wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:38 PM BST

PDB ID	:	2VOY
EMDB ID:	:	EMD-5004
Title	:	CryoEM model of CopA, the copper transporting ATPase from Archaeoglobus fulgidus
Authors	:	Wu, C.-C.; Rice, W.J.; Stokes, D.L.
Deposited on	:	2008-02-25
Resolution	:	18.00 Å(reported)
Based on PDB ID	:	2B8E,2HC8,2EAR

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

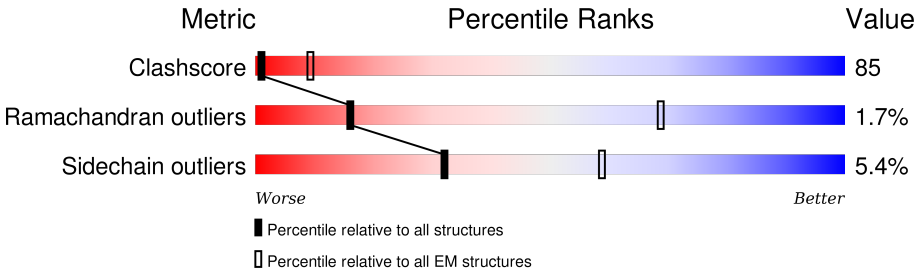
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 18.00 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	80	43% 49% 9%
2	B	42	26% 67% 5% •
3	C	22	82% 18%
4	D	23	13% 83% •
5	E	30	40% 60%
6	F	113	75% 22% • •
7	G	36	78% 19% •
8	H	48	33% 56% 10%
9	I	128	52% 45% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	118	53% 42% 5%
11	K	32	31% 63% 6%
12	L	21	24% 67% 10%

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 5270 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 POTENTIAL COPPER-TRANSPORTING ATPASE.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	80	615	382	104	126	3	0	0

- Molecule 2 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	42	341	227	52	61	1	0	0

- Molecule 3 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	22	176	123	25	27	1	0	0

- Molecule 4 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	23	176	121	28	26	1	0	0

- Molecule 5 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	E	30	234	155	38	41	0	0

- Molecule 6 is a protein called CATION-TRANSPORTING ATPASE, P-TYPE.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	113	Total	C	N	O	S	0	0
			829	525	138	164	2		

- Molecule 7 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	36	Total	C	N	O	S	0	0
			287	185	47	54	1		

- Molecule 8 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	48	Total	C	N	O	S	0	0
			356	233	64	57	2		

- Molecule 9 is a protein called CATION-TRANSPORTING ATPASE.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	128	Total	C	N	O	Se	0	0
			959	604	165	188	2		

- Molecule 10 is a protein called CATION-TRANSPORTING ATPASE.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	118	Total	C	N	O	Se	0	0
			887	553	157	176	1		

- Molecule 11 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	32	Total	C	N	O	S	0	0
			253	163	43	45	2		

- Molecule 12 is a protein called SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1.

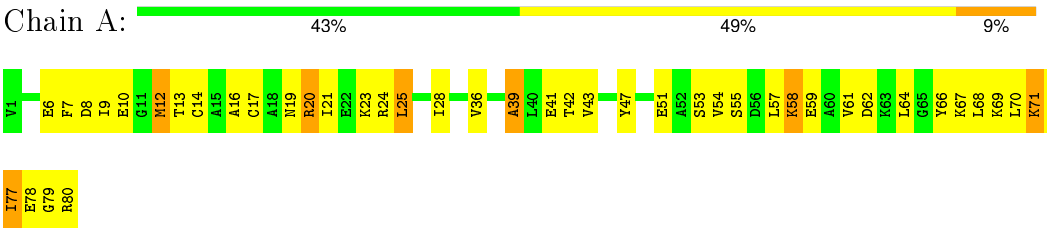
Mol	Chain	Residues	Atoms				AltConf	Trace
12	L	21	Total	C	N	O	0	0
			157	106	24	27		

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 errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

