wwPDB X-ray Structure Validation Summary Report (i)

Nov 7, 2023 – 03:50 am GMT

PDB ID	:	6GOY
Title	:	Structure of mEos4b in the green fluorescent state
Authors	:	De Zitter, E.; Adam, V.; Byrdin, M.; Van Meervelt, L.; Dedecker, P.; Bourgeois, D.
Deposited on	:	2018-06-04
Resolution	:	1.65 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
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)	:	2.36

PERCENTILES INFOmissingINFO

1 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2192 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 to red photoconvertible GFP-like protein EosFP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	221	1899	1199	326	363	11	0	15	0

There are 49 discrepancies between the modelled and reference sequences:

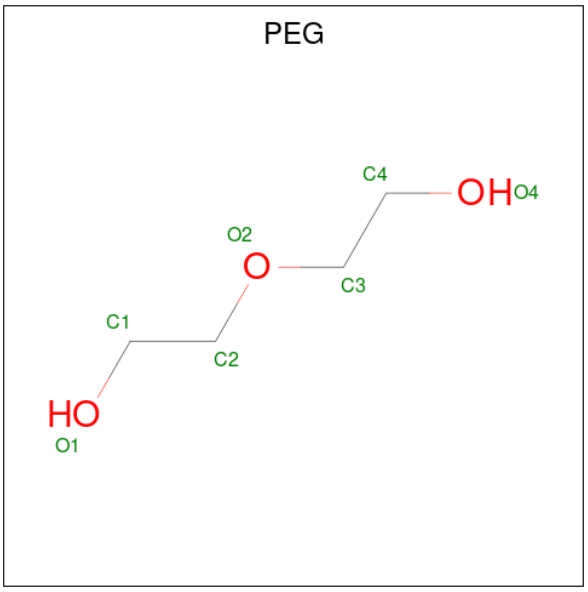
Chain	Residue	Modelled	Actual	Comment	Reference
A	-32	MET	-	initiating methionine	UNP Q5S6Z9
A	-31	ARG	-	expression tag	UNP Q5S6Z9
A	-30	GLY	-	expression tag	UNP Q5S6Z9
A	-29	SER	-	expression tag	UNP Q5S6Z9
A	-28	HIS	-	expression tag	UNP Q5S6Z9
A	-27	HIS	-	expression tag	UNP Q5S6Z9
A	-26	HIS	-	expression tag	UNP Q5S6Z9
A	-25	HIS	-	expression tag	UNP Q5S6Z9
A	-24	HIS	-	expression tag	UNP Q5S6Z9
A	-23	HIS	-	expression tag	UNP Q5S6Z9
A	-22	GLY	-	expression tag	UNP Q5S6Z9
A	-21	MET	-	expression tag	UNP Q5S6Z9
A	-20	ALA	-	expression tag	UNP Q5S6Z9
A	-19	SER	-	expression tag	UNP Q5S6Z9
A	-18	MET	-	expression tag	UNP Q5S6Z9
A	-17	THR	-	expression tag	UNP Q5S6Z9
A	-16	GLY	-	expression tag	UNP Q5S6Z9
A	-15	GLY	-	expression tag	UNP Q5S6Z9
A	-14	GLN	-	expression tag	UNP Q5S6Z9
A	-13	GLN	-	expression tag	UNP Q5S6Z9
A	-12	MET	-	expression tag	UNP Q5S6Z9
A	-11	GLY	-	expression tag	UNP Q5S6Z9
A	-10	ARG	-	expression tag	UNP Q5S6Z9
A	-9	ASP	-	expression tag	UNP Q5S6Z9
A	-8	LEU	-	expression tag	UNP Q5S6Z9
A	-7	TYR	-	expression tag	UNP Q5S6Z9
A	-6	ASP	-	expression tag	UNP Q5S6Z9

Continued on next page...

Continued from previous page...

