



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:17 AM GMT

PDB ID : 3OOC
Title : Crystal structure of the membrane fusion protein CusB from Escherichia coli
Authors : Su, C.-C.; Yang, F.; Long, F.; Reyon, D.; Routh, M.D.; Kuo, D.W.; Mokhtari, A.K.; Van Ornam, J.D.; Rabe, K.L.; Hoy, J.A.; Lee, Y.J.; Rajashankar, K.R.; Yu, E.W.
Deposited on : 2010-08-30
Resolution : 3.40 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
<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

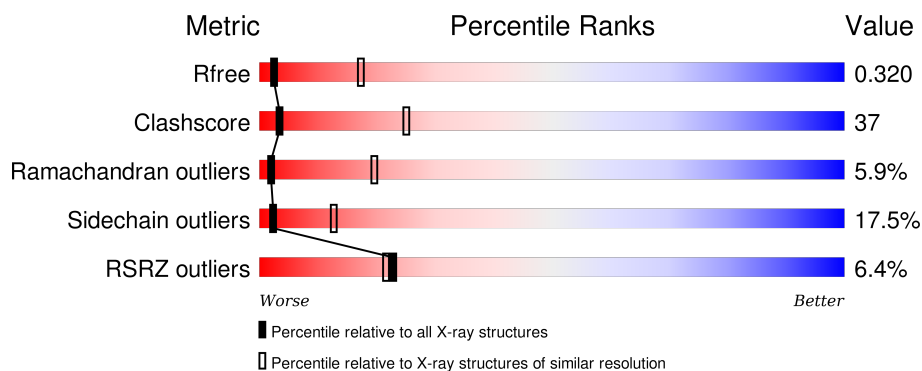
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

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. 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	413	 4% 28% 35% 10% 28%
1	B	413	 5% 28% 33% 9% 28%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 4548 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 efflux system protein cusB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	0	0
			2274	1448	392	429	5			
1	B	297	Total	C	N	O	S	0	0	0
			2274	1448	392	429	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	408	HIS	-	EXPRESSION TAG	UNP P77239
A	409	HIS	-	EXPRESSION TAG	UNP P77239
A	410	HIS	-	EXPRESSION TAG	UNP P77239
A	411	HIS	-	EXPRESSION TAG	UNP P77239
A	412	HIS	-	EXPRESSION TAG	UNP P77239
A	413	HIS	-	EXPRESSION TAG	UNP P77239
B	408	HIS	-	EXPRESSION TAG	UNP P77239
B	409	HIS	-	EXPRESSION TAG	UNP P77239
B	410	HIS	-	EXPRESSION TAG	UNP P77239
B	411	HIS	-	EXPRESSION TAG	UNP P77239
B	412	HIS	-	EXPRESSION TAG	UNP P77239
B	413	HIS	-	EXPRESSION TAG	UNP P77239

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 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.

[illegible]

Chain B:

5% 28% 33% 9% 28%

MET LYS LYS PHE ILE ALA LEU ILE ILE GLY SER MET ILE ILE ALA GLY GLY ILE ILE SER ALA ALA GLY PHE THR TRP VAL ALA LYS GLU PRO PRO PRO GLU LYS THR SER THR ALA GLU THR ARG LYS ILE LEU PHE TRP TTR ASP PRO MET TYR ASN ARG PHE ASP LYS PRO GLY

SER PRO PHE MET ASP MET ASP VAL PRO TYR ALA ASP GLU SER ALA GLY VAL ARG ILE ASP PRO THR GLN T89 Q90 P91 L92 G93 V94 K95 T96 A97 V98 T99 L100 R101 G102 P103 L104 T105 F106 A107 S115 M116 Y117 Q120 Y121 A122 I123 V124 Q125 A126 P127

A128 A129 G130 F131 I132 D133 K134 V135 V136 P137 L138 T139 V140 G141 D142 K143 V144 G145 T147 T148 P149 L150 D152 L153 T154 I155 P156 D157 W158 E159 E160 A161 Q162 S163 E164 Y165 L166 L167 L168 T171 T174 A175 T176 Q177 T178 L181 E183 R184 L185 R186 L187 E192 I195

