



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 6, 2023 – 11:42 pm GMT

PDB ID : 6GP0
Title : Structure of mEos4b in the red fluorescent state
Authors : De Zitter, E.; Adam, V.; Byrdin, M.; Van Meervelt, L.; Dedecker, P.; Bourgeois, D.
Deposited on : 2018-06-04
Resolution : 1.50 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 ⓘ 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 : 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

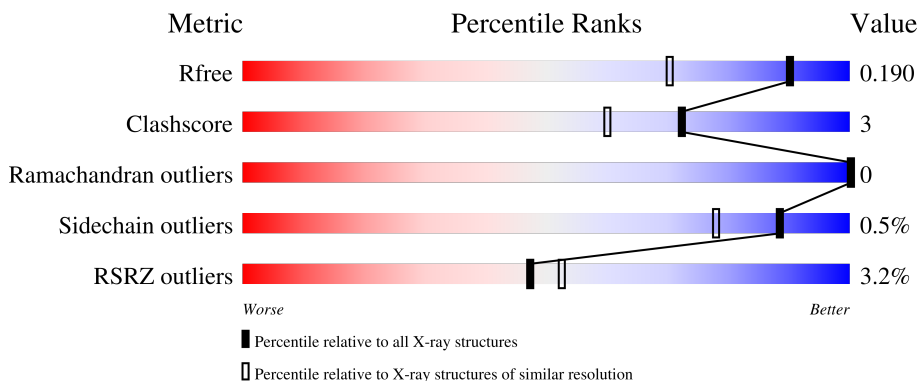
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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 $\leq 5\%$. 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	259	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2304 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	222	1957	1238	336	372	11	0	18	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	conflict	UNP Q5S6Z9
A	11	LYS	ASN	conflict	UNP Q5S6Z9
A	34	TYR	PHE	conflict	UNP Q5S6Z9
A	39	THR	SER	conflict	UNP Q5S6Z9
A	62	RC7	HIS	chromophore	UNP Q5S6Z9
A	62	RC7	TYR	chromophore	UNP Q5S6Z9
A	62	RC7	GLY	chromophore	UNP Q5S6Z9
A	67	VAL	ALA	conflict	UNP Q5S6Z9
A	68	LYS	GLU	conflict	UNP Q5S6Z9
A	72	ASN	HIS	conflict	UNP Q5S6Z9
A	100	ASN	ILE	conflict	UNP Q5S6Z9
A	119	TYR	HIS	conflict	UNP Q5S6Z9
A	121	THR	VAL	conflict	UNP Q5S6Z9
A	156	GLU	THR	conflict	UNP Q5S6Z9
A	187	ALA	TYR	conflict	UNP Q5S6Z9
A	193	ALA	CYS	conflict	UNP Q5S6Z9

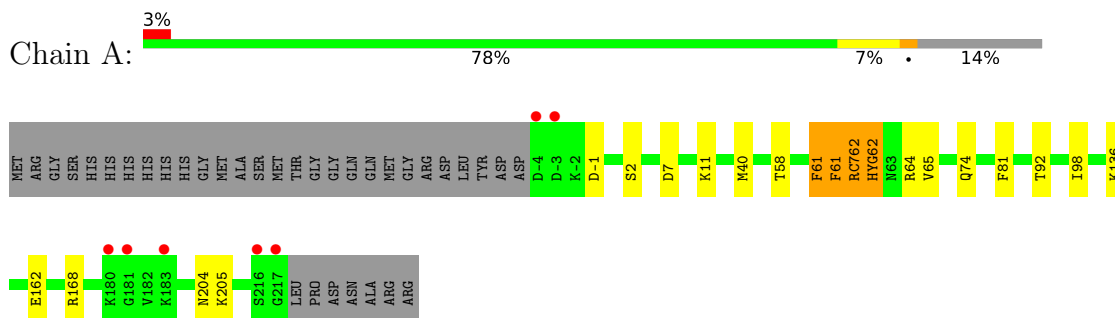
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	347	Total O 347 347	0	6

