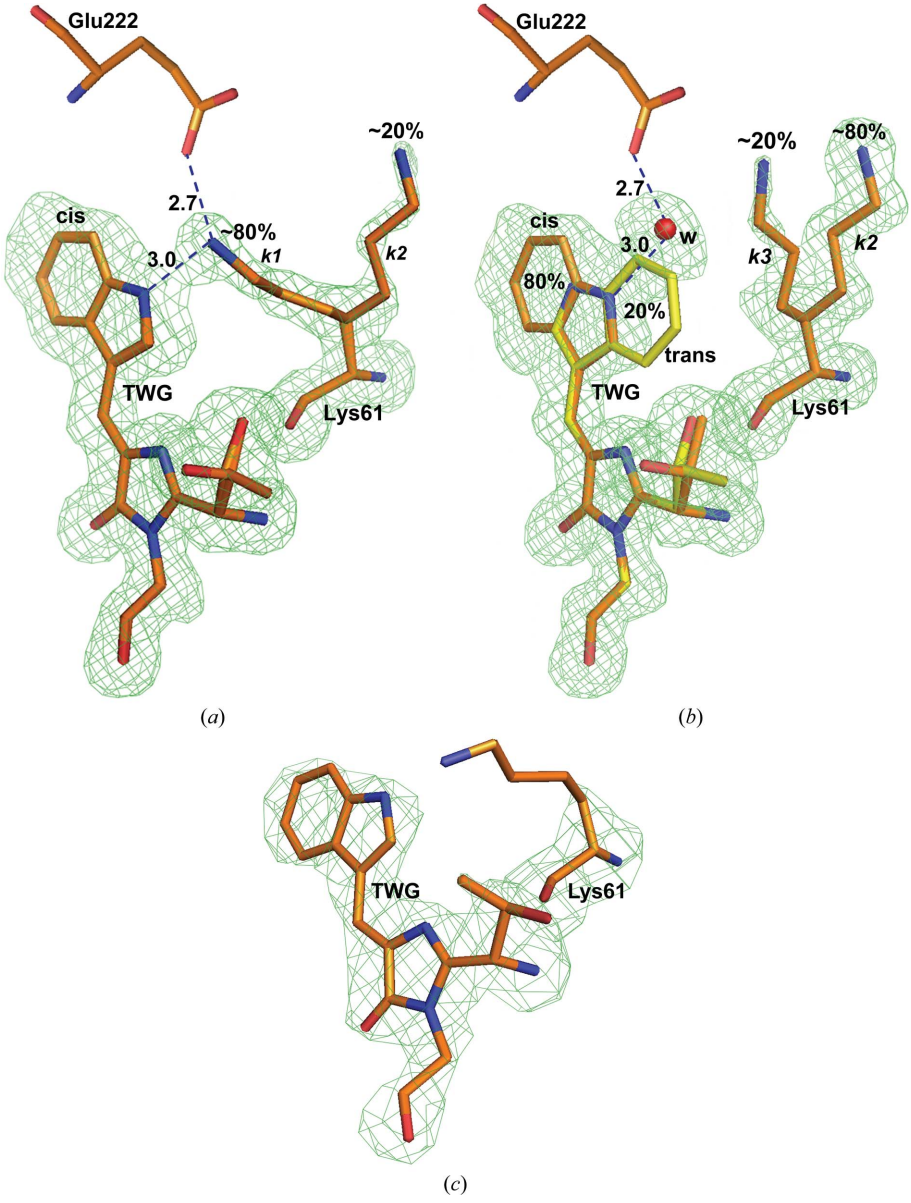


Figure 5
The chromophore Thr65-Trp66-Gly67 in the dominant *cis* state (shown in orange) and the proximal key residue Lys61 in $F_o - F_c$ electron density (cutoff $\rho = 3.0\sigma$) at pH ∼9.0 (*a*), pH ∼4.8 (*b*) and pH ∼7.0 in the photoconverted structure (*c*) (figure created with *PyMOL*; DeLano, 2002). The electron density reveals the alternative conformations of the Thr65 and Lys61 side chains, the presence of *trans*-state contamination of the Trp66 indole ring (shown in yellow; pH ∼4.8) and the decomposition of the Lys side chain after blue-light illumination.

