



Comparison of two crystal polymorphs of NowGFP reveals a new conformational state trapped by crystal packing

Jin Kyun Kim,^{a,†§} Hannah Jeong,^{a,†} Jeongwoo Seo,^{a,†} Seoyoon Kim,^b Kyung Hyun Kim,^c Duyoung Min^{b*} and Chae Un Kim^{a*}

Received 28 June 2024

Accepted 21 August 2024

Edited by B. Kobe, University of Queensland, Australia

† These authors made equal contributions.

§ Current address: Frontier of Extreme Physics, Korea Research Institute of Standards and Science, Daejeon 34113, Republic of Korea.

Keywords: fluorescent proteins; NowGFP; crystal polymorphism; protein crystal packing; conformational flexibility; protein crystallization.

PDB references: NowGFP, monoclinic form obtained at pH 4.8, 8xh0; orthorhombic form obtained at pH 9.0, 8xh1; orthorhombic form obtained at pH 6.0, 8xh2

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

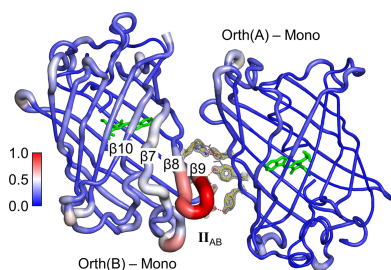
^aDepartment of Physics, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea, ^bDepartment of Chemistry, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea, and ^cDepartment of Biotechnology and Bioinformatics, Korea University, Sejong 30019, Republic of Korea. *Correspondence e-mail: dymin@unist.ac.kr, cukim@unist.ac.kr

Crystal polymorphism serves as a strategy to study the conformational flexibility of proteins. However, the relationship between protein crystal packing and protein conformation often remains elusive. In this study, two distinct crystal forms of a green fluorescent protein variant, NowGFP, are compared: a previously identified monoclinic form (space group $C2$) and a newly discovered orthorhombic form (space group $P2_12_12_1$). Comparative analysis reveals that both crystal forms exhibit nearly identical linear assemblies of NowGFP molecules interconnected through similar crystal contacts. However, a notable difference lies in the stacking of these assemblies: parallel in the monoclinic form and perpendicular in the orthorhombic form. This distinct mode of stacking leads to different crystal contacts and induces structural alteration in one of the two molecules within the asymmetric unit of the orthorhombic crystal form. This new conformational state captured by orthorhombic crystal packing exhibits two unique features: a conformational shift of the β -barrel scaffold and a restriction of pH-dependent shifts of the key residue Lys61, which is crucial for the pH-dependent spectral shift of this protein. These findings demonstrate a clear connection between crystal packing and alternative conformational states of proteins, providing insights into how structural variations influence the function of fluorescent proteins.

1. Introduction

Proteins can be crystallized in multiple forms, a phenomenon known as crystal polymorphism. A well-known example of protein crystal polymorphism is lysozyme, which exhibits six crystal polymorphs (Vaney *et al.*, 2001). The primary interest in discovering crystal polymorphs of proteins lies in finding a superior crystal form that enables the acquisition of detailed structural information through X-ray crystallography. This crystal form may either have high crystallinity for improved diffraction (Yamada *et al.*, 2017) or possess enhanced thermal and chemical durability (Gerlits *et al.*, 2019). Additionally, different crystal forms can reveal different conformational states of protein molecules which are trapped by crystal lattice packing (Jiang *et al.*, 2013). The availability of different crystal forms provides insight into the conformational flexibility of protein molecules and a detailed understanding of individual conformational states (Zhang *et al.*, 1995). In this study, we demonstrate a new case of protein crystal polymorphism with NowGFP, a variant of green fluorescent protein (GFP).

Fluorescent proteins have become essential tools in cell biology and biomedicine, serving as non-invasive methods for visualizing and tracking cellular and organism-wide processes



(Chudakov *et al.*, 2010). Most of these fluorescent proteins are characterized by the tyrosine (Tyr) at the centre of the three residues responsible for forming the chromophore, resulting in fluorescence that spans from green to far-red wavelengths. Recently, a fluorescent protein based on an anionic tryptophan (Trp) was developed and was named WasCFP (Sarkisyan *et al.*, 2012). Derived from the cyan-fluorescent mCerulean (Rizzo *et al.*, 2004), it was engineered by introducing a critical Val61Lys substitution and four additional mutations. WasCFP was further improved to enhance its stability across a broad range of pH levels and temperatures, incorporating 13 additional mutations. This improved variant, which exhibits strong green fluorescence under physiological conditions, has been named NowGFP (Sarkisyan *et al.*, 2015).

Spectral and structural studies of NowGFP consistently reveal pH-dependent changes. Spectrally, NowGFP exhibits green fluorescence ($\lambda_{\text{ex}}/\lambda_{\text{em}} = 493/502$ nm) at physiological and higher pH levels, contrasting with cyan fluorescence ($\lambda_{\text{ex}}/\lambda_{\text{em}} = 429/475$ nm) under acidic conditions (pH < 6). The intensity ratio between green and cyan fluorescence is highly sensitive to pH conditions. Structurally, NowGFP adopts two distinct pH-dependent conformations at pH 4.8 and pH 9.0. These conformations involve the chromophore and the key residue Lys61, which is believed to play a central role in chromophore ionization and the resultant shifts in the fluorescence spectrum (Pletnev *et al.*, 2015).

We present three crystal structures of NowGFP obtained under various pH conditions from two distinct crystal forms: a monoclinic crystal structure at pH 4.8 (PDB entry 8xh0, 1.45 Å resolution) and orthorhombic crystal structures at pH 9.0 (PDB entry 8xh1, 1.7 Å resolution) and pH 6.0 (PDB entry 8xh2, 1.8 Å resolution). Our study begins with a detailed comparative analysis of crystal contacts and stacking interactions between these two crystal forms. Following this, we identify major differences in the crystal contacts across the two crystal forms and determine how the crystal packing may alter the NowGFP structure. Lastly, we investigate the alternative conformations trapped by the orthorhombic crystal form by examining five NowGFP molecules: one from the monoclinic form at pH 4.8, two from the orthorhombic form at pH 9.0 and two from the orthorhombic form at pH 6.0. This comparison aims to deepen our understanding of how crystal packing influences the pH-dependent conformational changes in NowGFP.

2. Materials and methods

2.1. Macromolecule production

The fragment encoding NowGFP with an N-terminal His tag and TEV protease recognition sequence site was cloned into the pET-24a(+) vector and transformed into *Escherichia coli* strain BL21(DE3) (Invitrogen, USA). Bacterial cultures were grown overnight at 16°C. Isopropyl β -D-1-thiogalactopyranoside (IPTG) induction was necessary for effective protein expression. The cells were pelleted by centrifugation, resuspended in phosphate-buffered saline (PBS) pH 7.4,

Table 1

Macromolecule-production information for NowGFP.

DNA source	Synthetic DNA
Forward primer	T7 promoter
Reverse primer	T7 terminator
Expression vector	pET-24a(+)
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	SVSKGKELFTGVVPILEVELDGDVNGHKF SVSGEGEGDATY GKMSLKFICTTGKLP VPWPTLKTTLTWGMQCFARYPDHMKQH DFFKSAMPEGYVQERTIFFKDDGNYKT RAEVKFEEDTLVNRIELKGVDFKEDGN ILGHKLEYNAISGNANITADKQKNGIK AYFTIRHDVEDGSVLLADHYQQNTPIG DGPVLLPDNHYLSTQSKQSKDPNEKRD HMVLLFVTAAGIPLGADELYK

1 mM PMSF, 1 mM TCEP and lysed by an EmulsiFlex-C3 at 103–117 MPa. NowGFP was purified by immobilized metal-ion affinity chromatography using His60 Ni Superflow Resin (Clontech, USA) and then buffer-exchanged for cleavage by TEV protease on a Cytiva PD-10 DG column (1× TEV protease buffer: 25 mM Tris-HCl, 150 mM NaCl, 1 mM TCEP). TEV protease was applied (10 U μl^{-1} , 1:100 ratio) and incubated at 30°C for 1 h. After incubation, further purification was achieved by size-exclusion chromatography using FPLC (ÄKTApure 25) with a HiLoad 16/600 Superdex 75 pg column (Cytiva, USA). The purity of the sample was then confirmed by SDS-PAGE analysis (Supplementary Fig. S1). The purified NowGFP protein was concentrated using an Amicon Ultra-15 centrifugal filter unit. The amino-acid sequence and other details are given in Table 1.

2.2. Crystallization

Crystallization screening was conducted to identify optimal conditions for the formation of NowGFP crystals. Single crystals in a monoclinic space group ($C2$) were obtained using KH_2PO_4 and PEG 3350. Additionally, single crystals in an orthorhombic space group ($P2_12_12_1$) were obtained using sodium citrate and PEG 4000. Details of the crystallization screening, including other conditions that did not yield diffraction-quality single crystals, are provided in Supplementary Table S1.

