



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 12:33 PM UTC

PDB ID : 8G28 / pdb_00008g28
Title : Crystal Structure of the C-terminal Fragment of AAA ATPase from Streptococcus pneumoniae.
Authors : Minasov, G.; Shuvalova, L.; Brunzelle, J.S.; Kiryukhina, O.; Satchell, K.J.F.; Center for Structural Biology of Infectious Diseases (CSBID); Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on : 2023-02-03
Resolution : 1.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) 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	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
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)	:	2.49

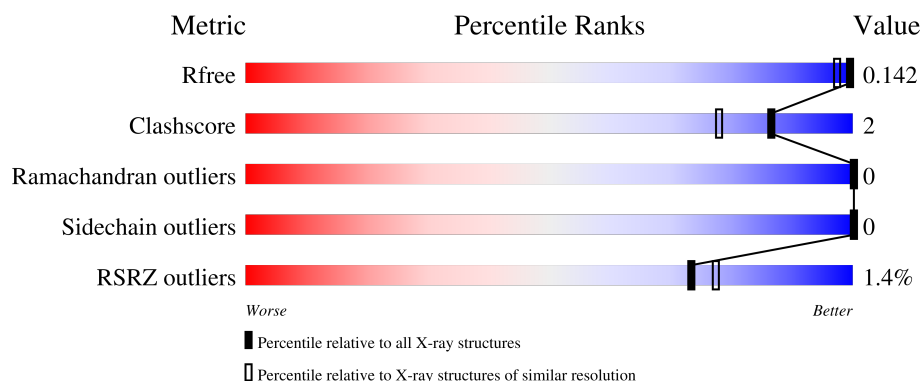
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.30 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}	180053	1553 (1.30-1.30)
Clashscore	190562	1595 (1.30-1.30)
Ramachandran outliers	187476	1551 (1.30-1.30)
Sidechain outliers	187428	1551 (1.30-1.30)
RSRZ outliers	180081	1549 (1.30-1.30)

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 $\leq 5\%$. 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	426	
1	B	426	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2260 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 ATPase, AAA family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	108	Total	C	N	O	Se	0	23	0
			1006	631	181	192	2			
1	B	104	Total	C	N	O	Se	0	12	0
			880	553	155	170	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP A0A0H2URL5
A	-1	ASN	-	expression tag	UNP A0A0H2URL5
A	0	ALA	-	expression tag	UNP A0A0H2URL5
B	-2	SER	-	expression tag	UNP A0A0H2URL5
B	-1	ASN	-	expression tag	UNP A0A0H2URL5
B	0	ALA	-	expression tag	UNP A0A0H2URL5

- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Cl	0	1
			2	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	152	Total	O	0	38
			178	178		
3	B	168	Total	O	0	39
			194	194		

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.

- [illegible]

- Chain B: 24% 76%

Sequence logo for Chain B, showing amino acid conservation across 100 positions. The y-axis represents the log-odds of an amino acid being at a position relative to the background frequency. The x-axis shows positions 1 to 100. The logo is color-coded by amino acid: Arg (red), Asn (blue), His (green), Tyr (orange), Phe (yellow), Glu (purple), Ala (pink), Gly (light blue), Ser (light green), Thr (light orange), Leu (light purple), Ile (light pink), Val (light green), Met (light blue), Lys (light orange), Asp (light yellow), Gln (light purple), Pro (light green), Trp (light orange), Cys (light pink), and others. Key positions include position 10 (Glu, high conservation), position 20 (Glu, high conservation), position 30 (Glu, high conservation), position 40 (Glu, high conservation), position 50 (Glu, high conservation), position 60 (Glu, high conservation), position 70 (Glu, high conservation), position 80 (Glu, high conservation), position 90 (Glu, high conservation), and position 100 (Glu, high conservation).

4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	71.53Å 71.53Å 151.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.18 – 1.30 29.18 – 1.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.18-1.30) 99.9 (29.18-1.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.25 (at 1.30Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.120 , 0.143 0.121 , 0.142	Depositor DCC
R_{free} test set	2879 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	13.7	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 57.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	2260	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 7.35% 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: CL

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	0.90	0/1021	0.91	0/1378
1	B	0.90	0/889	0.90	0/1202
All	All	0.90	0/1910	0.91	0/2580

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	1006	0	1021	8	0
1	B	880	0	904	5	0
2	A	2	0	0	0	0
3	A	178	0	0	1	0
3	B	194	0	0	0	0
All	All	2260	0	1925	8	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 2.

All (8) 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:250:ILE:O	1:A:328[A]:LYS:HE2	2.02	0.59
1:A:309[B]:PHE:HB3	1:A:310[B]:PRO:HA	1.90	0.53
1:A:336[B]:MSE:HE1	1:B:262:THR:CG2	2.40	0.51
1:A:336[B]:MSE:HE1	1:B:262:THR:HG21	1.94	0.49
1:A:336[B]:MSE:HB2	1:B:317:ALA:HB2	1.97	0.47
1:A:292[A]:GLU:HG2	3:A:657:HOH:O	2.18	0.44
1:A:276[A]:ARG:NH2	1:B:307[A]:ILE:HG13	2.34	0.43
1:A:336[A]:MSE:HB2	1:B:317:ALA:HB2	2.02	0.41

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	130/426 (30%)	130 (100%)	0	0	100	100
1	B	114/426 (27%)	112 (98%)	2 (2%)	0	100	100
All	All	244/852 (29%)	242 (99%)	2 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	108/352 (31%)	108 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	94/352 (27%)	94 (100%)	0	100	100
All	All	202/704 (29%)	202 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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	239	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 oligosaccharides 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	107/426 (25%)	-0.09	3 (2%) 55 61	6, 13, 28, 41	22 (20%)
1	B	103/426 (24%)	-0.18	0 100 100	6, 13, 25, 49	11 (10%)
All	All	210/852 (24%)	-0.13	3 (1%) 73 77	6, 13, 27, 49	33 (15%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	345	THR	3.7
1	A	238	GLY	2.2
1	A	241	ASP	2.1

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 oligosaccharides 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(Å ²)	Q<0.9
2	CL	A	501[A]	1/1	0.99	0.04	24,24,24,24	1
2	CL	A	501[B]	1/1	0.99	0.04	21,21,21,21	1

6.5 Other polymers [i](#)

There are no such residues in this entry.