Full wwPDB X-ray Structure Validation Report ⓘ

Jul 25, 2023 – 01:35 pm BST

PDB ID	:	8AM4
Title	:	Cl-rsEGFP2 Long Wavelength Structure
Authors	:	Orr, C.M.; Fadini, A.; van Thor, J.
Deposited on	:	2022-08-02
Resolution	:	2.02 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

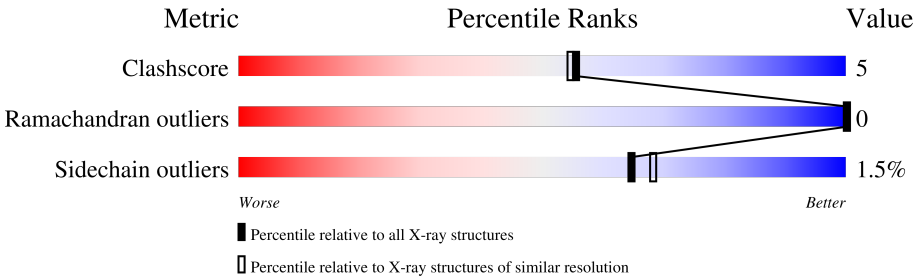
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.02 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))
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)

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%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	250	79% 12% 8%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1822 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 Green fluorescent protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	Cl	N	O	S			
1	A	230	1822	1159	1	304	351	7	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P42212
A	2	ARG	-	expression tag	UNP P42212
A	3	GLY	-	expression tag	UNP P42212
A	4	SER	-	expression tag	UNP P42212
A	5	HIS	-	expression tag	UNP P42212
A	6	HIS	-	expression tag	UNP P42212
A	7	HIS	-	expression tag	UNP P42212
A	8	HIS	-	expression tag	UNP P42212
A	9	HIS	-	expression tag	UNP P42212
A	10	HIS	-	expression tag	UNP P42212
A	11	THR	-	expression tag	UNP P42212
A	12	ASP	-	expression tag	UNP P42212
A	13	PRO	-	expression tag	UNP P42212
A	15	VAL	-	insertion	UNP P42212
A	78	LEU	PHE	engineered mutation	UNP P42212
A	?	-	SER	deletion	UNP P42212
A	?	-	TYR	deletion	UNP P42212
A	79	OHD	GLY	chromophore	UNP P42212
A	81	LEU	GLN	engineered mutation	UNP P42212
A	175	SER	VAL	engineered mutation	UNP P42212
A	218	LYS	ALA	engineered mutation	UNP P42212
A	243	LEU	HIS	engineered mutation	UNP P42212

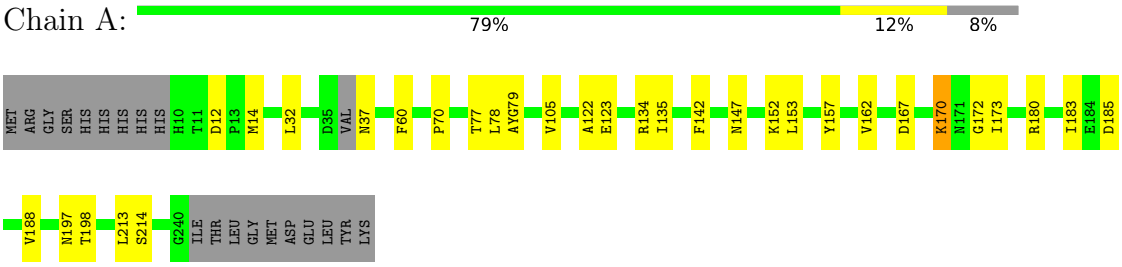
3 Residue-property plots (i)

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. 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. 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.

Note EDS failed to run properly.

