WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2025 – 06:04 PM EST

PDB ID	:	9BBJ / pdb_00009bbj
Title	:	M. tuberculosis ClpC1-NTD complexed with a click chemistry analog of Rufomycin
Authors	:	Abad-Zapatero, C.; Ratia, K.
Deposited on	:	2024-04-05
Resolution	:	2.56 Å(reported)

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

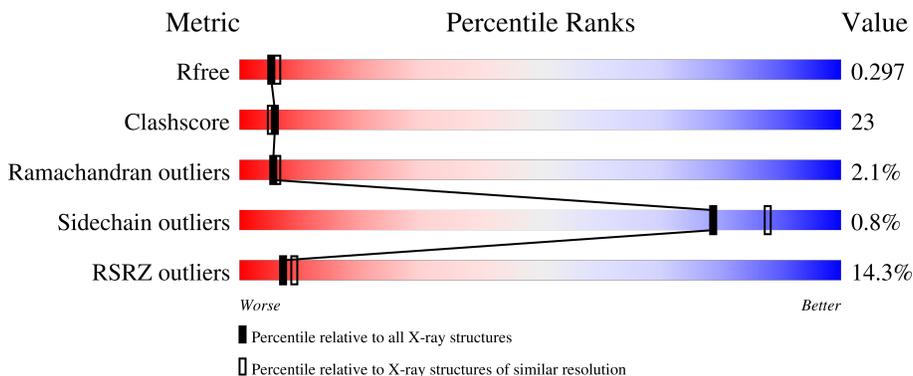|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

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 2.56 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		0.297
Clashscore		23
Ramachandran outliers		2.1%
Sidechain outliers		0.8%
RSRZ outliers		14.3%

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}	164625	1685 (2.58-2.54)
Clashscore	180529	1779 (2.58-2.54)
Ramachandran outliers	177936	1766 (2.58-2.54)
Sidechain outliers	177891	1766 (2.58-2.54)
RSRZ outliers	164620	1685 (2.58-2.54)

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	158	13% / 50% 41% • 8%
2	B	7	29% 14% 57%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 1191 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 ATP-dependent Clp protease ATP-binding subunit ClpC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	145	1125	709	207	207	2	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	146	LYS	-	expression tag	UNP P9WPC9
A	147	LEU	-	expression tag	UNP P9WPC9
A	148	ALA	-	expression tag	UNP P9WPC9
A	149	ALA	-	expression tag	UNP P9WPC9
A	150	ALA	-	expression tag	UNP P9WPC9
A	151	LEU	-	expression tag	UNP P9WPC9
A	152	GLU	-	expression tag	UNP P9WPC9
A	153	HIS	-	expression tag	UNP P9WPC9
A	154	HIS	-	expression tag	UNP P9WPC9
A	155	HIS	-	expression tag	UNP P9WPC9
A	156	HIS	-	expression tag	UNP P9WPC9
A	157	HIS	-	expression tag	UNP P9WPC9
A	158	HIS	-	expression tag	UNP P9WPC9

- Molecule 2 is a protein called Click chemistry analog of Rufomycin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	7	66	48	11	7	0	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: ATP-dependent Clp protease ATP-binding subunit ClpC1

Chain A:

• Molecule 2: Click chemistry analog of Rufomycin

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	57.97Å 57.97Å 131.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.15 – 2.56 39.15 – 2.56	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.15-2.56) 99.9 (39.15-2.56)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 2.55Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874, PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.259 , 0.294 0.264 , 0.297	Depositor DCC
R_{free} test set	393 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	72.1	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 89.9	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	0.93	EDS
Total number of atoms	1191	wwPDB-VP
Average B, all atoms (Å^2)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.25% 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: A1AL8, MLE, A1AMB, NLE, A1AMC

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	0.47	0/1139	0.81	0/1534
2	B	3.72	1/11 (9.1%)	1.66	0/12
All	All	0.59	1/1150 (0.1%)	0.82	0/1546

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
2	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	6	LEU	N-CA	-9.91	1.27	1.46

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	5	A1AMC	Mainchain

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	1125	0	1163	54	0
2	B	66	0	34	2	0
All	All	1191	0	1197	54	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 23.

