wwPDB EM Validation Summary Report

Mar 29, 2022 – 04:10 pm BST

PDB ID	:	7QGQ
EMDB ID	:	EMD-13957
Title	:	Extended H/L (SLPH/SLPL) complex from C. difficile (CD630 strain) fit into R20291 S-layer negative stain map
Authors	:	Banerji, O.; Wilson, J.S.; Bullough, P.A.
Deposited on	:	2021-12-09
Resolution	:	Not provided

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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

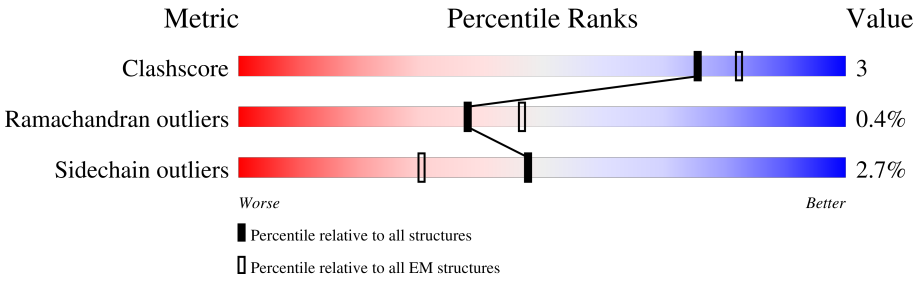
EMDB validation analysis	:	0.0.0.dev97
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance

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

The reported resolution of this entry is unknown.

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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	373	93% 6% •
1	D	373	92% 7%
1	J	373	94% 5% •
1	K	373	93% 6% •
1	L	373	93% 6% •
1	M	373	94% 5% •
1	N	373	94% 5% •
1	T	373	92% 7%
1	U	373	91% 8% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	V	373	91% 8%
1	W	373	92% 8%
1	X	373	90% 9%
2	A	318	88% 10%
2	C	318	90% 9%
2	E	318	84% 13%
2	F	318	88% 10%
2	G	318	85% 13%
2	H	318	88% 10%
2	I	318	89% 10%
2	O	318	93% 6%
2	P	318	88% 10%
2	Q	318	89% 8%
2	R	318	84% 14%
2	S	318	84% 14%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 124044 atoms, of which 62208 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 Precursor of the S-layer proteins.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	B	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	J	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	D	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	T	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	K	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	L	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	U	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	V	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	M	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	N	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	W	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		
1	X	373	Total	C	H	N	O	S	0	0
			5562	1711	2795	457	594	5		

- Molecule 2 is a protein called Precursor of the S-layer proteins.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	C	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	O	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	A	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		

Continued on next page...

Continued from previous page...

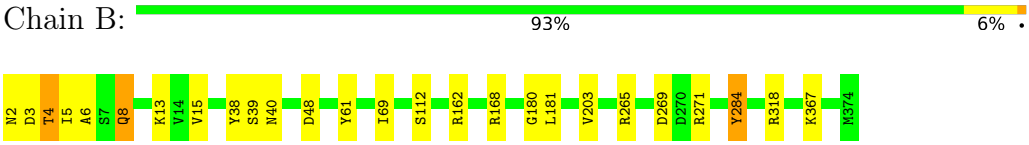
Mol	Chain	Residues	Atoms						AltConf	Trace
2	P	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	Q	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	E	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	F	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	R	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	S	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	G	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	H	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	2		
2	I	318	Total	C	H	N	O	S	0	0
			4775	1490	2389	394	500	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.

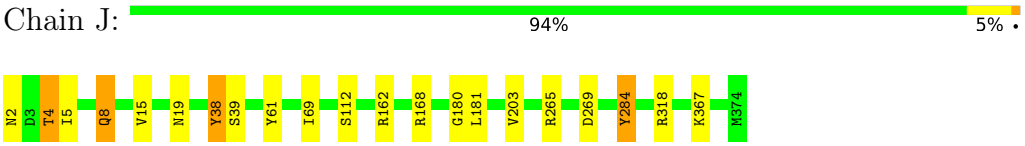
• Molecule 1: Precursor of the S-layer proteins

Chain B:



