WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 08:31 AM EDT

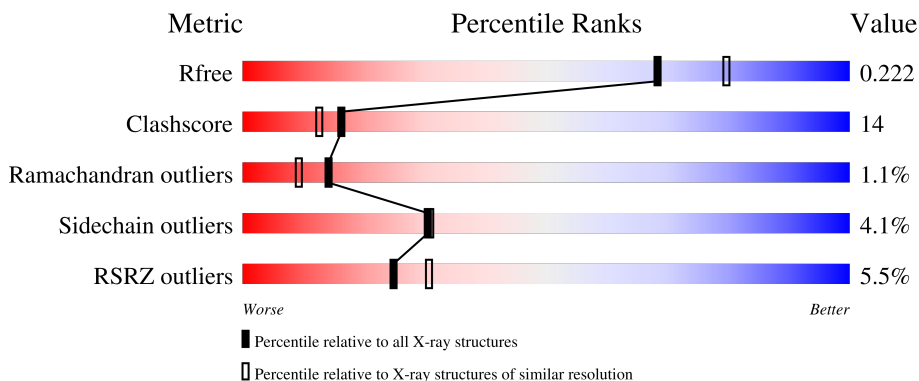|		
PDB ID	:	1X8C
Title	:	Crystal structure of the SeMet-derivative copper homeostasis protein (cutCm) with calcium binding from Shigella flexneri 2a str. 301
Authors	:	Wang, D.C.; Zhu, D.Y.
Deposited on	:	2004-08-18
Resolution	:	2.10 Å(reported)

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 2.10 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		0.222
Clashscore		14
Ramachandran outliers		1.1%
Sidechain outliers		4.1%
RSRZ outliers		5.5%

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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	256	4% / 68% 21% • 9%
1	B	256	6% / 72% 18% • 7%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3947 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 Copper homeostasis protein cutC.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	234	Total	C	N	O	S	Se		0	0	0
			1765	1108	317	324	6	10				
1	B	238	Total	C	N	O	S	Se		0	0	0
			1799	1128	322	333	6	10				

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	11	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	103	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	106	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	110	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	126	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	162	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	174	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	207	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	216	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	238	MSE	MET	MODIFIED RESIDUE	UNP P67825
A	249	LEU	-	EXPRESSION TAG	UNP P67825
A	250	GLU	-	EXPRESSION TAG	UNP P67825
A	251	HIS	-	EXPRESSION TAG	UNP P67825
A	252	HIS	-	EXPRESSION TAG	UNP P67825
A	253	HIS	-	EXPRESSION TAG	UNP P67825
A	254	HIS	-	EXPRESSION TAG	UNP P67825
A	255	HIS	-	EXPRESSION TAG	UNP P67825
A	256	HIS	-	EXPRESSION TAG	UNP P67825
B	1	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	11	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	103	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	106	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	110	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	126	MSE	MET	MODIFIED RESIDUE	UNP P67825

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	162	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	174	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	207	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	216	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	238	MSE	MET	MODIFIED RESIDUE	UNP P67825
B	249	LEU	-	EXPRESSION TAG	UNP P67825
B	250	GLU	-	EXPRESSION TAG	UNP P67825
B	251	HIS	-	EXPRESSION TAG	UNP P67825
B	252	HIS	-	EXPRESSION TAG	UNP P67825
B	253	HIS	-	EXPRESSION TAG	UNP P67825
B	254	HIS	-	EXPRESSION TAG	UNP P67825
B	255	HIS	-	EXPRESSION TAG	UNP P67825
B	256	HIS	-	EXPRESSION TAG	UNP P67825

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Ca 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	191	Total 191	O 191	0	0
3	B	191	Total 191	O 191	0	0

3 Residue-property plots ⓘ

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: Copper homeostasis protein cutC

Chain A:

• Molecule 1: Copper homeostasis protein cutC

Chain B:
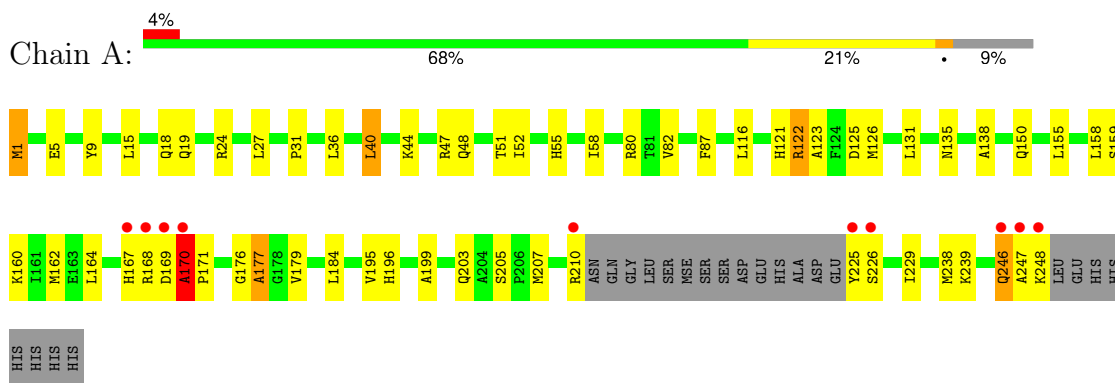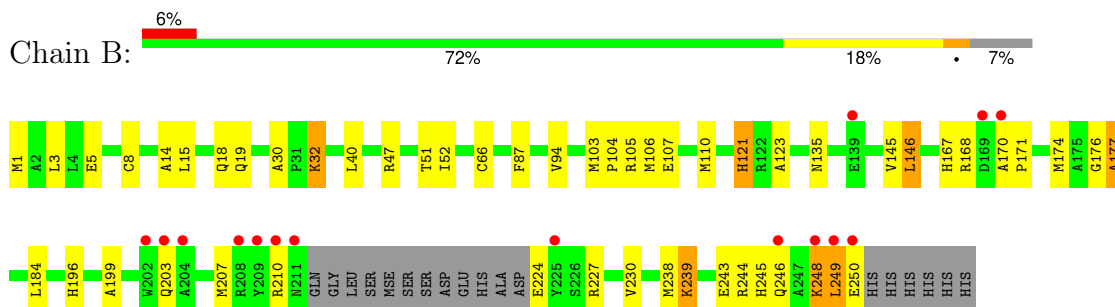
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	74.58Å 97.58Å 131.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	59.26 – 2.10 59.25 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.4 (59.26-2.10) 99.6 (59.25-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	8.73 (at 2.10Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.196 , 0.222 0.193 , 0.222	Depositor DCC
R_{free} test set	1377 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	14.9	Xtriage
Anisotropy	0.489	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 53.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3947	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.80% 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: CA

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	0.30	0/1786	0.60	1/2400 (0.0%)
1	B	0.30	0/1820	0.59	0/2446
All	All	0.30	0/3606	0.59	1/4846 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	170	ALA	C-N-CD	5.06	139.03	128.40

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	1765	0	1781	59	0
1	B	1799	0	1810	47	0
2	B	1	0	0	0	0
3	A	191	0	0	3	0
3	B	191	0	0	1	1
All	All	3947	0	3591	103	1

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 14.

The worst 5 of 103 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:B:18:GLN:HE22	1:B:51:THR:H	1.07	0.99
1:B:203:GLN:NE2	1:B:227:ARG:HH11	1.63	0.97
1:A:18:GLN:HE22	1:A:51:THR:H	1.18	0.89
1:A:168:ARG:NE	1:A:170:ALA:HB2	1.87	0.89
1:A:135:ASN:ND2	1:A:168:ARG:HH21	1.73	0.85

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:601:HOH:O	3:B:601:HOH:O[4_566]	1.79	0.41

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	230/256 (90%)	224 (97%)	3 (1%)	3 (1%)	12	7
1	B	234/256 (91%)	227 (97%)	5 (2%)	2 (1%)	17	12
All	All	464/512 (91%)	451 (97%)	8 (2%)	5 (1%)	14	9

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	170	ALA
1	A	177	ALA
1	B	177	ALA
1	A	122	ARG
1	B	249	LEU

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	183/192 (95%)	176 (96%)	7 (4%)	33	34
1	B	187/192 (97%)	179 (96%)	8 (4%)	29	29
All	All	370/384 (96%)	355 (96%)	15 (4%)	30	31

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	3	LEU
1	B	239	LYS
1	B	32	LYS
1	B	248	LYS
1	B	184	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	185	HIS
1	B	196	HIS
1	A	196	HIS
1	B	203	GLN
1	B	55	HIS

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 1 ligands modelled in this entry, 1 is 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 (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 ⓘ

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, 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	224/256 (87%)	0.10	10 (4%)	33	38	7, 14, 34, 60	0
1	B	228/256 (89%)	0.19	15 (6%)	18	23	6, 13, 43, 66	0
All	All	452/512 (88%)	0.14	25 (5%)	25	31	6, 13, 36, 66	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	225	TYR	9.4
1	A	248	LYS	6.5
1	B	210	ARG	5.5
1	B	250	GLU	5.2
1	A	247	ALA	5.0

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, 95th 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($\mathring{A}^2$)	Q<0.9
2	CA	B	501	1/1	0.99	0.07	26,26,26,26	0

6.5 Other polymers ⓘ

There are no such residues in this entry.