research papers

Structural analysis of the bright monomeric yellow-green fluorescent protein mNeonGreen obtained by directed evolution

Damien Clavel,[a,b,c,d]‡ Guillaume Gotthard,[e]‡ David von Stetten,[e] Daniele De Sanctis,[e] Hélène Pasquier,[d] Gerard G. Lambert,[f] Nathan C. Shaner[f] and Antoine Royant[a,b,c,e]*

[a]Université Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38044 Grenoble, France, [b]CNRS, IBS, F-38044 Grenoble, France, [c]CEA, IBS, F-38044 Grenoble, France, [d]CNRS, Laboratoire de Chimie Physique, UMR 8000 and Université Paris-Sud, F-91405 Orsay, France, [e]European Synchrotron Radiation Facility, F-38043 Grenoble, France, and [f]Department of Photobiology and Bioimaging, The Scintillon Institute, San Diego, California, USA. *Correspondence e-mail: antoine.royant@ibs.fr

Edited by M. Czjzek, Station Biologique de Roscoff, France

‡ These authors contributed equally to this work.

Until recently, genes coding for homologues of the autofluorescent protein GFP had only been identified in marine organisms from the phyla Cnidaria and Arthropoda. New fluorescent-protein genes have now been found in the phylum Chordata, coding for particularly bright oligomeric fluorescent proteins such as the tetrameric yellow fluorescent protein *lan*YFP from *Branchiostoma lanceolatum*. A successful monomerization attempt led to the development of the bright yellow-green fluorescent protein mNeonGreen. The structures of *lan*YFP and mNeonGreen have been determined and compared in order to rationalize the directed evolution process leading from a bright, tetrameric to a still bright, monomeric fluorescent protein. An unusual discolouration of crystals of mNeonGreen was observed after X-ray data collection, which was investigated using a combination of X-ray crystallography and UV–visible absorption and Raman spectroscopies, revealing the effects of specific radiation damage in the chromophore cavity. It is shown that X-rays rapidly lead to the protonation of the phenolate O atom of the chromophore and to the loss of its planarity at the methylene bridge.

1. Introduction

The discovery and study of green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* led researchers to give it a prime role in live cell biology as a gene reporter (Tsien, 1998). A visible light-absorbing group, the chromophore, auto-catalytically forms during protein folding from three consecutive amino-acid residues (Ser, Tyr and Gly in *Aequorea* GFP) in the core of the protein. The chromophore is fluorescent as a result of the favourable environment provided by the rigid 11-strand β-barrel structure of the protein.

Fluorescent proteins (FPs) homologous to GFP were subsequently found in Anthozoa species such as sea anemones and reef-building corals, which belong to the same phylum, Cnidaria, as jellyfish species (Matz *et al.*, 1999). The vast majority of these homologues form oligomers, with the tetramer as the preferred arrangement in Anthozoa (Alieva *et al.*, 2008). Tetrameric FPs are classically described as four protein chains, *A*, *B*, *C* and *D*, forming two pairs of interaction interfaces: a hydrophilic one between monomers *A* and *B* or *C* and *D* ('*A*–*B* interface'), and a hydrophobic one between monomers *A* and *C* or *B* and *D* ('*A*–*C* interface'). As oligomer-forming gene reporters would result in obvious artefacts in cell-imaging experiments, it is of paramount

importance to engineer monomeric variants of wild-type fluorescent proteins that retain the original fluorescence properties (Kremers *et al.*, 2011).

Recently, homologues of GFP have been identified in the marine organism lancelet, which belongs to the distinct phylum Chordata (Deheyn *et al.*, 2007). Some of these fluorescent proteins display very high values of fluorescence quantum efficiency (quantum yield; QY) and molar extinction coefficient (EC), resulting in particularly high levels of brightness. Notably, nearly all chordate fluorescent proteins have a GYG chromophore tripeptide sequence. Of the 21 GFP-like genes present in the lancelet *Branchiostoma lanceolatum*, the yellow fluorescent protein *bl*FP-Y3 (GenBank accession No. EU482389) attracted particular interest owing to its near-perfect QY (>0.95) and very high EC (150 000 M^{-1} cm^{-1}), and could be successfully evolved into the monomeric green fluorescent protein mNeonGreen, which retains 65% of the brightness of the parent protein (Shaner *et al.*, 2013). Monomerization of *lan*YFP was achieved in a series of rounds of mutagenesis based on rational design and directed evolution. The rational design approach was used to model *lan*YFP based on the known structure of TurboGFP, a solubility-improved variant of the closest primary-sequence homologue ppluGFP2 from *Pontellina plumata* (Evdokimov *et al.*, 2006). In an initial setup to ensure optimal folding and high expression yield, the first and last seven residues of enhanced green fluorescent protein (EGFP) were added to the sequence, with a four-residue linker before the N-terminus of the original protein. The next step consisted of breaking the *A*–*B* interface with the I118K mutation, which negatively affected the fluorescence properties. Addition of the folding mutation N174D, and introduction of the additional mutations A45D, S165N and V171A after several rounds of directed evolution, resulted in a dimeric FP with only slightly decreased QY and EC values (0.90 and 125 000 M^{-1} cm^{-1}, respectively), which was dubbed d*lan*YFP. Finally, 16 mutations were introduced in a series of alternate *A*–*C* interface-perturbation and fluorescence-restoration rounds: V140R, I144T, D156K, T158S, Q168R and F192Y at the *A*–*C* interface and F15I, R25Q, Q56H, F67H, K79V, S100V, F115A, T141S, M143K and I185Y elsewhere. The resulting FP, mNeonGreen, exhibits a small but significant blue shift of its excitation and emission maxima (hence the 'green' colour) and, once more, slightly decreased QY and EC values (0.80 and 116 000 M^{-1} cm^{-1}, respectively). In brief, mNeonGreen is a monomeric version of *lan*YFP, at the expense of a 35% loss in fluorescence, but with a particularly fast maturation speed, a probable gain resulting from the evolution process. Determining and comparing the structures of these proteins should help to pinpoint the various roles of the introduced mutations.

Determining the structure of proteins using high X-ray fluxes at synchrotron beamlines inevitably poses the question of the integrity of the structure, in particular at the active-site location (Weik *et al.*, 2000; Ravelli & McSweeney, 2000; Burmeister, 2000). The so-called 'specific' radiation damage notably affects the chromophore site in coloured proteins such as bacteriorhodopsin (Matsui *et al.*, 2002) or photoactive

yellow protein (Kort *et al.*, 2004), but also that of FPs such as EGFP (Royant & Noirclerc-Savoye, 2011) or IrisFP (Adam *et al.*, 2009). The observed changes include the decarboxylation of a conserved glutamate residue appearing on the dose scale of a classical X-ray data collection (~1 MGy) and the loss of electron conjugation on the chromophore, which appears well below the classical dose scale (~0.1 MGy). These results strongly suggest that a radiation-damage study should follow the structure elucidation of any new FP.

We have crystallized *lan*YFP at near-physiological pH and mNeonGreen at both acidic and near-physiological pH values, and obtained their structures at 2.05, 1.70 and 1.21 Å resolution, respectively. We compared the various structures in order to understand the evolution process of the bright variant mNeonGreen and the structural reasons for the loss of fluorescence at acidic pH. Finally, we performed a radiation-damage study of mNeonGreen using complementary optical spectroscopy.

