



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2017 – 07:54 PM EDT

PDB ID : 3BZB
Title : Crystal structure of uncharacterized protein CMQ451C from the primitive red alga *Cyanidioschyzon merolae*
Authors : McCoy, J.G.; Bitto, E.; Bingman, C.A.; Wesenberg, G.E.; Phillips Jr., G.N.; Center for Eukaryotic Structural Genomics (CESG)
Deposited on : unknown
Resolution : 2.79 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

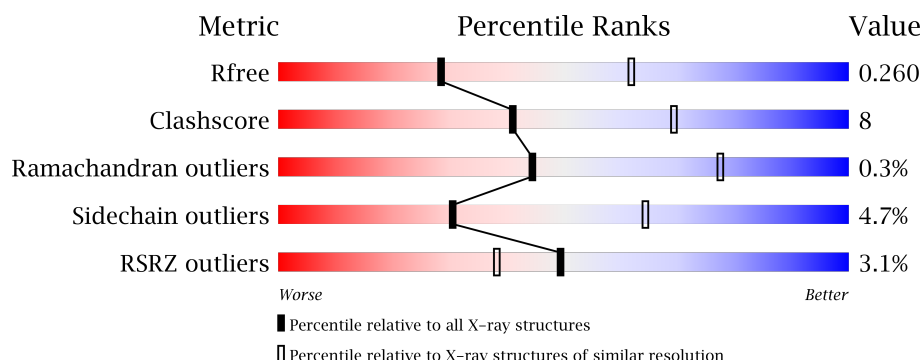
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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}	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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 ≥ 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	281	<div> <div>3%</div> <div>53%</div> <div>14%</div> <div>33%</div> </div>
1	B	281	<div> <div>%</div> <div>58%</div> <div>12%</div> <div>29%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3027 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	189	Total	C	N	O	S	Se	0	0	0
			1455	926	260	264	4	1			
1	B	200	Total	C	N	O	S	Se	0	0	0
			1559	989	286	278	4	2			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	O	0	0
			4	4		
2	B	9	Total	O	0	0
			9	9		

These plots are drawn for all protein, RNA and DNA 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 ($\text{RSRZ} > 2$). 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.

- Chain A:
-
- 3% 53% 14% 33%
- SER ASN ASN SER ARG ARG ALA ALA ASP GLN GLN ASP GLU LEU LEU ASP GLY LEU PRO PRO THR PRO ASP ASP PHE TYR ARG GLU ARG GLN GLN ARG ARG SER R29 P36 A37 G38 A39 P40 LEU Q42 C43 T50 GLN GLU HIS PRO PRO LEU LEU TRP THR SER HIS VAL TRP S62 G63 A64 R65 A66 L67 I78 F79 T80

- Chain B:
-
- 58% 12% 29%
- Legend:
- P111
 - E114
 - H126
 - THR
 - ASN
 - CYS
 - SER
 - ASP
 - VAL
 - ILE
 - ARG
 - GLY
 - Q256
 - V257
 - L262
 - R263
 - W264
 - R265
 - SER
 - ALA
 - SER
 - ALA
 - ASN
 - ILE
 - PRO
 - ALA
 - HIS
 - ALA
 - ARG
 - ASN
 - GLU
 - L183
 - R184
 - K187
 - M188
 - D196
 - F197
 - T198
 - A201
 - T204
 - R209
 - R215
 - V223

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	103.78Å 103.78Å 105.39Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.55 – 2.79 46.55 – 2.79	Depositor EDS
% Data completeness (in resolution range)	98.9 (46.55-2.79) 98.9 (46.55-2.79)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.209 , 0.261 0.212 , 0.260	Depositor DCC
R_{free} test set	842 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	65.2	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.029 for -h,-k,l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3027	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.30% 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

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	0/1484	0.56	0/2021
1	B	0.46	0/1592	0.60	0/2166
All	All	0.45	0/3076	0.58	0/4187

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	1455	0	1450	23	0
1	B	1559	0	1558	26	0
2	A	4	0	0	0	0
2	B	9	0	0	0	0
All	All	3027	0	3008	49	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.

All (49) 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:78:ILE:HD11	1:B:100:ALA:HB3	1.42	1.01

