



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:27 AM UTC

PDB ID : 2JJM / pdb_00002jjm
Title : Crystal Structure of a family GT4 glycosyltransferase from Bacillus anthracis ORF BA1558.
Authors : Ruane, K.M.; Davies, G.J.; Martinez-Fleites, C.
Deposited on : 2008-04-15
Resolution : 3.10 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

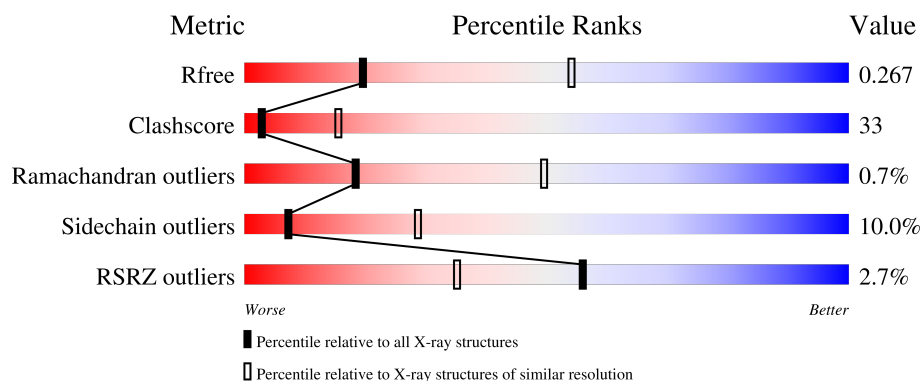
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1456 (3.10-3.10)
Clashscore	190562	1539 (3.10-3.10)
Ramachandran outliers	187476	1467 (3.10-3.10)
Sidechain outliers	187428	1467 (3.10-3.10)
RSRZ outliers	180081	1456 (3.10-3.10)

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	394	2% 49% 33% 8% 9%
1	B	394	2% 49% 33% 8% 9%
1	C	394	2% 49% 33% 8% 9%
1	D	394	2% 49% 33% 7% 9%
1	E	394	2% 49% 33% 7% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	394	2% 49% 32% 8% 9%
1	G	394	2% 49% 32% 8% 9%
1	H	394	7% 49% 33% 8% 9%
1	I	394	3% 49% 32% 8% 9%
1	J	394	2% 49% 33% 7% 9%
1	K	394	2% 50% 32% 8% 9%
1	L	394	3% 49% 32% 8% 9%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 33924 atoms, of which 0 are hydrogens and 0 are deuteriums.

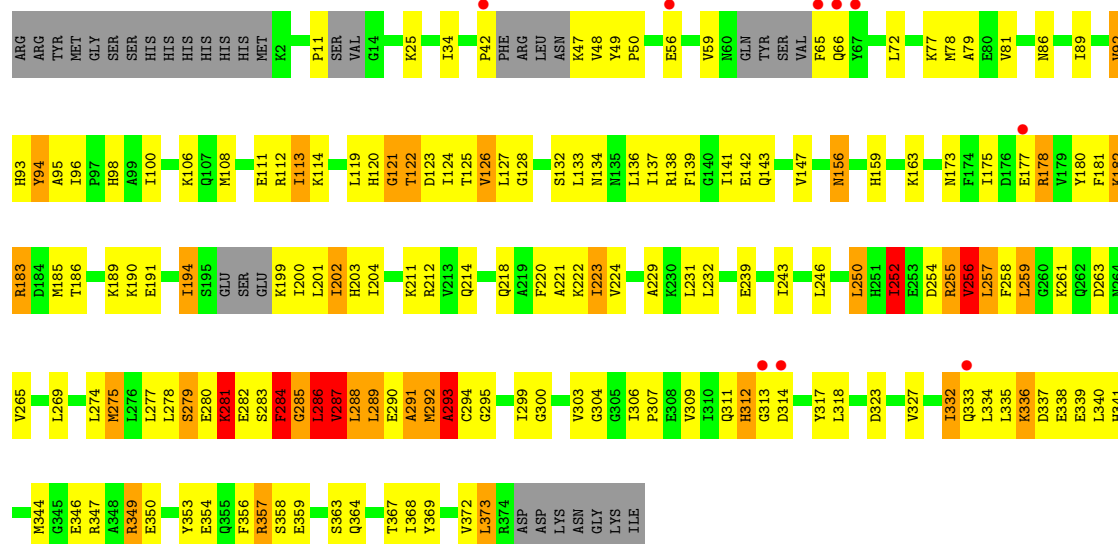
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN.