2. Methods

2.1. Protein expression and purification

Genes coding for *lan*YFP and mNeonGreen were cloned into a pNCST plasmid, a derivative of the constitutive expression plasmid described in Shaner *et al.* (2013). Overexpression of each protein in *Escherichia coli* was performed by growing bacterial cultures at 37°C for 14 h in 2×YT microbial medium. Bacterial cells were pelleted by centrifugation at 5000*g* and 4°C for 20 min and resuspended in a lysis buffer composed of 20 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 20 m*M* imidazole. Cells were disrupted with a high-pressure TS Series cell disruptor (Constant Systems Ltd, Daventry, England) at 7°C and 95 MPa. Cell debris was removed by centrifugation at 43 000*g* and 4°C for 30 min. The clarified lysate was purified by the successive use of nickel-affinity (HisTrap HP 5 ml, GE Healthcare, Little Chalfont, England) and size-exclusion chromatography columns (Superdex 200 10/300 GL, GE Healthcare). The purified protein was concentrated to 20 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0. Prior to crystallization, the purified proteins were submitted to tryptic digestion for 1 h at room temperature at a 1:10 trypsin:protein ratio.

2.2. Protein crystallization

All crystals were obtained using the vapour-diffusion technique at 20°C. *lan*YFP was crystallized in 0.19 m*M* CYMAL-7, 100 m*M* HEPES pH 7.5, 40%(*v*/*v*) PEG 400. The acidic form of mNeonGreen was crystallized as long needles in 6.8 m*M* CYMAL-7, 100 m*M* sodium citrate tribasic dihydrate pH 4.5, 14% PEG 20 000. The near-physiological pH form of mNeonGreen was crystallized as bipyramids in 100 m*M* HEPES pH 8.0, 20% PEG 8000.

2.3. X-ray data collection and reduction

Only the acidic crystalline form of mNeonGreen required cryoprotection before flash-cooling, which was performed by

Table 1
Data collection and processing.

Values in parentheses are for the outer shell.

Protein	*lan*YFP (pH 7.5)	mNeonGreen (pH 4.5)	mNeonGreen (pH 8.0)
PDB code	5ltq	5ltp	5ltr
Data collection			
Wavelength (Å)	0.976	0.976	0.954
Temperature (K)	100	100	100
pH	7.5	4.5	8.0
Space group	$P2_1$	$P2_12_12_1$	$P6_522$
Unit-cell parameters			
a (Å)	78.7	75.9	72.3
b (Å)	197.2	127.6	72.3
c (Å)	115.1	146.9	154.0
α (°)	90.0	90.0	90.0
β (°)	90.5	90.0	90.0
γ (°)	90.0	90.0	120.0
Resolution range† (Å)	49.29–2.05 (2.12–2.05)	46.34–1.70 (1.76–1.70)	48.56–1.21 (1.25–1.21)
Total reflections	835157 (86173)	670077 (65962)	1330789 (131417)
Unique reflections	217268 (21713)	155051 (15390)	72944 (7142)
Multiplicity	3.8 (4.0)	4.3 (4.3)	18.2 (18.4)
Completeness (%)	99.38 (99.55)	98.87 (99.05)	99.86 (99.92)
Mean $I/\sigma(I)$	13.5 (2.5)	19.4 (1.9)	18.8 (2.1)
Wilson B factor (Å^2)	28.5	22.4	12.9
R_{meas}‡	0.088 (0.659)	0.054 (0.844)	0.088 (1.705)
$CC_{1/2}$	0.998 (0.730)	1.000 (0.678)	1.000 (0.729)
Structure refinement			
No. of molecules in the asymmetric unit	16	6	1
R_{work}	0.216 (0.296)	0.181 (0.275)	0.158 (0.316)
R_{free}	0.244 (0.321)	0.205 (0.301)	0.184 (0.346)
No. of atoms			
Protein	27491	10879	1876
Chromophore	304	114	19
Water	553	911	230
B factors (Å^2)			
Protein	38.4	27.5	15.1
Chromophore	34.2	18.8	9.5
Water	27.8	32.5	25.8
R.m.s. deviations			
Bond lengths (Å)	0.006	0.008	0.009
Bond angles (°)	1.29	1.50	1.59

† The resolution cutoff is based on $CC_{1/2}$. ‡ $R_{\mathrm{meas}} = R_{\mathrm{merge}} \times [N/(N-1)]^{1/2}$, where N is the data multiplicity.

the addition of 20%(v/v) glycerol. All data sets were collected on beamline ID29 at the ESRF (de Sanctis *et al.*, 2012). The X-ray damage study on mNeonGreen was also performed on ID29. X-ray diffraction intensities were integrated, scaled and merged using the *XDS* package (Kabsch, 2010). Data-collection and reduction statistics are shown in Table 1.

2.4. Structure phasing and refinement

The structure of acidic mNeonGreen was solved by the molecular-replacement method with *Phaser* (McCoy *et al.*, 2007) using the structure of *lan*RFP, which has 69% sequence identity to mNeonGreen, as a search model (PDB entry 4jf9; Pletnev *et al.*, 2013). The structures of *lan*YFP and near-physiological pH mNeonGreen were both solved by molecular replacement in *Phaser* using the structure of acidic mNeon-Green. Models were manually built in *Coot* (Emsley *et al.*, 2010) and refined with *REFMAC*5 (Murshudov *et al.*, 2011). Structure-refinement statistics are shown in Table 1. The structures of *lan*YFP, acidic mNeonGreen and near-physio-logical pH mNeonGreen have been deposited in the Protein Data Bank with PDB codes 5ltq, 5ltp and 5ltr, respectively.

2.5. *In crystallo* UV–visible absorption spectroscopy

UV–Vis absorption spectra were recorded at the ID29S-Cryobench laboratory of the ESRF (von Stetten *et al.*, 2015) using a high-sensitivity fixed-grating QE65Pro spectrophotometer with a back-thinned CCD detector (Ocean Optics, Dunedin, Florida, USA) and a balanced deuterium–halogen DH2000-BAL light source (Ocean Optics). Spectra were averaged from ten 400 ms acquisitions from ∼50 μm thick crystals flash-cooled at 100 K.

2.6. Online Raman spectroscopy

Online Raman spectroscopy was performed on beamline ID29 as described previously (Bui *et al.*, 2014) using a setup specifically designed for the collection of X-ray and Raman data in an interleaved manner (von Stetten *et al.*, unpublished work). In brief, Raman spectra were recorded using an inVia Raman instrument (Renishaw PLC, Wotton-under-Edge, England) equipped with a near-infrared (785 nm) 300 mW diode laser source. Raman spectra were measured from the X-ray-exposed region of a static mNeonGreen crystal with a composite acquisition time of 10×10 s for the 300–1800 cm^{-1} spectral window. Spectra were corrected for background using the *WiRE* software v.3.4 (Renishaw PLC). X-ray burn cycles were performed in between Raman data sets, but no diffraction data were recorded.

3. Results and discussion

3.1. Structures of *lan*YFP and mNeonGreen at near-physiological pH

We solved the structure of *lan*YFP, modified with the two folding mutations V171A and N174T, at pH 7.5 at 2.05 Å resolution and that of mNeonGreen at pH 8.0 at 1.21 Å resolution. The asymmetric unit of the *lan*YFP crystals contains a 16-mer, which can be described as a nonphysio-logical tetramer of the physiological tetramer (Fig. 1*a*) after comparison of the various interaction surface areas. The eight *A*–*B*-type interfaces have an area ranging from 843 to 872 Å^2 and the *A*–*C*-type interfaces have an area ranging from 1020 to 1133 Å^2, with one outlier at 779 Å in the most agitated tetramer, while the maximal area value for other interfaces is only 346 Å^2, strongly suggesting that they correspond to crystal contacts. As an illustration of the monomeric state of mNeonGreen (yet not as a proof), the asymmetric unit of

mNeonGreen crystals obtained at pH 8.0 contains only one monomer (Fig. 1c), and the two largest interaction surface areas with symmetry-related monomers are 562 and 742 Å^2. A comparison of the structures highlights how lanYFP was successfully mutated into the monomeric yet still bright mNeonGreen.

