wwPDB X-ray Structure Validation Summary Report (i)

Nov 2, 2021 – 12:49 PM EDT

PDB ID	:	2A53
Title	:	fluorescent protein asFP595, A143S, off-state
Authors	:	Andresen, M.; Wahl, M.C.; Stiel, A.C.; Graeter, F.; Schaefer, L.; Trowitzsch, S.; Weber, G.; Eggeling, C.; Grubmueller, H.; Hell, S.W.; Jakobs, S.
Deposited on	:	2005-06-30
Resolution	:	1.45 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

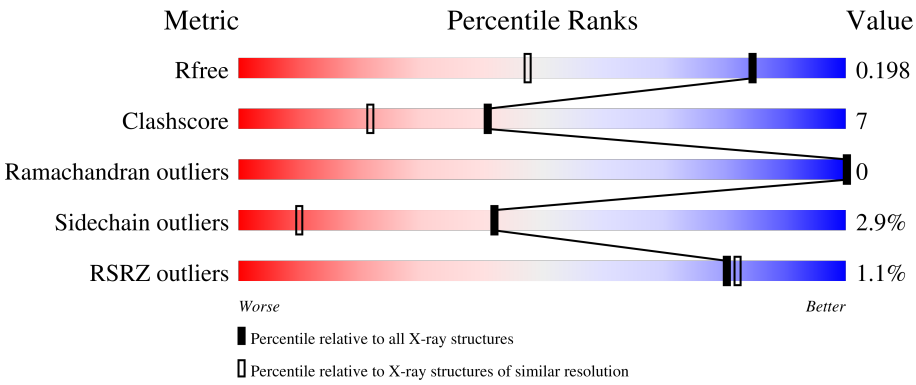
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.45 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	73	68% 18% • 12%
1	C	73	% 71% 10% 19%
2	B	168	% 80% 17% • •
2	D	168	% 89% 10% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4270 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called GFP-like non-fluorescent chromoprotein FP595 chain 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	64	501	320	81	96	4	0	2	0
			Total	C	N	O	S			
1	C	59	456	295	72	85	4	2	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MET	-	expression tag	UNP Q9GZ28
A	-9	ARG	-	expression tag	UNP Q9GZ28
A	-8	GLY	-	expression tag	UNP Q9GZ28
A	-7	SER	-	expression tag	UNP Q9GZ28
A	-6	HIS	-	expression tag	UNP Q9GZ28
A	-5	HIS	-	expression tag	UNP Q9GZ28
A	-4	HIS	-	expression tag	UNP Q9GZ28
A	-3	HIS	-	expression tag	UNP Q9GZ28
A	-2	HIS	-	expression tag	UNP Q9GZ28
A	-1	HIS	-	expression tag	UNP Q9GZ28
A	0	GLY	-	expression tag	UNP Q9GZ28
A	1	SER	-	expression tag	UNP Q9GZ28
C	-10	MET	-	expression tag	UNP Q9GZ28
C	-9	ARG	-	expression tag	UNP Q9GZ28
C	-8	GLY	-	expression tag	UNP Q9GZ28
C	-7	SER	-	expression tag	UNP Q9GZ28
C	-6	HIS	-	expression tag	UNP Q9GZ28
C	-5	HIS	-	expression tag	UNP Q9GZ28
C	-4	HIS	-	expression tag	UNP Q9GZ28
C	-3	HIS	-	expression tag	UNP Q9GZ28
C	-2	HIS	-	expression tag	UNP Q9GZ28
C	-1	HIS	-	expression tag	UNP Q9GZ28
C	0	GLY	-	expression tag	UNP Q9GZ28
C	1	SER	-	expression tag	UNP Q9GZ28

- Molecule 2 is a protein called GFP-like non-fluorescent chromoprotein FP595 chain 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	168	1369	869	234	256	10	9	3	0
2	D	168	1366	869	232	255	10	6	2	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	65	NRQ	MET	chromophore	UNP Q9GZ28
B	65	NRQ	TYR	chromophore	UNP Q9GZ28
B	65	NRQ	GLY	chromophore	UNP Q9GZ28
B	143	SER	ALA	engineered mutation	UNP Q9GZ28
D	65	NRQ	MET	chromophore	UNP Q9GZ28
D	65	NRQ	TYR	chromophore	UNP Q9GZ28
D	65	NRQ	GLY	chromophore	UNP Q9GZ28
D	143	SER	ALA	engineered mutation	UNP Q9GZ28

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
3	B	1	1	1	0	0

- Molecule 4 is water.

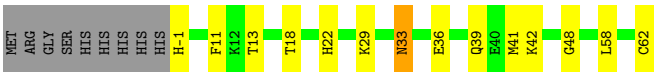
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	95	95	95	0	0
4	B	215	215	215	0	0
4	C	69	69	69	0	0
4	D	198	198	198	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

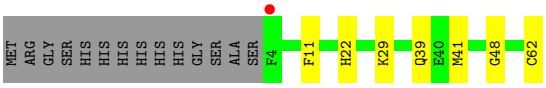
• Molecule 1: GFP-like non-fluorescent chromoprotein FP595 chain 1

Chain A:



• Molecule 1: GFP-like non-fluorescent chromoprotein FP595 chain 1

Chain C:

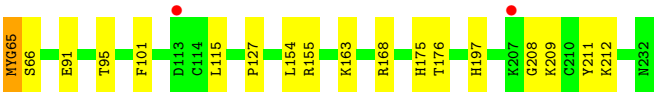


• Molecule 2: GFP-like non-fluorescent chromoprotein FP595 chain 2

Chain B:

