



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:45 AM GMT

PDB ID : 3PRU  
Title : Crystal Structure of Phycobilisome 32.1 kDa linker polypeptide, phycocyanin-associated, rod 1 (fragment 14-158) from *Synechocystis* sp. PCC 6803, Northeast Structural Genomics Consortium Target SgR182A  
Authors : Kuzin, A.; Su, M.; Patel, P.; Xiao, R.; Ciccocanti, C.; Lee, D.; Everett, J.K.; Nair, R.; Acton, T.B.; Rost, B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)  
Deposited on : 2010-11-30  
Resolution : 2.68 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

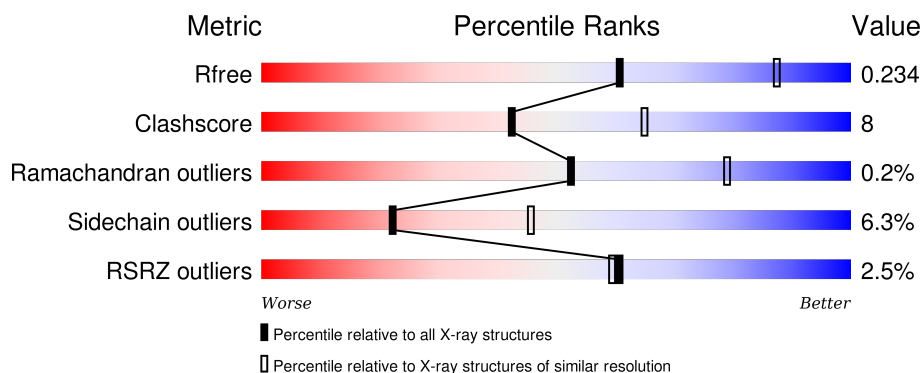
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.68 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))
$R_{free}$	91344	2780 (2.70-2.66)
Clashscore	102246	3138 (2.70-2.66)
Ramachandran outliers	100387	3089 (2.70-2.66)
Sidechain outliers	100360	3089 (2.70-2.66)
RSRZ outliers	91569	2789 (2.70-2.66)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	154	<div> <div>%</div> <div> <div></div> <div>69%</div> <div>20%</div> <div>• • 8%</div> </div> </div>
1	B	154	<div> <div>6%</div> <div> <div></div> <div>75%</div> <div>16%</div> <div>• 6%</div> </div> </div>
1	C	154	<div> <div></div> <div> <div>71%</div> <div>16%</div> <div>• 11%</div> </div> </div>
1	D	154	<div> <div>%</div> <div> <div></div> <div>77%</div> <div>10%</div> <div>• 12%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NA	A	1	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4729 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 Phycobilisome 32.1 kDa linker polypeptide, phycocyanin-associated, rod 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	137	Total	C	N	O	Se	0	0	0
			1142	729	192	219	2			
1	A	142	Total	C	N	O	Se	0	0	0
			1186	755	199	230	2			
1	B	144	Total	C	N	O	Se	0	0	0
			1205	768	208	227	2			
1	D	135	Total	C	N	O	Se	0	0	0
			1126	721	188	215	2			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	13	MSE	-	INITIATING METHIONINE	UNP P73203
C	159	LEU	-	EXPRESSION TAG	UNP P73203
C	160	GLU	-	EXPRESSION TAG	UNP P73203
C	161	HIS	-	EXPRESSION TAG	UNP P73203
C	162	HIS	-	EXPRESSION TAG	UNP P73203
C	163	HIS	-	EXPRESSION TAG	UNP P73203
C	164	HIS	-	EXPRESSION TAG	UNP P73203
C	165	HIS	-	EXPRESSION TAG	UNP P73203
C	166	HIS	-	EXPRESSION TAG	UNP P73203
A	13	MSE	-	INITIATING METHIONINE	UNP P73203
A	159	LEU	-	EXPRESSION TAG	UNP P73203
A	160	GLU	-	EXPRESSION TAG	UNP P73203
A	161	HIS	-	EXPRESSION TAG	UNP P73203
A	162	HIS	-	EXPRESSION TAG	UNP P73203
A	163	HIS	-	EXPRESSION TAG	UNP P73203
A	164	HIS	-	EXPRESSION TAG	UNP P73203
A	165	HIS	-	EXPRESSION TAG	UNP P73203
A	166	HIS	-	EXPRESSION TAG	UNP P73203
B	13	MSE	-	INITIATING METHIONINE	UNP P73203
B	159	LEU	-	EXPRESSION TAG	UNP P73203