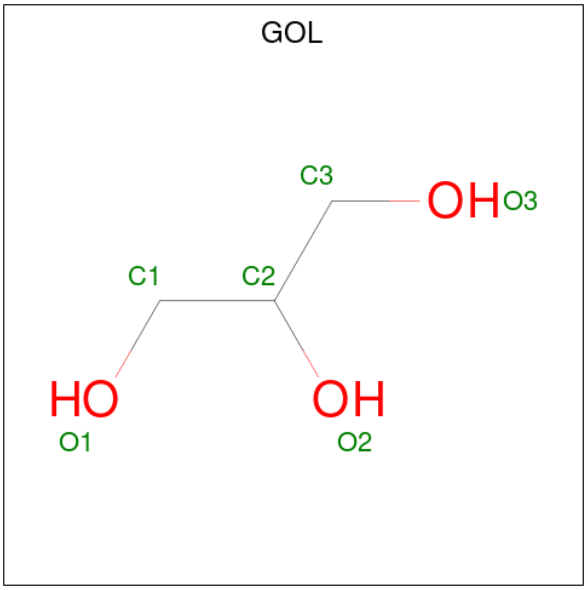
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	ASP	-	expression tag	UNP Q5S6Z9
A	-4	ASP	-	expression tag	UNP Q5S6Z9
A	-3	ASP	-	expression tag	UNP Q5S6Z9
A	-2	LYS	-	expression tag	UNP Q5S6Z9
A	-1	ASP	-	expression tag	UNP Q5S6Z9
A	0	PRO	-	expression tag	UNP Q5S6Z9
A	9	ARG	LYS	engineered mutation	UNP Q5S6Z9
A	11	LYS	ASN	engineered mutation	UNP Q5S6Z9
A	34	TYR	PHE	engineered mutation	UNP Q5S6Z9
A	39	THR	SER	engineered mutation	UNP Q5S6Z9
A	64	5SQ	HIS	chromophore	UNP Q5S6Z9
A	64	5SQ	TYR	chromophore	UNP Q5S6Z9
A	64	5SQ	GLY	chromophore	UNP Q5S6Z9
A	69	VAL	ALA	engineered mutation	UNP Q5S6Z9
A	70	LYS	GLU	engineered mutation	UNP Q5S6Z9
A	74	ASN	HIS	engineered mutation	UNP Q5S6Z9
A	102	ASN	ILE	engineered mutation	UNP Q5S6Z9
A	121	TYR	HIS	engineered mutation	UNP Q5S6Z9
A	123	THR	VAL	engineered mutation	UNP Q5S6Z9
A	158	GLU	THR	engineered mutation	UNP Q5S6Z9
A	189	ALA	TYR	engineered mutation	UNP Q5S6Z9
A	195	ALA	CYS	engineered mutation	UNP Q5S6Z9

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



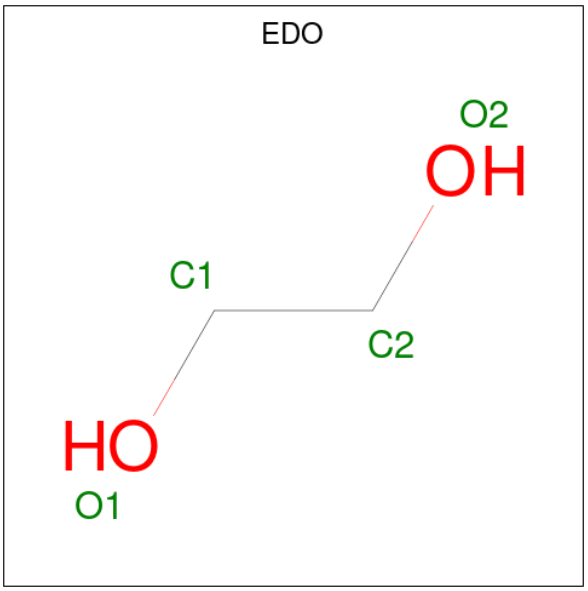
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	7	4	3	0	0
			Total	C	O		
2	A	1	7	4	3	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	6	3	3	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	269	269	269	0	4

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

2 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	38.60Å 58.09Å 103.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.59 – 1.65 38.59 – 1.65	Depositor EDS
% Data completeness (in resolution range)	97.1 (38.59-1.65) 97.1 (38.59-1.65)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 1.65Å)	Xtriage
Refinement program	PHENIX 1.11	Depositor
R, R_{free}	0.168 , 0.204 0.167 , 0.204	Depositor DCC
R_{free} test set	1397 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	16.9	Xtriage
Anisotropy	0.428	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2192	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.92% 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.

