wwPDB Geometry-Only Validation Summary Report (i)

May 27, 2024 – 08:06 PM EDT

PDB ID	:	6L26
Title	:	Neutron crystal structure of the mutant green fluorescent protein (EGFP)
Authors	:	Adachi, M.; Shimizu, R.; Shibazaki, C.; Kagotani, Y.; Ostermann, A.; Schrader, T.E.
Deposited on	:	2019-10-02
Resolution	:	1.44 Å(reported)

This is a wwPDB Geometry-Only 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

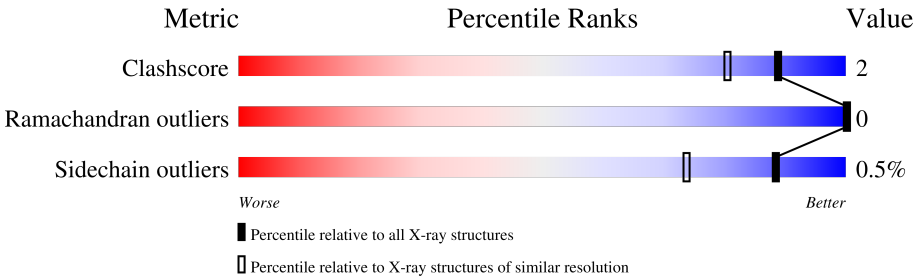
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
NEUTRON DIFFRACTION

The reported resolution of this entry is 1.44 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	2086 (1.46-1.42)
Ramachandran outliers	138981	2047 (1.46-1.42)
Sidechain outliers	138945	2047 (1.46-1.42)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	A	227	96% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4940 atoms, of which 1552 are hydrogens and 1152 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.

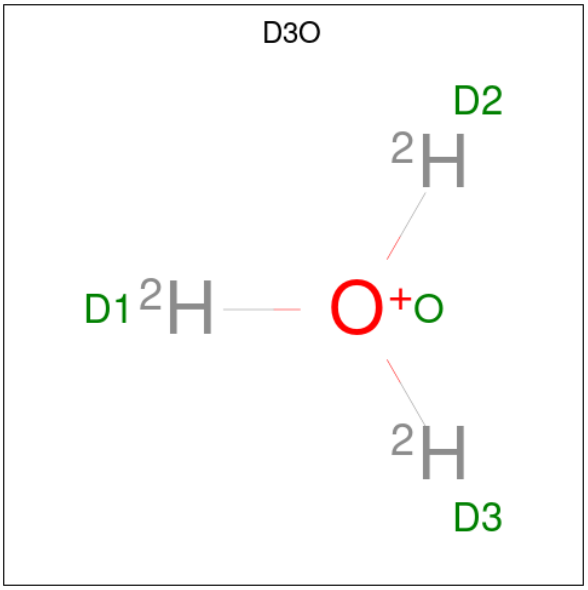
- Molecule 1 is a protein called Green fluorescent protein.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	D	H	N	O	S			
1	A	226	3804	1181	397	1552	310	357	7	0	117	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	MET	-	initiating methionine	UNP W6KDG8
A	48	SER	CYS	conflict	UNP W6KDG8
A	64	LEU	PHE	conflict	UNP W6KDG8
A	66	CRO	SER	chromophore	UNP W6KDG8
A	66	CRO	TYR	chromophore	UNP W6KDG8
A	66	CRO	GLY	chromophore	UNP W6KDG8
A	159	GLN	ASN	conflict	UNP W6KDG8
A	167	THR	ILE	conflict	UNP W6KDG8

- Molecule 2 is trideuteriooxidanium (three-letter code: D3O) (formula: D_3O) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total D O 4 3 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	372	Total D O 1132 752 380	0	9

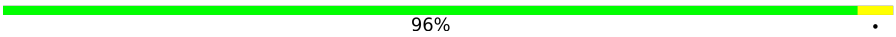
3 Residue-property plots (i)

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 was not executed.

- Molecule 1: Green fluorescent protein

Chain A:



4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CRO, DOD, D3O

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	0.66	6/2754 (0.2%)	0.83	2/3720 (0.1%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	17[A]	GLU	CD-OE1	-8.98	1.15	1.25
1	A	17[B]	GLU	CD-OE1	-8.98	1.15	1.25
1	A	17[A]	GLU	CD-OE2	6.42	1.32	1.25
1	A	17[B]	GLU	CD-OE2	6.42	1.32	1.25
1	A	222[A]	GLU	CD-OE1	-6.03	1.19	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	122[A]	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	122[B]	ARG	NE-CZ-NH1	5.32	122.96	120.30

There are no chirality outliers.

There are no planarity outliers.

4.2 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	2252	1552	986	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	4	0	0	0	0
3	A	1132	0	0	5	2
All	All	3388	1552	986	7	2

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 2.

The worst 5 of 7 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:105[B]:ASN:ND2	3:A:361:DOD:O	1.92	1.02
1:A:117[A]:ASP:OD1	3:A:491:DOD:O	2.14	0.65
1:A:80:ARG:NH2	3:A:576:DOD:O	2.34	0.61
1:A:133:ASP:OD1	1:A:133:ASP:N	2.44	0.47

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:485:DOD:O	3:A:502:DOD:O[4_455]	1.98	0.22
3:A:591:DOD:O	3:A:610:DOD:O[4_455]	2.18	0.02

4.3 Torsion angles (i)

4.3.1 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	335/227 (148%)	327 (98%)	8 (2%)	0	100 100

There are no Ramachandran outliers to report.

4.3.2 Protein sidechains

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	295/198 (149%)	293 (99%)	2 (1%)	84 64

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	159[A]	GLN
1	A	159[B]	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	157	GLN

4.3.3 RNA

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
1	CRO	A	66[B]	-	23,23,24	3.43	8 (34%)	30,32,34	3.01	9 (30%)
1	CRO	A	66[A]	-	23,23,24	3.41	7 (30%)	30,32,34	3.00	9 (30%)

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	CRO	A	66[B]	-	-	0/12/31/32	0/2/2/2
1	CRO	A	66[A]	-	-	0/12/31/32	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66[A]	CRO	CB2-CA2	12.79	1.45	1.35
1	A	66[B]	CRO	CB2-CA2	12.79	1.45	1.35
1	A	66[A]	CRO	C1-N2	-6.21	1.23	1.32
1	A	66[B]	CRO	C1-N2	-6.21	1.23	1.32
1	A	66[A]	CRO	CA1-C1	4.05	1.57	1.51

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	66[A]	CRO	N3-C1-N2	8.86	117.59	111.45
1	A	66[B]	CRO	N3-C1-N2	8.86	117.59	111.45
1	A	66[A]	CRO	C2-N3-C1	-6.45	104.70	107.97
1	A	66[B]	CRO	C2-N3-C1	-6.45	104.70	107.97
1	A	66[A]	CRO	CA1-C1-N3	-6.13	117.40	124.75

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	66[B]	CRO	1	0

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is modelled with single atom - 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.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.