



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 11:36 am GMT

PDB ID : 3KL9
Title : Crystal structure of PepA from Streptococcus pneumoniae
Authors : Kim, K.K.; Lee, S.; Kim, D.
Deposited on : 2009-11-07
Resolution : 2.70 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
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

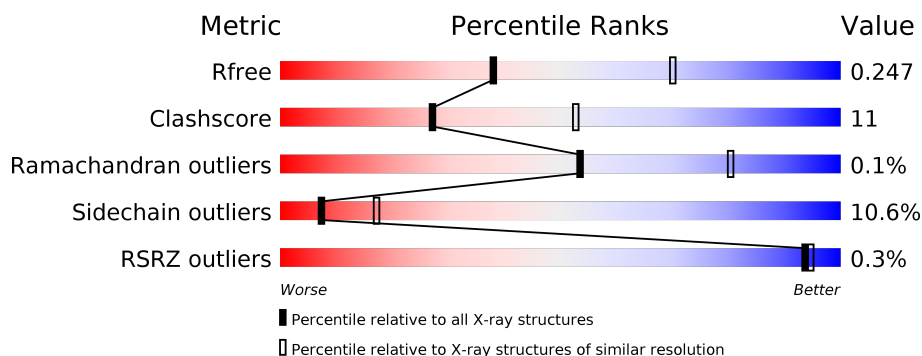
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	111664	2449 (2.70-2.70)
Clashscore	122126	2756 (2.70-2.70)
Ramachandran outliers	120053	2716 (2.70-2.70)
Sidechain outliers	120020	2716 (2.70-2.70)
RSRZ outliers	108989	2376 (2.70-2.70)

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. 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	355	
1	B	355	
1	C	355	
1	D	355	
1	E	355	
1	F	355	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	355	<div><div></div><div>70%19%5%5%</div></div>
1	H	355	<div><div></div><div>71%19%5%5%</div></div>
1	I	355	<div><div></div><div>70%20%5%5%</div></div>
1	J	355	<div><div>%<div></div></div><div>71%20%•5%</div></div>
1	K	355	<div><div></div><div>73%18%5%•</div></div>
1	L	355	<div><div>%<div></div></div><div>70%21%•5%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 31329 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 Glutamyl aminopeptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	340	Total	C	N	O	S	0	0	0
			2577	1635	434	498	10			
1	B	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	C	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	D	340	Total	C	N	O	S	0	0	0
			2577	1635	434	498	10			
1	E	340	Total	C	N	O	S	0	0	0
			2577	1635	434	498	10			
1	F	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	G	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	H	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	I	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	J	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			
1	K	340	Total	C	N	O	S	0	0	0
			2577	1635	434	498	10			
1	L	337	Total	C	N	O	S	0	0	0
			2556	1623	431	493	9			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
B	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
C	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
D	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
E	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
G	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
H	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
I	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
J	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
K	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7
L	0	SER	-	EXPRESSION TAG	UNP Q8DNJ7

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	2	Total 2	Zn 2	0	0
2	J	2	Total 2	Zn 2	0	0
2	D	2	Total 2	Zn 2	0	0
2	K	2	Total 2	Zn 2	0	0
2	E	2	Total 2	Zn 2	0	0
2	H	2	Total 2	Zn 2	0	0
2	B	2	Total 2	Zn 2	0	0
2	I	2	Total 2	Zn 2	0	0
2	C	2	Total 2	Zn 2	0	0
2	A	2	Total 2	Zn 2	0	0
2	L	2	Total 2	Zn 2	0	0
2	F	2	Total 2	Zn 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	46	Total 46	O 46	0	0
3	B	45	Total 45	O 45	0	0

Continued on next page...

