



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:58 PM UTC

PDB ID : 8BKE / pdb_00008bke
Title : Crystal structure of the Klebsiella phage KP34p57 capsular depolymerase
Authors : Berisio, R.; Squeglia, F.
Deposited on : 2022-11-09
Resolution : 2.00 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 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
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

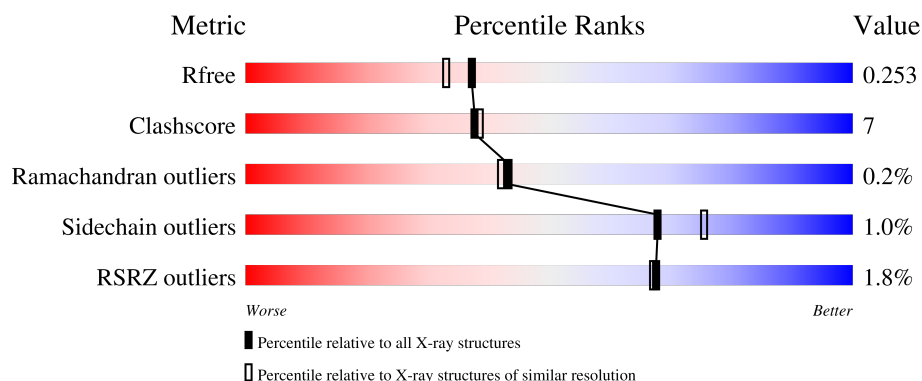
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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	AAA	638	57% 10% 33% 2%
1	BBB	638	56% 11% 33% 2%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6893 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 Uncharacterized protein 57.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	428	Total	C	N	O	S	0	0	0
			3237	2033	573	620	11			
1	BBB	428	Total	C	N	O	S	0	0	0
			3240	2034	573	622	11			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	631	LYS	-	expression tag	UNP D1L302
AAA	632	GLY	-	expression tag	UNP D1L302
AAA	633	HIS	-	expression tag	UNP D1L302
AAA	634	HIS	-	expression tag	UNP D1L302
AAA	635	HIS	-	expression tag	UNP D1L302
AAA	636	HIS	-	expression tag	UNP D1L302
AAA	637	HIS	-	expression tag	UNP D1L302
AAA	638	HIS	-	expression tag	UNP D1L302
BBB	631	LYS	-	expression tag	UNP D1L302
BBB	632	GLY	-	expression tag	UNP D1L302
BBB	633	HIS	-	expression tag	UNP D1L302
BBB	634	HIS	-	expression tag	UNP D1L302
BBB	635	HIS	-	expression tag	UNP D1L302
BBB	636	HIS	-	expression tag	UNP D1L302
BBB	637	HIS	-	expression tag	UNP D1L302
BBB	638	HIS	-	expression tag	UNP D1L302

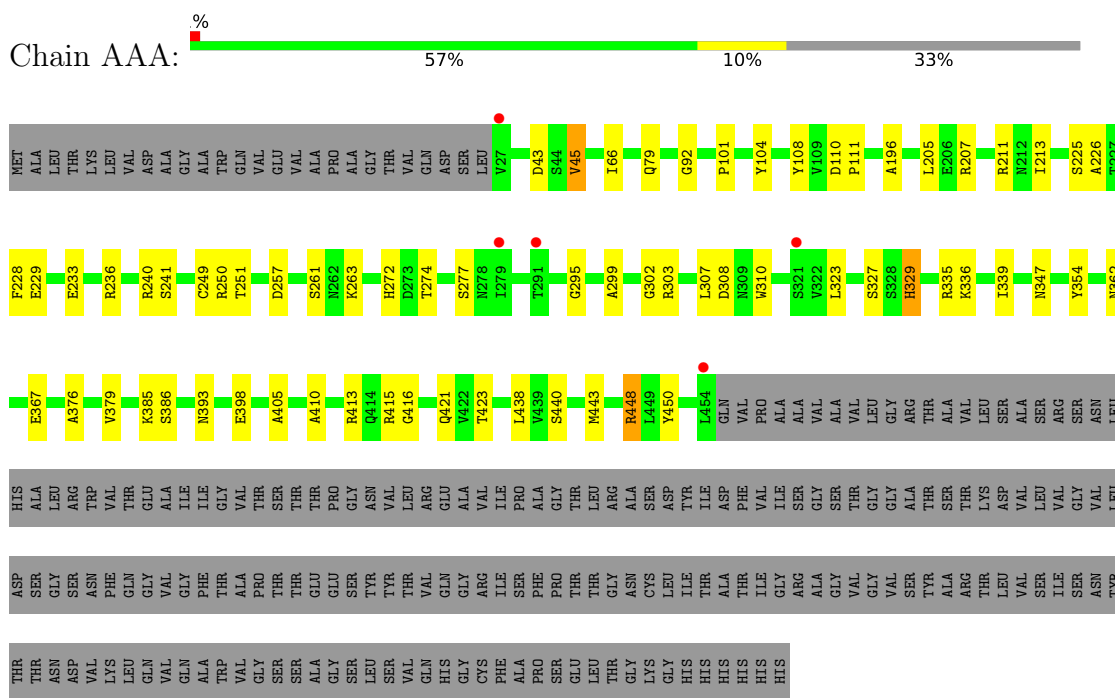
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	254	Total	O	0	0
			254	254		
2	BBB	162	Total	O	0	0
			162	162		

