wwPDB X-ray Structure Validation Summary Report (i)

Nov 6, 2023 – 08:22 am GMT

PDB ID	:	2BTJ
Title	:	Fluorescent Protein EosFP - red form
Authors	:	Nar, H.; Nienhaus, K.; Wiedenmann, J.; Nienhaus, G.U.
Deposited on	:	2005-06-01
Resolution	:	2.00 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:

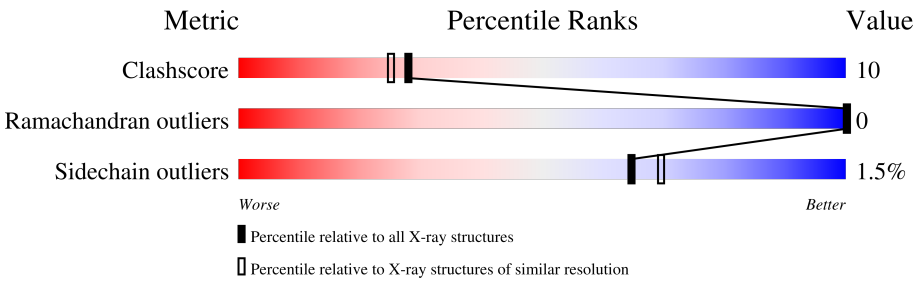
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	aa	61	100%
1	bb	61	100%
1	cc	61	100%
1	dd	61	98% ·
2	A	159	81% 18% ·
2	B	159	78% 22%
2	C	159	79% 21% ·
2	D	159	81% 19% ·

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7796 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	aa	61	460	292	78	87	3	13	0	0
1	bb	61	460	292	78	87	3	13	0	0
1	cc	61	460	292	78	87	3	13	0	0
1	dd	61	460	292	78	87	3	13	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
aa	1	GLY	-	expression tag	UNP Q5S6Z9
aa	2	GLY	-	expression tag	UNP Q5S6Z9
bb	1	GLY	-	expression tag	UNP Q5S6Z9
bb	2	GLY	-	expression tag	UNP Q5S6Z9
cc	1	GLY	-	expression tag	UNP Q5S6Z9
cc	2	GLY	-	expression tag	UNP Q5S6Z9
dd	1	GLY	-	expression tag	UNP Q5S6Z9
dd	2	GLY	-	expression tag	UNP Q5S6Z9

- Molecule 2 is a protein called Green to red photoconvertible GFP-like protein EosFP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	159	1315	843	225	240	7	10	0	0
2	B	159	1315	843	225	240	7	7	0	0
2	C	159	1315	843	225	240	7	10	0	0
2	D	159	1315	843	225	240	7	5	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	64	RC7	HIS	chromophore	UNP Q5S6Z9
A	64	RC7	TYR	chromophore	UNP Q5S6Z9
A	64	RC7	GLY	chromophore	UNP Q5S6Z9
B	64	RC7	HIS	chromophore	UNP Q5S6Z9
B	64	RC7	TYR	chromophore	UNP Q5S6Z9
B	64	RC7	GLY	chromophore	UNP Q5S6Z9
C	64	RC7	HIS	chromophore	UNP Q5S6Z9
C	64	RC7	TYR	chromophore	UNP Q5S6Z9
C	64	RC7	GLY	chromophore	UNP Q5S6Z9
D	64	RC7	HIS	chromophore	UNP Q5S6Z9
D	64	RC7	TYR	chromophore	UNP Q5S6Z9
D	64	RC7	GLY	chromophore	UNP Q5S6Z9

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	176	176	176	0	0
			Total	O		
3	B	185	185	185	0	0
			Total	O		
3	C	173	173	173	0	0
			Total	O		
3	D	162	162	162	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. 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 to red photoconvertible GFP-like protein EosFP

Chain aa: 100%

There are no outlier residues recorded for this chain.

• Molecule 1: Green to red photoconvertible GFP-like protein EosFP

Chain bb: 100%

There are no outlier residues recorded for this chain.

• Molecule 1: Green to red photoconvertible GFP-like protein EosFP

Chain cc: 100%

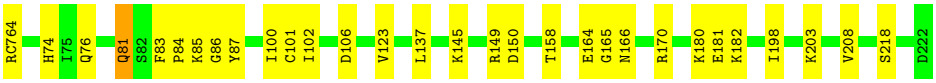
There are no outlier residues recorded for this chain.

• Molecule 1: Green to red photoconvertible GFP-like protein EosFP

Chain dd: 98%

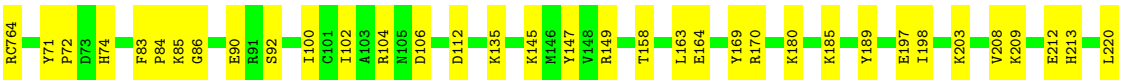
• Molecule 2: Green to red photoconvertible GFP-like protein EosFP

Chain A: 81% 18%



• Molecule 2: Green to red photoconvertible GFP-like protein EosFP

Chain B: 78% 22%



• Molecule 2: Green to red photoconvertible GFP-like protein EosFP

Chain C:

• Molecule 2: Green to red photoconvertible GFP-like protein EosFP

Chain D:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.59Å 105.51Å 119.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00	Depositor
% Data completeness (in resolution range)	99.9 (50.00-2.00)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.198 , 0.232	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7796	wwPDB-VP
Average B, all atoms ($Å^2$)	27.0	wwPDB-VP

