wwPDB EM Validation Summary Report

Mar 8, 2025 – 06:02 am GMT

PDB ID	:	9GJT
EMDB ID	:	EMD-51402
Title	:	Structure of Nipah Virus RNA Polymerase Complex - Apo state
Authors	:	Sala, F.; Ditter, K.; Dybkov, O.; Urlaub, H.; Hillen, H.S.
Deposited on	:	2024-08-22
Resolution	:	2.60 Å(reported)
Based on initial model	:	.

This is a wwPDB EM 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/EMValidationReportHelp
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:

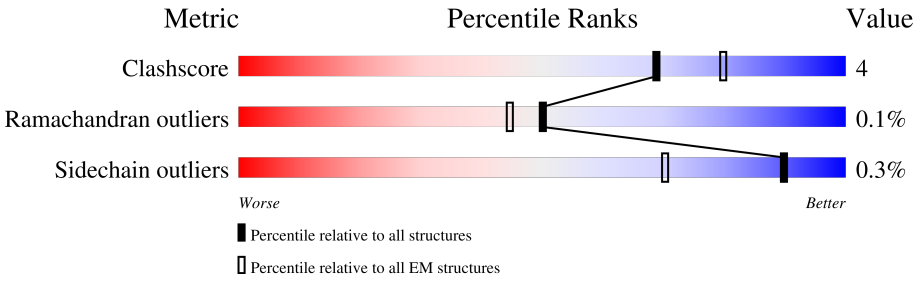
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 2.60 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	B	709	24% · 72%
1	C	709	14% · 85%
1	D	709	13% · 85%
1	E	709	13% · 85%
2	A	2246	50% 5% 45%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14017 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Phosphoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	C	105	Total	C	N	O	S	0	0
			839	531	138	163	7		
1	D	103	Total	C	N	O	S	0	0
			818	516	135	160	7		
1	E	104	Total	C	N	O	S	0	0
			827	522	137	161	7		
1	B	196	Total	C	N	O	S	0	0
			1572	982	264	318	8		

- Molecule 2 is a protein called RNA-directed RNA polymerase L.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	1233	Total	C	N	O	S	0	0
			9959	6352	1701	1840	66		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP Q997F0
A	0	ASN	-	expression tag	UNP Q997F0
A	1	ALA	-	expression tag	UNP Q997F0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