3.1.1. Mutations at the oligomer-interaction interfaces. Analysis of surface interactions was performed using the PISA server at http://www.ebi.ac.uk/pdbe/pisa/ (Krissinel &

Henrick, 2007). The A–B interface is composed of residues 14, 16, 17, 19, 85, 87, 89, 91, 95, 97, 99, 101, 103, 116, 118, 120–124, 146, 147, 150, 152, 172, 174 and 176, and the A–C interface of residues 137–140, 142, 144, 145, 154, 156, 158, 166, 168, 172, 189, 190, 192, 194–196 and 211–219. Two mutations are located at the A–B interface, with N174T essentially being a folding mutation. Ile118 faces His87, Asn101 and Arg103 at the dimer interface, and its replacement by a longer and positively charged lysine is sufficient to disrupt this interface (Fig. 2a). There are six mutations at the A–C interface, but most of the perturbation is ensured only by the V140R mutation, which destabilizes a large hydrophobic patch formed by six phenylalanines and two valines (Fig. 2b). A different mode of interface destabilization was obtained in the D156K and Q168R mutations, which insert positive charges at a location which in lanYFP accommodates a negatively charged area in front of a positively charged one.

3.1.2. Chromophore environment. Comparison of the chromophore environment in the structures of lanYFP and mNeonGreen allows us to propose the structural rationale for the successful propagation of fluorescence properties from one to the other, yet at the price of a 35% loss of overall brightness and of a 7 nm blue shift of both UV–Vis absorption and fluorescence emission maxima, the latter occurring in the evolution step from dlanYFP to mNeonGreen. The progressive loss of brightness suggests that the oligomerization state of an FP provides some of the rigidity necessary for efficient chromophore fluorescence within the protein scaffold.

There are two significant changes in the immediate vicinity of the chromophore of mNeonGreen relative to lanYFP, one on each side of the conjugated electron cloud. The first one is the localization of a strong spherical electron density at 4.5–4.8 Å from the phenolate ring of the chromophore in lanYFP (Fig. 3a) that we modelled as a chloride ion given the interaction distances
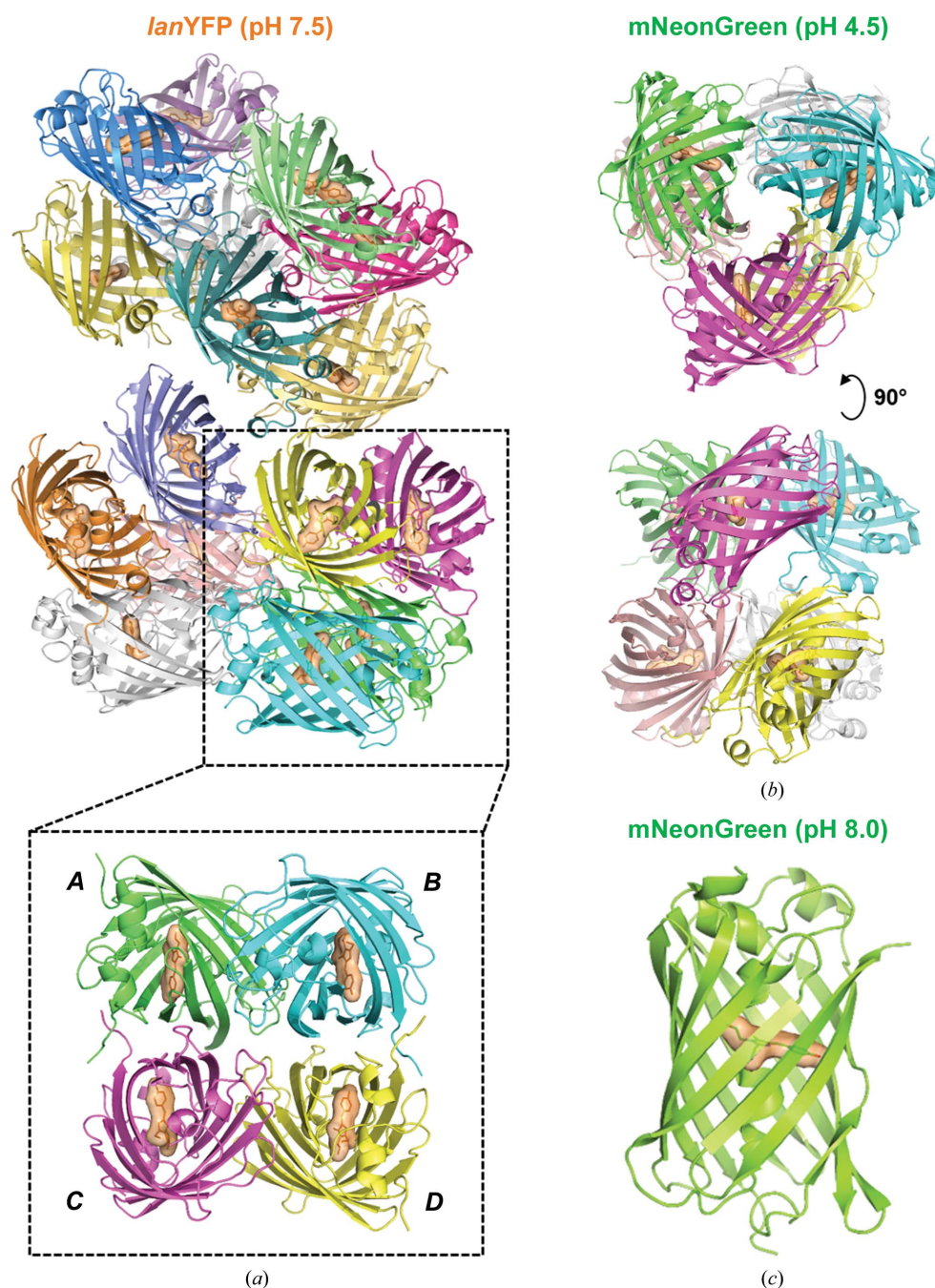
Figure 1
Asymmetric unit in the various crystal forms of lanYFP and mNeonGreen. (a) lanYFP tetramer of tetramers in crystals of space group P2$_1$ obtained at pH 7.5. (b) mNeonGreen dimer of trimers in crystals of space group P2$_1$2$_1$2$_1$ obtained at pH 4.5. (c) mNeonGreen monomer in crystals of space group P6$_5$22 obtained at pH 8.0.

with the six closest neighbouring groups (average distances of 2.9, 2.9, 3.3, 3.4, 3.6 and 4.0 Å for the most resolved tetramer), the fact that one of these groups is a positively charged arginine and the fact that the B factor of the refined chloride compares well with that of the average B factor of interacting atoms (32.8 versus 24.0 Å^2; Carugo, 2014). Given that there is no chloride in the mother-liquor composition, this suggests that lanYFP has a very strong affinity for chloride ions, a property that is also observed in the YFP variant of A. victoria GFP (Wachter & Remington, 1999). The presence of chloride at this location is greatly decreased in mNeonGreen thanks to the unexpected carboxylation of Lys143, which introduces a negative charge preventing the binding of any anion nearby (Fig. 3b). The carboxylation of lysines is a spontaneous post-translational modification proceeding from dissolved carbon dioxide at basic pH, and is believed to be present in up to 1.3% of large proteins (Jimenez-Morales et al., 2014). While it would be tempting to attribute a role to this totally unpredictable modification near the chromophore, it is unlikely to significantly control the main fluorescence properties of mNeonGreen, since the M143K mutation was inserted at a late stage in the evolution process, in particular after the blue shift of the fluorescence excitation and emission properties of the protein had occurred. However, this serendipitous modification seems to have greatly reduced the affinity of mNeonGreen for chloride when compared with lanYFP.