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:265:ARG:HB2	1:B:265:ARG:HH11	1.39	0.87
1:B:265:ARG:CB	1:B:265:ARG:HH11	1.91	0.84
1:B:63:GLY:O	1:B:204:THR:HG21	1.86	0.75
1:B:78:ILE:CD1	1:B:100:ALA:HB3	2.15	0.75
1:B:78:ILE:HD11	1:B:100:ALA:CB	2.18	0.74
1:B:264:TRP:O	1:B:265:ARG:HB2	1.89	0.73
1:B:78:ILE:HD13	1:B:97:ALA:HA	1.71	0.71
1:B:209:ARG:HH22	1:B:215:ARG:HH22	1.39	0.70
1:B:184:ARG:O	1:B:188:MSE:HG3	1.94	0.68
1:B:265:ARG:HB2	1:B:265:ARG:NH1	2.10	0.66
1:A:170:ALA:HA	1:A:204:THR:HG22	1.77	0.66
1:B:82:THR:HG22	1:B:165:GLN:HG3	1.78	0.65
1:B:209:ARG:HH22	1:B:215:ARG:NH2	1.96	0.63
1:A:223:VAL:O	1:A:226:ASP:O	2.17	0.62
1:A:63:GLY:N	1:A:171:ASP:OD2	2.30	0.62
1:A:78:ILE:HD11	1:A:100:ALA:HB3	1.82	0.60
1:A:43:CYS:SG	1:A:99:LEU:HD13	2.42	0.59
1:A:171:ASP:OD1	1:A:206:THR:HG22	2.04	0.58
1:B:78:ILE:CD1	1:B:97:ALA:HA	2.34	0.58
1:B:47:VAL:HG13	1:B:90:ALA:HB1	1.89	0.55
1:A:169:LEU:HB3	1:A:172:LEU:HD12	1.90	0.54
1:B:183:LEU:HD13	1:B:223:VAL:HG22	1.90	0.53
1:A:83:VAL:HG22	1:A:166:VAL:HB	1.90	0.53
1:A:201:ALA:HB3	1:A:262:LEU:HB2	1.91	0.53
1:B:196:ASP:OD1	1:B:198:THR:HG22	2.08	0.53
1:B:175:PHE:CE1	1:B:177:GLN:HB2	2.45	0.52
1:A:78:ILE:CD1	1:A:97:ALA:HA	2.42	0.50
1:B:82:THR:HG23	1:B:165:GLN:H	1.76	0.50
1:A:264:TRP:O	1:A:265:ARG:HB2	2.11	0.50
1:A:67:LEU:HB2	1:A:204:THR:HG21	1.95	0.49
1:A:86:LEU:HD12	1:A:169:LEU:HD22	1.95	0.49
1:A:168:LEU:C	1:A:169:LEU:HD23	2.33	0.48
1:A:63:GLY:O	1:A:204:THR:HG21	2.13	0.48
1:B:265:ARG:NH1	1:B:265:ARG:CB	2.71	0.47
1:A:112:ASP:HB3	1:A:115:ILE:HG12	1.96	0.47
1:A:116:LEU:HD11	1:A:145:PRO:HB3	1.98	0.45
1:B:223:VAL:HG11	1:B:262:LEU:HD22	1.99	0.44
1:A:234:TRP:HB2	1:A:261:ARG:HB2	1.99	0.44
1:B:240:MSE:HG3	1:B:257:VAL:HG23	2.00	0.44
1:B:82:THR:CG2	1:B:165:GLN:HG3	2.48	0.43
1:A:163:ARG:NH2	1:A:187:LYS:O	2.50	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:201:ALA:HB3	1:B:262:LEU:HB2	2.00	0.43
1:A:78:ILE:CD1	1:A:100:ALA:HB3	2.48	0.42
1:A:78:ILE:HD11	1:A:100:ALA:CB	2.47	0.42
1:A:110:TYR:CG	1:A:111:PRO:HD2	2.55	0.42
1:B:163:ARG:NH2	1:B:187:LYS:O	2.53	0.41
1:B:93:VAL:HG12	1:B:168:LEU:HD13	2.03	0.41
1:A:151:SER:HA	1:A:152:PRO:HD3	1.96	0.40

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	177/281 (63%)	171 (97%)	6 (3%)	0	100	100
1	B	192/281 (68%)	187 (97%)	4 (2%)	1 (0%)	32	67
All	All	369/562 (66%)	358 (97%)	10 (3%)	1 (0%)	44	77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	111	PRO

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	154/230 (67%)	149 (97%)	5 (3%)	44	78
1	B	165/230 (72%)	155 (94%)	10 (6%)	22	53
All	All	319/460 (69%)	304 (95%)	15 (5%)	30	64

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	169	LEU
1	A	176	HIS
1	A	180	ASP
1	A	182	LEU
1	A	232	GLU
1	B	76	GLU
1	B	114	GLU
1	B	140	SER
1	B	151	SER
1	B	171	ASP
1	B	176	HIS
1	B	209	ARG
1	B	215	ARG
1	B	232	GLU
1	B	265	ARG

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

Mol	Chain	Res	Type
1	A	34	GLN
1	B	239	GLN

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](#)

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	188/281 (66%)	0.50	8 (4%) 36 26	18, 39, 80, 82	0
1	B	198/281 (70%)	0.24	4 (2%) 65 56	19, 26, 38, 41	0
All	All	386/562 (68%)	0.37	12 (3%) 49 38	18, 32, 73, 82	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	78	ILE	3.2
1	A	79	ALA	3.0
1	B	161	LEU	3.0
1	A	36	PRO	2.9
1	B	152	PRO	2.6
1	B	157	ARG	2.4
1	B	162	GLN	2.4
1	A	38	GLY	2.3
1	A	65	ARG	2.2
1	A	219	PHE	2.2
1	A	92	LEU	2.2
1	A	37	ALA	2.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.