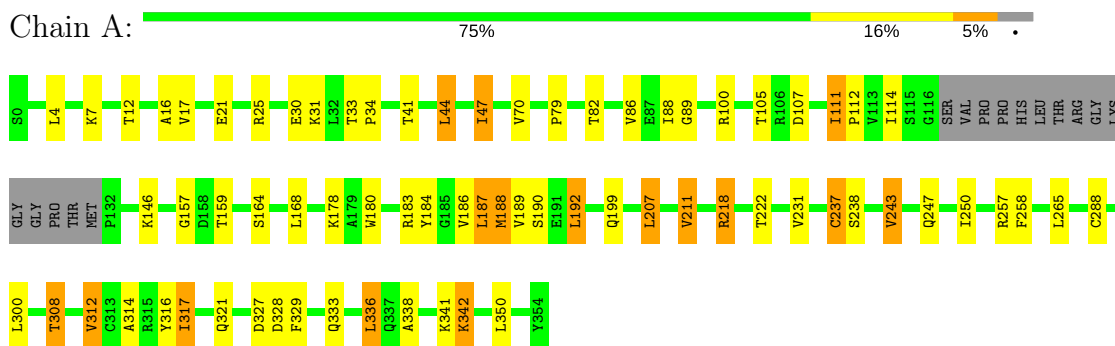
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	50	Total	O	0	0
			50	50		
3	D	44	Total	O	0	0
			44	44		
3	E	44	Total	O	0	0
			44	44		
3	F	45	Total	O	0	0
			45	45		
3	G	47	Total	O	0	0
			47	47		
3	H	36	Total	O	0	0
			36	36		
3	I	50	Total	O	0	0
			50	50		
3	J	45	Total	O	0	0
			45	45		
3	K	54	Total	O	0	0
			54	54		
3	L	43	Total	O	0	0
			43	43		

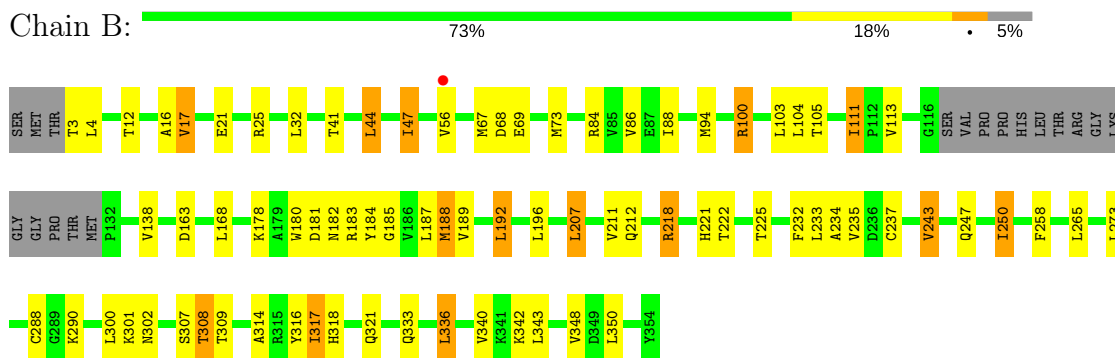
3 Residue-property plots

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.