3 Residue-property plots

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: Green to red photoconvertible GFP-like protein EosFP



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	38.61Å 57.90Å 102.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.14 – 1.50 36.14 – 1.50	Depositor EDS
% Data completeness (in resolution range)	90.1 (36.14-1.50) 96.3 (36.14-1.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.49 (at 1.50Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.154 , 0.190 0.154 , 0.190	Depositor DCC
R_{free} test set	1827 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	15.0	Xtrriage
Anisotropy	0.257	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	2304	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: 5SQ, NFA, RC7

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	0.30	0/1941	0.55	0/2615

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 [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	1957	0	1807	13	0
2	A	347	0	0	8	1
All	All	2304	0	1807	13	1

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 3.

The worst 5 of 13 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:168[B]:ARG:NH1	2:A:307:HOH:O	2.28	0.67
1:A:136:LYS:NZ	2:A:309:HOH:O	2.31	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:7[B]:ASP:OD2	2:A:302:HOH:O	2.17	0.57
1:A:205:LYS:NZ	2:A:315:HOH:O	2.36	0.56
1:A:162:GLU:OE2	2:A:303:HOH:O	2.19	0.51

All (1) 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)
2:A:306:HOH:O	2:A:553:HOH:O[4_455]	2.12	0.08

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	233/259 (90%)	232 (100%)	1 (0%)	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	202/217 (93%)	201 (100%)	1 (0%)	88 78

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	64	ARG

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	204	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

3 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	RC7	A	62[A]	1	24,26,27	4.03	7 (29%)	27,35,37	2.92	8 (29%)
1	NFA	A	61[A]	1	12,12,12	0.38	0	15,15,15	1.14	1 (6%)
1	5SQ	A	62[B]	1	23,27,28	4.38	6 (26%)	29,37,39	3.09	7 (24%)

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	RC7	A	62[A]	1	-	2/9/28/29	0/3/3/3
1	NFA	A	61[A]	1	-	2/8/8/8	0/1/1/1
1	5SQ	A	62[B]	1	-	6/12/31/32	0/3/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	62[B]	5SQ	CB2-CA2	16.82	1.49	1.35
1	A	62[A]	RC7	CB2-CA2	15.73	1.48	1.35
1	A	62[B]	5SQ	CA2-C2	-10.25	1.38	1.48
1	A	62[A]	RC7	CA2-C2	-9.34	1.39	1.48
1	A	62[B]	5SQ	OH-CZ1	-4.77	1.25	1.37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	62[B]	5SQ	CA2-C2-N3	10.16	108.18	103.37
1	A	62[B]	5SQ	O2-C2-CA2	-10.08	125.30	130.96
1	A	62[A]	RC7	CA2-C2-N3	8.50	107.39	103.37
1	A	62[A]	RC7	O2-C2-CA2	-7.24	126.89	130.96
1	A	62[A]	RC7	N3-C1-N2	-7.06	109.42	112.87

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	61[A]	NFA	O-C-CA-CB
1	A	61[A]	NFA	NXT-C-CA-CB
1	A	62[B]	5SQ	C3-CA3-N3-C2
1	A	62[B]	5SQ	CA1-CB1-CG1-C2H
1	A	62[B]	5SQ	CA1-CB1-CG1-N1H

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	62[A]	RC7	1	0
1	A	61[A]	NFA	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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	220/259 (84%)	-0.16	7 (3%) 47 52	11, 15, 31, 54	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	-4	ASP	4.8
1	A	217	GLY	3.3
1	A	181	GLY	2.8
1	A	183	LYS	2.6
1	A	216	SER	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
1	NFA	A	61[A]	12/12	0.95	0.10	12,13,15,19	12
1	RC7	A	62[A]	24/25	0.95	0.08	11,14,17,17	24
1	5SQ	A	62[B]	25/26	0.95	0.08	11,14,17,18	25

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.