wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 11:19 AM GMT

PDB ID	:	3OPO
Title	:	Crystal structure of the membrane fusion protein CusB from Escherichia coli
Authors	:	Su, C.-C.
Deposited on	:	2010-09-01
Resolution	:	3.85 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

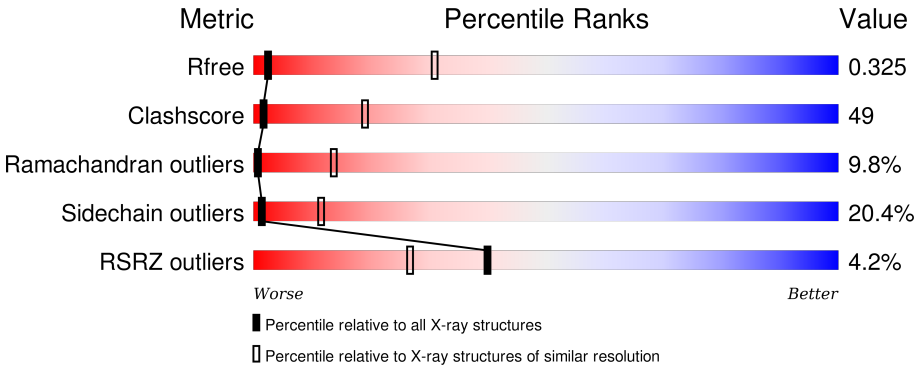
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.85 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	1334 (4.18-3.50)
Clashscore	102246	1036 (4.16-3.52)
Ramachandran outliers	100387	1415 (4.18-3.50)
Sidechain outliers	100360	1410 (4.18-3.50)
RSRZ outliers	91569	1342 (4.18-3.50)

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 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 <=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	2% 22% 36% 12% • 28%
1	B	413	4% 22% 36% 12% • 28%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	AG	A	414	-	-	-	X

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4543 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	296	Total	C	N	O	S	0	0	0
			2267	1444	391	427	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

- Molecule 2 is SILVER ION (three-letter code: AG) (formula: Ag).

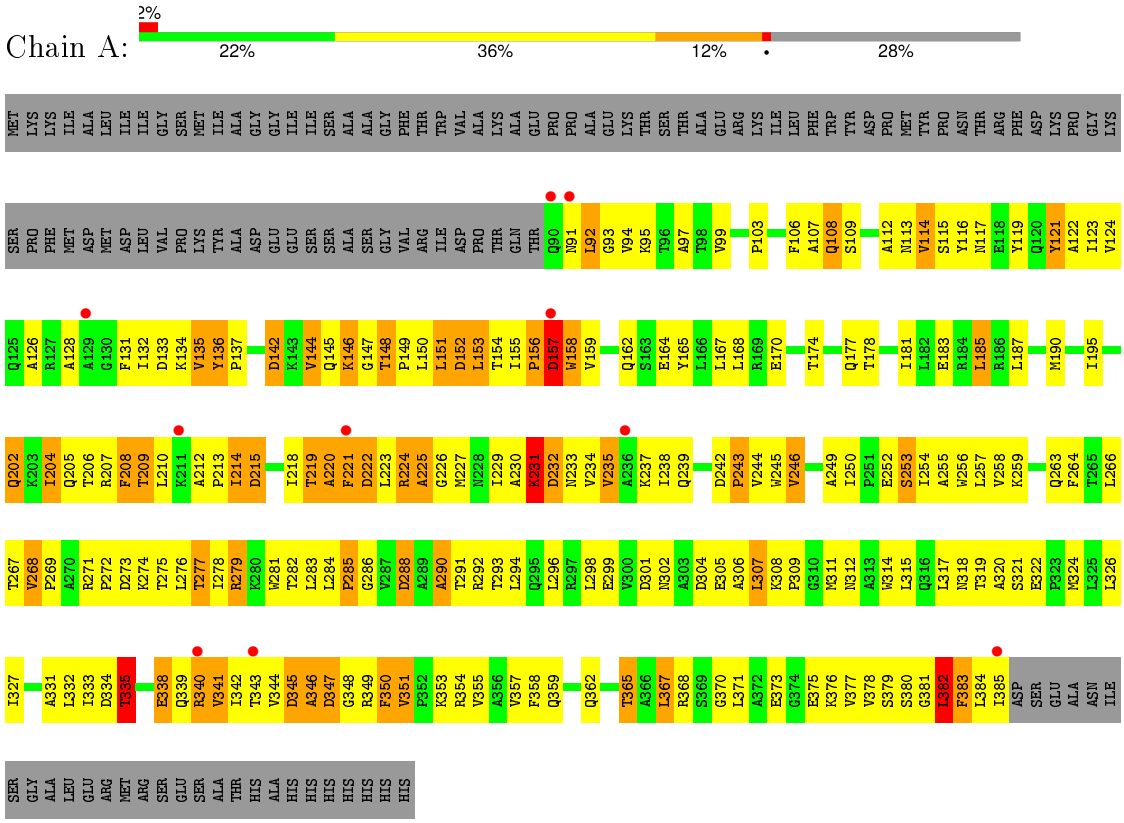
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ag	0	0
			1	1		
2	A	1	Total	Ag	0	0
			1	1		

