



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:41 PM BST

PDB ID : 3FOI
EMDB ID: : EMD-1086
Title : Fitting of gp18M crystal structure into 3D cryo-EM reconstruction of bacteriophage T4 contracted tail
Authors : Aksyuk, A.A.; Leiman, P.G.; Kurochkina, L.P.; Shneider, M.M.; Kostyuchenko, V.A.; Mesyanzhinov, V.V.; Rossmann, M.G.
Deposited on : 2008-12-30
Resolution : 16.00 Å(reported)
Based on PDB ID : 3FOA

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

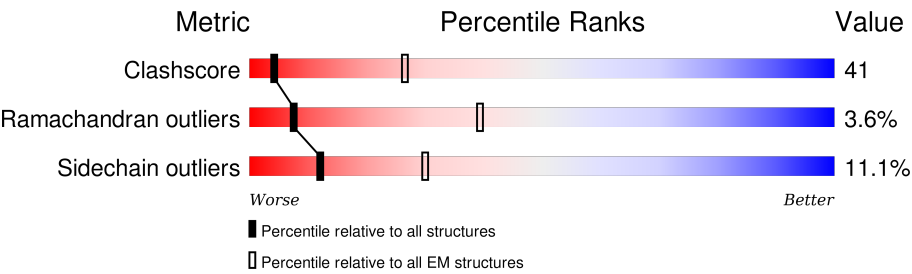
MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

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 16.00 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	510	<div><div>39%</div><div>44%</div><div>10%</div><div>• 6%</div></div>
1	B	510	<div><div>40%</div><div>44%</div><div>10%</div><div>• 6%</div></div>
1	C	510	<div><div>40%</div><div>44%</div><div>10%</div><div>• 6%</div></div>
1	D	510	<div><div>40%</div><div>44%</div><div>10%</div><div>• 6%</div></div>
1	E	510	<div><div>40%</div><div>44%</div><div>10%</div><div>• 6%</div></div>
1	F	510	<div><div>40%</div><div>43%</div><div>10%</div><div>• 6%</div></div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 21678 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 Tail sheath protein Gp18.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	479	Total	C	N	O	S	0	0
			3613	2279	600	726	8		
1	B	479	Total	C	N	O	S	0	0
			3613	2279	600	726	8		
1	C	479	Total	C	N	O	S	0	0
			3613	2279	600	726	8		
1	D	479	Total	C	N	O	S	0	0
			3613	2279	600	726	8		
1	E	479	Total	C	N	O	S	0	0
			3613	2279	600	726	8		
1	F	479	Total	C	N	O	S	0	0
			3613	2279	600	726	8		

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	GLU	ASP	SEE REMARK 999	UNP P13332
A	148	ALA	GLY	SEE REMARK 999	UNP P13332
A	150	ILE	ASN	SEE REMARK 999	UNP P13332
A	151	ILE	TYR	SEE REMARK 999	UNP P13332
A	301	GLY	GLU	SEE REMARK 999	UNP P13332
A	399	VAL	ALA	SEE REMARK 999	UNP P13332
A	454	TYR	HIS	SEE REMARK 999	UNP P13332
A	510	PRO	ARG	ENGINEERED	UNP P13332
B	100	GLU	ASP	SEE REMARK 999	UNP P13332
B	148	ALA	GLY	SEE REMARK 999	UNP P13332
B	150	ILE	ASN	SEE REMARK 999	UNP P13332
B	151	ILE	TYR	SEE REMARK 999	UNP P13332
B	301	GLY	GLU	SEE REMARK 999	UNP P13332
B	399	VAL	ALA	SEE REMARK 999	UNP P13332
B	454	TYR	HIS	SEE REMARK 999	UNP P13332
B	510	PRO	ARG	ENGINEERED	UNP P13332
C	100	GLU	ASP	SEE REMARK 999	UNP P13332
C	148	ALA	GLY	SEE REMARK 999	UNP P13332

Continued on next page...