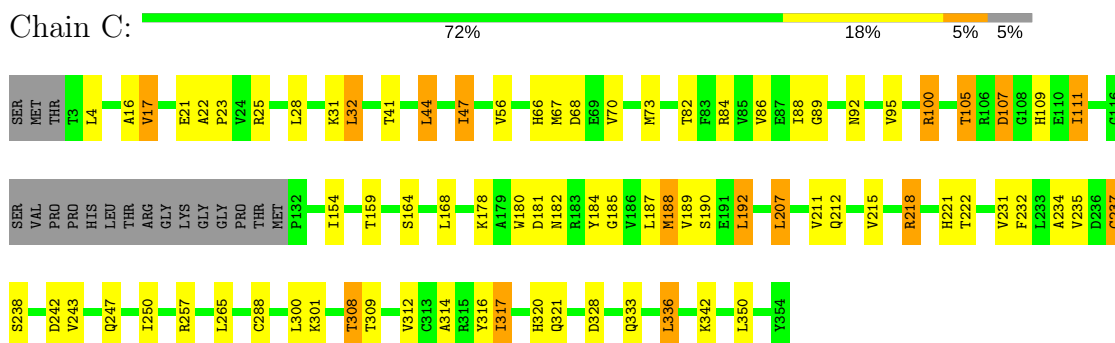
- Molecule 1: Glutamyl aminopeptidase



- Molecule 1: Glutamyl aminopeptidase

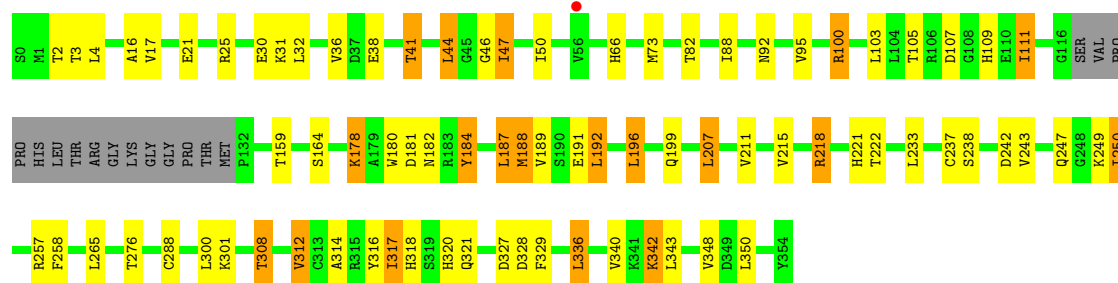


- Molecule 1: Glutamyl aminopeptidase



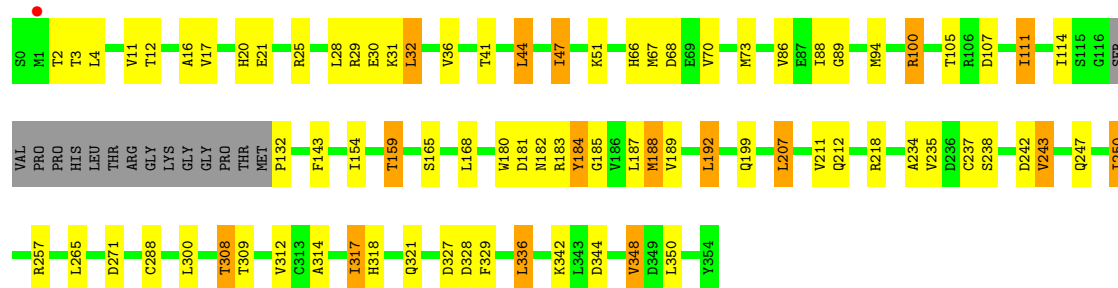
- Molecule 1: Glutamyl aminopeptidase

Chain D:  73% 17% 5% .



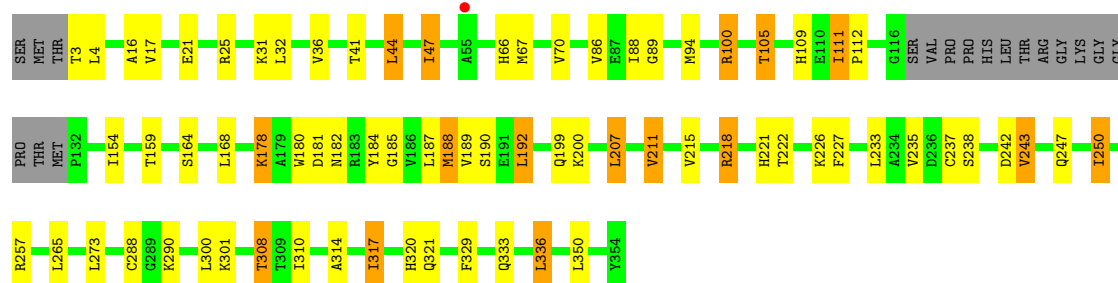
- Molecule 1: Glutamyl aminopeptidase

Chain E:  72% 19% 5% .