3 Model quality (i)

3.1 Standard geometry (i)

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

3.2 Too-close contacts (i)

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

3.3 Torsion angles (i)

3.3.1 Protein backbone (i)

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

3.3.2 Protein sidechains (i)

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

3.3.3 RNA (i)

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

3.4 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	5SQ	A	64	1	23,27,28	4.32	6 (26%)	29,37,39	2.54	6 (20%)

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	5SQ	A	64	1	-	4/12/31/32	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	64	5SQ	CB2-CA2	17.31	1.49	1.35
1	A	64	5SQ	CA2-C2	-8.79	1.39	1.48
1	A	64	5SQ	OH-CZ1	-4.57	1.26	1.37
1	A	64	5SQ	C1-N2	2.79	1.36	1.32
1	A	64	5SQ	O2-C2	2.69	1.28	1.23

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	64	5SQ	CA2-C2-N3	9.79	108.00	103.37
1	A	64	5SQ	O2-C2-CA2	-5.77	127.72	130.96
1	A	64	5SQ	N3-C1-N2	-3.72	108.88	111.45
1	A	64	5SQ	C2-CA2-N2	-3.42	106.54	108.93
1	A	64	5SQ	CA2-N2-C1	2.83	107.86	105.77

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	64	5SQ	C3-CA3-N3-C2
1	A	64	5SQ	CA1-CB1-CG1-C2H
1	A	64	5SQ	CA1-CB1-CG1-N1H
1	A	64	5SQ	N2-CA2-CB2-CG2

There are no ring outliers.

No monomer is involved in short contacts.

3.5 Carbohydrates (i)

There are no monosaccharides in this entry.

3.6 Ligand geometry (i)

4 ligands 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	PEG	A	302	-	6,6,6	0.46	0	5,5,5	0.20	0
3	GOL	A	303	-	5,5,5	0.35	0	5,5,5	0.56	0
2	PEG	A	301	-	6,6,6	0.47	0	5,5,5	0.29	0
4	EDO	A	304	-	3,3,3	0.46	0	2,2,2	0.31	0

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	PEG	A	302	-	-	3/4/4/4	-
3	GOL	A	303	-	-	2/4/4/4	-
2	PEG	A	301	-	-	2/4/4/4	-
4	EDO	A	304	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	302	PEG	C1-C2-O2-C3
2	A	302	PEG	O1-C1-C2-O2
3	A	303	GOL	O1-C1-C2-C3
3	A	303	GOL	O1-C1-C2-O2
2	A	302	PEG	O2-C3-C4-O4

There are no ring outliers.

No monomer is involved in short contacts.

3.7 Other polymers (i)

There are no such residues in this entry.

3.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

4 Fit of model and data (i)

4.1 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	220/257 (85%)	-0.23	6 (2%) 54 55	10, 16, 33, 51	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	185	LYS	2.6
1	A	182	LYS	2.4
1	A	-4	ASP	2.3
1	A	-3	ASP	2.2
1	A	218	SER	2.0

4.2 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
1	5SQ	A	64	25/26	0.95	0.08	8,11,15,16	0

4.3 Carbohydrates (i)

There are no monosaccharides in this entry.

4.4 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
2	PEG	A	302	7/7	0.82	0.12	24,32,43,43	0
3	GOL	A	303	6/6	0.82	0.15	27,38,46,52	0
4	EDO	A	304	4/4	0.85	0.14	31,40,44,47	0
2	PEG	A	301	7/7	0.89	0.10	31,37,44,44	0

4.5 Other polymers (i)

There are no such residues in this entry.