The worst 5 of 54 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:21:ARG:NH2	1:A:76:GLY:O	2.10	0.84
1:A:130:GLU:HB3	1:A:133:ARG:HB2	1.72	0.71
1:A:31:GLU:HG2	1:A:32:HIS:N	2.13	0.63
1:A:106:GLU:HG2	1:A:107:HIS:CD2	2.36	0.61
1:A:113:ILE:HG23	1:A:131:LEU:HB3	1.81	0.61

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	143/158 (90%)	129 (90%)	11 (8%)	3 (2%)	5	6
2	B	2/7 (29%)	2 (100%)	0	0	100	100
All	All	145/165 (88%)	131 (90%)	11 (8%)	3 (2%)	5	6

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	116	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	39	HIS
1	A	38	ILE

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	118/128 (92%)	118 (100%)	0	100	100
2	B	1/1 (100%)	0	1 (100%)	0	0
All	All	119/129 (92%)	118 (99%)	1 (1%)	79	88

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	6	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

5 non-standard protein/DNA/RNA residues 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	A1AL8	B	1	2	15,20,21	1.73	2 (13%)	17,29,31	1.46	3 (17%)
2	NLE	B	4	2	5,6,8	0.93	0	2,6,9	0.21	0
2	MLE	B	2	2	7,8,9	0.78	0	7,9,11	1.60	2 (28%)
2	A1AMC	B	5	2	6,7,9	3.74	3 (50%)	4,7,10	3.53	1 (25%)
2	A1AMB	B	7	2	8,10,11	1.04	1 (12%)	4,12,14	1.80	2 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	A1AL8	B	1	2	-	6/11/15/17	0/2/2/2
2	NLE	B	4	2	-	0/4/5/8	-
2	MLE	B	2	2	-	4/5/8/10	-
2	A1AMC	B	5	2	-	1/4/7/10	-
2	A1AMB	B	7	2	-	0/1/6/8	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	5	A1AMC	O-C	5.36	1.40	1.20
2	B	5	A1AMC	CB-CA	-5.26	1.41	1.53
2	B	5	A1AMC	CA-N	5.03	1.56	1.47
2	B	1	A1AL8	CD1-CG	4.97	1.43	1.37
2	B	1	A1AL8	CH2-CZ2	2.39	1.41	1.36

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	5	A1AMC	O-C-CA	-6.78	107.33	124.77
2	B	1	A1AL8	C2-C1-NE1	4.10	113.38	108.97
2	B	7	A1AMB	CD2-CE2-NE1	-2.72	105.56	109.20
2	B	2	MLE	CB-CA-C	-2.71	106.82	110.99
2	B	2	MLE	CN-N-CA	2.34	120.69	113.70

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	A1AL8	NE1-C1-C4-C5
2	B	2	MLE	C-CA-CB-CG
2	B	2	MLE	CA-CB-CG-CD2
2	B	2	MLE	CA-CB-CG-CD1
2	B	1	A1AL8	C2-C1-NE1-CE2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1	A1AL8	1	0
2	B	2	MLE	1	0

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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, 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	145/158 (91%)	1.13	21 (14%)	7	9	61, 101, 171, 205	0
2	B	2/7 (28%)	0.86	0	100	100	64, 64, 64, 66	0
All	All	147/165 (89%)	1.13	21 (14%)	7	9	61, 99, 171, 205	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	34	LEU	4.3
1	A	139	ILE	4.1
1	A	136	GLN	3.7
1	A	124	LEU	3.3
1	A	138	VAL	3.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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(Å^2)	Q<0.9
2	A1AMC	B	5	8/10	0.90	0.15	48,56,59,61	0
2	NLE	B	4	7/9	0.94	0.09	48,50,56,60	0
2	A1AMB	B	7	10/11	0.94	0.08	49,55,59,60	0
2	MLE	B	2	9/10	0.95	0.11	63,71,72,73	0
2	A1AL8	B	1	19/20	0.95	0.12	66,78,86,87	0

6.3 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.