



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 7, 2025 – 10:09 AM JST

PDB ID : 9J0R / pdb_00009j0r
Title : Structure of pcStar in the green fluorescent state
Authors : Zheng, S.P.; Shi, X.R.
Deposited on : 2024-08-02
Resolution : 1.66 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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.42

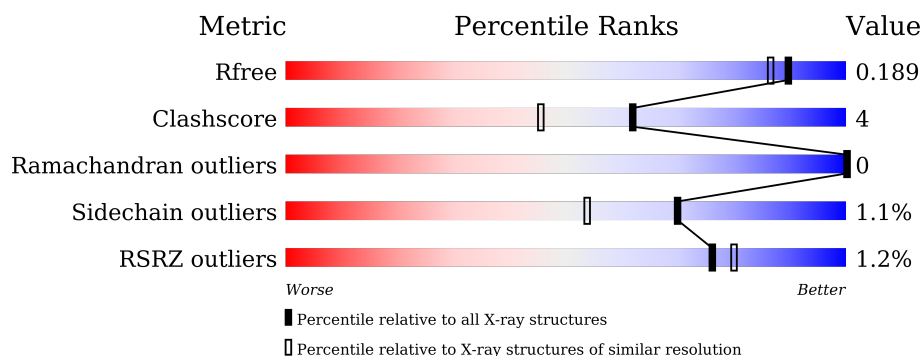
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.66 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}	164625	2328 (1.66-1.66)
Clashscore	180529	2515 (1.66-1.66)
Ramachandran outliers	177936	2475 (1.66-1.66)
Sidechain outliers	177891	2475 (1.66-1.66)
RSRZ outliers	164620	2328 (1.66-1.66)

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	228	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0; width: 100%; text-align: center;">%</div> <div style="position: absolute; top: 10px; left: 0; width: 100%; text-align: center;">86% 8% 6%</div> </div> </div>
1	B	228	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0; width: 100%; text-align: center;">%</div> <div style="position: absolute; top: 10px; left: 0; width: 100%; text-align: center;">86% 9% 6%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3957 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
1	A	215	Total	C	N	O	S	0	1	0
			1740	1112	296	321	11			
1	B	215	Total	C	N	O	S	0	2	0
			1743	1114	296	322	11			

There are 36 discrepancies between the modelled and reference sequences:

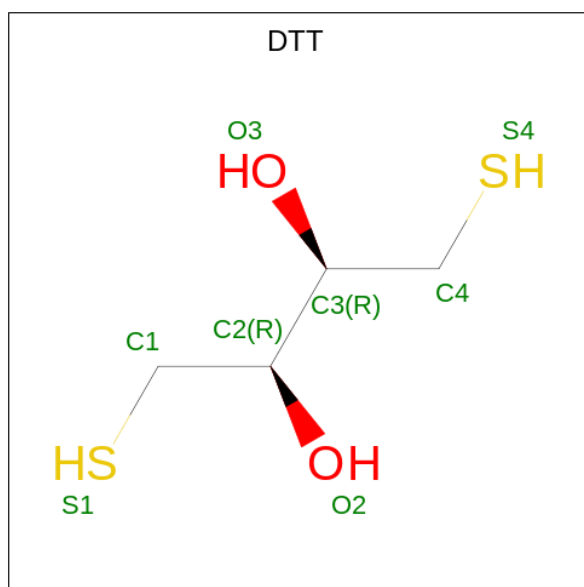
Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q5S6Z9
A	-2	PRO	-	expression tag	UNP Q5S6Z9
A	-1	GLU	-	expression tag	UNP Q5S6Z9
A	0	PHE	-	expression tag	UNP Q5S6Z9
A	11	LYS	ASN	engineered mutation	UNP Q5S6Z9
A	28	GLU	ASP	engineered mutation	UNP Q5S6Z9
A	64	5SQ	HIS	chromophore	UNP Q5S6Z9
A	64	5SQ	TYR	chromophore	UNP Q5S6Z9
A	64	5SQ	GLY	chromophore	UNP Q5S6Z9
A	70	LYS	GLU	engineered mutation	UNP Q5S6Z9
A	74	ASN	HIS	engineered mutation	UNP Q5S6Z9
A	93	MET	LEU	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	166	GLY	ASN	engineered mutation	UNP Q5S6Z9
A	189	ALA	TYR	engineered mutation	UNP Q5S6Z9
B	-3	GLY	-	expression tag	UNP Q5S6Z9
B	-2	PRO	-	expression tag	UNP Q5S6Z9
B	-1	GLU	-	expression tag	UNP Q5S6Z9
B	0	PHE	-	expression tag	UNP Q5S6Z9
B	11	LYS	ASN	engineered mutation	UNP Q5S6Z9
B	28	GLU	ASP	engineered mutation	UNP Q5S6Z9
B	64	5SQ	HIS	chromophore	UNP Q5S6Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	64	5SQ	TYR	chromophore	UNP Q5S6Z9
B	64	5SQ	GLY	chromophore	UNP Q5S6Z9
B	70	LYS	GLU	engineered mutation	UNP Q5S6Z9
B	74	ASN	HIS	engineered mutation	UNP Q5S6Z9
B	93	MET	LEU	engineered mutation	UNP Q5S6Z9
B	102	ASN	ILE	engineered mutation	UNP Q5S6Z9
B	121	TYR	HIS	engineered mutation	UNP Q5S6Z9
B	123	THR	VAL	engineered mutation	UNP Q5S6Z9
B	158	GLU	THR	engineered mutation	UNP Q5S6Z9
B	166	GLY	ASN	engineered mutation	UNP Q5S6Z9
B	189	ALA	TYR	engineered mutation	UNP Q5S6Z9