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	160	GLU	-	EXPRESSION TAG	UNP P73203
B	161	HIS	-	EXPRESSION TAG	UNP P73203
B	162	HIS	-	EXPRESSION TAG	UNP P73203
B	163	HIS	-	EXPRESSION TAG	UNP P73203
B	164	HIS	-	EXPRESSION TAG	UNP P73203
B	165	HIS	-	EXPRESSION TAG	UNP P73203
B	166	HIS	-	EXPRESSION TAG	UNP P73203
D	13	MSE	-	INITIATING METHIONINE	UNP P73203
D	159	LEU	-	EXPRESSION TAG	UNP P73203
D	160	GLU	-	EXPRESSION TAG	UNP P73203
D	161	HIS	-	EXPRESSION TAG	UNP P73203
D	162	HIS	-	EXPRESSION TAG	UNP P73203
D	163	HIS	-	EXPRESSION TAG	UNP P73203
D	164	HIS	-	EXPRESSION TAG	UNP P73203
D	165	HIS	-	EXPRESSION TAG	UNP P73203
D	166	HIS	-	EXPRESSION TAG	UNP P73203

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total Cl 1 1	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	19	Total O 19 19	0	0
4	A	11	Total O 11 11	0	0
4	B	30	Total O 30 30	0	0
4	D	8	Total O 8 8	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.35Å 91.53Å 124.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.77 – 2.68 47.77 – 2.68	Depositor EDS
% Data completeness (in resolution range)	98.2 (47.77-2.68) 98.3 (47.77-2.68)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.82 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, $R_{free}$	0.193 , 0.241 0.188 , 0.234	Depositor DCC
$R_{free}$ test set	1236 reflections (5.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.6	Xtriage
Anisotropy	0.790	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 46.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 24212 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4729	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: NA, CL

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	0.44	0/1209	0.61	0/1629
1	B	0.45	0/1231	0.58	0/1657
1	C	0.44	0/1164	0.58	0/1568
1	D	0.40	0/1148	0.56	0/1546
All	All	0.43	0/4752	0.58	0/6400

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	1186	0	1137	26	0
1	B	1205	0	1151	24	0
1	C	1142	0	1103	22	0
1	D	1126	0	1091	11	0
2	C	1	0	0	0	0
3	A	1	0	0	0	0
4	A	11	0	0	0	0
4	B	30	0	0	1	0
4	C	19	0	0	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	8	0	0	0	0
All	All	4729	0	4482	76	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 8.

The worst 5 of 76 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:24:ARG:HD3	1:B:24:ARG:H	1.26	0.98
1:A:19:ARG:HB3	1:A:20:PRO:HA	1.55	0.86
1:A:37:ARG:HH22	1:B:19:ARG:HH22	1.26	0.81
1:A:15:TYR:CE1	1:B:37:ARG:HG2	2.18	0.78
1:A:33:LYS:O	1:A:37:ARG:HG3	1.89	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	A	140/154 (91%)	134 (96%)	5 (4%)	1 (1%)	26	53
1	B	140/154 (91%)	137 (98%)	3 (2%)	0	100	100
1	C	135/154 (88%)	134 (99%)	1 (1%)	0	100	100
1	D	133/154 (86%)	125 (94%)	8 (6%)	0	100	100
All	All	548/616 (89%)	530 (97%)	17 (3%)	1 (0%)	52	79

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ARG

### 5.3.2 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	130/138 (94%)	115 (88%)	15 (12%)	7	15
1	B	131/138 (95%)	125 (95%)	6 (5%)	33	61
1	C	125/138 (91%)	119 (95%)	6 (5%)	31	59
1	D	123/138 (89%)	118 (96%)	5 (4%)	37	66
All	All	509/552 (92%)	477 (94%)	32 (6%)	22	46

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	115	ILE
1	A	132	SER
1	D	51	ASP
1	A	119	ASP
1	A	141	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	42	GLN
1	A	154	ASN
1	D	42	GLN
1	A	16	ASN
1	B	165	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - 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.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	140/154 (90%)	-0.16	2 (1%) 78 77	32, 48, 83, 108	0
1	B	142/154 (92%)	0.02	10 (7%) 19 17	28, 45, 77, 109	0
1	C	135/154 (87%)	-0.29	0 100 100	31, 44, 74, 107	0
1	D	133/154 (86%)	-0.15	2 (1%) 76 75	36, 54, 81, 107	0
All	All	550/616 (89%)	-0.14	14 (2%) 61 59	28, 48, 79, 109	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	164	HIS	2.9
1	B	165	HIS	2.6
1	B	154	ASN	2.4
1	B	137	VAL	2.4
1	D	109	PHE	2.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NA	A	1	1/1	0.98	0.36	7.21	57,57,57,57	0
2	CL	C	1	1/1	0.98	0.10	-1.94	40,40,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.