



wwPDB X-ray Structure Validation Summary Report ⓘ

May 19, 2020 – 09:01 pm BST

PDB ID : 3DA4
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 : 2008-05-28
Resolution : 1.70 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

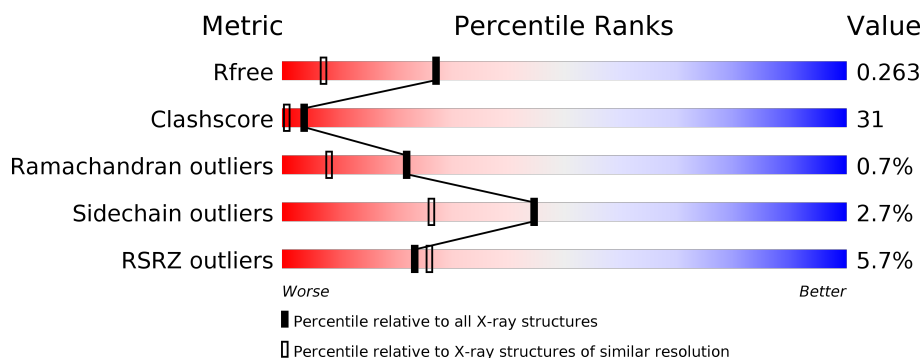
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 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 $\leq 5\%$. 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	279	<div> <div>5%</div> <div> <div></div> <div>53%</div> <div>41%</div> <div>• •</div> </div> </div>
1	B	279	<div> <div>6%</div> <div> <div></div> <div>49%</div> <div>45%</div> <div>• •</div> </div> </div>

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	NO3	A	282	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4801 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	1	0
			2074	1322	353	391	8			
1	B	270	Total	C	N	O	S	0	0	0
			2066	1316	352	390	8			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	272	ALA	-	EXPRESSION TAG	UNP P05820
A	273	ALA	-	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
A	279	HIS	-	EXPRESSION TAG	UNP P05820
B	272	ALA	-	EXPRESSION TAG	UNP P05820
B	273	ALA	-	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
B	279	HIS	-	EXPRESSION TAG	UNP P05820

- Molecule 2 is NITRATE ION (three-letter code: NO3) (formula: NO₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	N	O	0	0
			4	1	3		
2	A	1	Total	N	O	0	0
			4	1	3		
2	A	1	Total	N	O	0	0
			4	1	3		
2	A	1	Total	N	O	0	0
			4	1	3		
2	B	1	Total	N	O	0	0
			4	1	3		

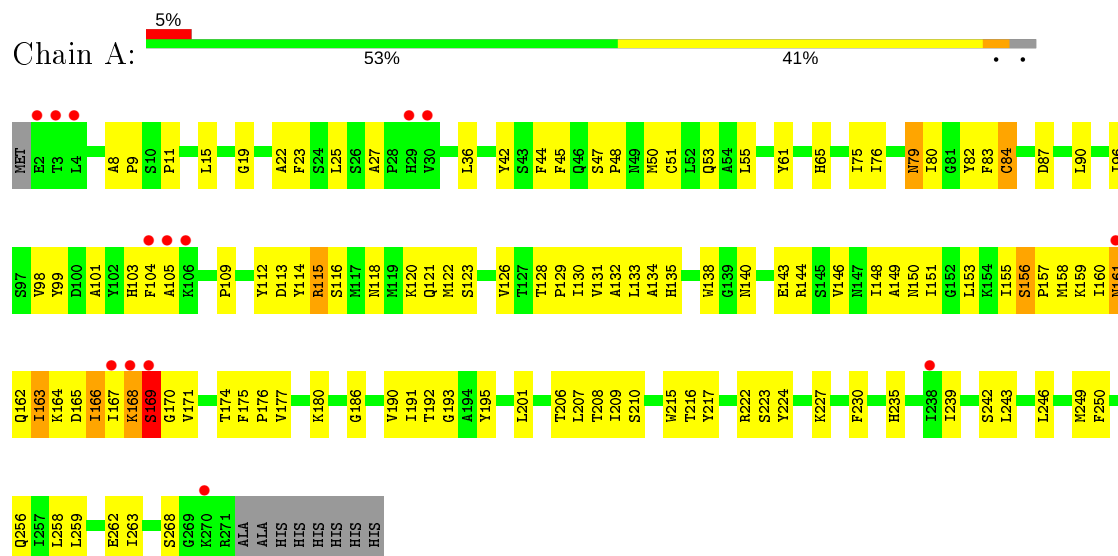
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	325	Total	O	0	0
			325	325		
3	B	316	Total	O	0	0
			316	316		