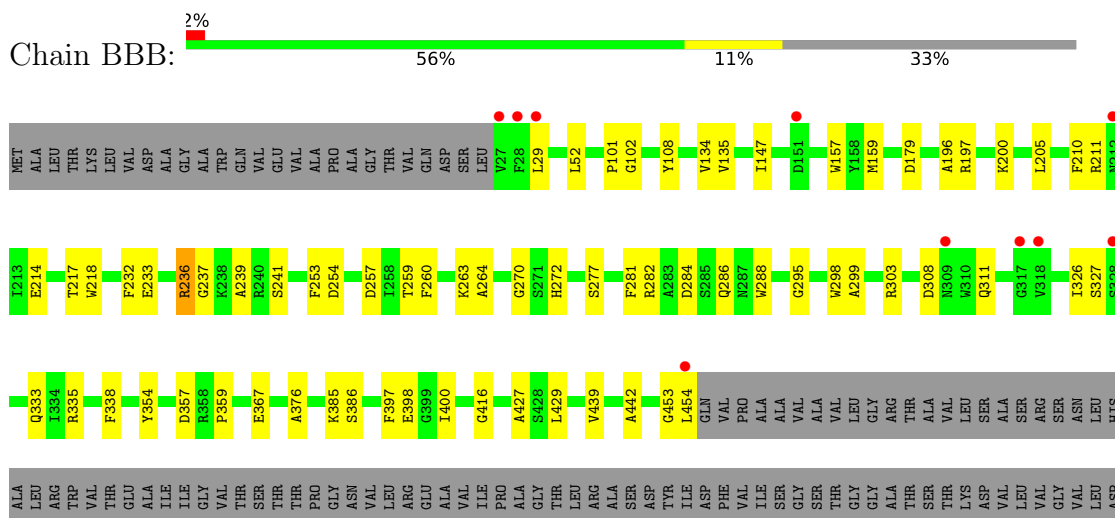
3 Residue-property plots [i](#)

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: Uncharacterized protein 57



• Molecule 1: Uncharacterized protein 57



[illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	38.34Å 140.99Å 81.92Å 90.00° 102.63° 90.00°	Depositor
Resolution (Å)	15.00 – 2.00 15.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	78.5 (15.00-2.00) 78.5 (15.00-2.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.13 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.204 , 0.247 0.211 , 0.253	Depositor DCC
R_{free} test set	746 reflections (1.31%)	wwPDB-VP
Wilson B-factor (Å ²)	25.4	Xtriage
Anisotropy	0.188	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 40.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.056 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6893	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.94% 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 [i](#)

5.1 Standard geometry [i](#)

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	AAA	1.06	0/3308	1.30	2/4506 (0.0%)
1	BBB	1.05	0/3311	1.32	3/4510 (0.1%)
All	All	1.05	0/6619	1.31	5/9016 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	92	GLY	CA-C-O	-5.94	118.36	122.22
1	BBB	102	GLY	N-CA-C	-5.62	104.50	112.81
1	BBB	102	GLY	CA-C-O	-5.17	117.82	122.57
1	AAA	45	VAL	N-CA-CB	5.16	116.59	110.55
1	BBB	281	PHE	CA-CB-CG	5.07	118.87	113.80

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	AAA	3237	0	3149	42	0
1	BBB	3240	0	3151	47	0
2	AAA	254	0	0	3	1
2	BBB	162	0	0	7	1
All	All	6893	0	6300	89	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 7.

The worst 5 of 89 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:BBB:397:PHE:HB3	1:BBB:400:ILE:HD12	1.60	0.83
1:BBB:299:ALA:HA	1:BBB:327:SER:O	1.83	0.78
1:AAA:43:ASP:OD1	1:AAA:45:VAL:HG23	1.83	0.78
1:BBB:397:PHE:HB3	1:BBB:400:ILE:CD1	2.14	0.78
1:BBB:236:ARG:HD3	1:BBB:237:GLY:N	1.99	0.77

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)
2:AAA:779:HOH:O	2:BBB:778:HOH:O[1_556]	2.18	0.02

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	AAA	426/638 (67%)	405 (95%)	20 (5%)	1 (0%)	43	42
1	BBB	426/638 (67%)	401 (94%)	24 (6%)	1 (0%)	43	42
All	All	852/1276 (67%)	806 (95%)	44 (5%)	2 (0%)	43	42

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	376	ALA
1	AAA	295	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	AAA	347/512 (68%)	342 (99%)	5 (1%)	59	66
1	BBB	348/512 (68%)	346 (99%)	2 (1%)	78	85
All	All	695/1024 (68%)	688 (99%)	7 (1%)	68	75

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	347	ASN
1	AAA	448	ARG
1	BBB	236	ARG
1	BBB	200	LYS
1	AAA	329	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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	AAA	428/638 (67%)	-0.22	5 (1%) 76 76	12, 28, 42, 71	6 (1%)
1	BBB	428/638 (67%)	0.13	10 (2%) 61 60	11, 35, 53, 77	2 (0%)
All	All	856/1276 (67%)	-0.05	15 (1%) 67 67	11, 31, 50, 77	8 (0%)

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	27	VAL	5.5
1	BBB	309	ASN	5.0
1	AAA	291	THR	3.6
1	BBB	212	ASN	2.9
1	AAA	27	VAL	2.8

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.