The most significant structural change is found on the other side of the chromophore (Fig. 3c). In lanYFP, Pro55, the residue just one residue before the chromophore in the protein sequence, is engaged in a symmetric (red dashed lines in Fig. 3c) lone pair–π interaction with all five atoms of the imidazolinone ring of the chromophore via its carbonyl O atom. In mNeonGreen, the equivalent carbonyl group is translated to the side of the imidazolinone ring by 1.2 Å, resulting in an asymmetric interaction (cyan dashed lines in Fig. 3c). The sideways change in polar interaction with the delocalized electron cloud must affect the absorption and emission transition dipole moments of the chromophore, effectively resulting in a shift of both the absorption and emission maxima. The origin of this displacement is to be found in a set of three neighbouring mutations that were introduced in order to restore the fluorescence of monomerized
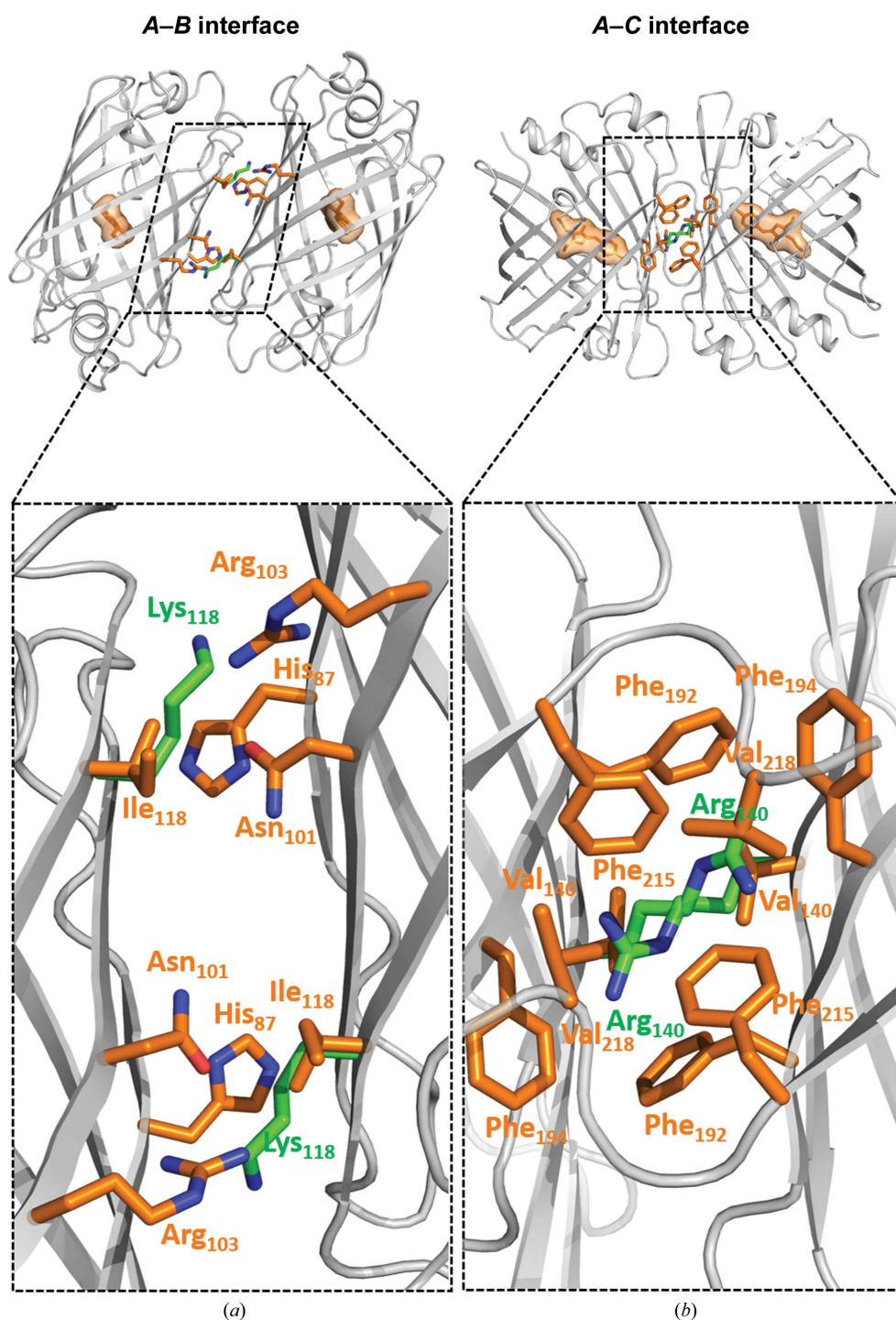
Figure 2
(a) A–B and (b) A–C interfaces in lanYFP (orange side chains) featuring two key mutations ensuring disruption of the interactions in mNeonGreen (green side chains).

d*lan*YFP: Q156H, S100V and F115A. In particular, these last two mutations allow the rotation of the side chain of Tyr102, which removes steric hindrance and leads to the formation of a very short hydrogen bond (2.6 Å) between its phenolate O atom and the carbonyl group of His56, the main and side chains of which are displaced (Fig. 3c) The displacement of Pro55 is made all the more visible by the fact that Trp157, with which it is interacting in a methyl–π interaction, is pushed away along the same direction (Fig. 3c). Most other mutations introduced to restore fluorescence after the first or the second step of interface disruption are located on loops that are distant from the chromophore, and must be considered as folding mutations at this stage.

3.2. Structure of mNeonGreen at acidic pH

We managed to crystallize mNeonGreen at acidic pH (pH 4.5), resulting in crystals with a distinct morphology from those grown at pH 8.0 (Fig. 4a). We were able to solve the structure of the acidic form of mNeonGreen at 1.70 Å resolution, and indeed the space group of this crystal form was

distinct from that we obtained at pH 8.0 ($P2_12_12_1$ versus $P6_522$). The asymmetric unit contains a hexamer, which can be described as two trimers (Fig. 1b), which are most likely to be nonphysiological given the two maximal interaction areas of 682 and 727 Å^2 within the trimer, which are well below the 840–1140 Å^2 seen in *lan*YFP tetramers. Using *in crystallo*

(a)