To obtain monoclinic crystals, NowGFP was transferred to 20 mM Tris pH 8.0, 200 mM NaCl buffer and concentrated to 12 mg ml^{-1} . Crystals suitable for data collection were obtained by the hanging-drop vapour-diffusion method (McPherson, 1982). Typically, 2 μl protein solution was mixed with an equal amount of reservoir solution and incubated at 4°C for a week. The best crystal was obtained from 16 mM KH_2PO_4 pH 4.8, 20% (w/v) PEG 3350.

To obtain orthorhombic crystals, NowGFP was transferred to 20 mM Tris pH 8.0, 200 mM NaCl buffer and concentrated to 17 mg ml^{-1} . 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 4°C for a week. The best crystal was obtained from 100 mM sodium citrate pH 6.0, 25% (w/v) PEG 4000. To obtain structures of NowGFP at

Table 2

Crystallization conditions for the two crystal forms of NowGFP.

Structure	Monoclinic pH 4.8	Orthorhombic pH 9.0	Orthorhombic pH 6.0
PDB code	8xh0	8xh1	8xh2
Method	Hanging drop	Hanging drop	Hanging drop
Plate type	24-well protein-crystallization plate	24-well protein-crystallization plate	24-well protein-crystallization plate
Temperature (K)	277	277	277
Protein concentration (mg ml ⁻¹)	12.1	17.4	17.4
Buffer composition of protein solution	20 mM Tris pH 8.0, 200 mM NaCl	20 mM Tris pH 8.0, 200 mM NaCl	20 mM Tris pH 8.0, 200 mM NaCl
Composition of reservoir solution	16 mM KH ₂ PO ₄ pH 4.8, 20%(w/v) PEG 3350	100 mM sodium citrate pH 6.0, 25%(w/v) PEG 4000	100 mM sodium citrate pH 6.0, 25%(w/v) PEG 4000
Composition of soaking solution†	—	100 mM ammonium sulfate, 50 mM Tris–HCl pH 9.0, 17.5%(w/v) PEG 4000	—
Volume and ratio of drop	2 µl:2 µl	2 µl:2 µl	2 µl:2 µl
Volume of reservoir (µl)	500	500	500

† The soaking solution is used to adjust crystals to different pH conditions.

Table 3

Data-collection statistics for NowGFP crystals.

Values in parentheses are for the outer shell.

Structure	Monoclinic pH 4.8	Orthorhombic pH 9.0	Orthorhombic pH 6.0
PDB code	8xh0	8xh1	8xh2
Diffraction source	PLS-II 7A	PLS-II 7A	PLS-II 7A
Wavelength (Å)	0.97934	0.97934	0.97934
Temperature (K)	100	100	100
Detector	ADSC Quantum 270	ADSC Quantum 270	ADSC Quantum 270
Crystal-to-detector distance (mm)	150	175	175
Rotation range per image (°)	1	1	1
Total rotation range (°)	360	360	360
Exposure time per image (s)	1	1	1
Space group	C2	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	110.55, 51.04, 55.52	51.24, 52.28, 195.30	51.17, 51.71, 196.24
α, β, γ (°)	90, 99.58, 90	90, 90, 90	90, 90, 90
Z/Z'	4/1	8/2	8/2
Estimated solvent content (%)	57.5	49.9	51.3
Resolution range (Å)	30–1.45	30–1.70	30–1.80
Total No. of reflections	401360	846596	709367
No. of unique reflections	54859	59042	49354
Completeness (%)	99.8	99.9	100.0
Multiplicity	7.3	14.3	14.4
$\langle I/\sigma(I) \rangle$	37.3 (2.5)	36.7 (3.5)	39.6 (3.6)
R _{r.i.m.}	0.072	0.079	0.071

pH 9.0, the crystal was transferred to 100 mM ammonium sulfate, 50 mM Tris–HCl pH 9.0, 17.5%(w/v) PEG 4000 and incubated for a week. Details of protein crystallization and photographic images of crystals are given in Table 2 and Supplementary Fig. S2, respectively.

2.3. Data collection and processing

X-ray diffraction data were collected on Pohang Light Source-II (PLS-II) beamline 7A at the Pohang Accelerator Laboratory, Pohang, Republic of Korea. Prior to data collection, the monoclinic crystals were briefly soaked in a cryoprotectant solution consisting of 30%(v/v) glycerol and 70%(v/v) reservoir, while the orthorhombic crystals were soaked in a cryoprotectant solution consisting of 20%(v/v) glycerol and 80%(v/v) reservoir, and then flash-cooled in a 100 K nitrogen stream. A total of 360 images with 1° oscillation angles were collected with sample-to-detector distances of 150 and 175 mm for monoclinic and orthorhombic crystals, respectively. All diffraction images were processed with *HKL-2000* (Otwinowski & Minor, 1997). The absorbed X-ray dose

for a single data set was less than 5×10^5 Gy, which is much lower than the Henderson dose limit of 1.45×10^7 Gy (Henderson, 1990). Data-processing statistics and diffraction images are given in Table 3 and Supplementary Fig. S3, respectively.

2.4. Structure solution and model refinement

The structures of NowGFP from both crystal forms were determined using the *CCP4* suite (Agirre *et al.*, 2023). The crystal structures were solved by the molecular-replacement method with *MOLREP* (Vagin & Teplyakov, 2010) using the previously solved structure of NowGFP (PDB entry 4rtc; Pletnev *et al.*, 2015) as a model. Crystallographic refinement was performed with *REFMAC5* (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*. The monoclinic crystal structure (C2) of NowGFP at pH 4.8 was determined at 1.45 Å resolution and orthorhombic crystal structures (P2₁2₁2₁) of NowGFP at pH 9.0 and 6.0 were solved

Table 4
Structure-refinement statistics for NowGFP crystals.

Values in parentheses are for the outer shell.

Structure	Monoclinic pH 4.8	Orthorhombic pH 9.0	Orthorhombic pH 6.0
PDB code	8xh0	8xh1	8xh2
Resolution	30.0–1.45 (1.48–1.45)	30.0–1.70 (1.73–1.70)	30.0–1.80 (1.83–1.80)
Completeness (%)	100.0	100.0	99.5
No. of reflections, working set	51316	55947	46765
No. of reflections, test set	2809	2870	2489
Final R_{cryst}	0.127	0.153	0.149
Final R_{free}	0.158	0.220	0.229
No. of non-H atoms			
Chain <i>A</i>	1893	1827	1818
Chain <i>B</i>	—	1809	1809
Glycerol	2	2	2
Water	266	229	159
R.m.s. deviations of bond lengths (Å)			
Chain <i>A</i>	0.0165	0.0158	0.0163
Chain <i>B</i>	—	0.0162	0.0164
R.m.s. deviations of angles (°)			
Chain <i>A</i>	2.01	1.92	2.02
Chain <i>B</i>	—	2.10	2.09
Average <i>B</i> factors (Å ²)			
Chain <i>A</i> (main/side chain)	22.69/25.30	26.75/33.02	27.00/33.26
Chain <i>B</i> (main/side chain)	—	30.98/37.49	29.47/36.52
Glycerol	37.92	55.35	42.38
Water	38.50	33.63	33.13
Ramachandran statistics (%)			
Preferred	97.76	98.43	97.76
Allowed	2.24	1.57	2.24
Outliers	—	—	—

at 1.7 and 1.8 Å resolution, respectively. The coordinates and structure factors for NowGFP in the monoclinic form at pH 4.8, the orthorhombic form at pH 9.0 and the orthorhombic form at pH 6.0 were deposited in the Protein Data Bank under accession codes 8xh0, 8xh1 and 8xh2, respectively. Structure-resolution and data-refinement statistics are given in Table 4. All structural figures were rendered with *PyMOL* (Schrödinger).

3. Results and discussion

3.1. Crystal structures of two crystal polymorphs

The assemblies of protein molecules within the unit cells of both orthorhombic and monoclinic crystals are illustrated in Fig. 1. For the comparative study of the two crystal forms, we chose the orthorhombic crystal structure obtained at pH 9.0 for comparison with the monoclinic structure since it has a higher resolution (1.7 Å). Comparison between the lower resolution orthorhombic crystal structure obtained at pH 6.0 (1.8 Å) and the monoclinic structure revealed nearly identical results to the higher pH (pH 9.0) counterpart. This comparison is detailed in Supplementary Figs. S4–S7 and Supplementary Tables S2 and S3.

In the case of the monoclinic (*C*2) crystal form there are four molecules inside the unit cell, and one molecule is in the asymmetric unit (chain *A*). We refer to four molecules, which are chains *A* from symmetry operations (x, y, z) , $(-x, y, -z)$, $(x + 1/2, y + 1/2, z)$ and $(-x + 1/2, y + 1/2, -z)$, as molecules 1, 2, 3 and 4, respectively (Figs. 1*a*–1*c*).