• Molecule 2: GFP-like non-fluorescent chromoprotein FP595 chain 2

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	76.00Å 126.26Å 93.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.45 65.11 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.00-1.45) 99.6 (65.11-1.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.173 , 0.199 0.171 , 0.198	Depositor DCC
R_{free} test set	3996 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	17.1	Xtriage
Anisotropy	0.275	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 46.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.010 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.018 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4270	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.49% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NRQ, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	0.53	0/514	0.64	0/690
1	C	0.89	1/468 (0.2%)	0.66	0/627
2	B	0.71	1/1382 (0.1%)	0.81	4/1863 (0.2%)
2	D	0.77	2/1380 (0.1%)	0.72	1/1862 (0.1%)
All	All	0.74	4/3744 (0.1%)	0.74	5/5042 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	209	LYS	CG-CD	14.82	2.02	1.52
1	C	29	LYS	CD-CE	13.98	1.86	1.51
2	B	207	LYS	CB-CG	11.26	1.82	1.52
2	D	212	LYS	CG-CD	8.62	1.81	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	203	GLU	N-CA-CB	10.64	129.75	110.60
2	D	209	LYS	CB-CG-CD	-7.72	91.53	111.60
2	B	162	LEU	CA-CB-CG	6.62	130.51	115.30
2	B	203	GLU	CB-CA-C	-5.93	98.54	110.40
2	B	207	LYS	CA-CB-CG	5.77	126.09	113.40

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	501	0	489	13	0
1	C	456	0	449	5	0
2	B	1369	0	1314	30	0
2	D	1366	0	1306	13	0
3	B	1	0	0	0	0
4	A	95	0	0	3	0
4	B	215	0	0	6	0
4	C	69	0	0	0	0
4	D	198	0	0	6	0
All	All	4270	0	3558	51	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 51 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:207:LYS:HB3	2:B:208:GLY:CA	1.64	1.26
2:B:207:LYS:CB	2:B:208:GLY:HA3	1.71	1.21
2:B:95[A]:THR:HG21	4:B:701:HOH:O	1.57	1.04
2:D:155:ARG:HD2	4:D:323:HOH:O	1.61	0.99
1:A:13[A]:THR:HG21	1:A:62:CYS:O	1.82	0.79

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	64/73 (88%)	64 (100%)	0	0	100	100
1	C	57/73 (78%)	56 (98%)	1 (2%)	0	100	100
2	B	168/168 (100%)	163 (97%)	5 (3%)	0	100	100
2	D	167/168 (99%)	158 (95%)	9 (5%)	0	100	100
All	All	456/482 (95%)	441 (97%)	15 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	55/61 (90%)	54 (98%)	1 (2%)	59	26
1	C	50/61 (82%)	50 (100%)	0	100	100
2	B	144/141 (102%)	136 (94%)	8 (6%)	21	1
2	D	143/141 (101%)	140 (98%)	3 (2%)	53	19
All	All	392/404 (97%)	380 (97%)	12 (3%)	42	9

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	228[A]	LYS
2	B	228[B]	LYS
2	D	168	ARG
2	D	127	PRO
2	B	155	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	197	HIS
2	D	232	ASN
2	B	231	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	232	ASN
1	C	22	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NRQ	D	65	2	23,24,25	4.44	6 (26%)	23,32,34	5.05	8 (34%)
2	NRQ	B	65	2	23,24,25	4.38	6 (26%)	23,32,34	4.98	8 (34%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NRQ	D	65	2	-	4/9/31/32	0/2/2/2
2	NRQ	B	65	2	-	2/9/31/32	0/2/2/2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	65	NRQ	CB2-CA2	18.26	1.50	1.35
2	D	65	NRQ	CB2-CA2	18.21	1.50	1.35
2	D	65	NRQ	CA2-C2	-9.03	1.39	1.48
2	B	65	NRQ	CA2-C2	-7.80	1.40	1.48
2	D	65	NRQ	OH-CZ	-3.43	1.29	1.37

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	65	NRQ	CA2-C2-N3	16.22	111.04	103.37
2	B	65	NRQ	CB2-CA2-C2	14.93	140.10	122.28
2	D	65	NRQ	CB2-CA2-C2	13.24	138.08	122.28
2	B	65	NRQ	CB2-CA2-N2	-11.95	112.25	128.83
2	B	65	NRQ	CA2-C2-N3	11.92	109.00	103.37

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	65	NRQ	N2-CA2-CB2-CG2
2	B	65	NRQ	C2-CA2-CB2-CG2
2	D	65	NRQ	N2-CA2-CB2-CG2
2	D	65	NRQ	C2-CA2-CB2-CG2
2	D	65	NRQ	CB1-CG1-SD-CE

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	65	NRQ	1	0
2	B	65	NRQ	3	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	64/73 (87%)	-0.11	0	100	100	13, 19, 24, 27	0
1	C	59/73 (80%)	-0.04	1 (1%)	70	70	15, 18, 24, 32	1 (1%)
2	B	167/168 (99%)	-0.17	2 (1%)	79	80	13, 16, 28, 41	2 (1%)
2	D	167/168 (99%)	-0.18	2 (1%)	79	80	14, 18, 28, 35	2 (1%)
All	All	457/482 (94%)	-0.15	5 (1%)	80	82	13, 17, 28, 41	5 (1%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	4	PHE	3.5
2	B	166	GLY	3.4
2	D	113	ASP	2.9
2	B	211	TYR	2.4
2	D	207	LYS	2.4

6.2 Non-standard residues in protein, DNA, RNA chains (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NRQ	D	65	23/24	0.93	0.08	17,20,22,27	0
2	NRQ	B	65	23/24	0.94	0.09	15,18,22,26	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CL	B	601	1/1	0.99	0.08	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.