(b)
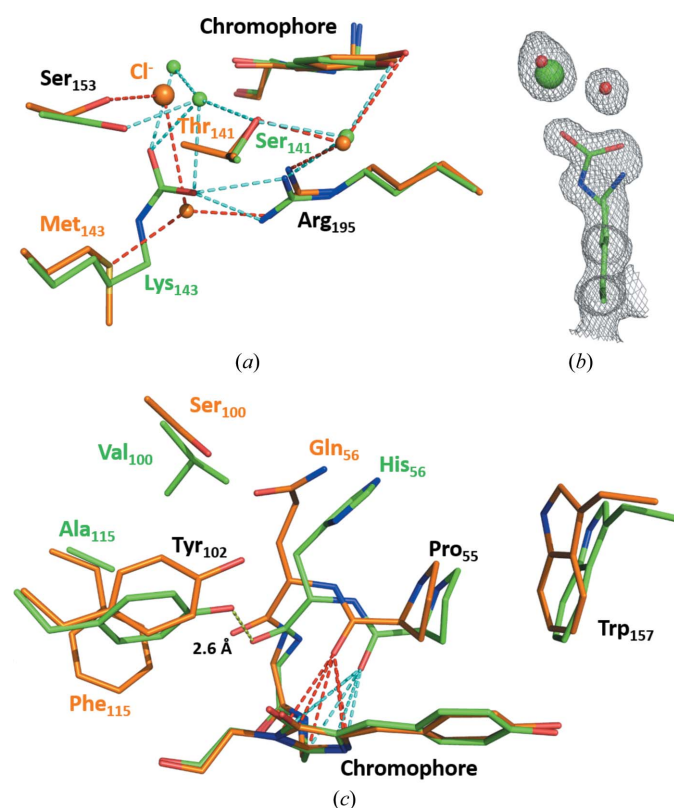
Figure 3
Comparison of the environment of the chromophore in *lan*YFP (orange) and in mNeonGreen (green) at pH 8.0. (a) Close-up of the chloride-binding site in *lan*YFP and the carboxylated lysine in mNeonGreen at physiological pH. (b) σ_A-weighted $2F_o - F_c$ electron-density map contoured at a 1.0σ level around Lys143 in mNeonGreen, the major conformation of which is carboxylated and the minor conformation of which allows the binding of a chloride ion. (c) Close-up of the differences located on the other side of the chromophore. The strong hydrogen bond between Tyr102 and the carbonyl group of His56 in mNeonGreen is represented as a yellow dashed line. The lone pair–π interaction between the carbonyl group of Pro55 and the imidazolinone ring of the chromophore is represented as red dashed lines in *lan*YFP and cyan dashed lines in mNeonGreen
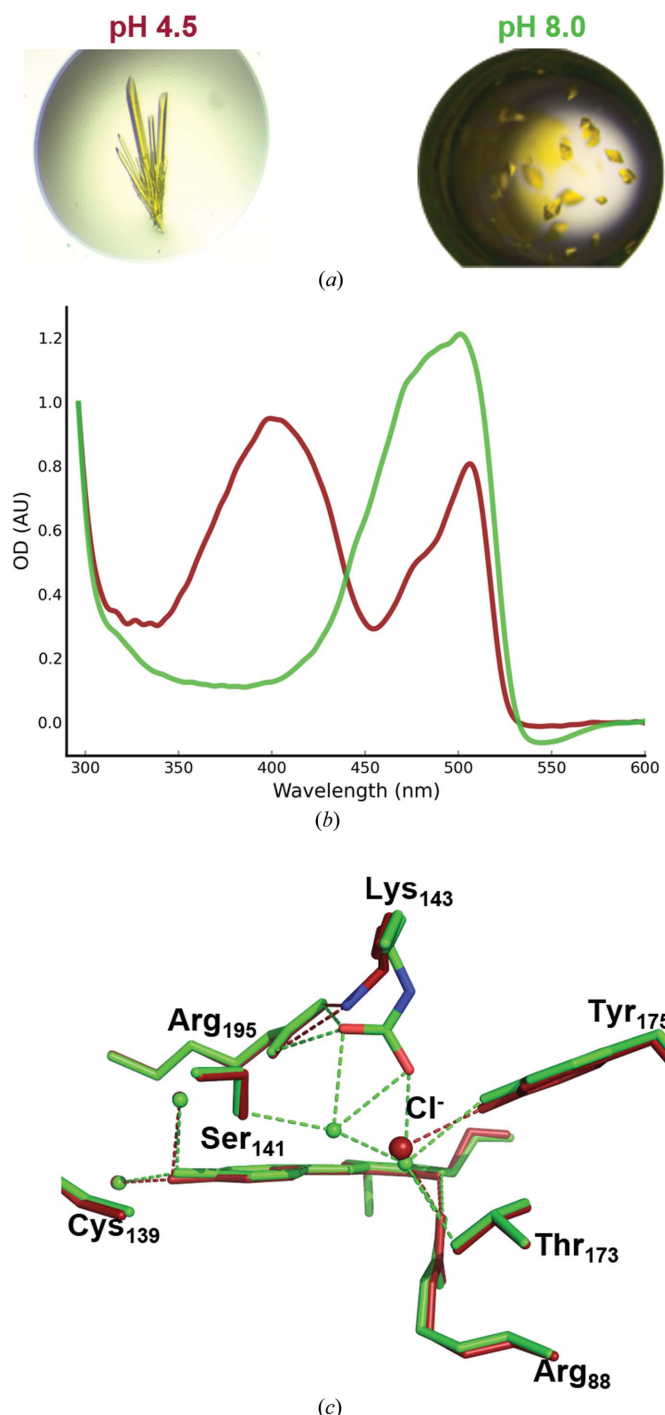

Figure 4
Spectroscopic and structural comparison of mNeonGreen at acidic and near-physiological pH values. (a) mNeonGreen crystal morphologies obtained at pH 4.5 (top) and pH 8.0 (bottom). (b) UV–visible absorption spectra of mNeonGreen crystals at pH 4.5 (dark red) and pH 8.0 (green). (c) Superposition of the chromophore environment in mNeonGreen at pH 4.5 (dark red) and pH 8.0 (green).

UV–visible absorption spectroscopy, we have verified that acidification of the mother liquor leads to an increase in the protonated form of the chromophore, which has an absorption maximum at 403 nm, while the deprotonated form corresponds to the 506 nm peak (Fig. 4b).

Comparison of the chromophore environment in both acidic and near-physiological pH conditions only reveals a change in the negatively charged carboxylated Lys143 (Fig. 4c). The lysine appears unmodified at acidic pH, leaving room for the binding of a chloride ion at 70% occupancy at the same location as seen for lanYFP (100% occupation) and for mNeonGreen (minor occupation of 30%) at pH 8.0. This can be explained by the instability of a carboxylated lysine under acidic conditions, which decarboxylates via a probable protonation mechanism (Golemi et al., 2001). However, since the pK_a of mNeonGreen can be measured at 5.6 or 5.7, whether special care is taken to use chloride-deprived buffers

or not (data not shown), this ion is unlikely to play a role in the protonation mechanism of the chromophore, which thus must relate to the change in hydrogen bonding within the chromophore cavity.

3.3. Specific radiation damage in mNeonGreen

While recording our first data set of an mNeonGreen crystal (near-physiological pH form) on beamline ID29 (de Sanctis et al., 2012), we observed on the MD2 video camera that the crystal volume irradiated by the X-ray beam had turned from a bright yellow-green to a dull orange (Fig. 5a). In order to understand the structural basis of this colour change, we performed a radiation-damage study correlating optical spectroscopy and X-ray crystallography.

3.3.1. Optical spectroscopy analysis. We performed an initial spectroscopic characterization after transporting one irradiated crystal from the goniometer of beamline ID29 to that of the optical spectroscopy facility ID29S-Cryobench (von Stetten et al., 2015). UV–visible absorption spectra could be measured from a non-irradiated area and an area from which an X-ray data set had been recorded, corresponding to a deposited dose of 389 kGy (Fig. 5b). The 0 kGy spectrum shows the expected absorption band between 420 and 520 nm, corresponding to a deprotonated form of the chromophore, yet distorted in shape compared with that in solution, owing to peak saturation resulting from the high optical density of the crystal (~2.0). The 389 kGy spectrum shows the rise of an absorption band centred around 384 nm, which resembles the absorption band of the protonated form of the chromophore observed at acidic pH and thus is likely to be representative of a chromophore that has been protonated upon X-ray irradiation at cryogenic temperature.