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: RC7, NFA

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	aa	0.33	0/457	0.60	0/613
1	bb	0.34	0/457	0.62	0/613
1	cc	0.34	0/457	0.60	0/613
1	dd	0.35	0/457	0.62	0/613
2	A	0.35	0/1328	0.66	0/1793
2	B	0.35	0/1328	0.64	0/1793
2	C	0.34	0/1328	0.63	0/1793
2	D	0.34	0/1328	0.66	0/1793
All	All	0.34	0/7140	0.64	0/9624

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	aa	460	0	451	0	0
1	bb	460	0	451	0	0
1	cc	460	0	451	0	0
1	dd	460	0	451	0	0
2	A	1315	0	1252	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1315	0	1252	29	0
2	C	1315	0	1252	28	0
2	D	1315	0	1252	32	0
3	A	176	0	0	2	0
3	B	185	0	0	10	0
3	C	173	0	0	8	0
3	D	162	0	0	12	0
All	All	7796	0	6812	101	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 10.

The worst 5 of 101 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:203:LYS:HE2	2:B:203:LYS:N	1.98	0.79
2:B:112:ASP:HB2	3:B:2092:HOH:O	1.85	0.76
2:C:112:ASP:HB2	3:C:2086:HOH:O	1.85	0.74
2:D:112:ASP:HB2	3:D:2081:HOH:O	1.86	0.74
2:D:110:GLU:HG2	3:D:2076:HOH:O	1.87	0.73

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	aa	59/61 (97%)	58 (98%)	1 (2%)	0	100	100
1	bb	59/61 (97%)	58 (98%)	1 (2%)	0	100	100
1	cc	59/61 (97%)	59 (100%)	0	0	100	100
1	dd	59/61 (97%)	57 (97%)	2 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	A	156/159 (98%)	153 (98%)	3 (2%)	0	100	100
2	B	156/159 (98%)	152 (97%)	4 (3%)	0	100	100
2	C	156/159 (98%)	152 (97%)	4 (3%)	0	100	100
2	D	156/159 (98%)	151 (97%)	5 (3%)	0	100	100
All	All	860/880 (98%)	840 (98%)	20 (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	aa	47/47 (100%)	47 (100%)	0	100	100
1	bb	47/47 (100%)	47 (100%)	0	100	100
1	cc	47/47 (100%)	47 (100%)	0	100	100
1	dd	47/47 (100%)	46 (98%)	1 (2%)	53	57
2	A	138/138 (100%)	135 (98%)	3 (2%)	52	55
2	B	138/138 (100%)	135 (98%)	3 (2%)	52	55
2	C	138/138 (100%)	136 (99%)	2 (1%)	67	72
2	D	138/138 (100%)	136 (99%)	2 (1%)	67	72
All	All	740/740 (100%)	729 (98%)	11 (2%)	65	69

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	145	LYS
1	dd	7	ASP
2	D	145	LYS
2	D	102	ILE
2	B	185	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	81	GLN
2	D	76	GLN
1	cc	38	GLN
2	D	74	HIS
2	B	213	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)

8 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	RC7	B	64	2	24,26,27	1.50	4 (16%)	27,35,37	2.25	8 (29%)
2	RC7	D	64	2	24,26,27	1.49	4 (16%)	27,35,37	2.37	9 (33%)
1	NFA	dd	61	1	12,12,12	0.40	0	15,15,15	0.46	0
2	RC7	C	64	2	24,26,27	1.50	4 (16%)	27,35,37	2.34	8 (29%)
1	NFA	aa	61	1	12,12,12	0.43	0	15,15,15	0.64	0
1	NFA	cc	61	1	12,12,12	0.48	0	15,15,15	0.49	0
2	RC7	A	64	2	24,26,27	1.55	4 (16%)	27,35,37	2.43	8 (29%)
1	NFA	bb	61	1	12,12,12	0.48	0	15,15,15	0.35	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	RC7	B	64	2	-	1/9/28/29	0/3/3/3
2	RC7	D	64	2	-	1/9/28/29	0/3/3/3
1	NFA	dd	61	1	-	4/8/8/8	0/1/1/1
2	RC7	C	64	2	-	1/9/28/29	0/3/3/3
1	NFA	aa	61	1	-	4/8/8/8	0/1/1/1
1	NFA	cc	61	1	-	4/8/8/8	0/1/1/1
2	RC7	A	64	2	-	1/9/28/29	0/3/3/3
1	NFA	bb	61	1	-	4/8/8/8	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	64	RC7	CB2-CA2	4.77	1.39	1.35
2	C	64	RC7	CB2-CA2	4.77	1.39	1.35
2	B	64	RC7	CB2-CA2	4.59	1.39	1.35
2	D	64	RC7	CB2-CA2	4.53	1.38	1.35
2	A	64	RC7	CD2-CG2	2.67	1.44	1.39

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	64	RC7	O2-C2-CA2	-7.55	126.72	130.96
2	D	64	RC7	O2-C2-CA2	-7.21	126.91	130.96
2	C	64	RC7	O2-C2-CA2	-7.10	126.97	130.96
2	B	64	RC7	O2-C2-CA2	-6.88	127.10	130.96
2	A	64	RC7	CA2-C2-N3	4.37	105.44	103.37

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	aa	61	NFA	NXT-C-CA-N
1	aa	61	NFA	O-C-CA-CB
1	aa	61	NFA	NXT-C-CA-CB
2	A	64	RC7	C3-CA3-N3-C2
1	bb	61	NFA	O-C-CA-N

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 (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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.