



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2017 – 11:15 AM EDT

PDB ID : 3DA3
Title : Crystal Structure of Colicin M, A Novel Phosphatase Specifically Imported by Escherichia Coli
Authors : Zeth, K.; Albrecht, R.; Romer, C.; Braun, V.
Deposited on : unknown
Resolution : 2.50 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

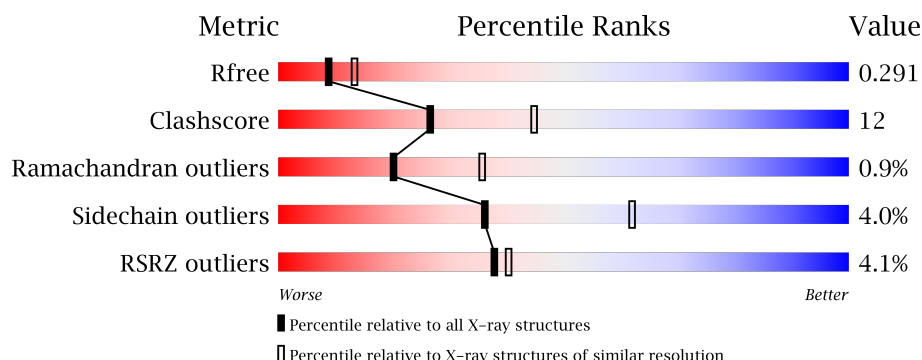
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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 $\leq 5\%$. 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	278	<div> <div>3%</div> <div> <div></div> <div>76%</div> <div>19%</div> <div>• •</div> </div> </div>
1	B	278	<div> <div>5%</div> <div> <div></div> <div>73%</div> <div>22%</div> <div>• •</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4280 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 Colicin-M.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total	C	N	O	S	0	0	0
			2067	1317	352	390	8			
1	B	270	Total	C	N	O	S	0	0	0
			2067	1317	352	390	8			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	272	ALA	-	EXPRESSION TAG	UNP P05820
A	273	HIS	-	EXPRESSION TAG	UNP P05820
A	274	HIS	-	EXPRESSION TAG	UNP P05820
A	275	HIS	-	EXPRESSION TAG	UNP P05820
A	276	HIS	-	EXPRESSION TAG	UNP P05820
A	277	HIS	-	EXPRESSION TAG	UNP P05820
A	278	HIS	-	EXPRESSION TAG	UNP P05820
B	272	ALA	-	EXPRESSION TAG	UNP P05820
B	273	HIS	-	EXPRESSION TAG	UNP P05820
B	274	HIS	-	EXPRESSION TAG	UNP P05820
B	275	HIS	-	EXPRESSION TAG	UNP P05820
B	276	HIS	-	EXPRESSION TAG	UNP P05820
B	277	HIS	-	EXPRESSION TAG	UNP P05820
B	278	HIS	-	EXPRESSION TAG	UNP P05820

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	68	Total 68	O 68	0	0
3	B	76	Total 76	O 76	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	119.88Å 119.88Å 96.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.98 – 2.50 19.98 – 2.40	Depositor EDS
% Data completeness (in resolution range)	94.2 (19.98-2.50) 94.2 (19.98-2.40)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.29 (at 2.41Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.235 , 0.295 0.235 , 0.291	Depositor DCC
R_{free} test set	1152 reflections (5.23%)	DCC
Wilson B-factor (Å ²)	28.3	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 38.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4280	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.74% 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, 0.333 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: MG

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	0/2118	0.59	0/2881
1	B	0.49	0/2118	0.60	0/2881
All	All	0.50	0/4236	0.59	0/5762

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	2067	0	2044	55	0
1	B	2067	0	2044	46	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	68	0	0	11	0
3	B	76	0	0	6	0
All	All	4280	0	4088	95	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 12.

The worst 5 of 95 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:172:VAL:HB	1:A:173:GLY:HA3	1.26	1.18
1:A:161:ASN:HA	3:A:331:HOH:O	1.44	1.14
1:B:172:VAL:HB	1:B:174:THR:H	1.25	0.99
1:A:172:VAL:HB	1:A:173:GLY:CA	2.01	0.90
1:B:19:GLY:H	1:B:79:ASN:HD21	1.12	0.88

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	268/278 (96%)	257 (96%)	9 (3%)	2 (1%)	25	43
1	B	268/278 (96%)	253 (94%)	12 (4%)	3 (1%)	17	29
All	All	536/556 (96%)	510 (95%)	21 (4%)	5 (1%)	20	36

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	THR
1	A	160	ILE
1	B	164	LYS
1	B	160	ILE
1	B	170	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 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	225/233 (97%)	216 (96%)	9 (4%)	36	62
1	B	225/233 (97%)	216 (96%)	9 (4%)	36	62
All	All	450/466 (97%)	432 (96%)	18 (4%)	36	62

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	270	LYS
1	B	13	THR
1	B	172	VAL
1	A	243	LEU
1	A	251	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	161	ASN
1	B	7	HIS
1	B	118	ASN
1	A	135	HIS
1	B	135	HIS

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 [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	270/278 (97%)	0.08	9 (3%) 47 50	15, 27, 46, 55	0
1	B	270/278 (97%)	0.03	13 (4%) 31 32	15, 27, 46, 56	0
All	All	540/556 (97%)	0.06	22 (4%) 38 40	15, 27, 47, 56	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	173	GLY	9.9
1	B	170	GLY	4.7
1	B	173	GLY	4.6
1	A	172	VAL	4.6
1	A	174	THR	4.0

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](#)

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, 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	LLDF	B-factors(\AA^2)	Q<0.9
2	MG	A	279	1/1	0.99	0.09	-1.21	14,14,14,14	0
2	MG	B	279	1/1	0.98	0.09	-1.40	15,15,15,15	0

6.5 Other polymers [i](#)

There are no such residues in this entry.