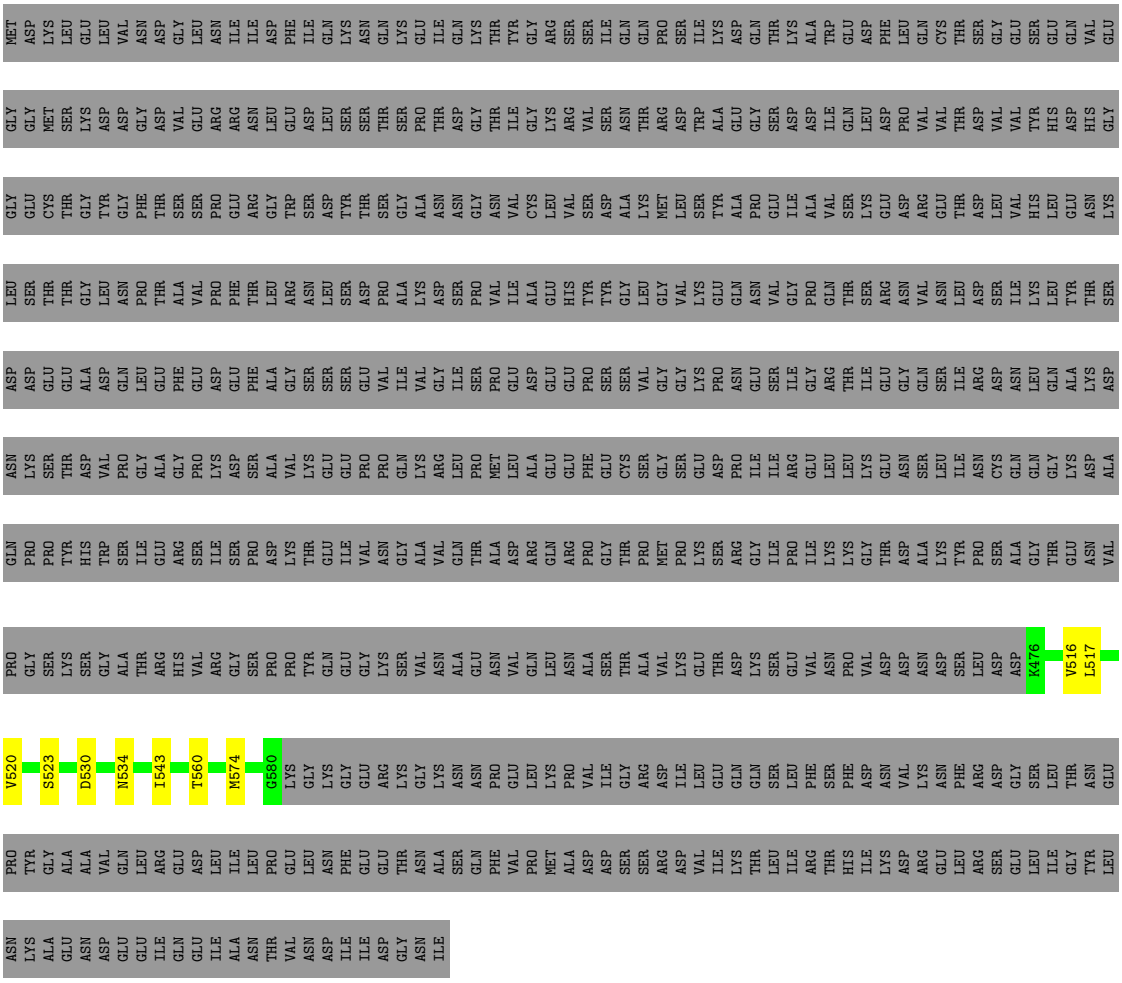
Mol	Chain	Residues	Atoms		AltConf
3	A	2	Total	Zn	0
			2	2	

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. 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. 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.