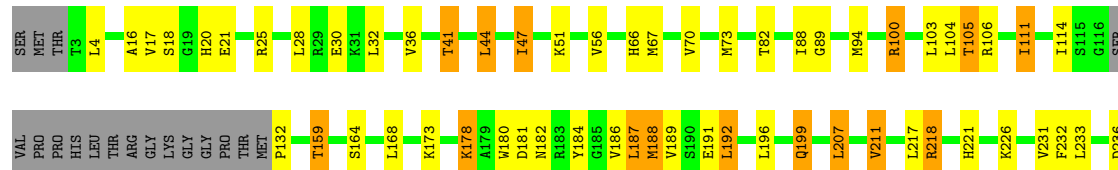
- Molecule 1: Glutamyl aminopeptidase

Chain F:  74% 16% 5% 5%



- Molecule 1: Glutamyl aminopeptidase

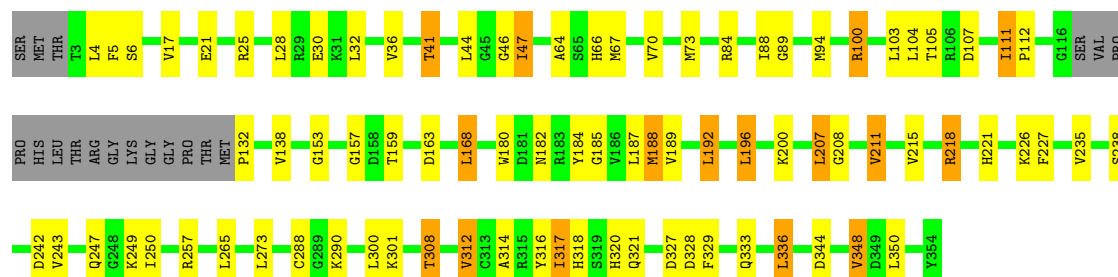
Chain G:  70% 19% 5% 5%





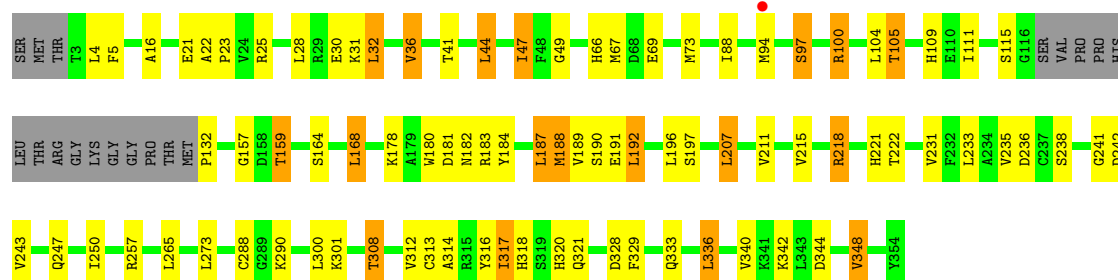
• Molecule 1: Glutamyl aminopeptidase

Chain H: 71% 19% 5% 5%



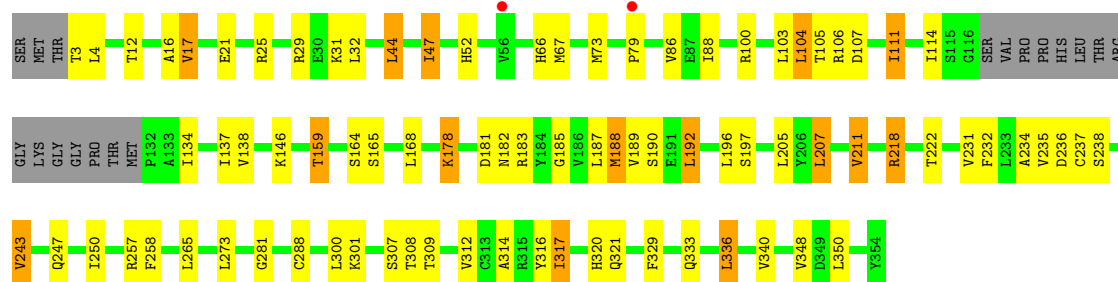
• Molecule 1: Glutamyl aminopeptidase

Chain I: 70% 20% 5% 5%



• Molecule 1: Glutamyl aminopeptidase

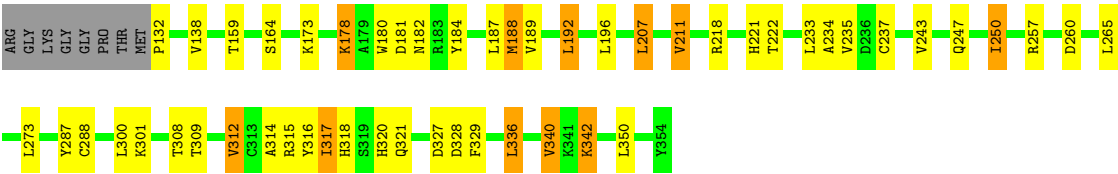
Chain J: 71% 20% 5%



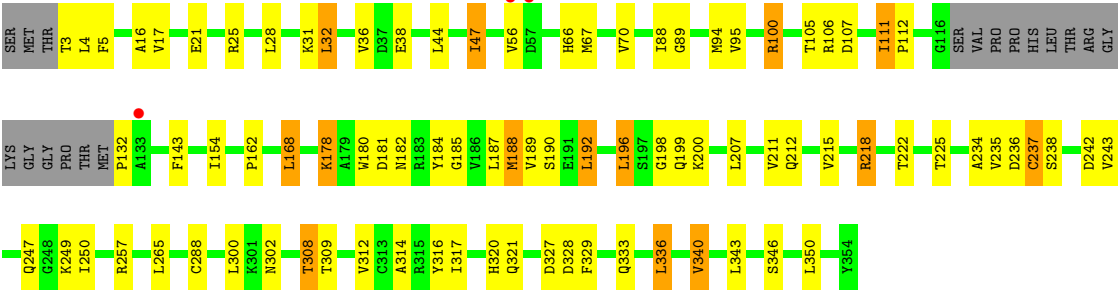
• Molecule 1: Glutamyl aminopeptidase

Chain K: 73% 18% 5%





● Molecule 1: Glutamyl aminopeptidase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	138.30Å 118.34Å 160.30Å 90.00° 106.35° 90.00°	Depositor
Resolution (Å)	44.43 – 2.70 44.43 – 2.69	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.43-2.70) 99.1 (44.43-2.69)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.60 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.200 , 0.247 0.202 , 0.247	Depositor DCC
R_{free} test set	6890 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	32.7	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 27.6	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	0.92	EDS
Total number of atoms	31329	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.32% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.43	0/2629	0.62	0/3558
1	B	0.42	0/2608	0.62	0/3530
1	C	0.41	0/2608	0.62	0/3530
1	D	0.42	0/2629	0.61	0/3558
1	E	0.41	0/2629	0.60	0/3558
1	F	0.43	0/2608	0.62	0/3530
1	G	0.42	0/2608	0.61	0/3530
1	H	0.42	0/2608	0.60	0/3530
1	I	0.41	0/2608	0.61	0/3530
1	J	0.43	0/2608	0.62	0/3530
1	K	0.42	0/2629	0.60	0/3558
1	L	0.42	0/2608	0.62	0/3530
All	All	0.42	0/31380	0.61	0/42472