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 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: Colicin-M



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	50.49Å 108.75Å 224.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.70 29.45 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-1.70) 97.8 (29.45-1.70)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.218 , 0.261 0.224 , 0.263	Depositor DCC
R_{free} test set	4692 reflections (7.00%)	wwPDB-VP
Wilson B-factor (Å ²)	14.7	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 50.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4801	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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: NO3

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	0.87	1/2128 (0.0%)	0.83	1/2895 (0.0%)
1	B	0.82	1/2117 (0.0%)	0.79	0/2880
All	All	0.84	2/4245 (0.0%)	0.81	1/5775 (0.0%)

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	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	84	CYS	CB-SG	-6.37	1.71	1.82
1	B	51	CYS	CB-SG	-5.43	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	169	SER	N-CA-C	-5.28	96.75	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	166	ILE	Peptide
1	A	168	LYS	Peptide

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	2074	0	2057	127	0
1	B	2066	0	2042	128	0
2	A	16	0	0	2	0
2	B	4	0	0	1	0
3	A	325	0	0	15	0
3	B	316	0	0	17	0
All	All	4801	0	4099	253	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 31.

The worst 5 of 253 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:162:GLN:HE21	1:A:177:VAL:HG11	0.98	1.09
1:A:163:ILE:O	1:A:167:ILE:HG22	1.55	1.07
1:B:162:GLN:HE21	1:B:177:VAL:HG11	1.25	1.00
1:A:160:ILE:O	1:A:163:ILE:HG22	1.62	0.99
1:B:162:GLN:NE2	1:B:177:VAL:HG11	1.76	0.98

There are no symmetry-related clashes.

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	269/279 (96%)	258 (96%)	11 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	268/279 (96%)	259 (97%)	5 (2%)	4 (2%)	10	2
All	All	537/558 (96%)	517 (96%)	16 (3%)	4 (1%)	22	8

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	169	SER
1	B	105	ALA
1	B	168	LYS
1	B	106	LYS

5.3.2 Protein sidechains ⓘ

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	227/233 (97%)	217 (96%)	10 (4%)	28	11
1	B	225/233 (97%)	222 (99%)	3 (1%)	69	56
All	All	452/466 (97%)	439 (97%)	13 (3%)	44	23

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	161[A]	ASN
1	A	161[B]	ASN
1	B	64	LYS
1	A	156	SER
1	A	210	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	235	HIS
1	B	49	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	111	GLN
1	A	140	ASN
1	B	103	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	NO3	A	281	-	1,3,3	3.59	1 (100%)	0,3,3	0.00	-
2	NO3	A	280	-	1,3,3	3.01	1 (100%)	0,3,3	0.00	-
2	NO3	A	283	-	1,3,3	3.63	1 (100%)	0,3,3	0.00	-
2	NO3	A	282	-	1,3,3	3.34	1 (100%)	0,3,3	0.00	-
2	NO3	B	280	-	1,3,3	3.64	1 (100%)	0,3,3	0.00	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	280	NO3	O1-N	3.64	1.40	1.24
2	A	283	NO3	O1-N	3.63	1.40	1.24
2	A	281	NO3	O1-N	3.59	1.40	1.24
2	A	282	NO3	O1-N	3.34	1.39	1.24
2	A	280	NO3	O1-N	3.01	1.37	1.24

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	282	NO3	2	0
2	B	280	NO3	1	0

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	270/279 (96%)	0.49	14 (5%) 27 30	6, 13, 22, 31	0
1	B	270/279 (96%)	0.44	17 (6%) 20 22	6, 13, 26, 34	0
All	All	540/558 (96%)	0.47	31 (5%) 23 26	6, 13, 24, 34	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	3	THR	6.9
1	A	167	ILE	6.1
1	B	168	LYS	5.1
1	A	168	LYS	4.4
1	A	29	HIS	4.4

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. 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(\AA^2)	Q<0.9
2	NO3	A	283	4/4	0.80	0.20	36,37,37,37	0
2	NO3	B	280	4/4	0.86	0.20	31,32,32,33	0
2	NO3	A	282	4/4	0.91	0.14	32,32,32,32	0
2	NO3	A	281	4/4	0.93	0.20	28,28,28,29	0
2	NO3	A	280	4/4	0.96	0.15	17,17,18,18	0

6.5 Other polymers [i](#)

There are no such residues in this entry.