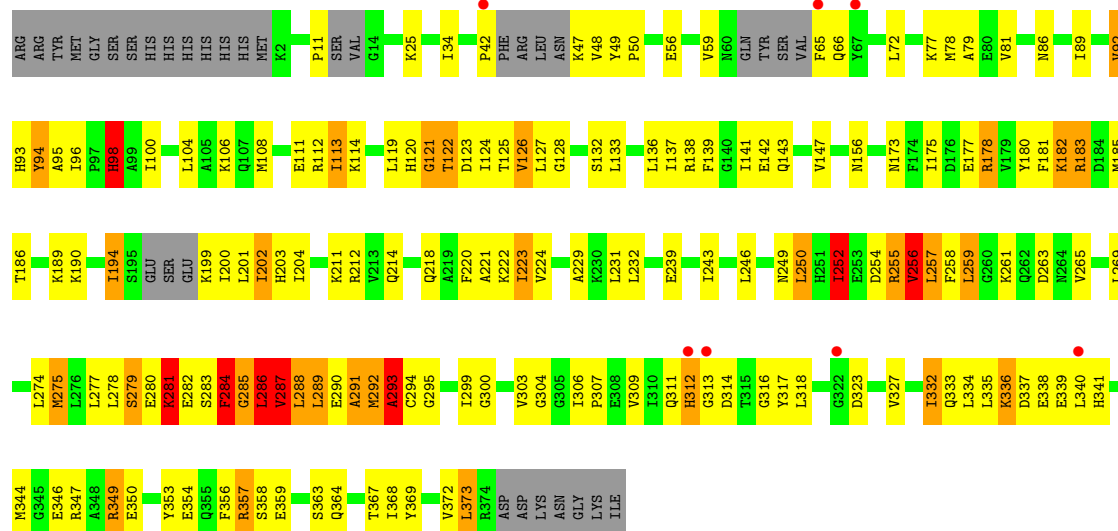
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	B	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	C	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	D	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	E	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	F	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	G	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	H	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	I	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	J	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	K	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			
1	L	360	Total	C	N	O	S	0	0	1
			2827	1806	477	531	13			



• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN

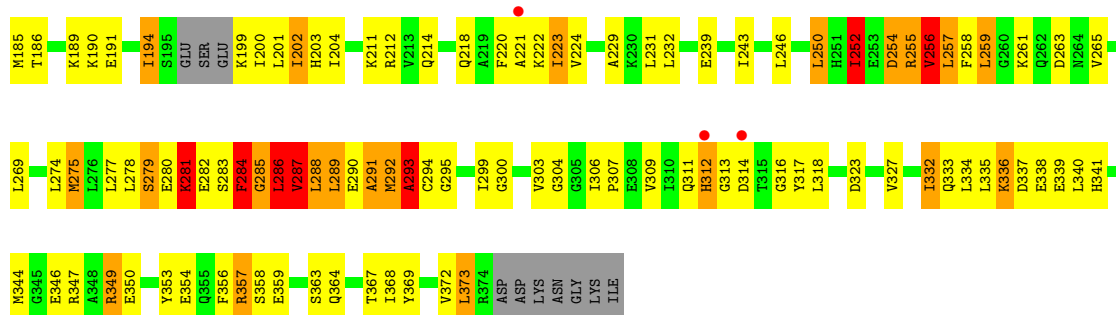


• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN

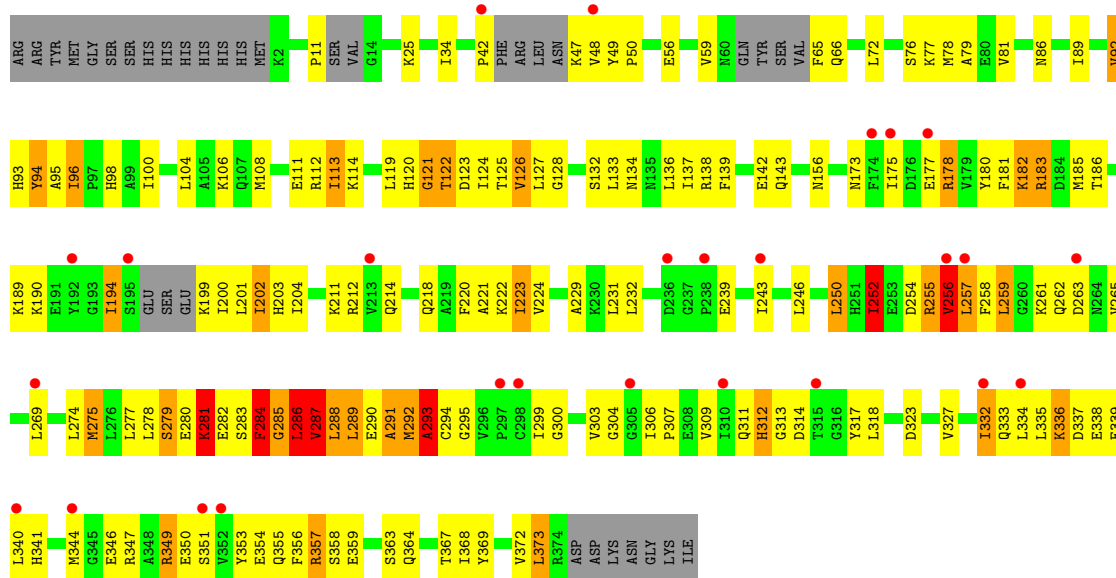


• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN

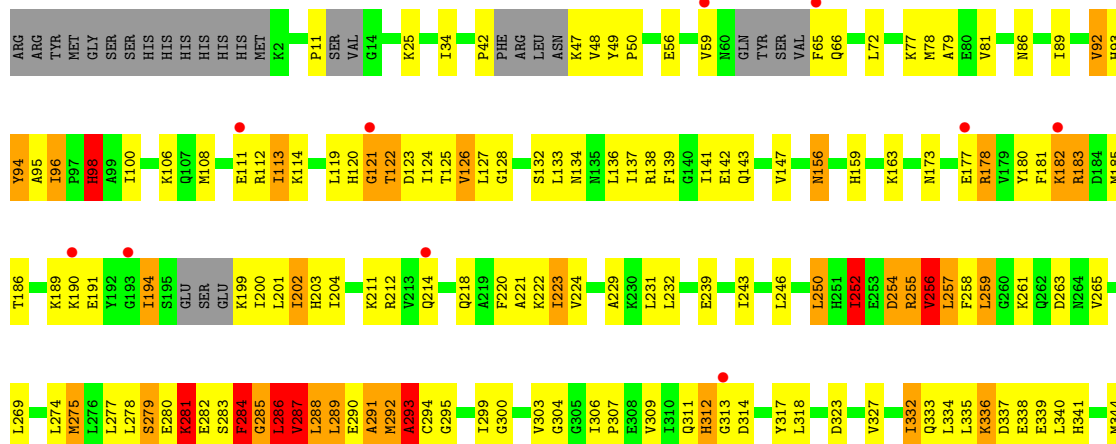




• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN

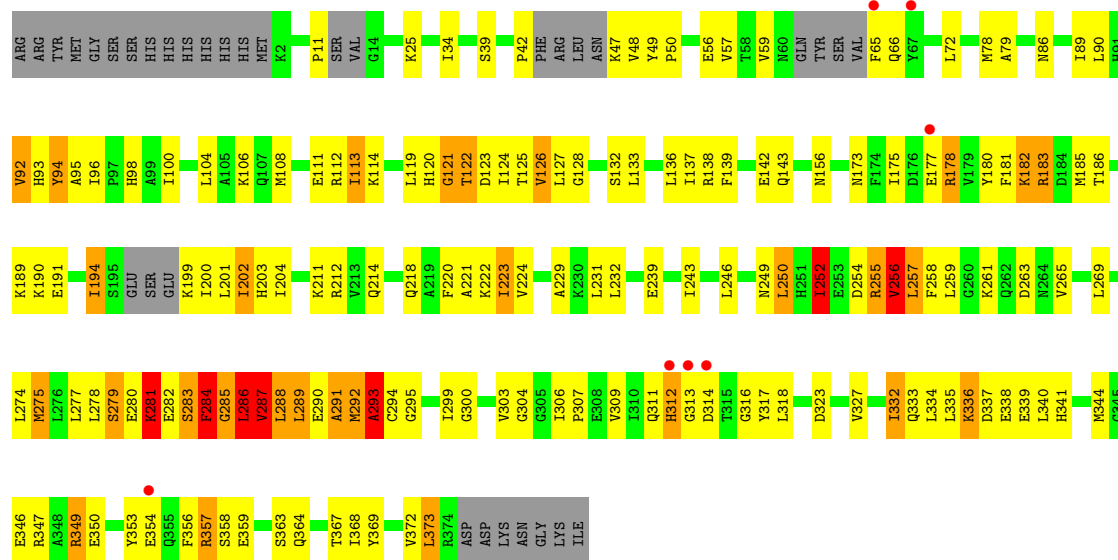


• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN

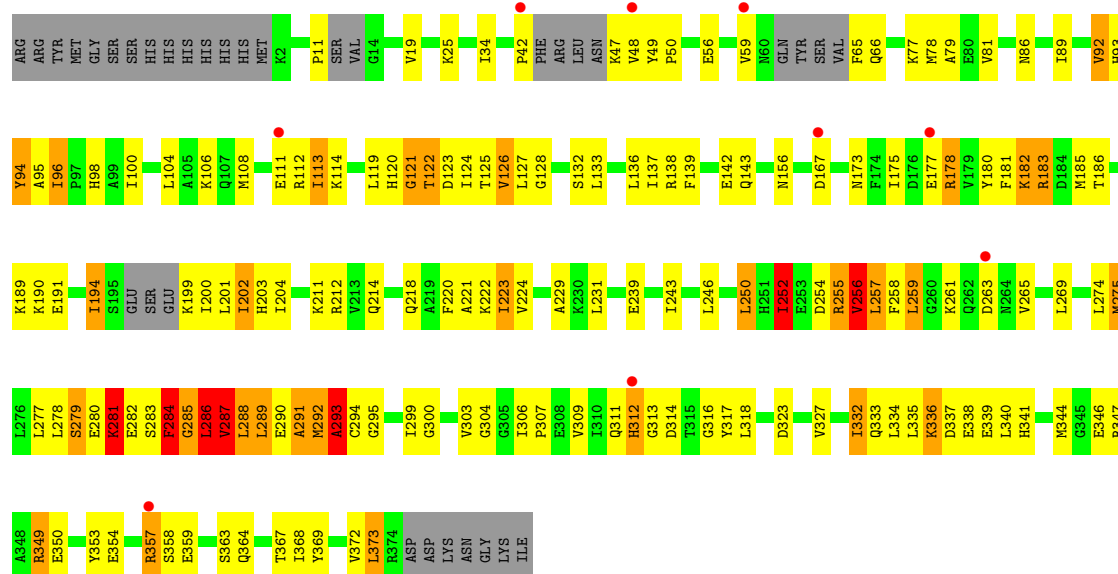




• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN

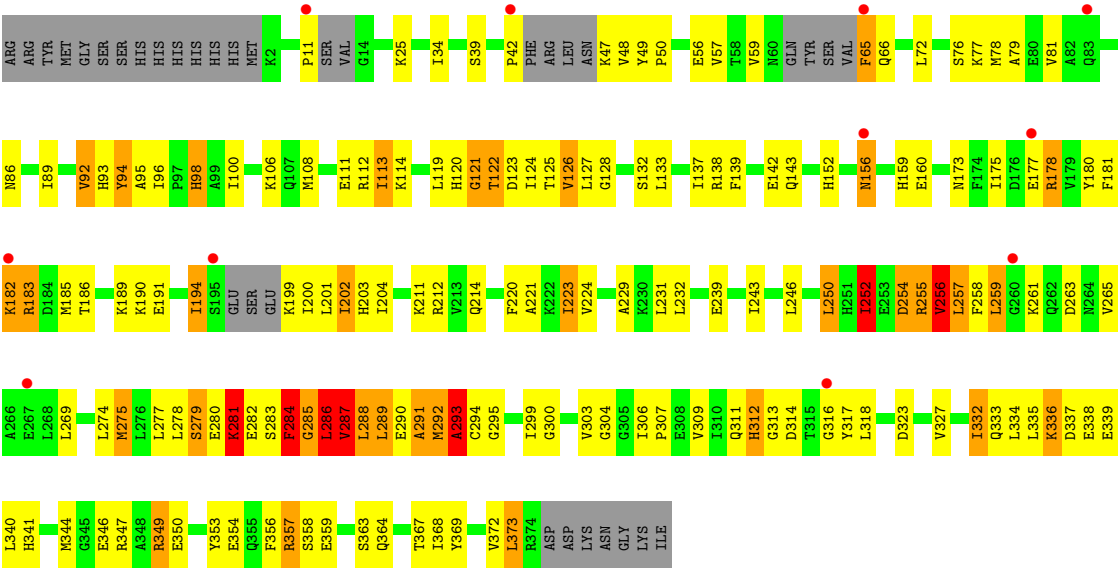


• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN



• Molecule 1: GLYCOSYL TRANSFERASE, GROUP 1 FAMILY PROTEIN





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	134.60Å 204.67Å 135.33Å 90.00° 115.49° 90.00°	Depositor
Resolution (Å)	20.00 – 3.10 20.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	98.3 (20.00-3.10) 97.9 (20.00-3.10)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 2.69Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.254 , 0.262 0.253 , 0.267	Depositor DCC
R_{free} test set	5915 reflections (3.66%)	wwPDB-VP