In the case of the orthorhombic (*P*2₁2₁2₁) crystal form there are eight molecules inside the unit cell, and two molecules are in the asymmetric unit (chains *A* and *B*). We refer to four molecules, which are chains *A* from symmetry operations (x, y, z) , $(x + 1/2, -y + 1/2, -z)$, $(-x, y + 1/2, -z + 1/2)$ and $(-x + 1/2, -y, z + 1/2)$, as molecules 1*A*, 2*A*, 3*A* and 4*A*, respectively. Similarly, we refer to four molecules, which are chains *B* from symmetry operations (x, y, z) , $(x + 1/2, -y + 1/2, -z)$, $(-x, y + 1/2, -z + 1/2)$ and $(-x + 1/2, -y, z + 1/2)$ as molecules 1*B*, 2*B*, 3*B* and 4*B*, respectively (Figs. 1*d*–1*f*).

For simplicity, we refer to chain *A* of the NowGFP structure in the monoclinic form as Mono. Similarly, we refer to chains *A* and *B* in the orthorhombic form as Orth(*A*) and Orth(*B*), respectively. Comparative analysis of Mono, Orth(*A*) and Orth(*B*) indicates that there are no significant differences in the overall protein structures, with the C^α – C^α r.m.s.d. being less than 0.4 Å. However, partial differences are observed in the r.m.s.d. from ideal angles, *B* factors and solvation free energy (Tables 4 and 5). These differences will be discussed in detail in Section 3.3.

3.2. Crystal contacts and related crystal packing

The term ‘crystal contacts’ denotes the interactions between protein molecules that emerge through the crystallization process (Janin & Rodier, 1995; Dasgupta *et al.*, 1997). In this section, the crystal contacts between NowGFP molecules are considered using *PISA* analysis (Krissinel & Henrick, 2005, 2007) and the protein crystal packing is discussed. All crystal contacts from each molecule are listed in Table 6.

In the monoclinic (*C*2) crystal form, the crystal contacts are categorized into three types, denoted as contacts I, II and III. Contact I is between molecules 1 and 4 (or 2 and 3) (Figs. 2*a* and 2*b*), which has the largest contact area for Mono. More specifically, contact I comes from two different symmetry-related molecules (Mono) at $(-x + 1/2, y - 1/2, -z - 1)$ and $(-x + 1/2, y + 1/2, -z - 1)$ and connects molecules 1 and 4 (or 2 and 3) to form a zigzag linear assembly along the *b* axis (Fig. 2*d*). Contact II is between molecules 1 and 2 (or 3 and 4) (Figs. 2*a* and 2*c*). More specifically, contact II comes from symmetry-related molecules at $(-x, y, -z - 1)$ and connects two types of linear assemblies (1–4 and 2–3) to form layers in the *ab* plane (Fig. 2*d*). The symmetry operation between two neighbouring linear assemblies is $(-x, y, -z)$, which is the same as the symmetry operation between molecules 1 and 2 or 3 and 4 (represented as opposite white arrows in Fig. 2*d*). Contact III is between molecules 1 and 3 (or 2 and 4), which has the smallest contact area (Figs. 2*a* and 2*b*). More specifically, contact III comes from two different symmetry-related molecules at $(-x + 1/2, y - 1/2, -z)$ and $(-x + 1/2, y + 1/2, -z)$ and results from the packing of sheets along the *c* axis (Fig. 2*d*).

In the orthorhombic (*P*2₁2₁2₁) crystal form, the crystal contacts are categorized into six types: contacts I_A, I_B, II_{AB}, III_{AB}, IV_{AB} and V_B. Contact I_A is between molecules 1*A* and 2*A* (or 3*A* and 4*A*) (Figs. 3*a* and 3*c*), which has the largest contact area for Orth(*A*). More specifically, contact I_A comes

from symmetry-related molecules chain *A* [Orth(*A*)] at $(x - 1/2, -y + 1/2, -z)$ and $(x + 1/2, -y + 1/2, -z)$ and connects molecules 1*A* and 2*A* (or 3*A* and 4*A*) to form a zigzag linear assembly along the *a* axis (Fig. 3*e*). Contact *I*_B is between molecules 1*B* and 3*B* (or 2*B* and 4*B*) (Figs. 3*b* and

3*c*), which has the largest contact area for Orth(*B*). More specifically, contact *I*_B comes from symmetry-related molecules chain *B* [Orth(*B*)] at $(-x + 1, y - 1/2, -z + 1/2)$ and $(-x + 1, y + 1/2, -z + 1/2)$ to form a zigzag linear assembly along the *b* axis (Fig. 3*e*). Contacts *II*_{AB} and *III*_{AB} are between

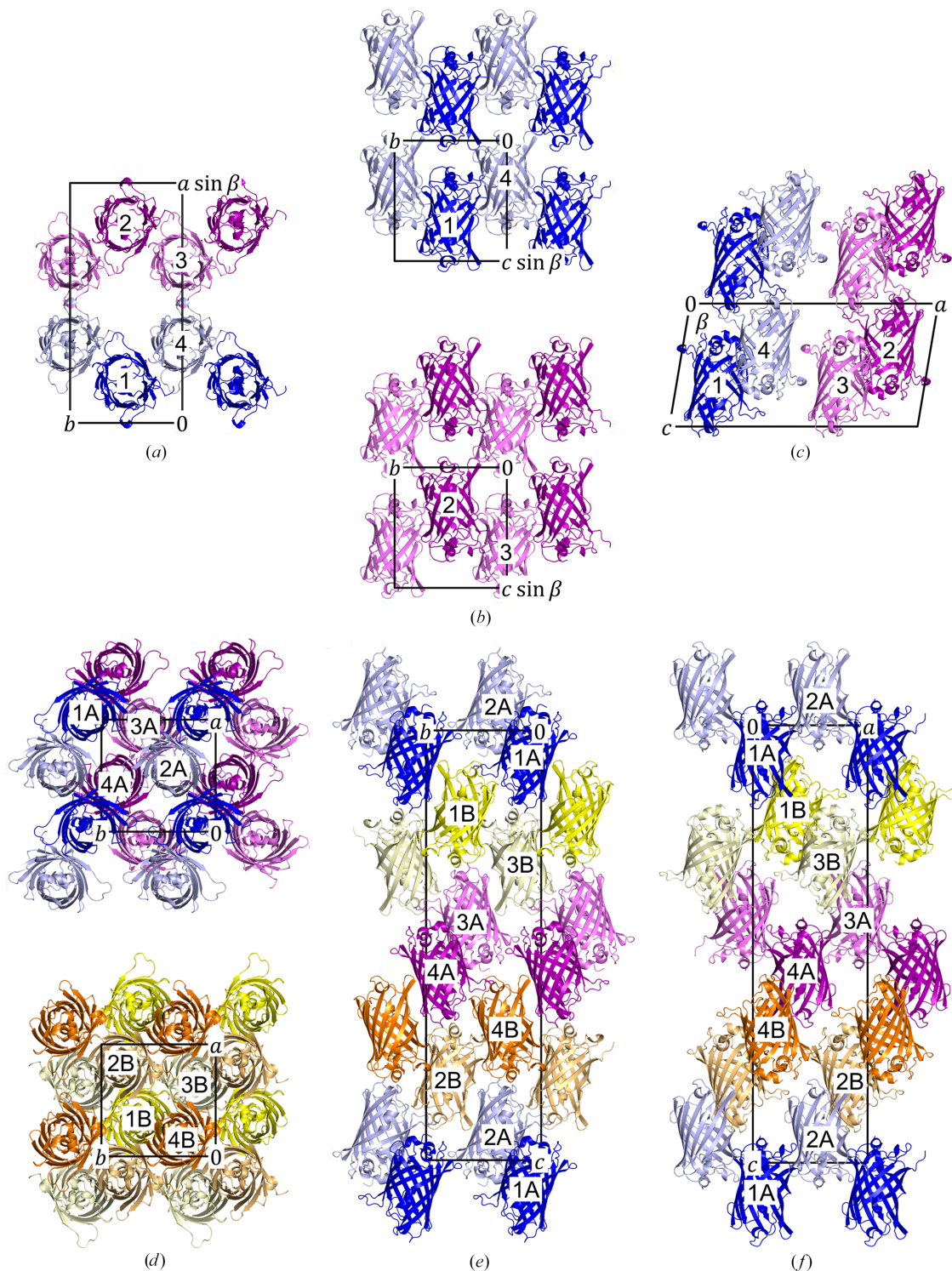


Figure 1
 Assembly of the protein molecules in the monoclinic (*a*–*c*) and orthorhombic (*d*–*f*) crystal forms. Crystal structures are viewed along the *c* axis (*a*, *d*), *a* axis (*b*, *e*) and *b* axis (*c*, *f*). The secondary structures of NowGFP are represented as ribbons, while the unit cell of each crystal form is outlined in a black box. For the monoclinic form, molecules 1, 2, 3 and 4 are coloured blue, purple, light purple and light blue, respectively. For the orthorhombic form, molecules 1*A*, 2*A*, 3*A*, 4*A*, 1*B*, 2*B*, 3*B* and 4*B* are coloured blue, light blue, light purple, purple, yellow, light orange, light yellow and orange, respectively.