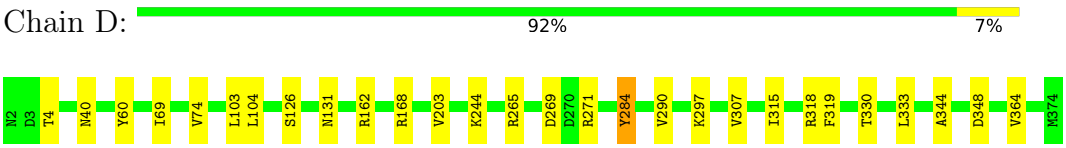
• Molecule 1: Precursor of the S-layer proteins

Chain J:



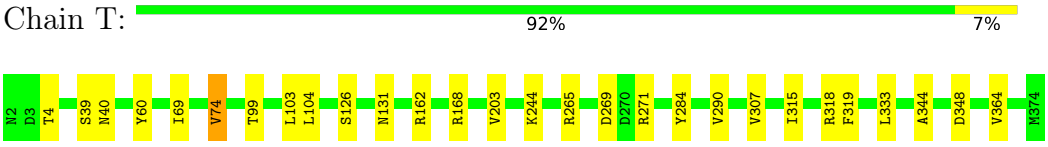
• Molecule 1: Precursor of the S-layer proteins

Chain D:



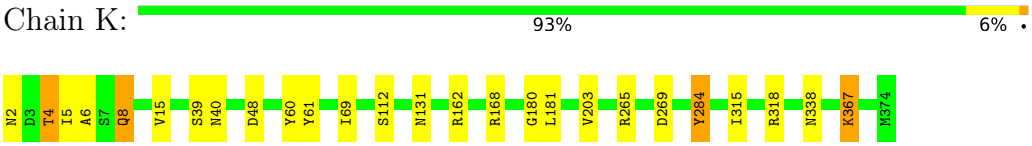
• Molecule 1: Precursor of the S-layer proteins

Chain T:



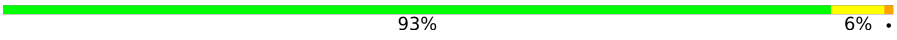
• Molecule 1: Precursor of the S-layer proteins

Chain K:



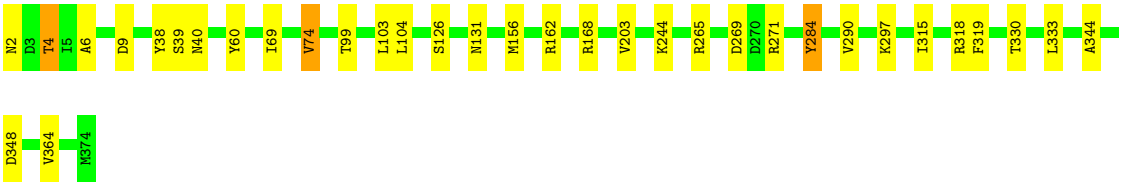
• Molecule 1: Precursor of the S-layer proteins

Chain L:



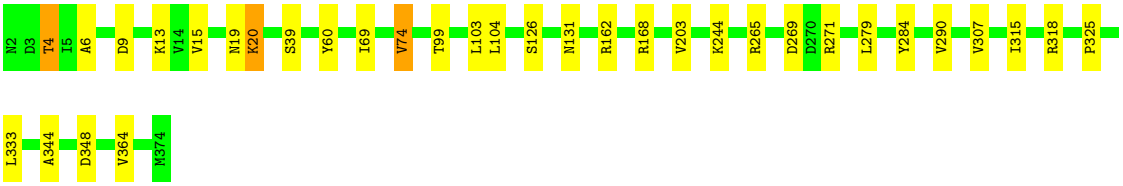
• Molecule 1: Precursor of the S-layer proteins

Chain U:



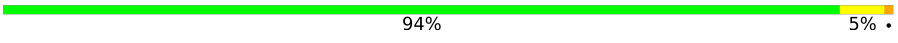
• Molecule 1: Precursor of the S-layer proteins

Chain V:



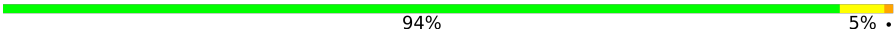
• Molecule 1: Precursor of the S-layer proteins