3.4. Structure-based site-directed mutagenesis

The identity of the residues that are potentially responsible for the observed characteristics of NowGFP has been verified by site-directed mutagenesis of individual residues. Residues Glu222, Lys61, Ser205 and Thr203, which form (together with the chromophore Trp66) the local hydrogen-bond network of the chromophore, have been chosen for examination. As expected, all of these sites were found to be important for the green fluorescence of NowGFP. A Thr203Ala replacement drastically shifted the equilibrium between an anionic (green) and a neutral (cyan) chromophore form towards the neutral form, so that the absorption spectrum of the mutant had a major 440 nm cyan band and only a minor 493 nm green band (Fig. 2*f*). The proportion of the 493 nm peak was further reduced in the Ser205Gly, Glu222Gln and Ser205Ala mutants. A Lys61Arg substitution completely suppressed the green fluorescence, highlighting the key role of Lys61 in stabilization of the anionic state of the NowGFP chromophore (Fig. 2*f*). An Ala150Val replacement in the vicinity of the chromophore indole ring resulted in a 2.5-fold increase in the photostability of the mutant, presumably owing to steric hindrance of the light-induced *cis–trans* isomerization of Trp66.

4. Discussion

The green-emitting variant NowGFP with a Trp66-based chromophore origi-

nates from the cyan precursor mCerulean ($\lambda_{ex}/\lambda_{em}$ = 434/474 nm). At physiological and higher pH NowGFP exhibits a bright green fluorescence ($\lambda_{ex}/\lambda_{em}$ = 493/502 nm) and at low pH a weak cyan fluorescence ($\lambda_{ex}/\lambda_{em}$ = 429/475 nm) with remnant green fluorescence (Figs. 2*a* and 2*b*). The brightness of the green fluorescence decreases more than 20-fold on a change in pH from 9 to 5 (Table 2). X-ray study of the parental Cerulean found that the chromophore Trp66 adopts a *cis* conformation at pH 7 (Lelimousin *et al.*, 2009) and a *trans* conformation at pH 5 (Malo *et al.*, 2007), and that these isomers differ by a 180° rotation around the $C^{\beta}-C^{\gamma}$ bond of Trp66. Note that in Cerulean, *cis–trans* isomerization of the chromophore results in a clearly detectable shift of the absorption band from 434/450 nm (*cis*) to 420 nm (*trans*) (Figs. 2*c* and 2*d*). At pH 7, *cis* Trp66 in Cerulean is stabilized by a hydrogen bond between the Trp66 indole N atom and the proximal Ser205. At pH 5, the side chain of Asp148 becomes protonated and changes its orientation from solvent-exposed to inwardly oriented, forming a stabilizing hydrogen bond to the indole N atom of *trans* Trp66. In NowGFP, position 148 occupied by Gly disables the stabilization effect of Asp and decreases the fraction of *trans* Trp66 at lower pH. The crystal structures of NowGFP at pH 9.0 and 4.8 show a dominant *cis* Trp66 state with occupancies of ∼100 and ∼80%, respectively (Fig. 5). However, comparison of the spectrum of NowGFP illuminated with blue light with those of Cerulean, which is known to show pH-dependent *cis–trans* isomerization of Trp66 (Figs. 2*c* and 2*d*), indicates that in acidic solutions the absorption spectra of NowGFP demonstrate a predominantly *trans* Trp66 (Fig. 2*a*), emphasizing a considerable influence of the crystal packing on the conformational mobility of the chromophore. Replacement of Val61 in Cerulean with Lys was shown to be the key mutation affecting the spectral properties of its successor, NowGFP (Figs. 1, 2*a* and 2*b*, Table 2), presumably owing to the pH-dependence of the Lys61 side-chain conformation.