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 [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	A	2577	0	2550	51	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2556	0	2526	59	0
1	C	2556	0	2526	68	0
1	D	2577	0	2550	58	0
1	E	2577	0	2550	63	0
1	F	2556	0	2526	49	0
1	G	2556	0	2526	65	0
1	H	2556	0	2526	68	0
1	I	2556	0	2526	74	0
1	J	2556	0	2526	65	0
1	K	2577	0	2550	63	0
1	L	2556	0	2526	71	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
2	F	2	0	0	0	0
2	G	2	0	0	0	0
2	H	2	0	0	0	0
2	I	2	0	0	0	0
2	J	2	0	0	0	0
2	K	2	0	0	0	0
2	L	2	0	0	0	0
3	A	46	0	0	0	0
3	B	45	0	0	0	0
3	C	50	0	0	1	0
3	D	44	0	0	0	0
3	E	44	0	0	1	0
3	F	45	0	0	1	0
3	G	47	0	0	1	0
3	H	36	0	0	1	0
3	I	50	0	0	0	0
3	J	45	0	0	3	0
3	K	54	0	0	3	0
3	L	43	0	0	2	0
All	All	31329	0	30408	664	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 11.

The worst 5 of 664 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:I:100:ARG:NH1	1:L:321:GLN:HE22	1.42	1.17
1:E:321:GLN:HE22	1:H:100:ARG:NH1	1.47	1.11
1:C:100:ARG:HH11	1:C:100:ARG:HG3	1.00	1.11
1:K:73:MET:HB2	1:K:88:ILE:HD11	1.35	1.08
1:J:25:ARG:HG3	1:J:47:ILE:HG12	1.31	1.08

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	336/355 (95%)	324 (96%)	12 (4%)	0	100	100
1	B	333/355 (94%)	323 (97%)	9 (3%)	1 (0%)	43	71
1	C	333/355 (94%)	321 (96%)	12 (4%)	0	100	100
1	D	336/355 (95%)	328 (98%)	8 (2%)	0	100	100
1	E	336/355 (95%)	321 (96%)	15 (4%)	0	100	100
1	F	333/355 (94%)	322 (97%)	11 (3%)	0	100	100
1	G	333/355 (94%)	324 (97%)	9 (3%)	0	100	100
1	H	333/355 (94%)	324 (97%)	9 (3%)	0	100	100
1	I	333/355 (94%)	320 (96%)	11 (3%)	2 (1%)	27	54
1	J	333/355 (94%)	322 (97%)	11 (3%)	0	100	100
1	K	336/355 (95%)	328 (98%)	8 (2%)	0	100	100
1	L	333/355 (94%)	320 (96%)	10 (3%)	3 (1%)	19	44
All	All	4008/4260 (94%)	3877 (97%)	125 (3%)	6 (0%)	53	80

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	97	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	199	GLN
1	B	56	VAL
1	I	197	SER
1	L	198	GLY