Chain M:



• Molecule 1: Precursor of the S-layer proteins

Chain N:

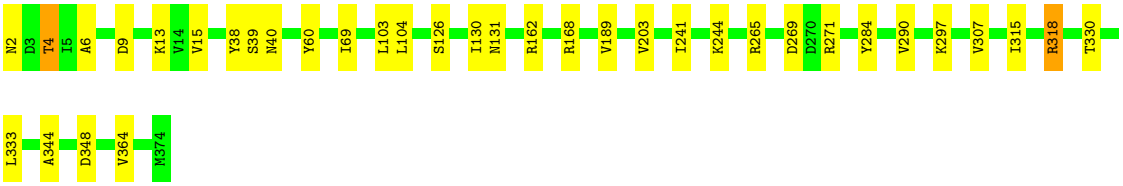


• Molecule 1: Precursor of the S-layer proteins

Chain W:

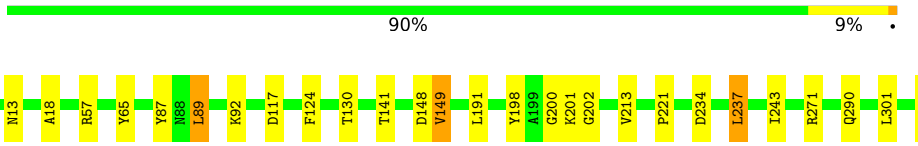
• Molecule 1: Precursor of the S-layer proteins

Chain X:



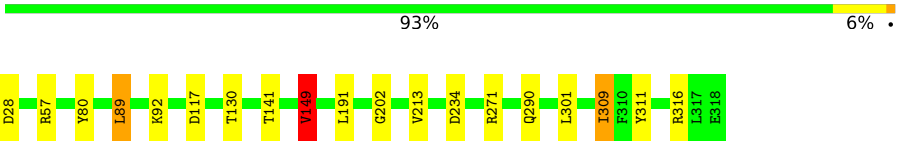
• Molecule 2: Precursor of the S-layer proteins

Chain C:



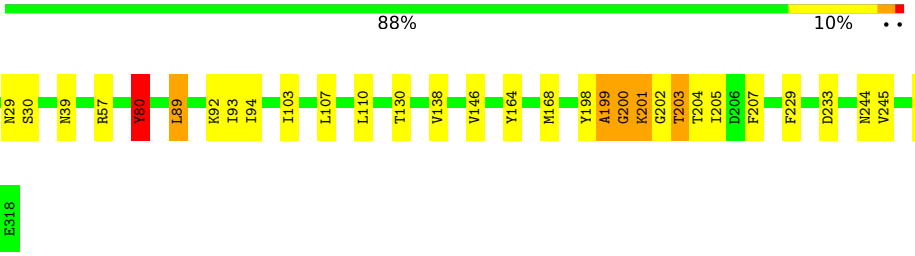
• Molecule 2: Precursor of the S-layer proteins

Chain O:



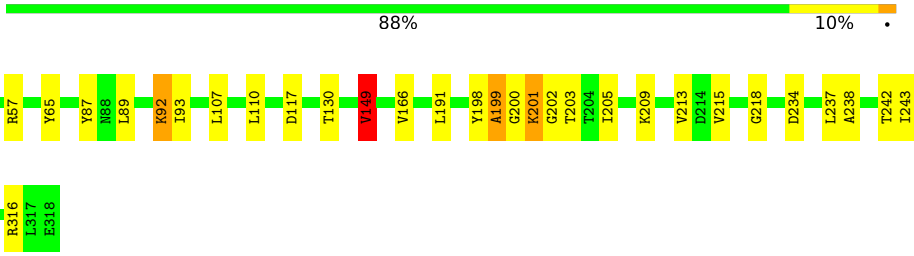
• Molecule 2: Precursor of the S-layer proteins

Chain A:



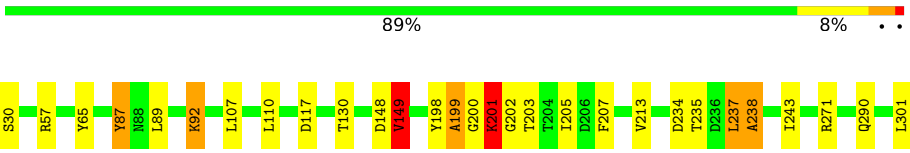
• Molecule 2: Precursor of the S-layer proteins

Chain P:



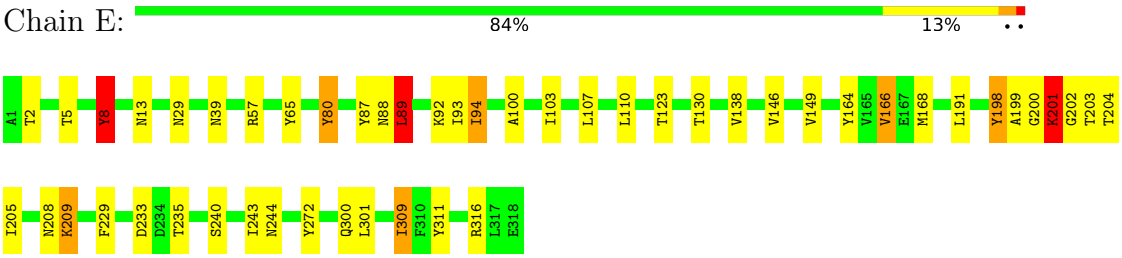
• Molecule 2: Precursor of the S-layer proteins

Chain Q:



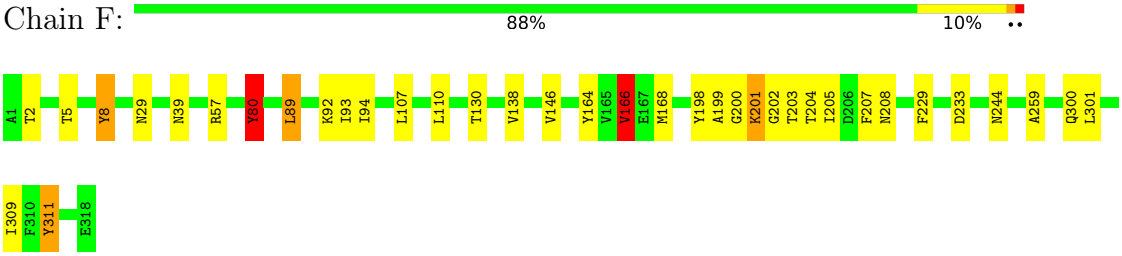
• Molecule 2: Precursor of the S-layer proteins

Chain E:



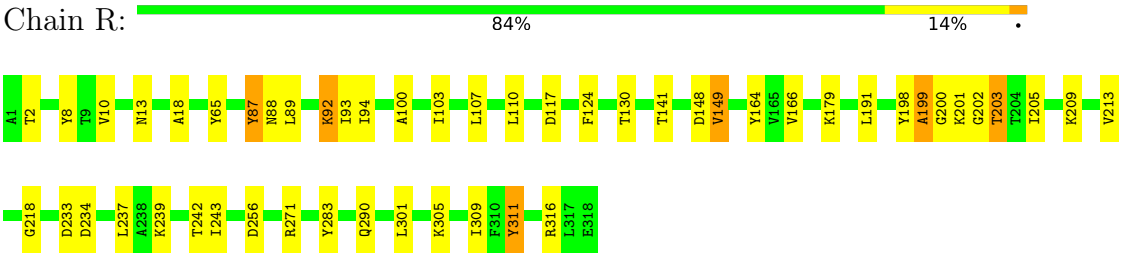
• Molecule 2: Precursor of the S-layer proteins

Chain F:



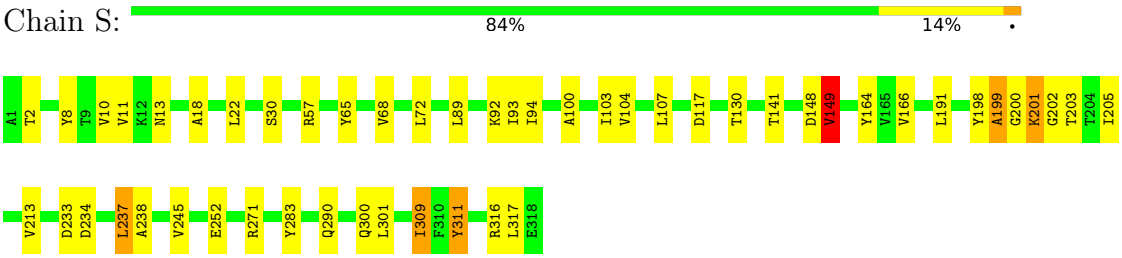
• Molecule 2: Precursor of the S-layer proteins

Chain R:



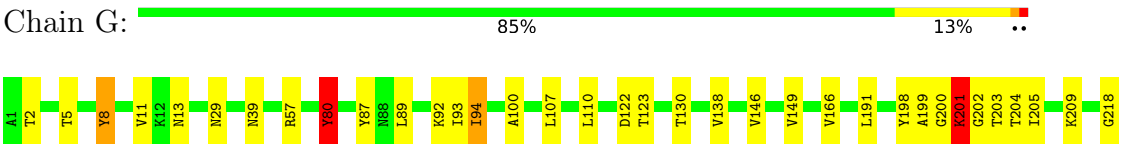
• Molecule 2: Precursor of the S-layer proteins

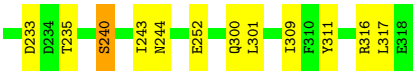
Chain S:



• Molecule 2: Precursor of the S-layer proteins

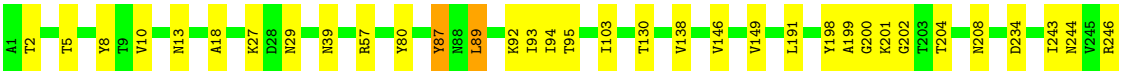
Chain G:

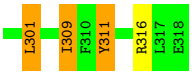




• Molecule 2: Precursor of the S-layer proteins

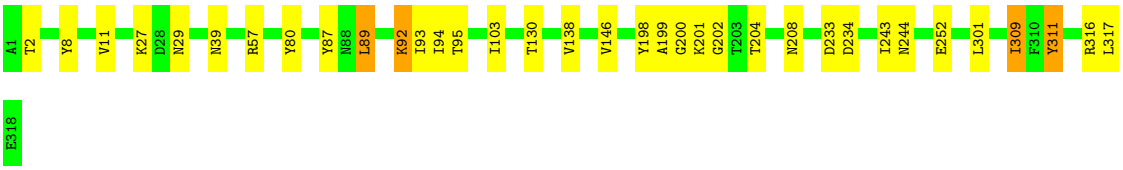
Chain H:





• Molecule 2: Precursor of the S-layer proteins

Chain I:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	CRYSTALLOGRAPHY	Depositor
Imposed symmetry	3D CRYSTAL, a=80 Å, b=80 Å, c=160 Å, α=90°, β=90°, γ=100°, space group=P 1 1 2	Depositor
Number of images used	Not provided	
Resolution determination method	DIFFRACTION PATTERN/LAYERLINES	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI/PHILIPS CM200FEG	Depositor
Voltage (kV)	200	Depositor
Electron dose (e^-/Å^2)	0.1	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN ULTRASCAN 4000 (4k x 4k)	Depositor