It has been shown that the Trp indole group in the synthetic TWG chromophore is more easily deprotonated (pK_a 12.4) than that of an isolated Trp residue ($pK_a > 20$) (Sarkisyan *et al.*, 2012). Hydrogen bonding of the chromophore Trp66 to the nearby basic residues (for example Lys61 in NowGFP) could decrease its pK_a even further, providing an efficient pathway for proton transfer from the Trp66 indole group to the ε-amino group of Lys61 and then, most likely, to Glu222, enabling deprotonation of the chromophore. The resulting anionic state of the chromophore provides the predominantly green fluorescence of NowGFP at physiological and higher pH. This conclusion was first made based on a biochemical study of WasCFP (Sarkisyan *et al.*, 2012), and the X-ray structures reported here provide evidence supporting this hypothesis. The minor conformation *k*2 (∼20%) of Lys61 observed at pH 9.0 becomes dominant at pH 4.8 (∼80%) and coexists with the minor conformation *k*3 (∼20%; Fig. 5*b*). In acidic media, the position occupied by the ε-amino group of Lys61 in the *k*1 conformation is occupied by a water molecule mediating a hydrogen bond between the Trp66 indole ring and the catalytic Glu222. At pH 4.8, the basicity of Glu222 is greatly reduced and the presence of water at the position of Lys61 sufficiently decreases the driving force for Trp66 deprotonation, enabling only marginal chromophore deprotonation and, as a consequence, a considerable decrease in green fluorescence.

The structure of NowGFP_conv suggests that the irradiation of NowGFP with blue light is accompanied by photodecomposition of the side chain of Lys61. This is in good agreement with the results of mass-spectrometric experiments, which suggest that cleavage of the Lys61 side chain predominantly occurs at the $C^{\gamma}-C^{\delta}$ bond. The absorption spectrum of photoconverted NowGFP suggest that blue light causes the same *cis–trans* isomerization of the Trp66 side chain as pH does in Cerulean (Figs. 2*c* and 2*d*). Light illumination of NowGFP is accompanied by an increase in the cyan fluorescence and a decrease in the green fluorescence, presumably owing to protonation of Trp66 (Figs. 2*c* and 2*d*). After illumination stops the protein exhibits a partial restoration of green fluorescence owing to partial recreation of the native *k*1 state of Lys61 from the intact Lys61 in the equilibrium minor *k*2 state (Figs. 2*d* and 5).

In silico modelling suggests that *cis–trans* isomerization of Trp66 enables efficient cation–π interactions between the ε-amino group of Lys61 and the *trans* Trp66 indole group. According to PDB data, ∼25% of all Trp residues in proteins experience an energetically significant cation–π interactions with an Arg or Lys side chain (Dougherty, 2013; Gallivan & Dougherty, 1999; Meyer *et al.*, 2003; Salonen *et al.*, 2011). We hypothesize that light-induced *cis–trans* isomerization of the Trp66 side chain triggers the photoconversion of NowGFP, whereas the catalytic Glu222 located on the other side of Lys61 promotes its photodecomposition.

The actual mechanism of the photodecomposition reaction remains unknown, but a decrease in its rate with a limited oxygen supply implies that molecular oxygen serves as one of the alternative electron acceptors (Fig. 2*e*). The hydrogen-bond network around the chromophore formed by Lys61, Glu222, Ser205 and Thr203 enables the release of protons and electrons accompanying deprotonation/protonation of Trp66 and photoconversion of NowGFP. Replacement of any of these residues was shown to cause a significant decrease in the green-emitting fraction of NowGFP (Fig. 2*f*). We were able to demonstrate that suppression of the *cis–trans* isomerization of the chromophore enables inhibition of the unwanted irreversible light-induced photoconversion of NowGFP. The decrease in the free volume around the chromophore indole ring achieved by an Ala150Val replacement created steric hindrance for chromophore isomerization, resulting in a 2.5-fold improvement in the photostability of the protein.