molecules 1A and 1B (or 2A and 2B, 3A and 3B, or 4A and 4B) (Fig. 3*d*). More specifically, contact II_{AB} comes from chain A and chain B in the same asymmetric unit, while contact III_{AB} comes from a symmetry-related chain B ($x, y + 1, z$) for chain A (x, y, z). Contact IV_{AB} is between molecules 1A and 3B (or 2A and 4B, 3A and 1B, or 4A and 2B) (Fig. 3*d*). More specifically, contact IV_{AB} comes from a symmetry-related

chain B ($-x, y + 1/2, -z + 1/2$) for chain A (x, y, z). Contacts II_{AB}, III_{AB} and IV_{AB} connect two types of linear assemblies which are orthogonal to each other (assembly of molecules 1A–2A and 1B–3B, 1A–2A and 2B–4B, 3A–4A and 1B–3B, or 3A–4A and 2B–4B; Fig. 3*e*), and II_{AB} has the largest contact area among them. The symmetry operation between two linear assemblies made of the same chains (chain A or B) is

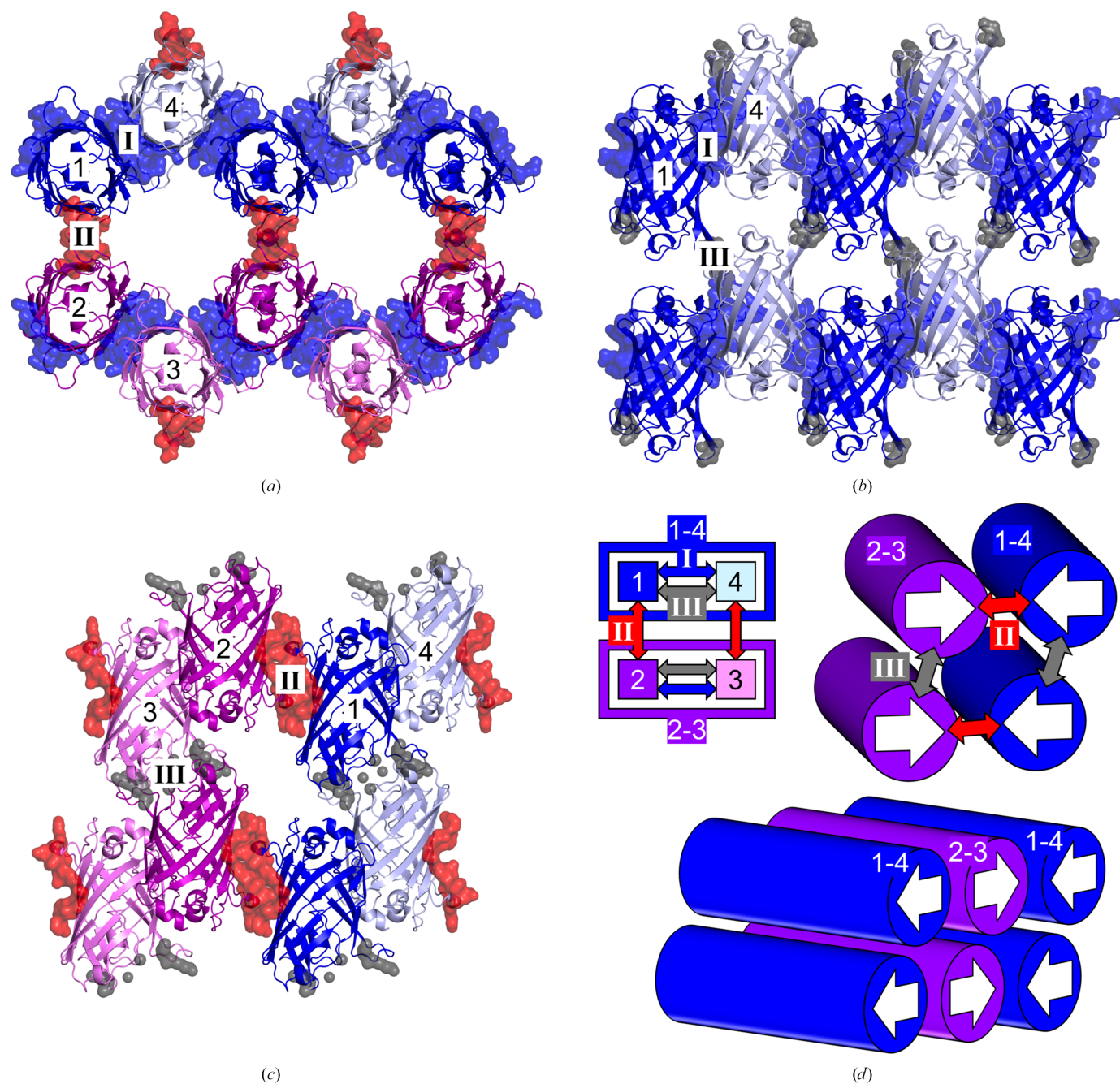


Figure 2

Crystal contacts and packing of NowGFP molecules in the monoclinic crystal structure. The secondary structures of NowGFP are represented as ribbons, and atoms belonging to crystal contacts are represented as surfaces. (a) Contacts I and II are viewed along the *c* axis. (b) Contacts I and III are viewed along the *b* axis. (c) Contacts II and III are viewed along the *a* axis. (d) Schematic diagram of crystal packing. Crystal contacts are represented by coloured arrows, and linear assemblies connected by contact I are represented as blue- or purple-coloured boxes. Molecules connected by the blue-coloured contact I are linked along the *b* axis. Linear assemblies are stacked parallel by contacts II and III along the *a* axis and *c* axis, respectively. The symmetry operation between assemblies 1–4 and 2–3, indicated by opposite white arrows, is $(-x, y, -z)$.

Table 5

Protein molecules from the monoclinic crystal form and the orthorhombic crystal form.

The total numbers of atoms and residues in the refined NowGFP molecule (residues 2–231) are 1809 and 228, respectively. H atoms are not considered. All values are from PISA analysis (Krissinel & Henrick, 2005, 2007).

Crystal form	Name of molecule	No. of atoms on the surface [†]	No. of residues on the surface [†]	Surface area [‡] (Å ²)	Solvation energy [‡] (kcal mol ⁻¹)
Monoclinic	Mono	1022	210	10614.4	-215.1
Orthorhombic	Orth(A)	998	211	10620.8	-214.9
Orthorhombic	Orth(B)	993	210	10438.4	-209.2

[†] Solvent-accessible surface area of the corresponding structure. [‡] Solvation energy gain upon protein folding.

Table 6

Crystal contacts of the monoclinic crystal form and the orthorhombic crystal form.

Symmetry operations for symmetry-related neighbouring molecules are given in parentheses. H atoms are not considered. All values are from PISA analysis (Krissinel & Henrick, 2005, 2007).

Contacting molecule	Name of contact	No. of atoms in the contact	No. of residues in the contact	Contact area [†] (Å ²)	Solvation energy [‡] (kcal mol ⁻¹)	No. of atom pairs closer than 3.0/3.5 Å	
Mono							
	($-x + 1/2, y - 1/2, -z - 1$)	I	52	15	525.3	-2.0	10/30
	($-x + 1/2, y + 1/2, -z - 1$)	I	61	19	502.8	-1.0	10/30
	($-x, y, -z - 1$)	II	61	14	512.3	1.2	0/18
	($-x + 1/2, y - 1/2, -z$)	III	13	6	123.0	-0.4	0/1
	($-x + 1/2, y + 1/2, -z$)	III	11	4	119.6	-1.0	0/1
Orth(A)							
	$A(x - 1/2, -y + 1/2, -z)$	I _A	52	17	492.9	0.7	5/21
	$A(x + 1/2, -y + 1/2, -z)$	I _A	50	12	475.8	-1.2	5/21
	$B(x, y, z)$	II _{AB}	50	12	404.6	0.0	4/10
	$B(x, y + 1, z)$	III _{AB}	26	10	186.4	-0.2	2/5
	$B(-x, y + 1/2, -z + 1/2)$	IV _{AB}	9	3	76.6	-0.7	0/1
Orth(B)							
	$B(-x + 1, y - 1/2, -z + 1/2)$	I _B	45	14	422.1	-0.6	3/13
	$B(-x + 1, y + 1/2, -z + 1/2)$	I _B	41	10	438.0	0.1	3/13
	$A(x, y, z)$	II _{AB}	45	16	411.0	-1.1	4/10
	$A(x, y - 1, z)$	III _{AB}	21	8	183.2	0.5	2/5
	$A(-x, y - 1/2, -z + 1/2)$	IV _{AB}	9	4	76.1	-0.6	0/1
	$B(-x, y - 1/2, -z + 1/2)$	V _B	8	2	77.3	0.1	0/1
	$B(-x, y + 1/2, -z + 1/2)$	V _B	6	2	76.5	-1.2	0/1

[†] Surface area buried by the intermolecular interface. [‡] Solvation energy gain upon formation of the interface. This value does not include the effect of satisfied hydrogen bonds and salt bridges across the interface.

