wwPDB X-ray Structure Validation Summary Report (i)

Sep 2, 2023 – 10:57 AM EDT

PDB ID	:	3Q17
Title	:	Structure of a slow CLC Cl-/H+ antiporter from a cyanobacterium in Bromide
Authors	:	Jayaram, H.; Robertson, J.L.; Fang, W.; Williams, C.; Miller, C.
Deposited on	:	2010-12-16
Resolution	:	3.60 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:

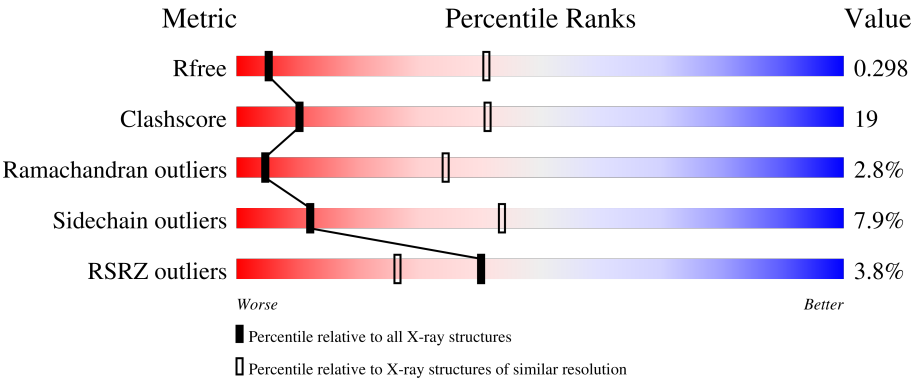
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 <=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	466	2% 52% 33% 6% 9%
1	B	466	5% 58% 29% . 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6400 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 Sll0855 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	425	Total	C	N	O	S	0	0	0
			3199	2110	533	536	20			
1	B	425	Total	C	N	O	S	0	0	0
			3199	2110	533	536	20			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	451	GLY	-	expression tag	UNP P73745
A	452	SER	-	expression tag	UNP P73745
A	453	LEU	-	expression tag	UNP P73745
A	454	VAL	-	expression tag	UNP P73745
A	455	PRO	-	expression tag	UNP P73745
A	456	ARG	-	expression tag	UNP P73745
A	457	GLY	-	expression tag	UNP P73745
A	458	SER	-	expression tag	UNP P73745
A	459	GLY	-	expression tag	UNP P73745
A	460	GLY	-	expression tag	UNP P73745
A	461	HIS	-	expression tag	UNP P73745
A	462	HIS	-	expression tag	UNP P73745
A	463	HIS	-	expression tag	UNP P73745
A	464	HIS	-	expression tag	UNP P73745
A	465	HIS	-	expression tag	UNP P73745
A	466	HIS	-	expression tag	UNP P73745
B	451	GLY	-	expression tag	UNP P73745
B	452	SER	-	expression tag	UNP P73745
B	453	LEU	-	expression tag	UNP P73745
B	454	VAL	-	expression tag	UNP P73745
B	455	PRO	-	expression tag	UNP P73745
B	456	ARG	-	expression tag	UNP P73745
B	457	GLY	-	expression tag	UNP P73745
B	458	SER	-	expression tag	UNP P73745
B	459	GLY	-	expression tag	UNP P73745

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	460	GLY	-	expression tag	UNP P73745
B	461	HIS	-	expression tag	UNP P73745
B	462	HIS	-	expression tag	UNP P73745
B	463	HIS	-	expression tag	UNP P73745
B	464	HIS	-	expression tag	UNP P73745
B	465	HIS	-	expression tag	UNP P73745
B	466	HIS	-	expression tag	UNP P73745

- Molecule 2 is BROMIDE ION (three-letter code: BR) (formula: Br).