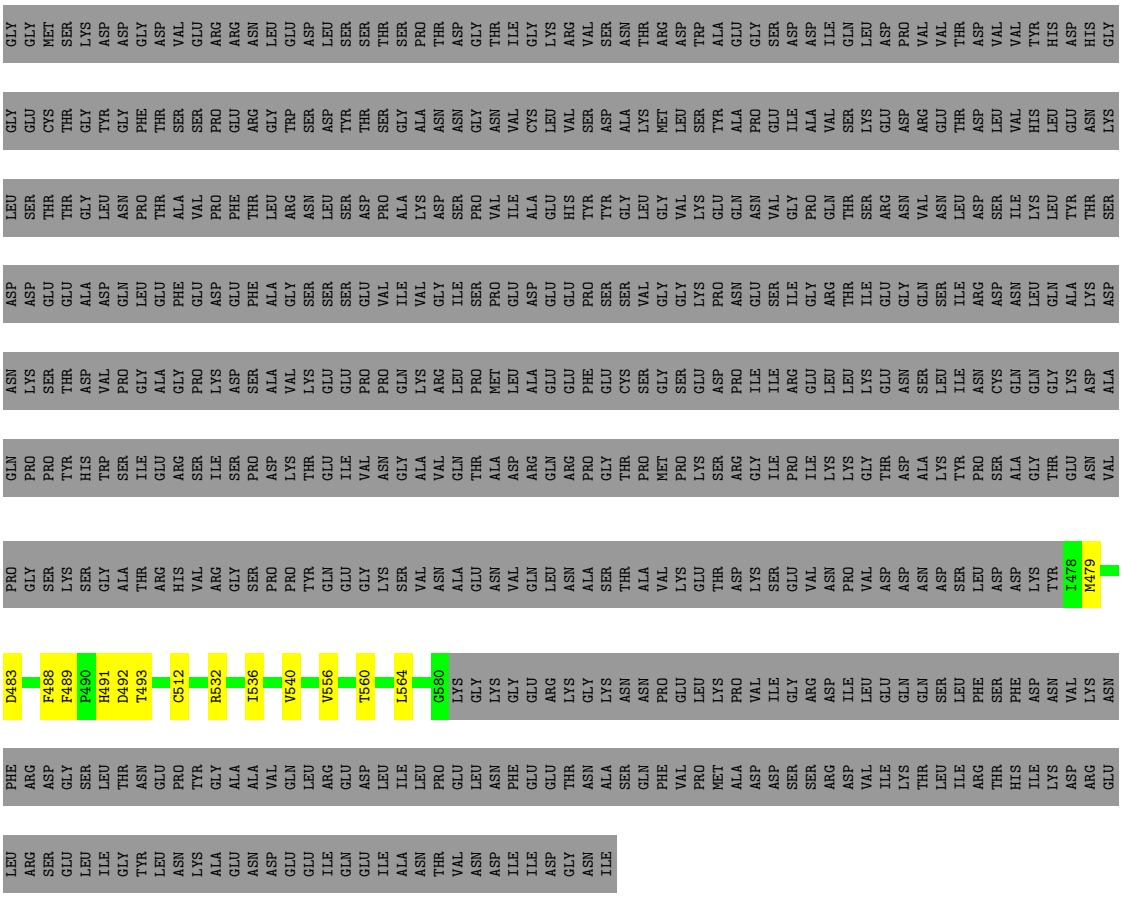
- Molecule 1: Phosphoprotein

Chain C:



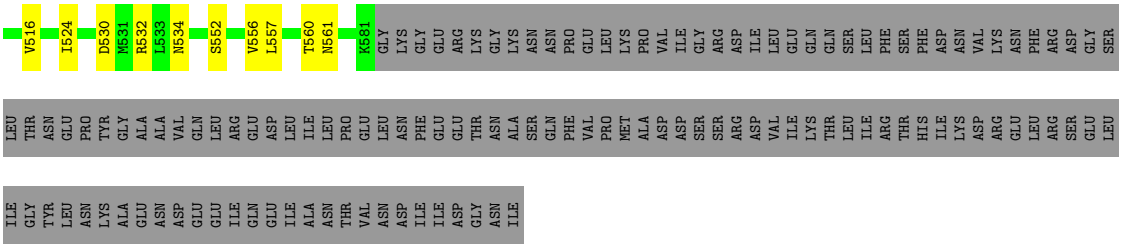
- Molecule 1: Phosphoprotein

Chain D:



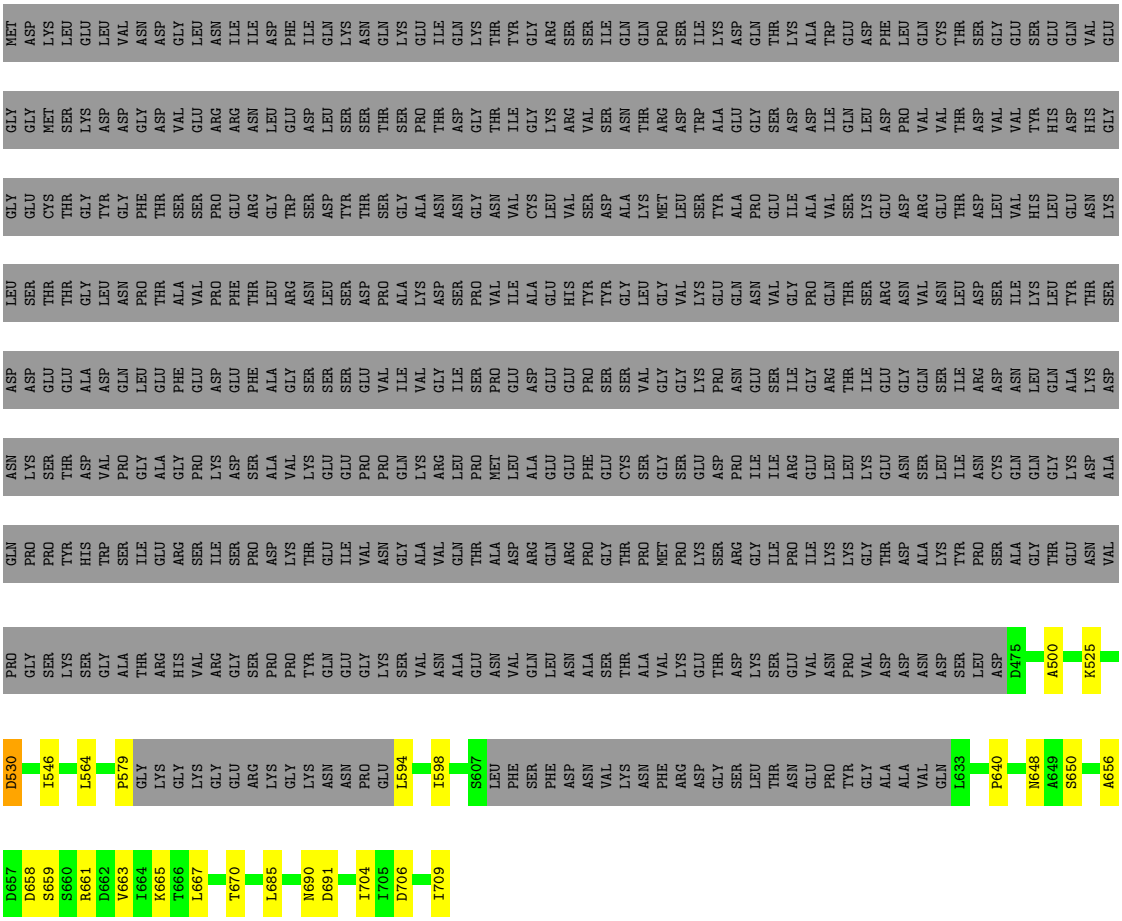
• Molecule 1: Phosphoprotein

Chain E:



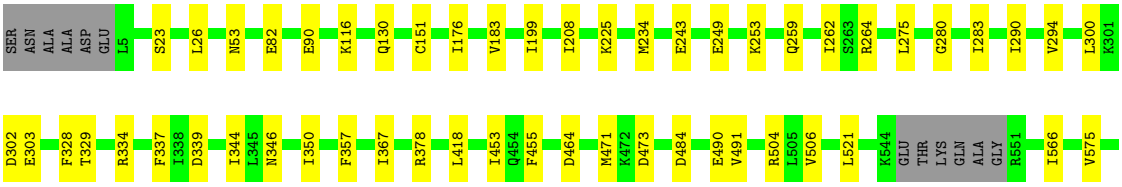
• Molecule 1: Phosphoprotein

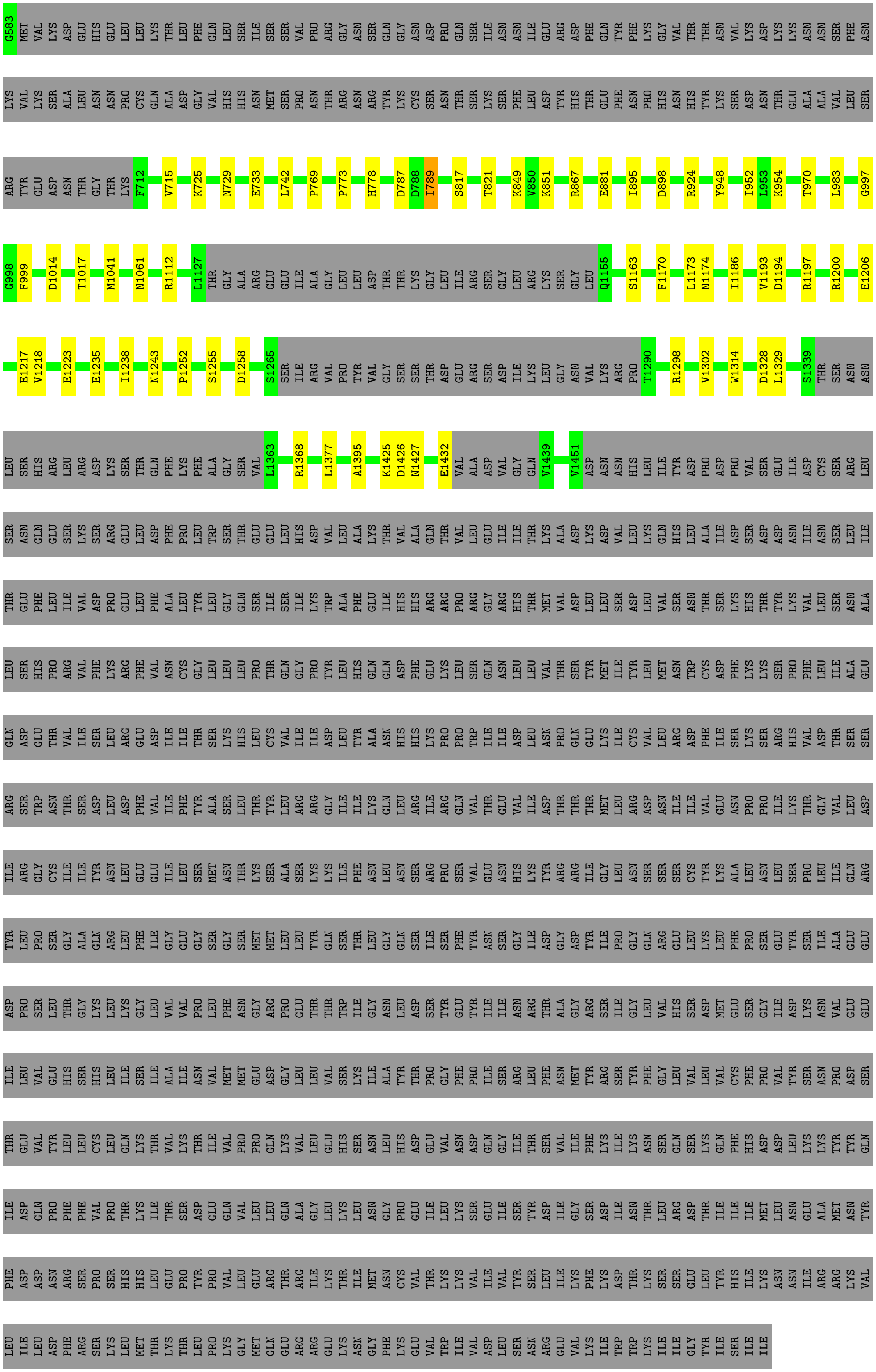
Chain B:



• Molecule 2: RNA-directed RNA polymerase L

Chain A:




G583 MET VAL LYS ASP GLU HIS GLU LEU LEU LYS THR LEU PHE GLN LEU SER ILE SER SER VAL PRO ARG GLY ASN SER GLN GLY ASN ASP PRO GLN SER ILE ASN ASN ILE GLU ARG ASP PHE GLN TYR PHE LYS GLY VAL THR THR ASN VAL LYS ASP LYS LYS ASN ASN SER PHE ASN
LYS VAL LYS SER ALA LEU ASN ASN PRO CYS GLN ALA ASP GLY VAL HIS HIS ASN MET SER PRO ASN THR ARG ASN ARG TYR LYS CYS SER ASN THR SER LYS SER PHE LEU ASP TYR HIS THR GLU PHE ASN PRO HIS ASN HIS TYR LYS SER ASP ASN THR GLU ALA ALA VAL LEU SER
ARG TYR GLU ASP ASN THR GLY THR LYS F712 V715 K725 N729 E733 L742 P769 P773 H778 D787 D788 I789 S817 T821 K849 V850 K851 R867 E881 I895 D898 R924 Y948 I952 L953 K954 T970 L983 G997
G998 F999 D1014 T1017 M1041 N1061 R1112 L1127 THR GLY ALA ARG GLU GLU ILE ALA GLY LEU LEU ASP THR THR LYS GLY LEU ILE ARG SER GLY LEU ARG LYS SER GLY LEU Q1155 S1163 F1170 L1173 N1174 I1186 V1193 D1194 R1197 R1200 E1206
E1217 V1218 E1223 E1235 I1238 N1243 P1252 S1255 D1258 S1265 SER ILE ARG VAL PRO TYR VAL GLY SER SER THR ASP GLU ARG SER ASP ILE LYS LEU GLY ASN VAL LYS ARG PRO T1290 R1298 V1302 W1314 D1328 L1329 S1339 THR SER ASN ASN
LEU SER HIS ARG LEU ARG ASP LYS SER THR GLN PHE LYS PHE ALA GLY SER VAL L1363 R1368 L1377 A1395 K1425 D1426 N1427 E1432 VAL ALA ASP VAL GLY GLN V1439 V1451 ASP ASN ASN HIS LEU ILE TYR ASP PRO ASP PRO VAL SER GLU ILE ASP CYS SER ARG LEU
SER ASN GLN GLU SER LYS SER ARG GLU LEU ASP PHE PRO LEU TRP SER THR GLU GLU LEU HIS ASP VAL LEU ALA LYS THR VAL ALA GLN THR VAL LEU GLU ILE ILE THR LYS ALA ASP LYS ASP VAL LEU LYS GLN HIS LEU ALA ILE ASP SER ASP ASP ASN ILE ASN SER LEU ILE
THR GLU PHE LEU ILE VAL ASP PRO GLU LEU PHE ALA LEU TYR LEU GLY GLN SER ILE SER ILE LYS TRP ALA PHE GLU ILE HIS HIS ARG ARG PRO ARG GLY ARG HIS THR MET VAL ASP LEU LEU SER ASP LEU VAL SER ASN THR SER LYS HIS THR TYR LYS VAL LEU SER ASN ALA
LEU SER HIS PRO ARG VAL PHE LYS ARG PHE VAL ASN CYS GLY LEU LEU LEU PRO THR GLN GLY PRO TYR LEU HIS GLN GLN ASP PHE GLU LYS LEU SER GLN ASN LEU LEU VAL THR SER TYR MET ILE TYR LEU MET ASN TRP CYS ASP PHE LYS LYS SER PRO PHE LEU ILE ALA GLU
GLN ASP GLU THR VAL ILE SER LEU ARG GLU ASP ILE ILE THR SER LYS HIS LEU CYS VAL ILE ILE ASP LEU TYR ALA ASN HIS HIS LYS PRO PRO TRP ILE ILE ASP LEU ASN PRO GLN GLU LYS ILE CYS VAL LEU ARG ASP PHE ILE SER LYS SER ARG HIS VAL ASP THR SER SER
ARG SER TRP ASN THR SER ASP LEU ASP PHE VAL ILE PHE TYR ALA SER LEU THR TYR LEU ARG ARG GLY ILE ILE LYS GLN LEU ARG ILE ARG GLN VAL THR GLU VAL ILE ASP THR THR THR MET LEU ARG ASP ASN ILE ILE VAL GLU ASN PRO PRO ILE LYS THR GLY VAL LEU ASP
ILE ARG GLY CYS ILE ILE TYR ASN LEU GLU GLU ILE LEU SER MET ASN THR LYS SER ALA SER LYS LYS ILE PHE ASN LEU ASN SER ARG PRO SER VAL GLU ASN HIS LYS TYR ARG ARG ILE GLY LEU ASN SER SER CYS TYR LYS ALA LEU ASN LEU SER PRO LEU ILE GLN ARG
TYR LEU PRO SER GLY ALA GLN ARG LEU PHE ILE GLY GLU GLY SER GLY SER MET MET LEU LEU TYR GLN SER THR LEU GLY GLN SER ILE SER PHE TYR ASN SER GLY ILE ASP GLY ASP TYR ILE PRO GLY GLN ARG GLU LEU LYS LEU PHE PRO SER GLU TYR SER ILE ALA GLU GLU
ASP PRO SER LEU THR GLY LYS LEU LYS GLY LEU VAL VAL PRO LEU PHE ASN GLY ARG PRO GLU THR THR TRP ILE GLY ASN LEU ASP SER TYR GLU TYR ILE ILE ASN ARG THR ALA GLY ARG SER ILE GLY LEU VAL HIS SER ASP MET GLU SER GLY ILE ASP LYS ASN VAL GLU GLU
ILE LEU VAL GLU HIS SER HIS LEU ILE SER ILE ALA ILE ASN VAL MET MET GLU ASP GLY LEU LEU VAL SER LYS ILE ALA TYR THR PRO GLY PHE PRO ILE SER ARG LEU PHE ASN ASN MET TYR ARG SER TYR PHE GLY LEU VAL LEU VAL CYS PHE PRO VAL TYR SER ASN PRO ASP SER
THR GLU VAL TYR LEU LEU CYS LEU GLN LYS THR VAL LYS THR ILE VAL PRO PRO GLN LYS VAL LEU GLU HIS SER ASN LEU HIS ASP GLU VAL ASN ASP GLN GLY ILE THR SER VAL ILE PHE LYS ILE LYS ASN SER GLN SER LYS GLN PHE HIS ASP ASP LEU LYS LYS TYR TYR GLN
ILE ASP GLN PRO PHE PHE VAL PRO THR LYS ILE THR SER ASP GLU GLN VAL LEU LEU GLN ALA GLY LEU LYS LEU ASN GLY PRO GLU ILE LEU LYS SER GLU ILE SER TYR ASP ILE GLY SER ASP ILE ASN THR LEU ARG ASP THR ILE ILE ILE MET LEU ASN GLU ALA MET ASN TYR
PHE ASP ASP ASN ARG SER PRO SER HIS HIS LEU GLU PRO TYR PRO VAL LEU GLU ARG THR ARG ILE LYS THR ILE MET ASN CYS VAL THR LYS LYS VAL ILE VAL TYR SER LEU ILE LYS PHE LYS ASP THR LYS SER SER GLU LEU TYR HIS ILE LYS ASN ASN ILE ARG ARG LYS VAL
LEU ILE LEU ASP PHE ARG SER LYS LEU MET THR LYS THR LEU PRO LYS GLY MET GLN GLU ARG ARG GLU LYS ASN GLY PHE LYS GLU VAL TRP ILE VAL ASP LEU SER ASN ARG GLU VAL LYS ILE TRP TRP LYS ILE ILE GLY TYR ILE SER ILE ILE