In conclusion, we have presented X-ray structures of intact NowGFP at pH 9.0 and pH 4.8 and of its photoconverted variant, NowGFP_conv, representing a sparse family of GFPs with a Trp-based chromophore. It was demonstrated that Lys61, which adopts two distinct pH-dependent conformations, plays a central role in chromophore ionization and the spectral properties of the protein. Another important factor affecting the spectral properties of NowGFP is the integrity of

the hydrogen-bond network comprising Lys61, Glu222, Thr203 and Ser205 and connected to the indole ring of the chromophore. The absorption spectrum of photoconverted NowGFP suggest that illumination with blue light causes *cis–trans* isomerization of the Trp66 side chain that initiates further chemical transformations. The photoconversion of NowGFP has an irreversible character and is accompanied by at least partial decomposition of the Lys61 side chain. Suppression of the light-induced *cis–trans* isomerization of the chromophore was shown to decelerate the unwanted photoconversion, providing an improvement in photostability.

Acknowledgements

This project has been supported by grants from the MCB RAS and the Russian Foundation for Basic Research (14-04-00004 and 13-04-01878-a). This research was partially carried out using the equipment provided by the IBCH core facility (CKP IBCH). Diffraction experiments were carried out on beamlines 22-ID and 22-BM of the Southeast Regional Collaborative Access Team (SER-CAT) located at the Advanced Photon Source, Argonne National Laboratory. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. This project has been supported in part with Federal funds from the National Cancer Institute, National Institutes of Health (NIH) contract No. HHSN261200800001E and by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. (2010). *Physiol. Rev.* **90**, 1103–1163.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Dougherty, D. A. (2013). *Acc. Chem. Res.* **46**, 885–893.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Gallivan, J. P. & Dougherty, D. A. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 9459–9464.
Lam, A. J., St-Pierre, F., Gong, Y., Marshall, J. D., Cranfill, P. J., Baird, M. A., McKeown, M. R., Wiedenmann, J., Davidson, M. W., Schnitzer, M. J., Tsien, R. Y. & Lin, M. Z. (2012). *Nature Methods*, **9**, 1005–1012.
Lelimousin, M., Noirclerc-Savoye, M., Lazareno-Saez, C., Paetzold, B., Le Vot, S., Chazal, R., Macheboeuf, P., Field, M. J., Bourgeois, D. & Royant, A. (2009). *Biochemistry*, **48**, 10038–10046.
Malo, G. D., Pouwels, L. J., Wang, M., Weichsel, A., Montfort, W. R., Rizzo, M. A., Piston, D. W. & Wachter, R. M. (2007). *Biochemistry*, **46**, 9865–9873.
Matsumoto, A. & Itoh, T. Q. (2011). *Biotechniques*, **51**, 55–56.
McDonald, I. K. & Thornton, J. M. (1994). *J. Mol. Biol.* **238**, 777–793.
Meyer, E. A., Castellano, R. K. & Diederich, F. (2003). *Angew. Chem. Int. Ed.* **42**, 1210–1250.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. (2004). *Nature Biotechnol.* **22**, 445–449.
Salonen, L. M., Ellermann, M. & Diederich, F. (2011). *Angew. Chem. Int. Ed.* **50**, 4808–4842.
Sarkisyan, K. S. *et al.* (2015). Submitted.
Sarkisyan, K. S., Yampolsky, I. V., Solntsev, K. M., Lukyanov, S. A., Lukyanov, K. A. & Mishin, A. S. (2012). *Sci. Rep.* **2**, 608.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng. Des. Sel.* **8**, 127–134.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Wu, B., Piatkevich, K. D., Lionnet, T., Singer, R. H. & Verkhusha, V. V. (2011). *Curr. Opin. Cell Biol.* **23**, 310–317.