($-x + 1/2, -y, z + 1/2$). This is the same as the symmetry operation between molecules 1A and 4A (or 1B or 4B, 2A and 3A, or 2B and 3B) and is represented as opposite white arrows in Fig. 3(e). Contact V_B is between molecules 1B and 3B (or 2B and 4B), which has the smallest contact area (Fig. 3b). More specifically, contact V_B comes from a symmetry-related chain B at ($-x, y - 1/2, -z + 1/2$) and ($-x, y + 1/2, -z + 1/2$).

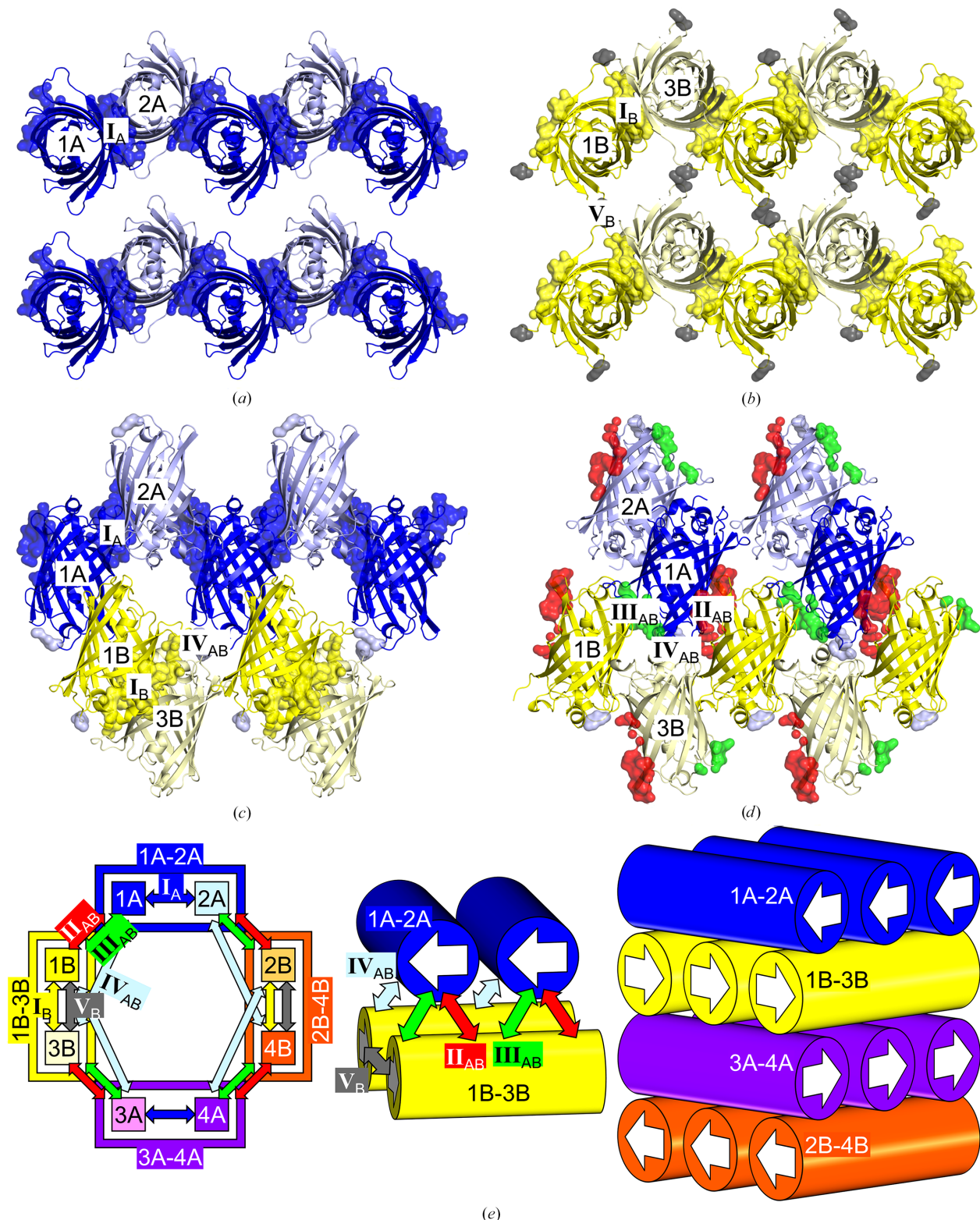
3.3. The structural alteration of NowGFP results from crystal contacts

Comparison of the crystal contacts of three types of molecules [Mono, Orth(A) and Orth(B)] from two different crystal forms shows that contacts I, I_A and I_B are similar in structure and sequence (blue and yellow in Fig. 4). Notably, linear assemblies derived from contacts I, I_A and I_B are also almost identical (Figs. 2a, 2b and 3a–3c). However, there is a significant difference between contacts II and II_{AB} (red in Fig. 4). This difference is responsible for the packing mode of the linear assemblies: parallel for the monoclinic form and perpendicular for the orthorhombic form. There are minor differences in the other contacts, including contacts III, III_{AB}, IV_{AB} and V_B, but those have a much smaller contact area compared with contacts I, I_A, I_B, II and II_{AB} (Table 6). In

addition, the histogram of the distances between atom pairs for each crystal contact indicates that contacts I, I_A, I_B, II and II_{AB} are more closely interacting compared with others (Figs. 4e and 4f). We thus focused on analysing the differences between contacts II and II_{AB} as represented in Fig. 5.

For contact II in the monoclinic crystal form, there are no atom pairs closer than 3.0 Å (Table 6). Phe99, Tyr182, Gln157 and Lys156 from two molecules contact each other without any water molecule in between. Gln157 also contacts the main chain of Arg96 and Gln183. Additionally, residues corresponding to contact II show no significant deviation from Mono to Orth(A) and Orth(B) (Fig. 5a). This indicates that contact II does not significantly influence the overall protein structure.

For contact II_{AB} in the orthorhombic crystal form, there are four atom pairs closer than 3.0 Å (Table 6). Two of them come from the arginine–aspartic acid salt bridge between Asp180 of Orth(A) and Arg168 of Orth(B). The other two are from Tyr182 and Tyr164 of Orth(A), which are connected to the main chain of Val176 and Asp173 of Orth(B) through hydrogen bonding. Compared with Mono, a significant shift of the main chain is observed for residues 170–176 of Orth(B), which mainly involves contact II_{AB} and the loop between the β8 and β9 strands (Figs. 5b and 5c). This indicates that contact


Figure 3

Crystal contacts and packing of NowGFP molecules in the orthorhombic crystal structure. The secondary structures of NowGFP are represented as ribbons, and atoms belonging to crystal contacts are represented as surfaces. (a) Contact I_A is viewed along the *c* axis. (b) Contacts I_B and V_B are viewed along the *c* axis. (c) Contacts I_A and I_B are viewed along the *b* axis. (d) Contacts II_{AB}, III_{AB} and IV_{AB} are viewed along the *a* axis. Contacts I_A, I_B, II_{AB}, III_{AB}, IV_{AB} and V_B are coloured blue, yellow, red, green, light blue and grey, respectively. (e) Schematic diagram of protein packing. Crystal contacts are represented by coloured arrows and linear assemblies connected by contact I_A are represented as blue- or purple-coloured boxes, while linear assemblies connected by contact I_B are represented as yellow- or orange-coloured boxes. The linear assemblies connected by contacts I_A and I_B are linked along the *a* axis and the *b* axis, respectively. Linear assemblies composed of chains A and B are stacked perpendicular by contacts II_{AB}, III_{AB} and IV_{AB} along the *c* axis. Linear assemblies composed of chain B are stacked parallel by contact V_B along the *a* axis. The symmetry operation between two assemblies composed of the same chains, 1A–2A and 3A–4A (or 1B–3B and 2B–4B), indicated by opposite white arrows, is $(-x + 1/2, -y, z + 1/2)$.