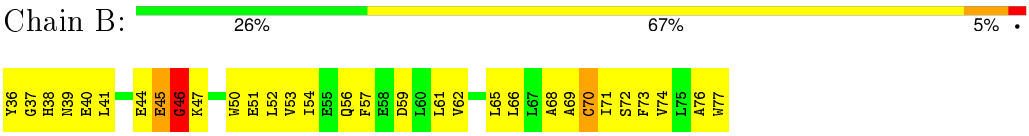
- Molecule 1: POTENTIAL COPPER-TRANSPORTING ATPASE

Chain A:



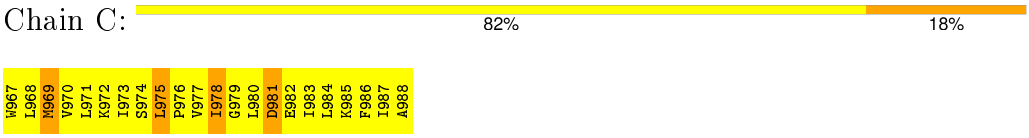
- Molecule 2: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain B:



- Molecule 3: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain C:

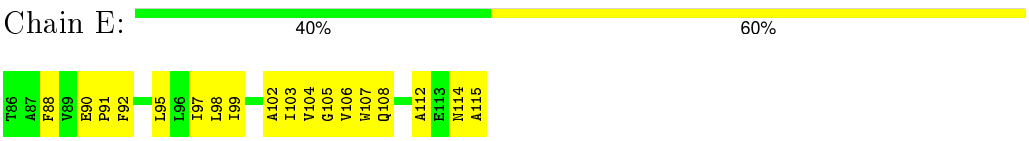


- Molecule 4: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain D:

- Molecule 5: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain E:



- Molecule 6: CATION-TRANSPORTING ATPASE, P-TYPE

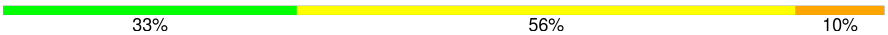
Chain F:

• Molecule 7: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain G:

• Molecule 8: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain H:

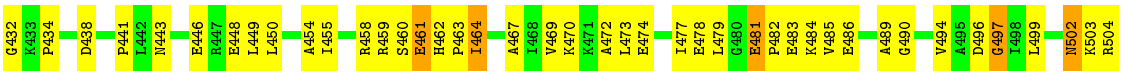


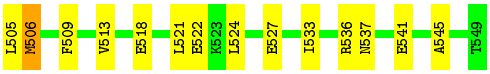
• Molecule 9: CATION-TRANSPORTING ATPASE

Chain I:

• Molecule 10: CATION-TRANSPORTING ATPASE

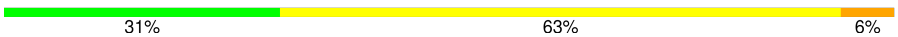
Chain J:





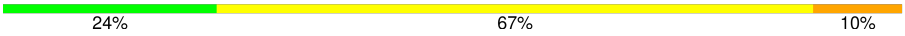
• Molecule 11: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

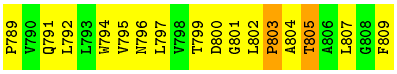
Chain K:



• Molecule 12: SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE 1

Chain L:





4 Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	INDIVIDUAL TUBES	Depositor
Microscope	FEI CM200FEG ST	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	900.00	Depositor
Maximum defocus (nm)	2500.00	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 FILM	Depositor

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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.50	0/620	0.87	0/831
10	J	0.49	1/892 (0.1%)	0.82	3/1202 (0.2%)
11	K	0.48	0/256	0.65	0/344
12	L	0.44	0/161	0.72	0/222
2	B	0.30	0/349	0.49	0/475
3	C	0.52	0/179	0.71	0/242
4	D	0.44	0/183	0.56	0/249
5	E	0.31	0/238	0.52	0/326
6	F	0.73	1/835 (0.1%)	0.77	1/1129 (0.1%)
7	G	0.29	0/291	0.49	0/393
8	H	0.30	0/360	0.54	0/487
9	I	0.52	1/962 (0.1%)	0.80	1/1293 (0.1%)
All	All	0.51	3/5326 (0.1%)	0.73	5/7193 (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
2	B	0	3
6	F	0	1
8	H	0	2
9	I	0	1
All	All	0	8

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	F	215	ALA	C-N	7.47	1.51	1.34
9	I	570	MSE	SE-CE	-5.54	1.62	1.95
10	J	506	MSE	SE-CE	-5.52	1.62	1.95