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	591312	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	48.94	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	B	0.24	0/1591	0.48	0/2147
1	C	0.25	0/853	0.45	0/1153
1	D	0.26	0/831	0.44	0/1124
1	E	0.24	0/840	0.43	0/1135
2	A	0.26	0/10165	0.50	0/13738
All	All	0.26	0/14280	0.49	0/19297

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	B	1572	0	1558	20	0
1	C	839	0	839	7	0
1	D	818	0	817	11	0
1	E	827	0	830	11	0
2	A	9959	0	9982	69	0
3	A	2	0	0	0	0
All	All	14017	0	14026	99	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 4.

The worst 5 of 99 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:D:532:ARG:NH1	1:E:530:ASP:OD1	2.19	0.75
2:A:53:ASN:ND2	2:A:490:GLU:O	2.20	0.75
2:A:898:ASP:O	2:A:1368:ARG:NH1	2.20	0.74
2:A:881:GLU:N	2:A:881:GLU:OE1	2.24	0.71
1:D:489:PHE:O	1:D:493:THR:OG1	2.06	0.71

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 PDB entries followed by that with respect to all EM entries.

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	B	190/709 (27%)	188 (99%)	2 (1%)	0	100	100
1	C	103/709 (14%)	101 (98%)	2 (2%)	0	100	100
1	D	101/709 (14%)	100 (99%)	1 (1%)	0	100	100
1	E	102/709 (14%)	101 (99%)	1 (1%)	0	100	100
2	A	1219/2246 (54%)	1191 (98%)	27 (2%)	1 (0%)	48	71
All	All	1715/5082 (34%)	1681 (98%)	33 (2%)	1 (0%)	50	71

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	789	ILE

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 PDB entries followed by that with respect to all EM entries.

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	B	182/625 (29%)	181 (100%)	1 (0%)	86 95
1	C	98/625 (16%)	98 (100%)	0	100 100
1	D	96/625 (15%)	95 (99%)	1 (1%)	73 88
1	E	97/625 (16%)	97 (100%)	0	100 100
2	A	1112/2047 (54%)	1109 (100%)	3 (0%)	91 97
All	All	1585/4547 (35%)	1580 (100%)	5 (0%)	90 97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	512	CYS
2	A	464	ASP
2	A	1112	ARG
2	A	1314	TRP
1	B	530	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	502	HIS
1	D	539	GLN
2	A	130	GLN
1	C	534	ASN
1	C	522	ASN

5.3.3 RNA (i)

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.