In an attempt to further characterize the effects of X-rays on mNeonGreen, we performed an on-line Raman spectroscopy experiment (von Stetten et al., unpublished work) on beamline ID29, in which four Raman spectra were recorded sequentially with increasing doses (Fig. 5c). Each spectrum is represented with a difference spectrum corrected for the initial zero-dose spectrum (dark grey). At 2 kGy, the spectrum has not changed. However, at 17 kGy a few strong peaks have started to decrease, and they are significantly smaller at 208 kGy. Examination of the
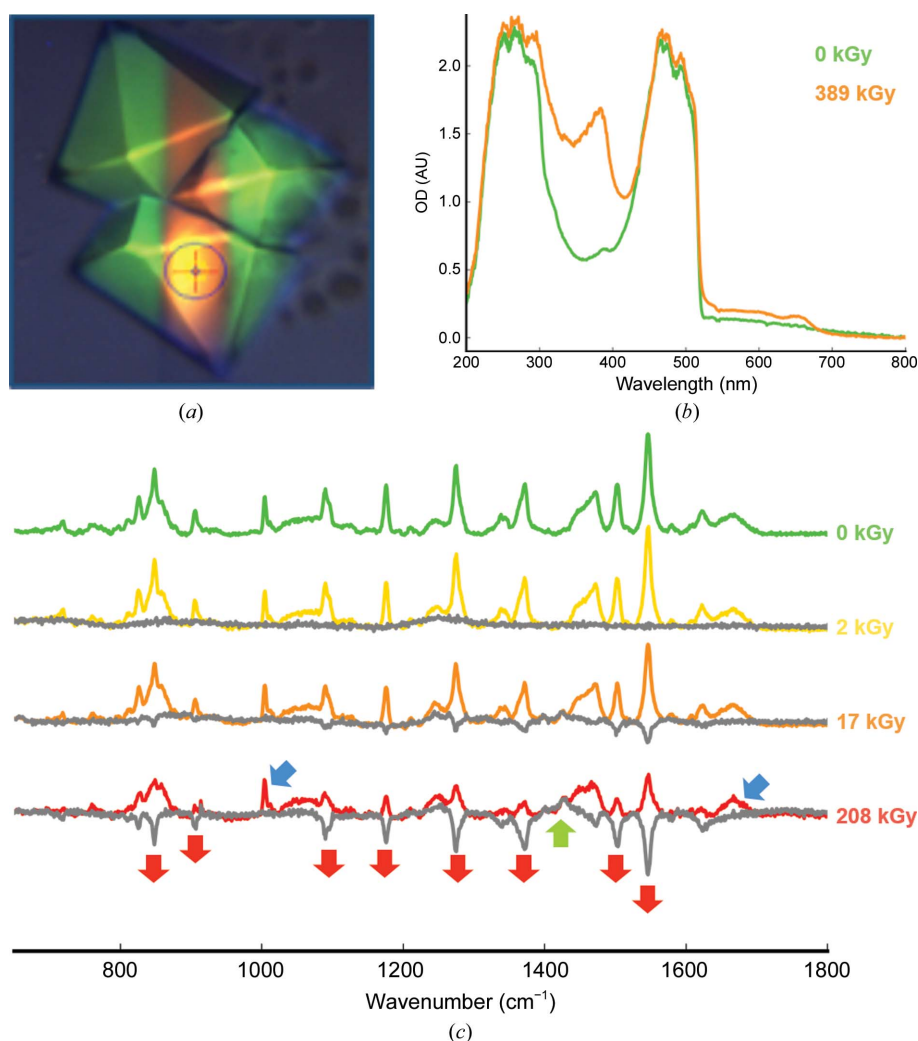
Figure 5
X-ray-induced spectroscopic changes of mNeonGreen. (a) Picture of mNeonGreen crystals after a 730 kGy X-ray data collection. (b) UV–visible absorption spectra of the irradiated (orange trace) and non-irradiated (green trace) areas of an mNeonGreen crystal. (c) Series of online Raman spectra measured on an mNeonGreen crystal with increasing X-ray dose. Grey traces correspond to spectra subtracted for the zero-dose spectrum. Blue arrows indicate invariant peaks characteristics of proteins, red arrows indicate decreasing peaks and the green arrow indicates an increasing peak.

Table 2
Principal Raman modes of mNeonGreen.

Decreasing (−), increasing (+) or invariant (=) modes with increasing X-ray dose are listed.

Mode description	ν (cm^{-1})	Change	Reference
Chromophore fingerprint	848	−	Schellenberg *et al.* (2001)
Chromophore fingerprint	906	−	Schellenberg *et al.* (2001)
Phenylalanine ring	1005	=	von Stetten *et al.* (2015)
Chromophore fingerprint	1090	−	Bell *et al.* (2000)
C—H phenyl bending	1176	−	Bell *et al.* (2000)
Chromophore fingerprint	1275	−	Bell *et al.* (2000)
Phenol C—H deformation	1371	−	He *et al.* (2002)
CH$_2$ scissor motion	1428	+	Bell *et al.* (2000)
Phenol 3, C—C stretch of tyrosine ring	1503	−	Schellenberg *et al.* (2001)
Chromophore exocyclic C═C stretch	1545	−	Schellenberg *et al.* (2001)
Amide I	1667	=	von Stetten *et al.* (2015)

usually strong protein stretching modes at 1005 cm^{-1} (phenylalanine ring band) and 1667 cm^{-1} (amide I band) (blue arrows in Fig. 5c; von Stetten *et al.*, 2015) shows that the

protein is not affected, and also that most of the strong bands correspond to the pre-resonance Raman signal of the chromophore [even though the excitation wavelength (785 nm) is distant from the absorption maximum of the chromophore at 506 nm]. The eight strongest peaks that decrease upon irradiation (red arrows in Fig. 5c) can be assigned to chromophore modes (Table 2). We have identified one band that increases with dose at 1428 cm^{-1} (green arrow in Fig. 5c), which we have assigned to a –CH2 scissoring motion. Together, these results suggest that only the chromophore is affected, that the delocalized π-electron cloud loses resonance with the excitation laser, leading to a decrease in all its modes, and that the only band that increases in intensity could be related to a modification of the methylene bridge.