Wilson B-factor (Å ²)	88.0	Xtriage
Anisotropy	0.099	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 56.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.022 for l,-k,h	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	33924	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.93% 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

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	1.14	19/2873 (0.7%)	1.29	30/3884 (0.8%)
1	B	1.14	19/2873 (0.7%)	1.30	30/3884 (0.8%)
1	C	1.15	19/2873 (0.7%)	1.29	30/3884 (0.8%)
1	D	1.15	18/2873 (0.6%)	1.29	33/3884 (0.8%)
1	E	1.16	19/2873 (0.7%)	1.29	31/3884 (0.8%)
1	F	1.15	19/2873 (0.7%)	1.29	30/3884 (0.8%)
1	G	1.15	19/2873 (0.7%)	1.29	32/3884 (0.8%)
1	H	1.13	18/2873 (0.6%)	1.30	31/3884 (0.8%)
1	I	1.16	20/2873 (0.7%)	1.29	31/3884 (0.8%)
1	J	1.19	23/2873 (0.8%)	1.29	28/3884 (0.7%)
1	K	1.16	20/2873 (0.7%)	1.28	32/3884 (0.8%)
1	L	1.15	19/2873 (0.7%)	1.29	30/3884 (0.8%)
All	All	1.15	232/34476 (0.7%)	1.29	368/46608 (0.8%)

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
1	B	0	2
1	C	0	2
1	D	0	2
1	E	0	2
1	F	0	2
1	G	0	2
1	H	0	2
1	I	0	3
1	J	0	2
1	K	0	2
1	L	0	2
All	All	0	25

