WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 01:16 am BST

PDB ID	:	2B8E
Title	:	CopA ATP Binding Domain
Authors	:	Sazinsky, M.H.; Mandal, A.K.; Arguello, J.M.; Rosenzweig, A.C.
Deposited on	:	2005-10-06
Resolution	:	2.30 Å(reported)

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

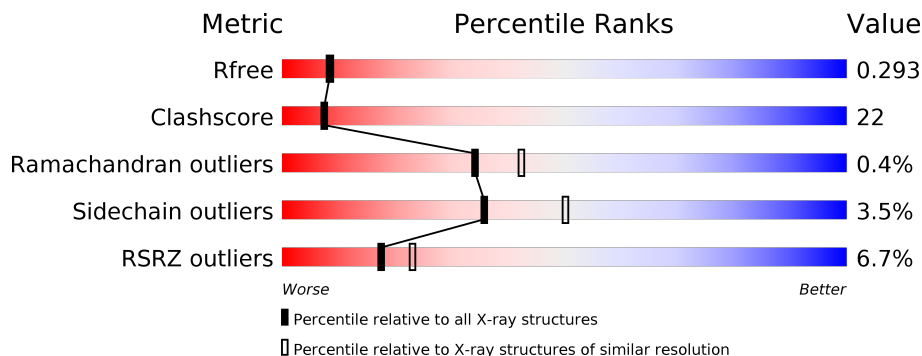|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.30 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	Percentile Ranks	Value
Rfree		0.293
Clashscore		22
Ramachandran outliers		0.4%
Sidechain outliers		3.5%
RSRZ outliers		6.7%

Worse *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 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	273	5% 56% 32% • 10%
1	B	273	7% 57% 30% • 12%
1	C	273	6% 51% 41% • 7%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 5747 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 cation-transporting ATPase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	246	Total	C	N	O	Se	0	0	0
			1787	1124	310	350	3			
1	B	241	Total	C	N	O	Se	0	0	0
			1747	1097	303	344	3			
1	C	254	Total	C	N	O	Se	0	0	0
			1859	1166	322	367	4			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	506	MSE	MET	MODIFIED RESIDUE	UNP O29777
A	564	MSE	MET	MODIFIED RESIDUE	UNP O29777
A	570	MSE	MET	MODIFIED RESIDUE	UNP O29777
A	669	MSE	MET	MODIFIED RESIDUE	UNP O29777
A	672	TRP	-	CLONING ARTIFACT	UNP O29777
A	673	SER	-	CLONING ARTIFACT	UNP O29777
A	674	HIS	-	CLONING ARTIFACT	UNP O29777
A	675	PRO	-	CLONING ARTIFACT	UNP O29777
A	676	GLN	-	CLONING ARTIFACT	UNP O29777
A	677	PHE	-	CLONING ARTIFACT	UNP O29777
A	678	GLU	-	CLONING ARTIFACT	UNP O29777
A	679	LYS	-	CLONING ARTIFACT	UNP O29777
B	506	MSE	MET	MODIFIED RESIDUE	UNP O29777
B	564	MSE	MET	MODIFIED RESIDUE	UNP O29777
B	570	MSE	MET	MODIFIED RESIDUE	UNP O29777
B	669	MSE	MET	MODIFIED RESIDUE	UNP O29777
B	672	TRP	-	CLONING ARTIFACT	UNP O29777
B	673	SER	-	CLONING ARTIFACT	UNP O29777
B	674	HIS	-	CLONING ARTIFACT	UNP O29777
B	675	PRO	-	CLONING ARTIFACT	UNP O29777
B	676	GLN	-	CLONING ARTIFACT	UNP O29777
B	677	PHE	-	CLONING ARTIFACT	UNP O29777
B	678	GLU	-	CLONING ARTIFACT	UNP O29777

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	679	LYS	-	CLONING ARTIFACT	UNP O29777
C	506	MSE	MET	MODIFIED RESIDUE	UNP O29777
C	564	MSE	MET	MODIFIED RESIDUE	UNP O29777
C	570	MSE	MET	MODIFIED RESIDUE	UNP O29777
C	669	MSE	MET	MODIFIED RESIDUE	UNP O29777
C	672	TRP	-	CLONING ARTIFACT	UNP O29777
C	673	SER	-	CLONING ARTIFACT	UNP O29777
C	674	HIS	-	CLONING ARTIFACT	UNP O29777
C	675	PRO	-	CLONING ARTIFACT	UNP O29777
C	676	GLN	-	CLONING ARTIFACT	UNP O29777
C	677	PHE	-	CLONING ARTIFACT	UNP O29777
C	678	GLU	-	CLONING ARTIFACT	UNP O29777
C	679	LYS	-	CLONING ARTIFACT	UNP O29777