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	273/285 (96%)	247 (90%)	26 (10%)	9	22
1	B	270/285 (95%)	241 (89%)	29 (11%)	7	17
1	C	270/285 (95%)	243 (90%)	27 (10%)	8	20
1	D	273/285 (96%)	240 (88%)	33 (12%)	5	12
1	E	273/285 (96%)	243 (89%)	30 (11%)	7	16
1	F	270/285 (95%)	240 (89%)	30 (11%)	7	16
1	G	270/285 (95%)	235 (87%)	35 (13%)	4	11
1	H	270/285 (95%)	240 (89%)	30 (11%)	7	16
1	I	270/285 (95%)	243 (90%)	27 (10%)	8	20
1	J	270/285 (95%)	244 (90%)	26 (10%)	9	21
1	K	273/285 (96%)	243 (89%)	30 (11%)	7	16
1	L	270/285 (95%)	248 (92%)	22 (8%)	13	29
All	All	3252/3420 (95%)	2907 (89%)	345 (11%)	7	17

5 of 345 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	159	THR
1	G	199	GLN
1	K	327	ASP
1	F	188	MET
1	G	4	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 106 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	247	GLN
1	G	320	HIS
1	K	337	GLN
1	F	320	HIS
1	G	20	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](#)

Of 24 ligands modelled in this entry, 24 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 [i](#)

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	340/355 (95%)	-0.38	0 100 100	20, 27, 37, 44	0
1	B	337/355 (94%)	-0.35	1 (0%) 93 94	20, 26, 35, 39	0
1	C	337/355 (94%)	-0.36	0 100 100	20, 27, 37, 42	0
1	D	340/355 (95%)	-0.32	1 (0%) 93 94	21, 27, 36, 44	0
1	E	340/355 (95%)	-0.35	1 (0%) 93 94	19, 27, 36, 47	0
1	F	337/355 (94%)	-0.34	1 (0%) 93 94	21, 27, 36, 42	0
1	G	337/355 (94%)	-0.41	0 100 100	21, 26, 34, 39	0
1	H	337/355 (94%)	-0.32	0 100 100	22, 29, 38, 45	0
1	I	337/355 (94%)	-0.37	1 (0%) 93 94	22, 28, 36, 42	0
1	J	337/355 (94%)	-0.34	2 (0%) 89 90	20, 27, 36, 44	0
1	K	340/355 (95%)	-0.40	1 (0%) 93 94	20, 26, 35, 41	0
1	L	337/355 (94%)	-0.22	3 (0%) 84 85	23, 30, 39, 46	0
All	All	4056/4260 (95%)	-0.35	11 (0%) 93 94	19, 27, 37, 47	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	56	VAL	3.5
1	J	56	VAL	3.2
1	D	56	VAL	3.2
1	L	57	ASP	3.2
1	E	1	MET	3.1

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 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	ZN	K	355	1/1	0.87	0.09	69,69,69,69	0
2	ZN	G	355	1/1	0.87	0.09	81,81,81,81	0
2	ZN	D	356	1/1	0.88	0.11	73,73,73,73	0
2	ZN	I	356	1/1	0.90	0.06	69,69,69,69	0
2	ZN	F	355	1/1	0.91	0.07	73,73,73,73	0
2	ZN	H	355	1/1	0.92	0.07	78,78,78,78	0
2	ZN	L	356	1/1	0.93	0.06	74,74,74,74	0
2	ZN	B	355	1/1	0.94	0.08	77,77,77,77	0
2	ZN	K	356	1/1	0.94	0.09	65,65,65,65	0
2	ZN	C	356	1/1	0.95	0.06	66,66,66,66	0
2	ZN	C	355	1/1	0.95	0.06	65,65,65,65	0
2	ZN	D	355	1/1	0.95	0.06	61,61,61,61	0
2	ZN	J	356	1/1	0.96	0.07	68,68,68,68	0
2	ZN	L	355	1/1	0.96	0.05	73,73,73,73	0
2	ZN	I	355	1/1	0.97	0.06	63,63,63,63	0
2	ZN	H	356	1/1	0.97	0.04	70,70,70,70	0
2	ZN	F	356	1/1	0.97	0.04	63,63,63,63	0
2	ZN	J	355	1/1	0.97	0.05	67,67,67,67	0
2	ZN	A	355	1/1	0.97	0.05	63,63,63,63	0
2	ZN	G	356	1/1	0.97	0.05	65,65,65,65	0
2	ZN	E	356	1/1	0.97	0.06	75,75,75,75	0
2	ZN	E	355	1/1	0.98	0.06	63,63,63,63	0
2	ZN	A	356	1/1	0.98	0.04	64,64,64,64	0
2	ZN	B	356	1/1	0.98	0.04	57,57,57,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.