The worst 5 of 232 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	287	VAL	CA-CB	-11.38	1.39	1.54
1	G	256	VAL	CA-CB	-11.36	1.40	1.54
1	E	256	VAL	CA-CB	-11.31	1.40	1.54
1	H	256	VAL	CA-CB	-11.29	1.40	1.54
1	L	256	VAL	CA-CB	-11.28	1.40	1.54

The worst 5 of 368 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	284	PHE	N-CA-C	11.73	135.78	110.80
1	E	284	PHE	N-CA-C	11.69	135.69	110.80
1	K	284	PHE	N-CA-C	11.63	135.57	110.80
1	F	284	PHE	N-CA-C	11.57	135.44	110.80
1	B	284	PHE	N-CA-C	11.53	135.36	110.80

There are no chirality outliers.

5 of 25 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	284	PHE	Peptide,Mainchain
1	B	284	PHE	Peptide,Mainchain
1	C	284	PHE	Peptide,Mainchain
1	D	284	PHE	Peptide,Mainchain
1	E	284	PHE	Mainchain

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	2827	0	2856	186	1
1	B	2827	0	2856	194	0
1	C	2827	0	2856	192	0
1	D	2827	0	2856	185	0
1	E	2827	0	2856	186	0
1	F	2827	0	2856	190	0
1	G	2827	0	2856	198	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	2827	0	2856	189	0
1	I	2827	0	2856	192	0
1	J	2827	0	2856	187	0
1	K	2827	0	2856	187	1
1	L	2827	0	2856	194	0
All	All	33924	0	34272	2220	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 33.