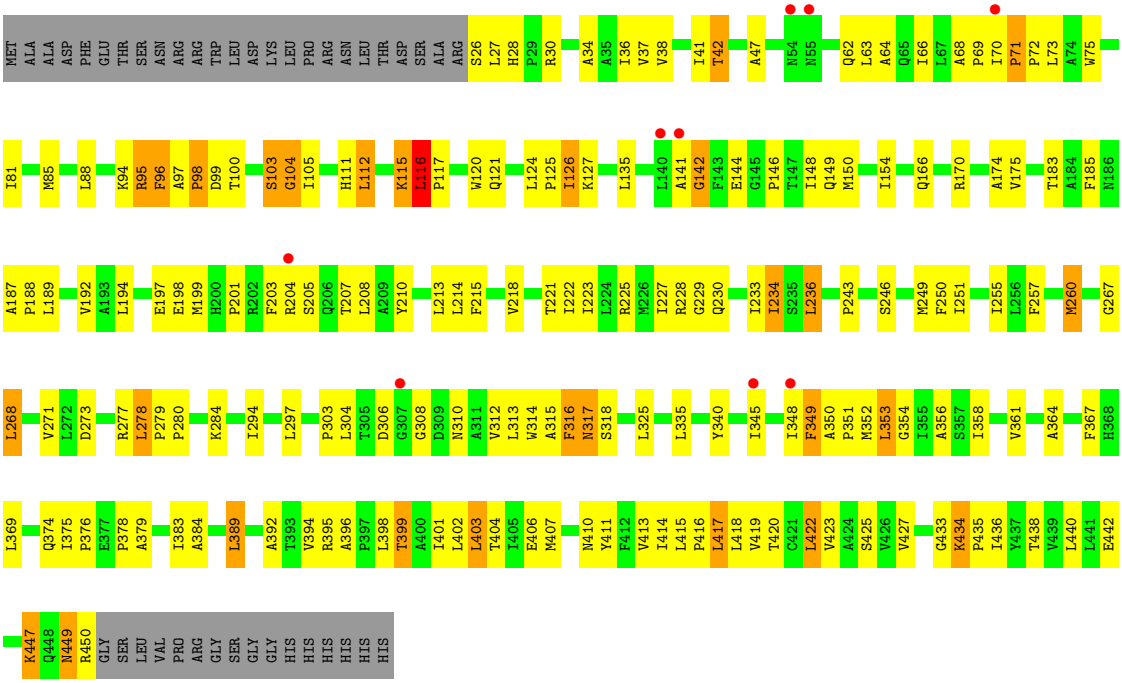
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Br 1 1	0	0
2	B	1	Total Br 1 1	0	0

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 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.

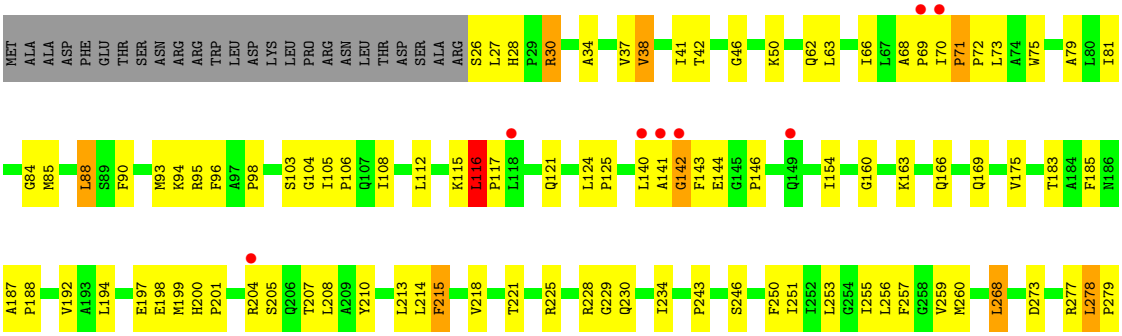
• Molecule 1: Sll0855 protein

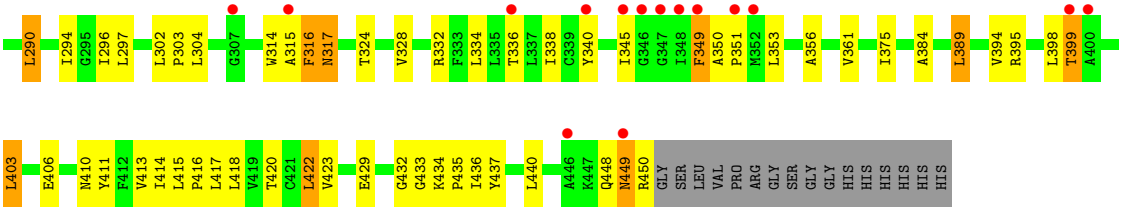
Chain A:



• Molecule 1: Sll0855 protein

Chain B:




L290
I294
G295
I296
L297
L302
P303
L304
G307
W314
A315
F316
N317
T324
V328
R332
F333
L334
L335
T336
L337
I338
C339
Y340
I345
G346
G347
I348
F349
A350
P351
M352
L353
A356
V361
I375
A384
L389
V394
R395
L398
T399
A400
L403
E406
N410
Y411
F412
V413
I414
L415
P416
L417
L418
V419
T420
C421
L422
V423
E429
G432
G433
K434
P435
I436
Y437
L440
A446
K447
Q448
N449
R450
GLY
SER
LEU
VAL
PRO
ARG
GLY
SER
GLY
GLY
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	207.21Å 207.21Å 105.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	68.48 – 3.60 68.48 – 3.60	Depositor EDS
% Data completeness (in resolution range)	97.3 (68.48-3.60) 97.4 (68.48-3.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.26	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 3.58Å)	Xtriage
Refinement program	PHENIX 1.6.1_357, REFMAC 5.5.0109	Depositor
R, R_{free}	0.281 , 0.310 0.267 , 0.298	Depositor DCC
R_{free} test set	1495 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	95.0	Xtriage
Anisotropy	0.697	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 58.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.059 for -h,-k,l	Xtriage
F_o,F_c correlation	0.80	EDS
Total number of atoms	6400	wwPDB-VP
Average B, all atoms (Å^2)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: BR

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	0.23	0/3276	0.42	0/4457
1	B	0.23	0/3276	0.41	0/4457
All	All	0.23	0/6552	0.42	0/8914

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	3199	0	3360	146	0
1	B	3199	0	3360	114	0
2	A	1	0	0	1	0
2	B	1	0	0	0	0
All	All	6400	0	6720	245	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 19.

The worst 5 of 245 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:225:ARG:HD3	1:A:230:GLN:HG3	1.27	1.10
1:B:225:ARG:HD3	1:B:230:GLN:HG3	1.35	1.09
1:B:71:PRO:HB2	1:B:72:PRO:HD2	1.34	1.08
1:A:278:LEU:HB3	1:A:279:PRO:HA	1.38	1.01
1:B:278:LEU:HB3	1:B:279:PRO:HA	1.41	1.00

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	423/466 (91%)	367 (87%)	43 (10%)	13 (3%)	4	32
1	B	423/466 (91%)	374 (88%)	38 (9%)	11 (3%)	5	35
All	All	846/932 (91%)	741 (88%)	81 (10%)	24 (3%)	5	34

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	71	PRO
1	B	71	PRO
1	A	142	GLY
1	A	278	LEU
1	A	449	ASN

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	334/368 (91%)	307 (92%)	27 (8%)	11	43
1	B	334/368 (91%)	308 (92%)	26 (8%)	12	44
All	All	668/736 (91%)	615 (92%)	53 (8%)	12	44

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	30	ARG
1	B	121	GLN
1	B	403	LEU
1	B	42	THR
1	B	95	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	449	ASN
1	B	206	GLN
1	B	230	GLN
1	A	230	GLN
1	A	206	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides 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

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

6.1 Protein, DNA and RNA chains

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^{th} 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	425/466 (91%)	-0.01	9 (2%)	63	48	68, 106, 159, 262	0
1	B	425/466 (91%)	0.23	23 (5%)	25	16	66, 126, 186, 252	0
All	All	850/932 (91%)	0.11	32 (3%)	40	26	66, 114, 180, 262	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	346	GLY	5.1
1	A	141	ALA	4.9
1	B	348	ILE	4.9
1	B	141	ALA	4.3
1	A	55	ASN	4.1

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

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, 95^{th} 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	BR	A	1	1/1	0.91	0.35	147,147,147,147	0
2	BR	B	1	1/1	0.95	0.74	176,176,176,176	0

6.5 Other polymers (i)

There are no such residues in this entry.