- Molecule 1: Green fluorescent protein

Chain A:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.91Å 63.25Å 70.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.11 – 2.02	Depositor
% Data completeness (in resolution range)	80.7 (47.11-2.02)	Depositor
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.20 (at 2.03Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.196 , 0.257	Depositor
Wilson B-factor (Å^2)	30.8	Xtriage
Anisotropy	1.348	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1822	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.15% 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

5.1 Standard geometry

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

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	0.29	0/1841	0.48	0/2488

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1822	0	1736	18	0
All	All	1822	0	1736	18	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 5.

All (18) 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)
1:A:167:ASP:HB3	1:A:170:LYS:HD2	1.44	0.98
1:A:170:LYS:HG3	1:A:198:THR:HG22	1.65	0.78
1:A:70:PRO:HG2	1:A:153:LEU:HD23	1.84	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:173:ILE:HG13	1:A:197:ASN:HB2	1.86	0.56
1:A:147:ASN:HA	1:A:152:LYS:HD3	1.88	0.56
1:A:173:ILE:CG1	1:A:197:ASN:HB2	2.36	0.56
1:A:162:VAL:HG22	1:A:213:LEU:HB2	1.91	0.53
1:A:37:ASN:ND2	1:A:142:PHE:O	2.42	0.52
1:A:167:ASP:HB3	1:A:170:LYS:CD	2.28	0.50
1:A:77:THR:HG22	1:A:135:ILE:HG21	1.94	0.49
1:A:32:LEU:HD12	1:A:135:ILE:HD12	1.94	0.49
1:A:183:ILE:HG22	1:A:185:ASP:H	1.79	0.48
1:A:105:VAL:HG22	1:A:123:GLU:HG2	1.97	0.46
1:A:122:ALA:HA	1:A:134:ARG:O	2.16	0.46
1:A:12:ASP:OD1	1:A:14:MET:N	2.51	0.43
1:A:180:ARG:HG2	1:A:188:VAL:CG1	2.50	0.42
1:A:60:PHE:CE1	1:A:78:LEU:HB3	2.55	0.42
1:A:170:LYS:HD3	1:A:172:GLY:CA	2.51	0.40

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	223/250 (89%)	218 (98%)	5 (2%)	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	197/220 (90%)	194 (98%)	3 (2%)	65 68

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	157	TYR
1	A	170	LYS
1	A	214	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
1	OHD	A	79	1	22,22,23	5.15	9 (40%)	29,31,33	3.86	9 (31%)

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
1	OHD	A	79	1	-	2/8/27/28	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	79	OHD	CB2-CA2	16.82	1.49	1.35
1	A	79	OHD	O2-C2	12.05	1.48	1.23
1	A	79	OHD	CZ-CE2	9.67	1.48	1.39
1	A	79	OHD	O3-C3	4.17	1.43	1.19
1	A	79	OHD	C1-N2	3.86	1.37	1.32
1	A	79	OHD	C2-N3	-2.83	1.33	1.39
1	A	79	OHD	CE2-CL	2.56	1.79	1.73
1	A	79	OHD	CA2-C2	-2.16	1.46	1.48
1	A	79	OHD	CA1-C1	-2.05	1.48	1.51

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	79	OHD	CA2-C2-N3	12.62	109.34	103.37
1	A	79	OHD	O2-C2-CA2	-11.55	124.47	130.96
1	A	79	OHD	CB2-CA2-C2	5.72	129.11	122.28
1	A	79	OHD	O3-C3-CA3	-5.59	109.50	126.39
1	A	79	OHD	C2-N3-C1	-4.44	105.72	107.97
1	A	79	OHD	C2-CA2-N2	-4.18	106.00	108.93
1	A	79	OHD	CD2-CE2-CZ	-3.16	119.08	120.91
1	A	79	OHD	CG2-CB2-CA2	-2.94	126.34	129.94
1	A	79	OHD	CB2-CA2-N2	-2.84	124.89	128.83

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	79	OHD	C3-CA3-N3-C2
1	A	79	OHD	C3-CA3-N3-C1

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.