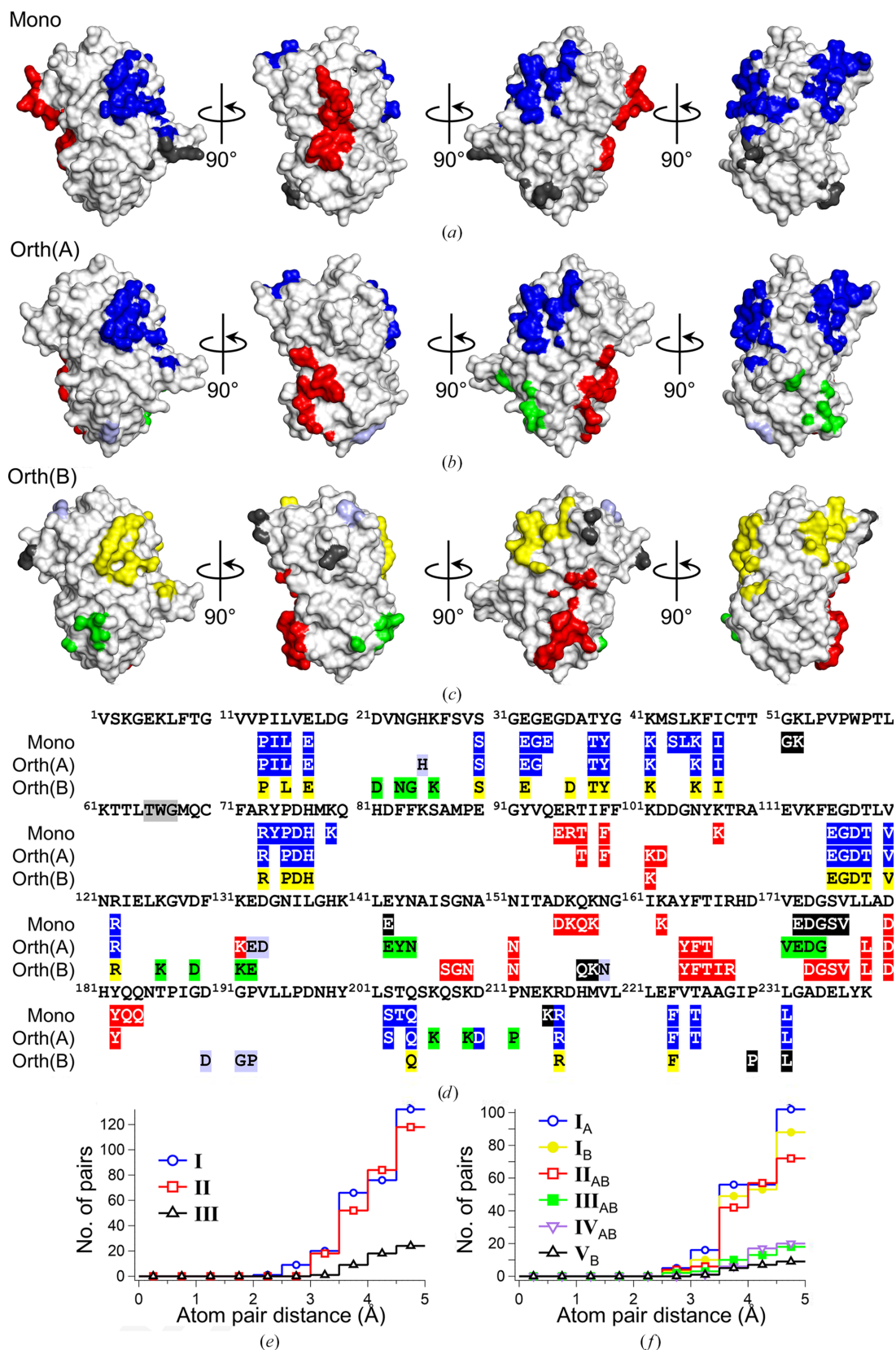


Figure 4

Comparison of crystal contacts for three types of molecules. (a) Chain A of the monoclinic structure, Orth(A). 1A, 2A, 3A and 4A are identical. (b) Chain A of the orthorhombic structure, Orth(A). 1A, 2A, 3A and 4A are identical. (c) Chain B of the orthorhombic structure, Orth(B). 1B, 2B, 3B and 4B are identical. (d) Amino-acid sequence representation of the three molecules. Thr65, Trp66 and Gly67 were replaced by the chromophore. For Mono, atoms belonging to contacts I, II and III are coloured blue, red and grey, respectively. For Orth(A) and Orth(B), atoms belonging to contacts I_A, I_B, II_{AB}, III_{AB}, IV_{AB} and V_B are coloured blue, yellow, red, green, light blue and grey, respectively. Histograms of the distances between atom pairs in each crystal contact from (e) the monoclinic and (f) the orthorhombic crystal forms are shown.

Π_{AB} might be responsible for the structural shifts in Orth(B). Additionally, this shift causes residues 141–147, which are part of the $\beta 7$ strand, to move away from the $\beta 10$ strand (Fig. 5). Specifically, the C^α – C^α distances between Asn144 in the $\beta 7$

strand and Gln207 in the $\beta 10$ strand are 6.3, 6.1 and 6.7 Å for Mono, Orth(A) and Orth(B), respectively.

The evidence that Orth(B) has a more unstable structure compared with Orth(A) or Mono is that the r.m.s.d. from ideal

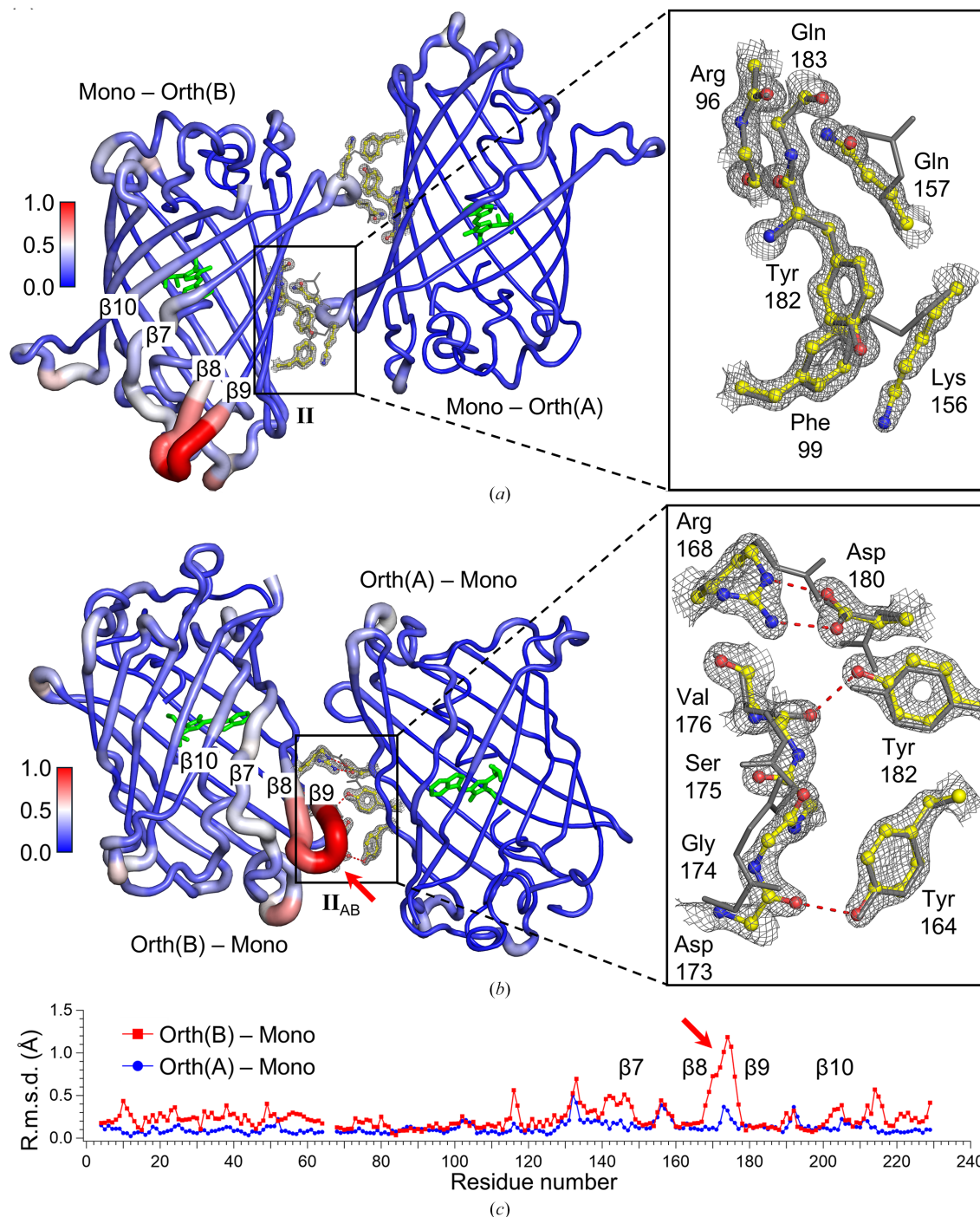


Figure 5

Comparison of contacts II and Π_{AB} . (a) Two molecules in the monoclinic crystal structure connected by contact II are presented. The Orth(A) and Orth(B) structures are superposed with the Mono structure and coloured grey. For clarity, only the main chains of Arg96 and Gln183 are represented. (b) Two molecules from the orthorhombic structure connected by contact Π_{AB} are presented. The Mono structures are superposed with Orth(A) and Orth(B) and coloured grey. For clarity, only the main chains of Val176, Ser175, Gly174 and Asp173 are represented. The r.m.s.d. between the main chains of two superposed molecules is shown as a putty representation, with a blue–white–red colour profile. Crystal contacts less than 3.0 Å are represented as red dotted lines. Stick representations of chromophores are coloured green. The electron-density map ($2F_o - F_c$ map) has a cutoff level of 2.0σ . (c) The r.m.s.d. plot of the main-chain atoms of two structures from different crystal forms. Residues close to the N-terminus and C-terminus are ignored, and only the residues numbered from 4 to 229 are considered. Residue numbers 65, 66 and 67 are ignored since these are substituted with the chromophore. The region with the most significant r.m.s.d., spanning residues 170–176, is indicated by a red arrow.