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	I	622	ASP	CB-CG-OD1	9.26	126.63	118.30
10	J	497	GLY	O-C-N	6.38	132.91	122.70
10	J	497	GLY	CA-C-N	-5.57	104.95	117.20
10	J	496	ASP	O-C-N	-5.26	114.26	123.20
6	F	215	ALA	CA-C-N	-5.01	106.17	117.20

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	47	TYR	Sidechain
2	B	45	GLU	Peptide
2	B	46	GLY	Peptide
2	B	70	CYS	Mainchain
6	F	215	ALA	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	615	0	608	285	0
2	B	341	0	336	267	0
3	C	176	0	196	165	0
4	D	176	0	164	143	0
5	E	234	0	239	89	0
6	F	829	0	874	218	0
7	G	287	0	294	40	0
8	H	356	0	397	74	0
9	I	959	0	999	149	0
10	J	887	0	915	127	0
11	K	253	0	257	40	0
12	L	157	0	163	18	0
All	All	5270	0	5442	913	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 85.

The worst 5 of 913 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)
3:C:968:LEU:HD21	8:H:296:PHE:CZ	1.17	1.68
3:C:985:LYS:CG	4:D:838:MET:HG2	1.24	1.66
1:A:58:LYS:HB3	6:F:298:THR:CG2	1.19	1.64
1:A:12:MET:SD	10:J:459:ARG:HB3	1.29	1.62
7:G:247:THR:HG21	9:I:410:ASP:CB	1.30	1.61

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 PDB entries followed by that with respect to all EM entries.

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	78/80 (98%)	65 (83%)	10 (13%)	3 (4%)	4 37
2	B	40/42 (95%)	29 (72%)	10 (25%)	1 (2%)	7 46
3	C	20/22 (91%)	13 (65%)	5 (25%)	2 (10%)	1 14
4	D	21/23 (91%)	19 (90%)	2 (10%)	0	100 100
5	E	28/30 (93%)	16 (57%)	11 (39%)	1 (4%)	4 38
6	F	111/113 (98%)	109 (98%)	2 (2%)	0	100 100
7	G	34/36 (94%)	28 (82%)	6 (18%)	0	100 100
8	H	46/48 (96%)	39 (85%)	4 (9%)	3 (6%)	1 25
9	I	122/128 (95%)	118 (97%)	4 (3%)	0	100 100
10	J	116/118 (98%)	110 (95%)	6 (5%)	0	100 100
11	K	30/32 (94%)	23 (77%)	7 (23%)	0	100 100
12	L	19/21 (90%)	13 (68%)	5 (26%)	1 (5%)	2 29
All	All	665/693 (96%)	582 (88%)	72 (11%)	11 (2%)	16 55

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	ALA
1	A	77	ILE
8	H	300	VAL
8	H	309	GLU
2	B	46	GLY

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 PDB entries followed by that with respect to all EM entries.

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	66/66 (100%)	56 (85%)	10 (15%)	3 22
2	B	36/36 (100%)	36 (100%)	0	100 100
3	C	20/20 (100%)	18 (90%)	2 (10%)	9 38
4	D	13/13 (100%)	12 (92%)	1 (8%)	16 52
5	E	24/24 (100%)	23 (96%)	1 (4%)	36 70
6	F	91/91 (100%)	89 (98%)	2 (2%)	60 83
7	G	33/33 (100%)	32 (97%)	1 (3%)	48 77
8	H	35/35 (100%)	33 (94%)	2 (6%)	25 62
9	I	102/100 (102%)	100 (98%)	2 (2%)	63 85
10	J	93/92 (101%)	88 (95%)	5 (5%)	27 64
11	K	27/27 (100%)	24 (89%)	3 (11%)	8 34
12	L	17/17 (100%)	16 (94%)	1 (6%)	24 61
All	All	557/554 (100%)	527 (95%)	30 (5%)	32 64

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	108	GLN
7	G	247	THR
11	K	767	SER
6	F	320	LEU
8	H	295	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
9	I	599	GLN
9	I	608	GLN
10	J	537	ASN
9	I	587	ASN
11	K	759	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

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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.