3.3.2. Structural analysis. A series of 23 X-ray data sets were collected one after another from the same crystal, resulting in a 'dose–lapse' movie of the specific radiation damage occurring within mNeonGreen. Fourier difference maps $(n - 1)$ were calculated between data set n and data set 1, using the structure factors of each data set and the
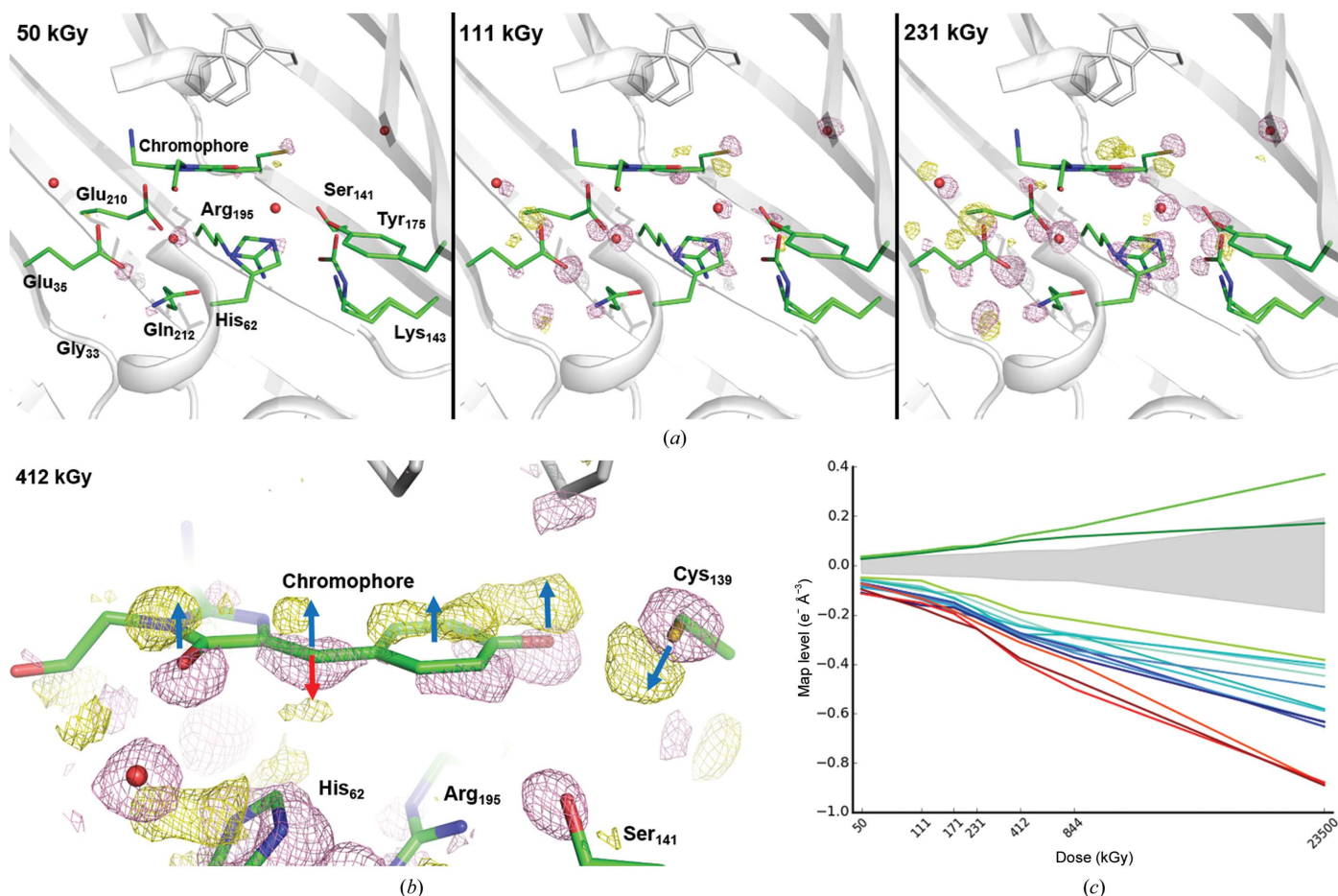
Figure 6
X-ray-induced structural changes in the chromophore cavity of mNeonGreen. (*a*) Fourier difference maps $F_o(n) - F_o(1)$ at increasing X-ray doses contoured at a 6.0σ level (pink, negative; yellow, positive). (*b*) Close-up on the chromophore at a 4.0σ level after 412 kGy. Dark blue arrows indicate the upward movement of the whole chromophore and the opposite movement of Cys139. The red arrow indicates the concomitant formation of a kinked configuration of the chromophore upon loss of conjugation at the exocyclic linkage. (*c*) X-ray-induced electron loss or gain displayed as a function of X-ray dose on a logarithmic scale. Three ensembles of affected groups can be distinguished: damage to Glu35, Glu210 and the water in between is represented in red, damage to the methylene bridge in green and strong damage to other groups listed in Table 3 in blue. The noise level (corresponding to peak heights with a σ level between −3.0 and +3.0) at each dose is represented by a light grey shade.

Table 3
X-ray-induced electron loss (or gain), expressed in $e^- \text{ Å}^{-3}$, in mNeonGreen at 100 K.

Dose Fourier difference map	50 kGy $(F_{o2} - F_{o1})$	111 kGy $(F_{o3} - F_{o1})$	171 kGy $(F_{o4} - F_{o1})$	231 kGy $(F_{o5} - F_{o1})$	412 kGy $(F_{o8} - F_{o1})$	844 kGy $(F_{o13} - F_{o1})$	23 500 kGy $(F_{o23} - F_{o1})$
1.0σ level	0.011	0.013	0.015	0.016	0.020	0.021	0.064
12 strongest negative peaks							
Glu35	−0.097	−0.174	−0.225	−0.256	−0.377	−0.464	−0.891
Wat(Glu35–Glu210–His62)	−0.114	−0.152	−0.195	−0.255	−0.389	−0.500	−0.879
Glu210	−0.074	−0.134	−0.184	−0.231	−0.312	−0.392	−0.885
Cys139	−0.110	−0.162	−0.171	−0.211	−0.294	−0.373	−0.634
Wat(Thr173–Ser153)	−0.071	−0.139	−0.190	−0.208	−0.286	−0.344	−0.492
His62	−0.084	−0.128	−0.161	−0.203	−0.281	−0.354	−0.634
Gly33	−0.071	−0.122	−0.151	−0.198	−0.275	−0.333	−0.653
Gln212	−0.089	−0.120	−0.150	−0.194	−0.248	−0.285	−0.402
Chromophore (phenolate)	−0.077	−0.134	−0.165	−0.194	−0.275	−0.325	−0.589
Wat(chromophore phenolate–Tyr175)	−0.060	−0.095	−0.131	−0.176	−0.280	−0.281	−0.582
Lys143	−0.055	−0.086	−0.136	−0.173	−0.242	−0.279	−0.415
Chromophore (imidazolinone)	−0.057	−0.103	−0.111	−0.148	−0.219	−0.294	−0.447
Peaks at the methylene-bridge location							
Negative	−0.049	−0.062	−0.108	−0.124	−0.188	−0.221	−0.383
Positive (sp^2 configuration)	0.035	0.058	0.076	0.079	0.120	0.154	0.370
Positive (sp^3 kinked configuration)	0.026	0.050	0.066	0.076	0.099	0.117	0.171

calculated phases of a mNeonGreen structure refined against data set 1. The three maps $(2 − 1)$, $(3 − 1)$ and $(5 − 1)$ are represented in Fig. 6(a), corresponding to accumulated doses of 50, 111 and 231 kGy, respectively. The height of the peaks located on the 12 most-affected groups of the protein (including water molecules) are listed in Table 3 and plotted in Fig. 6(c). The timing of the appearance and the various strengths of the peaks allow us to propose a scenario for the structural description of specific radiation damage occurring to mNeonGreen at early X-ray doses. All peaks are found on one side of the chromophore-binding cavity (Fig. 6a). The initial event appears to be the decarboxylation (negative peak in the maps) of a glutamate residue close to the chromophore, Glu35, which is different from that classically observed to be affected in other UV-damaged or X-ray-damaged fluorescent proteins (van Thor *et al.*, 2002; Royant & Noirclerc-Savoye, 2011; Adam *et al.*, 2009). The next residue affected is actually Glu210, the very residue that was damaged in the previous studies, which appears to be decarboxylated as well, with a negative peak of similar magnitude. However, the number of positive peaks around these two Glu residues suggests that each of the two side chains may reorientate. The most likely scenario is that one of the two is decarboxylated (with a slight preference for Glu35 given that its negative peak is consistently higher in all maps; Table 3), which triggers the reorientation of the other one. The ensuing events are the translation of the whole chromophore in one direction (dark blue arrows) and the movement of Cys139 in the opposite direction, maintaining an interaction with the phenolate O atom (Fig. 6b). These changes are accompanied by the displacement of water molecules and the reorientation of His62, while Arg195 and Lys143 move as well. This results in a drastic change in hydrogen bonding, which leads to the protonation of the chromophore, as observed by UV–Vis absorption spectroscopy (Fig. 5b). However, the loss of resonance is not solely explained by the translation of the whole chromophore and its ensuing protonation. Indeed, further examination of the difference map around the chromophore