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 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.

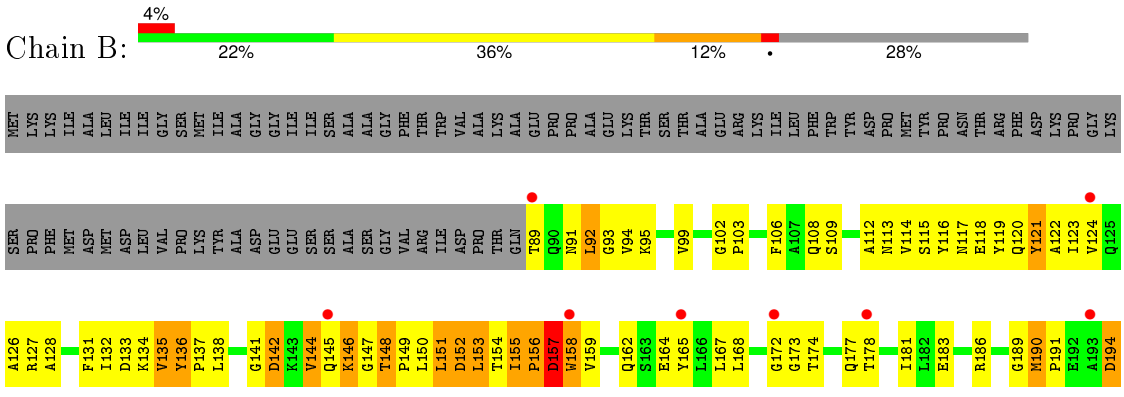
- Molecule 1: Cation efflux system protein cusB

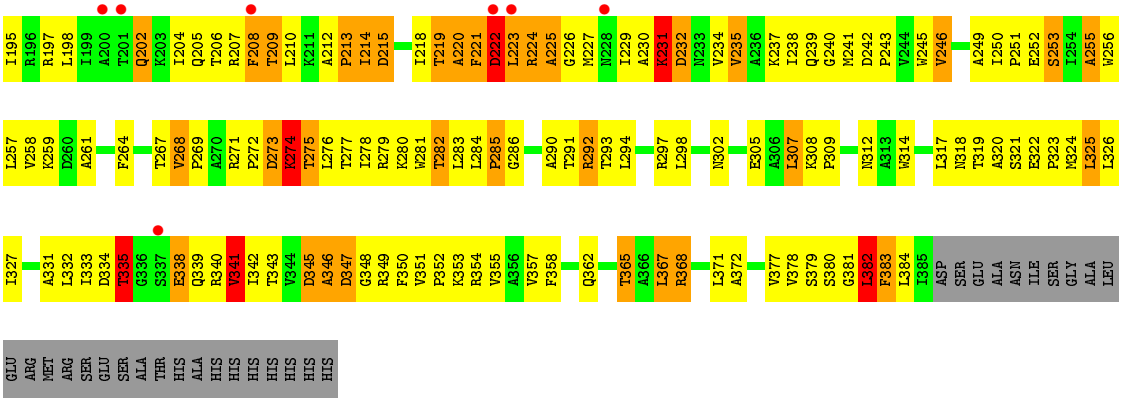
Chain A:



- Molecule 1: Cation efflux system protein cusB

Chain B:




I195 R196 R197 L198 I199 A200 T201 Q202 K203 I204 Q205 T206 R207 F208 T209 L210 K211 A212 P213 I214 D215 I218 T219 A220 F221 D222 L223 R224 A225 G226 M227 W228 I229 A230 K231 D232 N233 V234 V235 A236 K237 I238 Q239 G240 M241 D242 P243 V244 W245 V246 A249 I250 P251 E252 S253 I254 A255 W256
L257 V258 K259 D260 A261 F264 T267 V268 P269 A270 R271 P272 D273 K274 T275 L276 T277 I278 R279 K280 W281 T282 L283 L284 P285 G286 A290 T291 R292 T293 L294 R297 L298 N302 E305 A306 I307 K308 P309 N312 A313 W314 L317 N318 T319 A320 S321 E322 P323 M324 L325 L326
I327 A331 L332 I333 D334 T335 G336 S337 E338 Q339 R340 V341 I342 T343 V344 D345 A346 D347 G348 R349 F350 V351 P352 K353 R354 V355 A356 V357 F358 Q362 T365 A366 L367 R368 L371 A372 V377 V378 S379 S380 G381 L382 F383 L384 I385 ASP SER GLU ALA ASN ILE SER GLY ALA LEU
GLU ARG MET ARG SER GLU SER ALA THR HIS ALA HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	84.80Å 114.68Å 259.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.31 – 3.85 39.31 – 3.85	Depositor EDS
% Data completeness (in resolution range)	90.9 (39.31-3.85) 98.8 (39.31-3.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.88 (at 3.87Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.295 , 0.340 0.275 , 0.325	Depositor DCC
R_{free} test set	585 reflections (4.79%)	DCC
Wilson B-factor (Å^2)	159.3	Xtriage
Anisotropy	0.686	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 238.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Outliers	0 of 12235 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4543	wwPDB-VP
Average B, all atoms (Å^2)	222.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.04% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: AG

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	0.32	1/2306 (0.0%)	0.53	0/3142
1	B	0.27	0/2313	0.54	0/3152
All	All	0.30	1/4619 (0.0%)	0.53	0/6294

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	322	GLU	CB-CG	5.26	1.62	1.52

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	2267	0	2336	233	0
1	B	2274	0	2343	227	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
All	All	4543	0	4679	449	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 49.

The worst 5 of 449 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:332:LEU:HA	1:A:341:VAL:HG11	1.32	1.08
1:A:117:ASN:HD21	1:A:243:PRO:HG2	1.21	1.01
1:A:122:ALA:HB3	1:A:214:ILE:HD12	1.49	0.94
1:A:106:PHE:CD2	1:B:253:SER:HA	2.06	0.89
1:A:249:ALA:HB1	1:A:293:THR:HG21	1.52	0.89

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	294/413 (71%)	220 (75%)	45 (15%)	29 (10%)	1 13
1	B	295/413 (71%)	218 (74%)	48 (16%)	29 (10%)	1 14
All	All	589/826 (71%)	438 (74%)	93 (16%)	58 (10%)	1 14

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	135	VAL
1	A	146	LYS
1	A	156	PRO
1	A	157	ASP
1	A	206	THR

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	242/338 (72%)	192 (79%)	50 (21%)	1 11
1	B	243/338 (72%)	194 (80%)	49 (20%)	1 12
All	All	485/676 (72%)	386 (80%)	99 (20%)	1 11

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	367	LEU
1	B	142	ASP
1	B	345	ASP
1	A	375	GLU
1	B	92	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	330	GLN
1	A	339	GLN
1	B	162	GLN
1	A	318	ASN
1	B	91	ASN

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

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	296/413 (71%)	0.13	10 (3%)	49	34	112, 202, 318, 446	0
1	B	297/413 (71%)	0.30	15 (5%)	32	21	120, 214, 362, 581	0
All	All	593/826 (71%)	0.22	25 (4%)	40	27	112, 209, 343, 581	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	201	THR	7.6
1	B	200	ALA	3.4
1	A	90	GLN	3.2
1	B	158	TRP	3.1
1	B	89	THR	2.8

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
2	AG	A	414	1/1	0.77	1.36	11.75	400,400,400,400	0
2	AG	B	414	1/1	0.80	1.08	-	268,268,268,268	0

6.5 Other polymers (i)

There are no such residues in this entry.