HIS	R340	L266	R196
HIS	V341		R197
HIS	I342	P269	L198
HIS	T343	A270	I199
HIS	V344	R271	
HIS	D345	P272	Q202
HIS	A346	D273	K203
	D347	K274	
	G348	T275	L204
	R349	L276	Q205
	F350	T277	T206
	V351	I278	R207
	P352	I279	F208
	K353	K280	T209
	R354	W281	L210
		T282	K211
		L283	A212
	Q359	L284	P213
	A360	P285	
	S361		L218
	Q362	D288	T219
	V364		A220
	T365	R292	F221
	A366	T293	D222
	L367	L294	L223
			R224
	L371	R297	
	A372	L298	M227
		E299	N228
	E375	V300	I229
	K376	D301	A230
	V377	N302	K231
			D232
	S379	L307	N233
	S380	K308	
	G381	P309	I238
	L382	G310	Q239
		N311	G240
		N312	M241
	I385		D242
ASP		P243	P243
SER		L315	V244
GLU		Q316	W245
ALA		L317	
ASN		N318	
ILE		T319	A249
SER		A320	I250
GLY		S321	P251
ALA		E322	E252
LEU		P323	S253
GLU			I254
ARG		L326	A255
MET		I327	W256
ARG		P328	L257
SER		S329	V258
GLU			K259
SER		L332	D260
SER			A261
ALA		S337	S262
THR		E338	Q263
HIS			F264
ALA		Q339	T265

4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	84.96Å 113.24Å 258.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.10 – 3.40 47.10 – 3.40	Depositor EDS
% Data completeness (in resolution range)	89.3 (47.10-3.40) 98.0 (47.10-3.40)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.261 , 0.318 0.272 , 0.320	Depositor DCC
R_{free} test set	867 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	132.2	Xtriage
Anisotropy	0.664	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 145.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 17209 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4548	wwPDB-VP
Average B, all atoms (Å ²)	189.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.85% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.24	0/2313	0.50	0/3152
1	B	0.25	0/2313	0.50	0/3152
All	All	0.25	0/4626	0.50	0/6304

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	2274	0	2343	166	0
1	B	2274	0	2343	180	0
All	All	4548	0	4686	339	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 37.

The worst 5 of 339 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:261:ALA:HB2	1:A:281:TRP:CD1	2.04	0.93
1:B:107:ALA:HB2	1:B:362:GLN:HE21	1.35	0.89

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:229:ILE:HG23	1:A:233:ASN:HB2	1.54	0.89
1:B:166:LEU:HD22	1:B:202:GLN:HG3	1.52	0.89
1:B:340:ARG:HB3	1:B:353:LYS:O	1.76	0.85

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	295/413 (71%)	218 (74%)	60 (20%)	17 (6%)	2	20
1	B	295/413 (71%)	231 (78%)	46 (16%)	18 (6%)	2	18
All	All	590/826 (71%)	449 (76%)	106 (18%)	35 (6%)	2	19

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	135	VAL
1	A	140	VAL
1	A	149	PRO
1	A	150	LEU
1	A	221	PHE

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	243/338 (72%)	201 (83%)	42 (17%)	2	13
1	B	243/338 (72%)	200 (82%)	43 (18%)	2	12
All	All	486/676 (72%)	401 (82%)	85 (18%)	2	13

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	351	VAL
1	B	138	LEU
1	B	284	LEU
1	A	365	THR
1	B	98	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	202	GLN
1	B	362	GLN
1	B	233	ASN
1	A	239	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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	297/413 (71%)	0.57	17 (5%) 27 25	136, 169, 291, 372	0
1	B	297/413 (71%)	0.55	21 (7%) 19 18	133, 182, 279, 427	0
All	All	594/826 (71%)	0.56	38 (6%) 23 21	133, 175, 284, 427	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	385	ILE	7.7
1	B	90	GLN	5.4
1	B	337	SER	4.4
1	B	343	THR	3.7
1	B	158	TRP	3.5

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.