Table 7

Two alternative conformations of the key residue Lys61.

Name of molecule	pH	χ_1/χ_2 of <i>k1</i> conformation (°)†	χ_1/χ_2 of <i>k2</i> conformation (°)†
Orth(A) pH 9.0	9.0	-158/-152 (~80%)	-74/-175 (~20%)
Orth(B) pH 9.0	9.0	-169/-103 (~100%)	—
Orth(A) pH 6.0	6.0	-166/-132 (~50%)	-84/-179 (~50%)
Orth(B) pH 6.0	6.0	-165/-101 (~100%)	—
Mono pH 4.8	4.8	-169/-103 (~20%)	-72/-181 (~80%)

† χ_1/χ_2 are the torsion angles corresponding to $C^\alpha-C^\beta/C^\beta-C^\gamma$. The partial occupancies of the alternative conformational states are given in parentheses.

angles, average *B* factor and solvation free energy all show the highest values for Orth(B). In detail, the r.m.s.d.s from ideal

angles for Mono, Orth(A) and Orth(B) are $\sim 2.0^\circ$, $\sim 1.9^\circ$ and $\sim 2.1^\circ$, respectively, the average *B* factors of the main chain from Mono, Orth(A) and Orth(B) are ~ 22.7 , ~ 26.8 and 31.0 \AA^2 , respectively, and the solvation free energy from the isolated structure of each molecule is $\sim -215 \text{ kcal mol}^{-1}$ for Orth(A) and Mono and $\sim -209 \text{ kcal mol}^{-1}$ for Orth(B) (Tables 4 and 5). These results indicate that Orth(A) exhibits similar conformational behaviour to Mono, while Orth(B) represents a relatively unstable conformational state.

Moreover, protein crystals are highly hydrated, with those with lower solvent content typically having better diffraction quality (Matthews, 1968; Zhang *et al.*, 1995). However, solvent-content estimation reveals that the orthorhombic

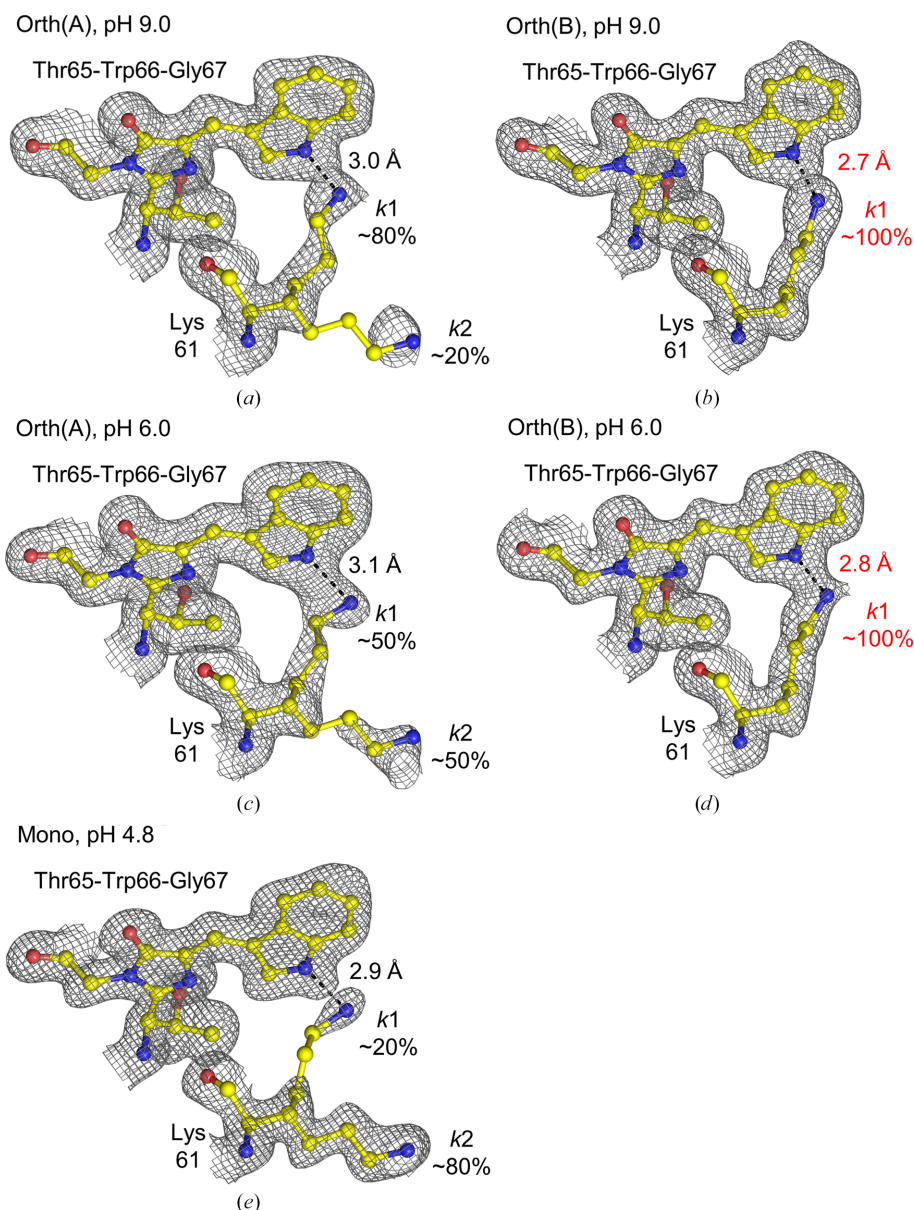


Figure 6

The chromophore Thr65-Trp66-Gly67 and key residue Lys61 under various pH conditions. The structures are from (a, b) the orthorhombic crystal at pH 9.0, (c, d) the orthorhombic crystal at pH 6.0 and (e) the monoclinic crystal at pH 4.8. The electron-density map ($2F_o - F_c$ map) has a cutoff level of 0.8σ . Black dotted lines between the N^ϵ atom of the Trp66 indole and the N^ζ atom of Lys61 represent hydrogen bonding. Orth(B), which is a new conformational state of NowGFP caused by an orthorhombic crystal lattice, shows the dominant *k1* conformation of Lys61, while both Mono and Orth(A) exhibit pH-dependent conformational shifts of Lys61.

crystal has less solvent (~50%) yet results in a lower resolution (1.7 Å) compared with the monoclinic crystal, which contains more solvent (~58%) but diffracts to a higher resolution (1.45 Å) (Table 3). This discrepancy could be additional evidence that the crystallinity of the orthorhombic crystal is poorer due to the instability of Orth(B).

3.4. New conformational state of NowGFP trapped by crystal lattice packing

In a previous study by Pletnev and coworkers, it was revealed that the key residue Lys61 plays a central role in the ionization process of the chromophore, showing significant pH-dependent conformational changes. Specifically, at high pH the N^ε atom of Lys61 makes two hydrogen bonds: one to the N^ε atom of Trp66 and the other to the O^{ε1} atom of Glu222 (*k1* conformation). This dual bonding helps to form a non-covalent connection and promotes the deprotonation of Trp66. In contrast, at lower pH the orientation of Lys61 shifts away from Trp66, instead forming a hydrogen bond to the N^{ε2} atom of Gln207 (*k2* conformation; Pletnev *et al.*, 2015).

To elucidate the alternative conformational state trapped by the orthorhombic crystal lattice, a total of five NowGFP molecules are compared (one molecule from the monoclinic crystal structure at pH 4.8 and four molecules from the orthorhombic crystal structures at pH 6.0 and pH 9.0).