- Molecule 2 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (CCD ID: DTT) (formula: $C_4H_{10}O_2S_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			8	4	2	2		

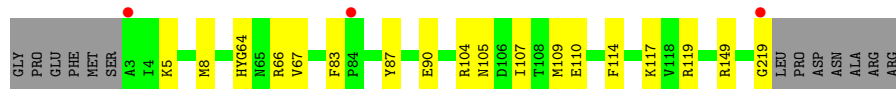
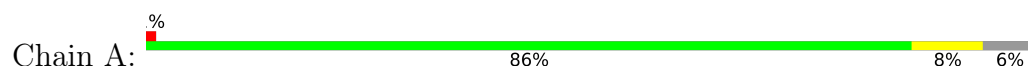
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	215	Total	O	0	0
			215	215		
3	B	251	Total	O	0	0
			251	251		

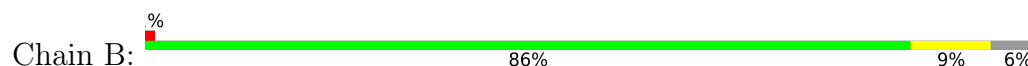
3 Residue-property plots [i](#)

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 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



- 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	50.85Å 76.77Å 124.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.15 – 1.66 28.15 – 1.66	Depositor EDS
% Data completeness (in resolution range)	99.5 (28.15-1.66) 99.4 (28.15-1.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.17 (at 1.66Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.161 , 0.190 0.161 , 0.189	Depositor DCC
R_{free} test set	55971 reflections (3.44%)	wwPDB-VP
Wilson B-factor (Å ²)	20.3	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 38.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3957	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.67% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 5SQ, DTT

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	0.38	0/1762	0.62	0/2369
1	B	0.40	0/1768	0.64	0/2377
All	All	0.39	0/3530	0.63	0/4746

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	1740	0	1662	10	0
1	B	1743	0	1667	17	0
2	A	8	0	8	0	0
3	A	215	0	0	4	3
3	B	251	0	0	8	2
All	All	3957	0	3337	27	3

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 4.

The worst 5 of 27 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:B:3:ALA:N	3:B:302:HOH:O	2.01	0.93
1:B:212:GLU:OE1	3:B:301:HOH:O	1.85	0.93
1:B:209:LYS:HD3	3:B:500:HOH:O	1.68	0.92
1:A:83:PHE:O	3:A:401:HOH:O	1.92	0.86
1:A:90:GLU:OE2	1:A:104:ARG:NH1	2.23	0.71

All (3) 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:550:HOH:O	3:B:482:HOH:O[3_646]	1.90	0.30
3:A:571:HOH:O	3:B:482:HOH:O[3_646]	1.92	0.28
3:A:580:HOH:O	3:A:593:HOH:O[4_556]	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	211/228 (92%)	209 (99%)	2 (1%)	0	100	100
1	B	212/228 (93%)	212 (100%)	0	0	100	100
All	All	423/456 (93%)	421 (100%)	2 (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	182/192 (95%)	179 (98%)	3 (2%)	58	37
1	B	183/192 (95%)	182 (100%)	1 (0%)	86	80
All	All	365/384 (95%)	361 (99%)	4 (1%)	70	54

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	ARG
1	A	109	MET
1	A	117	LYS
1	B	66	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	76	GLN
1	B	133	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 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	5SQ	B	64	1	23,27,28	2.47	7 (30%)	29,37,39	5.76	15 (51%)
1	5SQ	A	64	1	23,27,28	2.54	7 (30%)	29,37,39	6.11	16 (55%)

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

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	64	5SQ	C1-N2	6.85	1.42	1.32
1	A	64	5SQ	C1-N2	6.65	1.42	1.32
1	B	64	5SQ	CA2-C2	5.93	1.54	1.48
1	A	64	5SQ	CA2-C2	5.66	1.54	1.48
1	A	64	5SQ	CA2-N2	5.07	1.49	1.38

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	64	5SQ	CB2-CA2-C2	17.66	143.36	122.28
1	A	64	5SQ	O2-C2-CA2	-16.39	121.76	130.96
1	B	64	5SQ	CB2-CA2-C2	16.38	141.83	122.28
1	B	64	5SQ	O2-C2-CA2	-15.93	122.02	130.96
1	A	64	5SQ	CB2-CA2-N2	-14.79	108.31	128.83

There are no chirality outliers.

5 of 8 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	B	64	5SQ	C3-CA3-N3-C2
1	B	64	5SQ	CA1-CB1-CG1-C2H

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry

1 ligand 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	$\# Z > 2$
2	DTT	A	301	1	7,7,7	0.53	0	4,8,8	0.98	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	DTT	A	301	1	-	0/8/8/8	-

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.

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 [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	214/228 (93%)	-0.27	3 (1%) 73 77	14, 26, 46, 71	1 (0%)
1	B	214/228 (93%)	-0.37	2 (0%) 81 84	13, 22, 41, 65	2 (0%)
All	All	428/456 (93%)	-0.32	5 (1%) 76 80	13, 24, 45, 71	3 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	PRO	3.5
1	A	219	GLY	3.2
1	A	3	ALA	3.2
1	B	219	GLY	2.9
1	B	121	TYR	2.1

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	5SQ	A	64	25/26	0.97	0.06	15,18,22,23	0
1	5SQ	B	64	25/26	0.97	0.05	13,17,20,20	0

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands

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	DTT	A	301	8/8	0.91	0.14	32,38,45,47	0

6.5 Other polymers

There are no such residues in this entry.