The worst 5 of 2220 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:G:167:ASP:OD2	1:L:160:GLU:HA	1.15	1.27
1:A:286:LEU:HD12	1:A:286:LEU:N	1.55	1.17
1:B:286:LEU:H	1:B:286:LEU:CD1	1.54	1.16
1:K:280:GLU:O	1:K:281:LYS:HG2	1.46	1.16
1:G:280:GLU:O	1:G:281:LYS:HG2	1.47	1.15

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)
1:A:160:GLU:CA	1:K:167:ASP:OD2[1_455]	2.08	0.12

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	350/394 (89%)	323 (92%)	24 (7%)	3 (1%)	14 44
1	B	350/394 (89%)	322 (92%)	26 (7%)	2 (1%)	21 52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	350/394 (89%)	323 (92%)	25 (7%)	2 (1%)	21	52
1	D	350/394 (89%)	322 (92%)	25 (7%)	3 (1%)	14	44
1	E	350/394 (89%)	323 (92%)	24 (7%)	3 (1%)	14	44
1	F	350/394 (89%)	322 (92%)	26 (7%)	2 (1%)	21	52
1	G	350/394 (89%)	322 (92%)	26 (7%)	2 (1%)	21	52
1	H	350/394 (89%)	322 (92%)	26 (7%)	2 (1%)	21	52
1	I	350/394 (89%)	322 (92%)	25 (7%)	3 (1%)	14	44
1	J	350/394 (89%)	322 (92%)	26 (7%)	2 (1%)	21	52
1	K	350/394 (89%)	322 (92%)	26 (7%)	2 (1%)	21	52
1	L	350/394 (89%)	323 (92%)	25 (7%)	2 (1%)	21	52
All	All	4200/4728 (89%)	3868 (92%)	304 (7%)	28 (1%)	18	49

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	293	ALA
1	A	304	GLY
1	B	293	ALA
1	B	304	GLY
1	C	293	ALA

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	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	B	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	C	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	D	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	E	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	F	312/346 (90%)	281 (90%)	31 (10%)	7	29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	H	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	I	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	J	312/346 (90%)	280 (90%)	32 (10%)	7	27
1	K	312/346 (90%)	281 (90%)	31 (10%)	7	29
1	L	312/346 (90%)	281 (90%)	31 (10%)	7	29
All	All	3744/4152 (90%)	3371 (90%)	373 (10%)	7	29

5 of 373 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	202	ILE
1	J	122	THR
1	H	275	MET
1	I	183	ARG
1	J	275	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 122 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	312	HIS
1	K	364	GLN
1	G	364	GLN
1	K	333	GLN
1	L	312	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 [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	A	360/394 (91%)	0.22	7 (1%) 66 45	62, 92, 116, 125	0
1	B	360/394 (91%)	0.36	6 (1%) 69 48	63, 94, 116, 126	0
1	C	360/394 (91%)	0.32	9 (2%) 58 37	57, 93, 116, 125	0
1	D	360/394 (91%)	0.27	7 (1%) 66 45	58, 93, 115, 126	0
1	E	360/394 (91%)	0.44	8 (2%) 62 41	56, 92, 116, 125	0
1	F	360/394 (91%)	0.24	7 (1%) 66 45	57, 93, 116, 125	0
1	G	360/394 (91%)	0.37	8 (2%) 62 41	62, 93, 116, 125	0
1	H	360/394 (91%)	0.79	26 (7%) 21 11	63, 95, 117, 126	0
1	I	360/394 (91%)	0.39	11 (3%) 51 30	57, 92, 116, 125	0
1	J	360/394 (91%)	0.27	7 (1%) 66 45	38, 87, 116, 124	0
1	K	360/394 (91%)	0.39	9 (2%) 58 37	55, 90, 116, 124	0
1	L	360/394 (91%)	0.35	11 (3%) 51 30	61, 92, 115, 124	0
All	All	4320/4728 (91%)	0.37	116 (2%) 56 35	38, 92, 116, 126	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	312	HIS	4.1
1	A	42	PRO	4.0
1	H	315	THR	3.9
1	K	312	HIS	3.8
1	D	313	GLY	3.7

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.