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	B	0.62	0/2790	1.10	13/3774 (0.3%)
1	D	0.61	0/2790	1.04	10/3774 (0.3%)
1	J	0.61	0/2790	1.06	9/3774 (0.2%)
1	K	0.61	0/2790	1.08	13/3774 (0.3%)
1	L	0.61	0/2790	1.05	10/3774 (0.3%)
1	M	0.62	0/2790	1.08	13/3774 (0.3%)
1	N	0.63	0/2790	1.08	12/3774 (0.3%)
1	T	0.60	0/2790	1.01	8/3774 (0.2%)
1	U	0.60	0/2790	1.04	12/3774 (0.3%)
1	V	0.60	0/2790	1.03	10/3774 (0.3%)
1	W	0.61	0/2790	1.05	10/3774 (0.3%)
1	X	0.61	0/2790	1.07	12/3774 (0.3%)
2	A	0.65	0/2412	1.15	11/3261 (0.3%)
2	C	0.63	0/2412	1.03	7/3261 (0.2%)
2	E	0.65	0/2412	1.18	20/3261 (0.6%)
2	F	0.63	0/2412	1.12	9/3261 (0.3%)
2	G	0.65	0/2412	1.15	11/3261 (0.3%)
2	H	0.64	0/2412	1.06	8/3261 (0.2%)
2	I	0.64	0/2412	1.05	7/3261 (0.2%)
2	O	0.63	0/2412	1.03	6/3261 (0.2%)
2	P	0.66	0/2412	1.14	11/3261 (0.3%)
2	Q	0.64	0/2412	1.11	11/3261 (0.3%)
2	R	0.66	0/2412	1.17	13/3261 (0.4%)
2	S	0.65	0/2412	1.14	11/3261 (0.3%)
All	All	0.63	0/62424	1.08	257/84420 (0.3%)

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	B	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1
1	J	0	2
1	K	0	1
1	L	0	1
1	M	0	1
1	T	0	1
1	U	0	1
2	A	0	4
2	C	0	4
2	E	0	3
2	F	0	3
2	G	0	2
2	H	0	3
2	I	0	2
2	O	0	2
2	P	0	1
2	Q	0	2
2	R	0	2
2	S	0	2
All	All	0	39

There are no bond length outliers.

The worst 5 of 257 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	R	199	ALA	CB-CA-C	12.01	128.11	110.10
1	N	265	ARG	NE-CZ-NH2	-11.42	114.59	120.30
1	N	265	ARG	NE-CZ-NH1	10.94	125.77	120.30
2	S	199	ALA	CB-CA-C	10.84	126.37	110.10
1	L	265	ARG	NE-CZ-NH1	10.83	125.71	120.30

There are no chirality outliers.

5 of 39 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	61	TYR	Sidechain
2	C	311	TYR	Sidechain
2	C	316	ARG	Sidechain
2	C	8	TYR	Sidechain
2	C	87	TYR	Sidechain

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	B	2767	2795	2792	19	0
1	D	2767	2795	2792	9	0
1	J	2767	2795	2792	17	0
1	K	2767	2795	2792	17	0
1	L	2767	2795	2792	16	0
1	M	2767	2795	2792	11	0
1	N	2767	2795	2792	12	0
1	T	2767	2795	2792	10	0
1	U	2767	2795	2792	17	0
1	V	2767	2795	2792	22	0
1	W	2767	2795	2792	13	0
1	X	2767	2795	2792	19	0
2	A	2386	2389	2389	36	0
2	C	2386	2389	2389	11	0
2	E	2386	2389	2389	41	0
2	F	2386	2389	2389	28	0
2	G	2386	2389	2389	37	0
2	H	2386	2389	2389	22	0
2	I	2386	2389	2389	17	0
2	O	2386	2389	2389	4	0
2	P	2386	2389	2389	24	0
2	Q	2386	2389	2389	25	0
2	R	2386	2389	2389	36	0
2	S	2386	2389	2389	31	0
All	All	61836	62208	62172	384	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 3.

The worst 5 of 384 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)
2:P:237:LEU:HD13	2:P:243:ILE:HG22	1.51	0.89
2:S:309:ILE:HD13	1:X:4:THR:H	1.45	0.80
1:B:4:THR:H	2:A:309:ILE:HD13	1.47	0.79

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:4:THR:H	2:F:309:ILE:HD13	1.49	0.78
2:Q:309:ILE:HD13	1:U:4:THR:H	1.49	0.77