shows that there is a second positive peak of similar magnitude next to the exocyclic link (Fig. 6b). This is compatible with the $sp^2 \rightarrow sp^3$ dehybridization of the central C atom that had been transiently observed during the low-dose X-ray bleaching of the fluorescent protein IrisFP (Adam *et al.*, 2009). It is noteworthy that the peak heights of both positive peaks are similar at low X-ray doses and then progressively differ, with the peak height associated with the dehybridized chromophore falling below the $\pm 3.0\sigma$ noise level at high doses, in line with the observation that the dehybridization is only transient, thanks to an X-ray-induced reduction/repair mechanism (Adam *et al.*, 2009; Carpentier *et al.*, 2010).

Cumulatively, these observations lead us to propose that X-rays induce fluorescence loss of the chromophore resulting, on one hand, from the translation of the chromophore and its subsequent protonation leading to a blue shift in the absorption maximum which negatively affects the pre-resonance Raman effect and, on the other, from a loss of conjugation at the methylene bridge, separating the two resonant electron clouds of the phenolate and imidazolinone rings and leading to a loss of absorbance at the absorption maximum.

4. Conclusion

Our structural comparison of *lan*YFP and mNeonGreen confirms the rationale used for the successful evolution process using a combination of rational design and directed evolution. It highlights the key mutations resulting in oligomer interface disruption. It also explains how some of the mutations introduced to help restore fluorescence lead to a concerted rearrangement at the back of the chromophore, resulting in a more constrained environment favouring fluorescence. This rearrangement also causes a displacement of the key interaction between a proline residue and the imidazolinone ring of the chromophore, which must explain the small but significant blue shift of the fluorescence excitation and emission maxima of mNeonGreen when compared with *lan*YFP.

This correlative X-ray crystallography, UV–Vis absorption and Raman spectroscopy radiation-damage study leads us to interpret the colour change of mNeonGreen crystals along the path of the X-ray beam as an X-ray-induced loss of fluorescence. Compared with other fluorescent proteins, mNeonGreen presents the particularity of having two decarboxylated glutamate residues next to the chromophore instead of one. The colour change originates from the subsequent protonation of the chromophore phenolate group, which leads to a large blue shift of the absorption maximum into the UV region. It also results from the concomitant loss of conjugation of the chromophore at the methylene bridge, effectively separating the two resonant electron clouds of the phenolate and imidazolinone rings, which also leads to a loss of absorbance in the visible, but in this case without the appearance of a distinct absorption band.

References

Adam, V., Carpentier, P., Violot, S., Lelimousin, M., Darnault, C., Nienhaus, G. U. & Bourgeois, D. (2009). *J. Am. Chem. Soc.* **131**, 18063–18065.

Alieva, N. O., Konzen, K. A., Field, S. F., Meleshkevitch, E. A., Hunt, M. E., Beltran-Ramirez, V., Miller, D. J., Wiedenmann, J., Salih, A. & Matz, M. V. (2008). *PLoS One*, **3**, e2680.

Bell, A. F., He, X., Wachter, R. M. & Tonge, P. J. (2000). *Biochemistry*, **39**, 4423–4431.

Bui, S., von Stetten, D., Jambrina, P. G., Prangé, T., Colloc'h, N., de Sanctis, D., Royant, A., Rosta, E. & Steiner, R A. (2014). *Angew. Chem. Int. Ed. Engl.* **53**, 13710–13714.

Burmeister, W. P. (2000). *Acta Cryst.* D**56**, 328–341.

Carpentier, P., Royant, A., Weik, M. & Bourgeois, D. (2010). *Structure*, **18**, 1410–1419.

Carugo, O. (2014). *BMC Struct. Biol.* **14**, 19.

Deheyn, D. D., Kubokawa, K., McCarthy, J. K., Murakami, A., Porrachia, M., Rouse, G. W. & Holland, N. D. (2007). *Biol. Bull.* **213**, 95–100.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evdokimov, A. G., Pokross, M. E., Egorov, N. S., Zaraisky, A. G., Yampolsky, I. V., Merzlyak, E. M., Shkoporov, A. N., Sander, I., Lukyanov, K. A. & Chudakov, D. M. (2006). *EMBO Rep.* **7**, 1006–1012.

Golemi, D., Maveyraud, L., Vakulenko, S., Samama, J.-P. & Mobashery, S. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 14280–14285.

He, X., Bell, A. F. & Tonge, P. J. (2002). *J. Phys. Chem. B*, **106**, 6056–6066.

Jimenez-Morales, D., Adamian, L., Shi, D. & Liang, J. (2014). *Acta Cryst.* D**70**, 48–57.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kort, R., Hellingwerf, K. J. & Ravelli, R. B. G. (2004). *J. Biol. Chem.* **279**, 26417–26424.

Kremers, G., Gilbert, S. G., Cranfill, P. J., Davidson, M. W. & Piston, D. W. (2011). *J. Cell Sci.* **124**, 157–160.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Matsui, Y., Sakai, K., Murakami, M., Shiro, Y., Adachi, S.-I., Okumura, H. & Kouyama, T. (2002). *J. Mol. Biol.* **324**, 469–481.

Matz, M. V., Fradkov, A. F., Labas, Y., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L. & Lukyanov, S. A. (1999). *Nature Biotechnol.* **17**, 969–973.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

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.

Pletnev, V. Z., Pletneva, N. V., Lukyanov, K. A., Souslova, E. A., Fradkov, A. F., Chudakov, D. M., Chepurnykh, T., Yampolsky, I. V., Wlodawer, A., Dauter, Z. & Pletnev, S. (2013). *Acta Cryst.* D**69**, 1850–1860.

Ravelli, R. B. G. & McSweeney, S. M. (2000). *Structure*, **8**, 315–328.

Royant, A. & Noirclerc-Savoye, M. (2011). *J. Struct. Biol.* **174**, 385–390.

Sanctis, D. de *et al.* (2012). *J. Synchrotron Rad.* **19**, 455–461.

Schellenberg, P., Johnson, E., Esposito, A. P., Reid, P. J. & Parson, W. W. (2001). *J. Phys. Chem. B*, **105**, 5316–5322.

Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M., Davidson, M. W. & Wang, J. (2013). *Nature Methods*, **10**, 407–409.

Stetten, D. von, Giraud, T., Carpentier, P., Sever, F., Terrien, M., Dobias, F., Juers, D. H., Flot, D., Mueller-Dieckmann, C., Leonard, G. A., de Sanctis, D. & Royant, A. (2015). *Acta Cryst.* D**71**, 15–26.

Thor, J. J. van, Gensch, T., Hellingwerf, K. J. & Johnson, L. N. (2002). *Nature Struct. Biol.* **9**, 37–41.

Tsien, R. Y. (1998). *Annu. Rev. Biochem.* **67**, 509–544.

Wachter, R. M. & Remington, S. J. (1999). *Curr. Biol.* **9**, R628–R629.

Weik, M., Ravelli, R. B. G., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.