In the case of the Orth(A) and Mono structures, the key residue Lys61 has ~80% *k1* conformation and ~20% *k2* conformation under a high-pH condition (pH 9.0). This shifts to ~50% *k1* conformation and ~50% *k2* conformation as the pH decreases to 6.0, and finally to ~20% *k1* conformation and ~80% *k2* conformation under an acidic pH condition (pH 4.8) (Figs. 6*a*, 6*c* and 6*e*). This trend of conformational shifts from *k1* to *k2* as the pH decreases is consistent with previously published structures: the monoclinic structure at pH 9.0 (PDB entry 4rtc, 1.18 Å resolution) and pH 4.8 (PDB entry 4rys, 1.35 Å resolution) (Supplementary Fig. S8; Pletnev *et al.*, 2015). Even though the high-resolution structure of NowGFP shows a partial *trans* conformation of the chromophore at pH 4.8, this is difficult to observe in our moderate-resolution structure (1.45 Å).

In the case of the Orth(B) structure, however, the conformation of Lys61 is almost 100% *k1* at both pH 9.0 and pH 6.0, and the distance between the N^ε atoms of the Trp66 indole and the N^ε atom of Lys61 in the *k1* conformation is ~0.3 Å closer compared with the Orth(A) structure (Figs. 6*a*–6*d*). It appears that strong hydrogen bonding between the Trp66 indole and the *k1* conformation of Lys61 is responsible for the high occupancies of the *k1* conformation in the Orth(B) structure. Considering that Orth(B) has an unstable structure compared with Orth(A) and Mono, and this instability is linked to crystal contact II_{AB} as detailed in Section 3.3, it is evident that the alteration of the Lys61 conformation in the Orth(B) structure is a result of orthorhombic crystal packing. Table 7 shows the χ_1 and χ_2 angles of two alternative conformations of the key residue Lys61.

While this alternative conformational state could be dismissed as merely an artefact of crystal packing, a previous study of crystal polymorphs of photoactive yellow protein indicates that the crystal lattice does not simply impose arbitrary conformational changes on the protein molecule (van Aalten *et al.*, 2000). Instead, the protein shifts along essential eigenvectors to adapt to different lattice environments. Therefore, different protein structures from crystal polymorphs may represent the inherent conformational flexibility of the molecule (van Aalten *et al.*, 1997).

Previous studies, including the unfolding and backbone dynamics of GFP, reveal that β -strands 7, 8, 9 and 10 exhibit greater flexibility (Huang *et al.*, 2007; Seifert *et al.*, 2003). Additionally, the parental WasCFP shows conformational flexibility in the $\beta7$ and $\beta10$ strands, which opens the solvent channel between these two β -strands. This flexibility is expected to play a crucial role in managing solvent access to the chromophore (Laricheva *et al.*, 2015). These observations are consistent with our finding that the r.m.s.d. of the main-chain atoms between two structures from crystal polymorphs shows higher values at residues corresponding to β -strands 7, 8, 9 and 10. The most significant deviations are noted in the loop between the $\beta8$ and $\beta9$ strands, resulting in opening of the solvent channel between the $\beta7$ and $\beta10$ strands (Fig. 5). Based on these correspondences, the new conformational state of NowGFP captured by crystal packing in this study is likely to represent one of the potential conformational states inherent to this protein.

4. Conclusions

In this study, we report the discovery of a novel orthorhombic crystal form of NowGFP and conduct a detailed comparison with the known monoclinic crystal form. Our investigations primarily focused on the crystal contacts, revealing that both forms exhibit similar zigzag linear assemblies of protein molecules resulting from crystal contact I. The key distinction between the two forms lies in their stacking modes: parallel stacking for the monoclinic form and perpendicular stacking for the orthorhombic form. This difference in packing correlates with a specific crystal contact, referred to as crystal contact II (or II_{AB}), and results in an alteration of one molecule in the symmetry unit of the orthorhombic crystal form, designated as Orth(B). Given that these structural shifts are predominantly concentrated between β -strands 7–10, which are known for their partial flexibility, we propose that this altered molecule represents an alternative conformational state of NowGFP. In contrast, the other molecule in the orthorhombic form, Orth(A), remains unchanged and is similar to that found in the monoclinic form.

Significantly, this new conformational state of NowGFP captured in the orthorhombic crystal packing exhibits a different functional behaviour: the key residue Lys61, which is known for its pH-dependent shift from the *k1* to the *k2* conformation, appears to be locked in the *k1* configuration regardless of pH conditions. This contrasts with the unaltered molecule, in which Lys61 exhibits pH-dependent movement as

expected. These observations provide valuable insights into how crystal lattice packing influences the conformational states of protein molecules, enhancing our understanding of the conformational flexibility of protein structures.

Acknowledgements

The authors would like to thank the staff of beamline 7A at Pohang Light Source-II for their support during data collection. Author contributions were as follows. JKK and CUK conceived the research. HJ, JS, SK and DM ran the experiments. JKK analyzed the data and wrote the manuscript. KHK, DM and CUK contributed to the overall scientific interpretation and edited the manuscript.

Funding information

This research was supported by the National Research Foundation of Korea (award Nos. NRF-2022R1A2C2091815 to CUK and NRF-2020R1C1C1003937 to DM).

References

- Aalten, D. M. F. van, Conn, D. A., de Groot, B. L., Berendsen, H. J. C., Findlay, J. B. C. & Amadei, A. (1997). *Biophys. J.* **73**, 2891–2896.
- Aalten, D. M. F. van, Crielaard, W., Hellingwerf, K. J. & Joshua-Tor, L. (2000). *Protein Sci.* **9**, 64–72.
- Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. É., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. & Yamashita, K. (2023). *Acta Cryst.* **D79**, 449–461.
- Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. (2010). *Physiol. Rev.* **90**, 1103–1163.
- Dasgupta, S., Iyer, G. H., Bryant, S. H., Lawrence, C. E. & Bell, J. A. (1997). *Proteins*, **28**, 494–514.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Gerlits, O., Ho, K. Y., Cheng, X., Blumenthal, D., Taylor, P., Kovalevsky, A. & Radić, Z. (2019). *Chem. Biol. Interact.* **309**, 108698.
- Henderson, R. (1990). *Proc. R. Soc. Lond. B*, **241**, 6–8.
- Huang, J. R., Craggs, T. D., Christodoulou, J. & Jackson, S. E. (2007). *J. Mol. Biol.* **370**, 356–371.
- Janin, J. & Rodier, F. (1995). *Proteins*, **23**, 580–587.
- Jiang, Y., Lu, G., Trescott, L. R., Hou, Y., Guan, X., Wang, S., Stamenkovich, A., Brunzelle, J., Sirinupong, N., Li, C. & Yang, Z. (2013). *PLoS One*, **8**, e81904.
- Krissinel, E. & Henrick, K. (2005). *Computational Life Sciences*, edited by M. R. Berthold, R. C. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 163–174. Berlin, Heidelberg: Springer.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Laricheva, E. N., Goh, G. B., Dickson, A. & Brooks, C. L. III (2015). *J. Am. Chem. Soc.* **137**, 2892–2900.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*. Chichester: John Wiley & Sons.
- 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.* **D67**, 355–367.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pletnev, V. Z., Pletneva, N. V., Sarkisyan, K. S., Mishin, A. S., Lukyanov, K. A., Goryacheva, E. A., Ziganshin, R. H., Dauter, Z. & Pletnev, S. (2015). *Acta Cryst.* **D71**, 1699–1707.
- Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. (2004). *Nat. Biotechnol.* **22**, 445–449.
- Sarkisyan, K. S., Goryashchenko, A. S., Lidsky, P. V., Gorbachev, D. A., Bozhanova, N. G., Gorokhovatsky, A. Y., Pereverzeva, A. R., Ryumina, A. P., Zherdeva, V. V., Savitsky, A. P., Solntsev, K. M., Bommarius, A. S., Sharonov, G. V., Lindquist, J. R., Drobizhev, M., Hughes, T. E., Rebane, A., Lukyanov, K. A. & Mishin, A. S. (2015). *Biophys. J.* **109**, 380–389.
- Sarkisyan, K. S., Yampolsky, I. V., Solntsev, K. M., Lukyanov, S. A., Lukyanov, K. A. & Mishin, A. S. (2012). *Sci. Rep.* **2**, 608.
- Seifert, M. H., Georgescu, J., Ksiazek, D., Smialowski, P., Rehm, T., Steipe, B. & Holak, T. A. (2003). *Biochemistry*, **42**, 2500–2512.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Vaney, M. C., Broutin, I., Retailleau, P., Douangamath, A., Lafont, S., Hamiaux, C., Prangé, T., Ducruix, A. & Riès-Kautt, M. (2001). *Acta Cryst.* **D57**, 929–940.
- Yamada, K. D., Kunishima, N., Matsuura, Y., Nakai, K., Naitow, H., Fukasawa, Y. & Tomii, K. (2017). *Acta Cryst.* **D73**, 757–766.
- Zhang, X. J., Wozniak, J. A. & Matthews, B. W. (1995). *J. Mol. Biol.* **250**, 527–552.