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	371/373 (100%)	363 (98%)	6 (2%)	2 (0%)	29	29
1	D	371/373 (100%)	365 (98%)	6 (2%)	0	100	100
1	J	371/373 (100%)	361 (97%)	8 (2%)	2 (0%)	29	29
1	K	371/373 (100%)	364 (98%)	6 (2%)	1 (0%)	41	41
1	L	371/373 (100%)	364 (98%)	6 (2%)	1 (0%)	41	41
1	M	371/373 (100%)	363 (98%)	7 (2%)	1 (0%)	41	41
1	N	371/373 (100%)	364 (98%)	6 (2%)	1 (0%)	41	41
1	T	371/373 (100%)	363 (98%)	8 (2%)	0	100	100
1	U	371/373 (100%)	363 (98%)	7 (2%)	1 (0%)	41	41
1	V	371/373 (100%)	361 (97%)	10 (3%)	0	100	100
1	W	371/373 (100%)	361 (97%)	9 (2%)	1 (0%)	41	41
1	X	371/373 (100%)	363 (98%)	7 (2%)	1 (0%)	41	41
2	A	316/318 (99%)	308 (98%)	7 (2%)	1 (0%)	41	41
2	C	316/318 (99%)	309 (98%)	5 (2%)	2 (1%)	25	25
2	E	316/318 (99%)	308 (98%)	7 (2%)	1 (0%)	41	41
2	F	316/318 (99%)	308 (98%)	7 (2%)	1 (0%)	41	41
2	G	316/318 (99%)	309 (98%)	6 (2%)	1 (0%)	41	41
2	H	316/318 (99%)	309 (98%)	6 (2%)	1 (0%)	41	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	I	316/318 (99%)	308 (98%)	7 (2%)	1 (0%)	41	41
2	O	316/318 (99%)	309 (98%)	6 (2%)	1 (0%)	41	41
2	P	316/318 (99%)	309 (98%)	4 (1%)	3 (1%)	17	17
2	Q	316/318 (99%)	309 (98%)	4 (1%)	3 (1%)	17	17
2	R	316/318 (99%)	308 (98%)	6 (2%)	2 (1%)	25	25
2	S	316/318 (99%)	309 (98%)	5 (2%)	2 (1%)	25	25
All	All	8244/8292 (99%)	8058 (98%)	156 (2%)	30 (0%)	38	34

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	149	VAL
2	O	149	VAL
2	A	201	LYS
2	P	149	VAL
2	P	201	LYS

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	312/312 (100%)	306 (98%)	6 (2%)	57	57
1	D	312/312 (100%)	305 (98%)	7 (2%)	52	52
1	J	312/312 (100%)	307 (98%)	5 (2%)	62	62
1	K	312/312 (100%)	306 (98%)	6 (2%)	57	57
1	L	312/312 (100%)	306 (98%)	6 (2%)	57	57
1	M	312/312 (100%)	306 (98%)	6 (2%)	57	57
1	N	312/312 (100%)	307 (98%)	5 (2%)	62	62
1	T	312/312 (100%)	305 (98%)	7 (2%)	52	52
1	U	312/312 (100%)	305 (98%)	7 (2%)	52	52
1	V	312/312 (100%)	305 (98%)	7 (2%)	52	52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	W	312/312 (100%)	306 (98%)	6 (2%)	57	57
1	X	312/312 (100%)	306 (98%)	6 (2%)	57	57
2	A	259/259 (100%)	252 (97%)	7 (3%)	44	44
2	C	259/259 (100%)	250 (96%)	9 (4%)	36	36
2	E	259/259 (100%)	249 (96%)	10 (4%)	32	32
2	F	259/259 (100%)	250 (96%)	9 (4%)	36	36
2	G	259/259 (100%)	251 (97%)	8 (3%)	40	40
2	H	259/259 (100%)	249 (96%)	10 (4%)	32	32
2	I	259/259 (100%)	249 (96%)	10 (4%)	32	32
2	O	259/259 (100%)	249 (96%)	10 (4%)	32	32
2	P	259/259 (100%)	250 (96%)	9 (4%)	36	36
2	Q	259/259 (100%)	249 (96%)	10 (4%)	32	32
2	R	259/259 (100%)	250 (96%)	9 (4%)	36	36
2	S	259/259 (100%)	248 (96%)	11 (4%)	30	30
All	All	6852/6852 (100%)	6666 (97%)	186 (3%)	48	44

5 of 186 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	R	2	THR
1	M	8	GLN
2	R	141	THR
2	S	234	ASP
1	N	112	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
2	R	91	ASN
1	W	131	ASN
2	G	13	ASN
1	N	19	ASN
2	H	111	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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 Map visualisation

This section contains visualisations of the EMDB entry EMD-13957. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal surface views

This section was not generated.

6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.