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	140	Total 140	O 140	0	0
2	B	112	Total 112	O 112	0	0
2	C	102	Total 102	O 102	0	0

3 Residue-property plots (i)

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 (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.
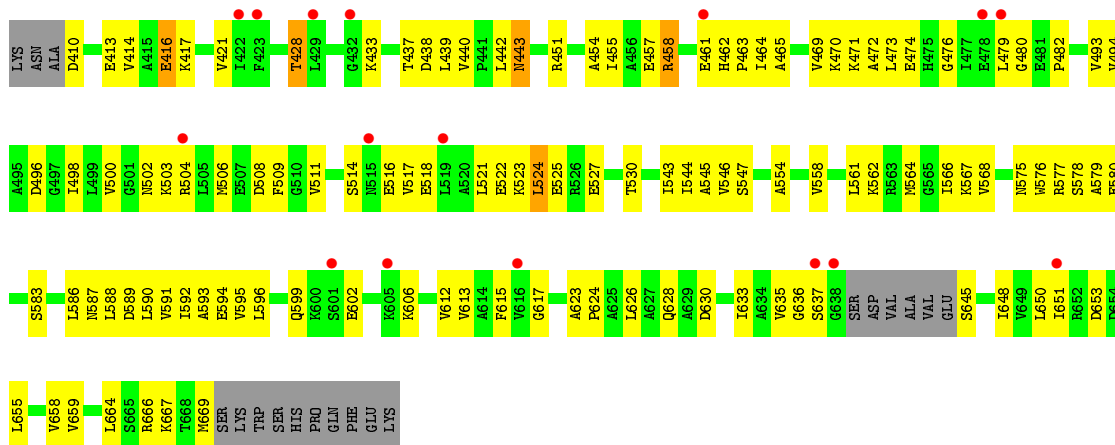
- Molecule 1: cation-transporting ATPase
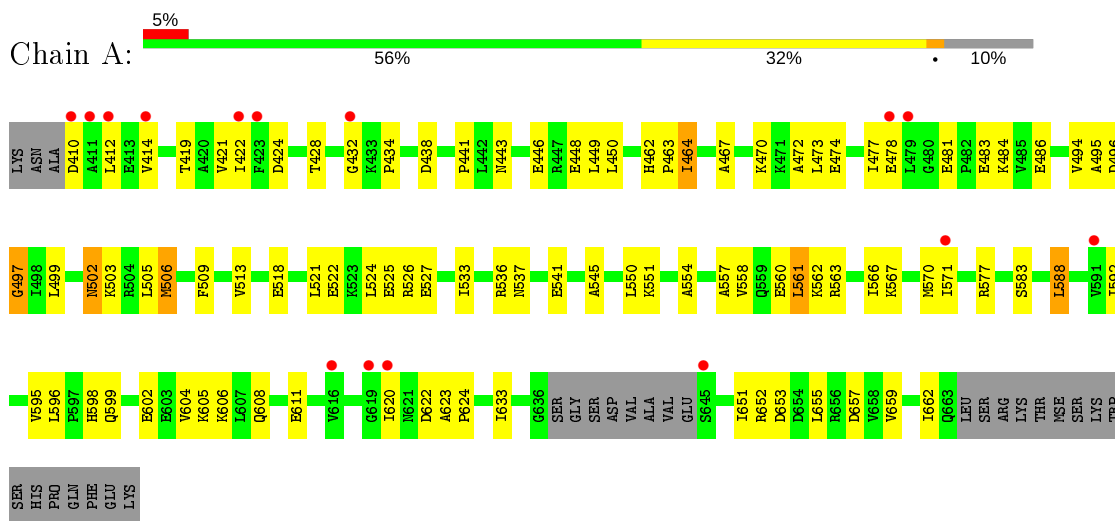
Chain A:

- Molecule 1: cation-transporting ATPase

Chain B:

- Molecule 1: cation-transporting ATPase

Chain C:
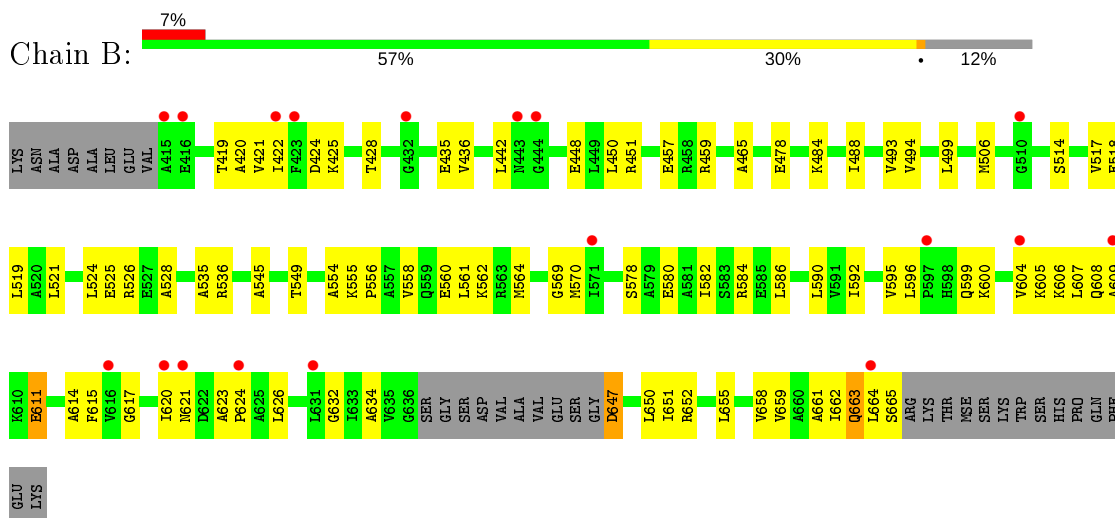
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	80.78Å 80.78Å 105.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	23.66 – 2.30 26.59 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.3 (23.66-2.30) 98.4 (26.59-2.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	4.62 (at 2.29Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.238 , 0.297 0.248 , 0.293	Depositor DCC
R_{free} test set	3343 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å^2)	38.7	Xtriage
Anisotropy	0.001	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 52.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.008 for -h,-k,l 0.026 for h,-h-k,-l 0.013 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5747	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.25% 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 (i)

5.1 Standard geometry (i)

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	0.50	2/1797 (0.1%)	0.77	4/2432 (0.2%)
1	B	0.43	1/1757 (0.1%)	0.62	0/2380
1	C	0.46	1/1868 (0.1%)	0.62	1/2521 (0.0%)
All	All	0.46	4/5422 (0.1%)	0.68	5/7333 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	496	ASP	C-N	-5.74	1.22	1.33
1	B	506	MSE	SE-CE	-5.53	1.62	1.95
1	A	506	MSE	SE-CE	-5.51	1.62	1.95
1	A	570	MSE	SE-CE	-5.51	1.62	1.95

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	622	ASP	CB-CG-OD1	9.23	126.61	118.30
1	A	497	GLY	O-C-N	6.41	132.95	122.70
1	A	497	GLY	CA-C-N	-5.59	104.89	117.20
1	C	612	VAL	N-CA-C	-5.48	96.20	111.00
1	A	496	ASP	O-C-N	-5.28	114.22	123.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	620	ILE	Mainchain

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	1787	0	1820	76	0
1	B	1747	0	1766	78	0
1	C	1859	0	1901	94	0
2	A	140	0	0	7	0
2	B	112	0	0	8	0
2	C	102	0	0	6	0
All	All	5747	0	5487	242	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 22.

The worst 5 of 242 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:605:LYS:HA	1:A:608:GLN:HE21	1.35	0.87
1:A:470:LYS:O	1:A:474:GLU:HG3	1.77	0.84
1:B:419:THR:OG1	1:B:611:GLU:HG3	1.79	0.82
1:B:494:VAL:HG22	1:B:499:LEU:HD22	1.64	0.79
1:B:421:VAL:HG22	1:B:614:ALA:HB3	1.64	0.79

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	242/273 (89%)	230 (95%)	11 (4%)	1 (0%)	34	42
1	B	237/273 (87%)	223 (94%)	13 (6%)	1 (0%)	34	42
1	C	250/273 (92%)	237 (95%)	12 (5%)	1 (0%)	34	42
All	All	729/819 (89%)	690 (95%)	36 (5%)	3 (0%)	34	42

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	432	GLY
1	B	663	GLN
1	C	463	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	181/215 (84%)	175 (97%)	6 (3%)	38	53
1	B	177/215 (82%)	171 (97%)	6 (3%)	37	51
1	C	191/215 (89%)	184 (96%)	7 (4%)	34	48
All	All	549/645 (85%)	530 (96%)	19 (4%)	36	50

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	536	ARG
1	B	611	GLU
1	C	458	ARG
1	B	478	GLU
1	C	461	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	443	ASN
1	B	462	HIS
1	C	537	ASN
1	A	608	GLN
1	A	628	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 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	243/273 (89%)	0.37	15 (6%)	20	26	19, 38, 58, 70	0
1	B	238/273 (87%)	0.43	18 (7%)	13	18	22, 45, 67, 79	0
1	C	250/273 (91%)	0.53	16 (6%)	19	25	18, 46, 65, 77	0
All	All	731/819 (89%)	0.45	49 (6%)	17	23	18, 43, 65, 79	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	609	ALA	4.0
1	A	410	ASP	3.9
1	B	444	GLY	3.9
1	C	637	SER	3.5
1	B	604	VAL	3.4

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.