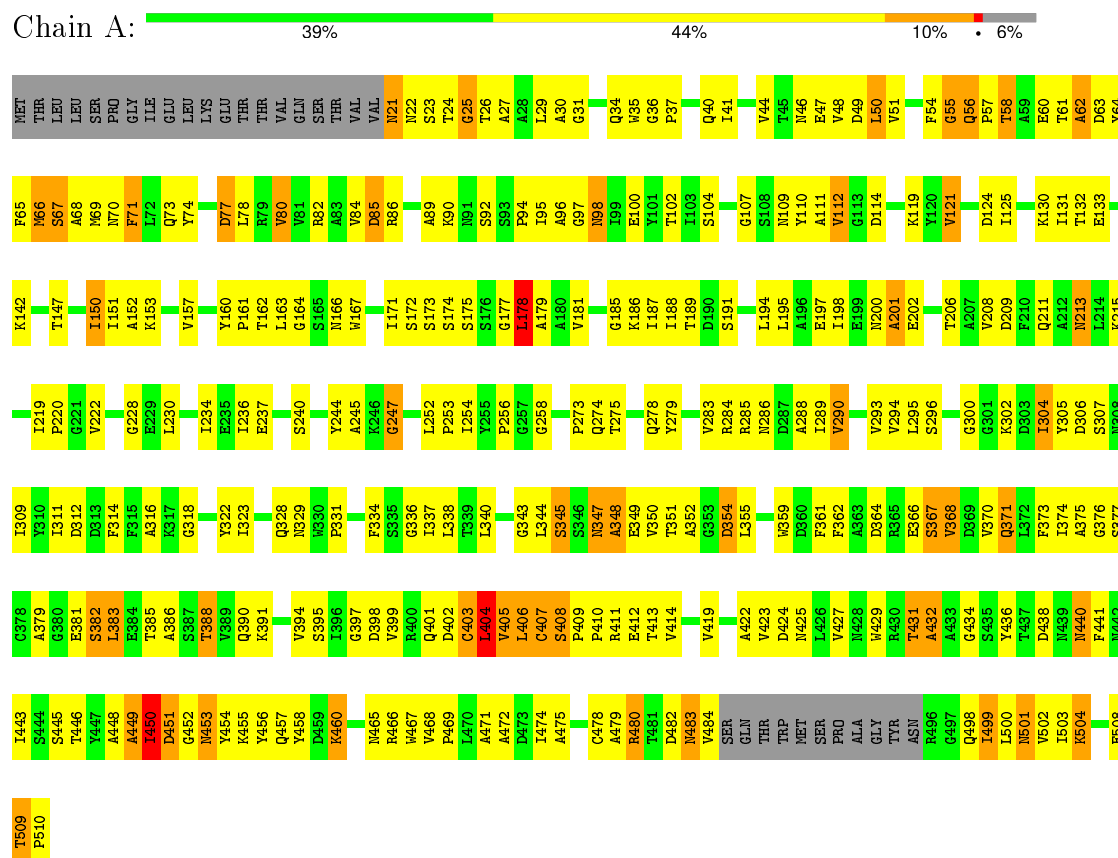
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	150	ILE	ASN	SEE REMARK 999	UNP P13332
C	151	ILE	TYR	SEE REMARK 999	UNP P13332
C	301	GLY	GLU	SEE REMARK 999	UNP P13332
C	399	VAL	ALA	SEE REMARK 999	UNP P13332
C	454	TYR	HIS	SEE REMARK 999	UNP P13332
C	510	PRO	ARG	ENGINEERED	UNP P13332
D	100	GLU	ASP	SEE REMARK 999	UNP P13332
D	148	ALA	GLY	SEE REMARK 999	UNP P13332
D	150	ILE	ASN	SEE REMARK 999	UNP P13332
D	151	ILE	TYR	SEE REMARK 999	UNP P13332
D	301	GLY	GLU	SEE REMARK 999	UNP P13332
D	399	VAL	ALA	SEE REMARK 999	UNP P13332
D	454	TYR	HIS	SEE REMARK 999	UNP P13332
D	510	PRO	ARG	ENGINEERED	UNP P13332
E	100	GLU	ASP	SEE REMARK 999	UNP P13332
E	148	ALA	GLY	SEE REMARK 999	UNP P13332
E	150	ILE	ASN	SEE REMARK 999	UNP P13332
E	151	ILE	TYR	SEE REMARK 999	UNP P13332
E	301	GLY	GLU	SEE REMARK 999	UNP P13332
E	399	VAL	ALA	SEE REMARK 999	UNP P13332
E	454	TYR	HIS	SEE REMARK 999	UNP P13332
E	510	PRO	ARG	ENGINEERED	UNP P13332
F	100	GLU	ASP	SEE REMARK 999	UNP P13332
F	148	ALA	GLY	SEE REMARK 999	UNP P13332
F	150	ILE	ASN	SEE REMARK 999	UNP P13332
F	151	ILE	TYR	SEE REMARK 999	UNP P13332
F	301	GLY	GLU	SEE REMARK 999	UNP P13332
F	399	VAL	ALA	SEE REMARK 999	UNP P13332
F	454	TYR	HIS	SEE REMARK 999	UNP P13332
F	510	PRO	ARG	ENGINEERED	UNP P13332

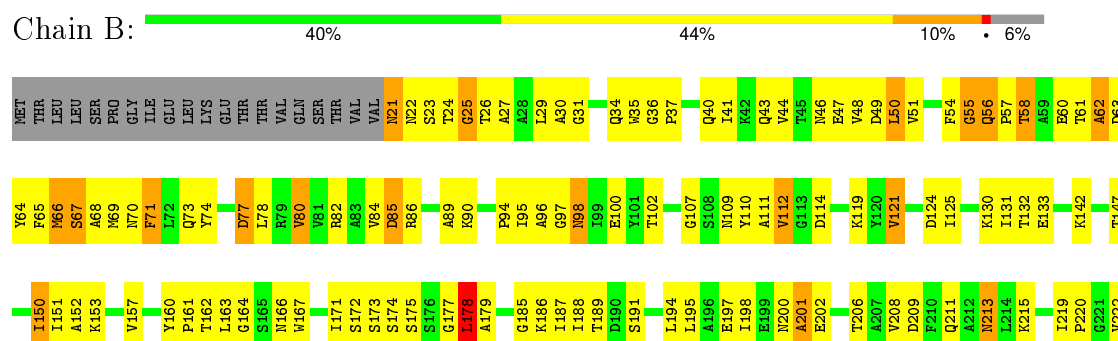
3 Residue-property plots

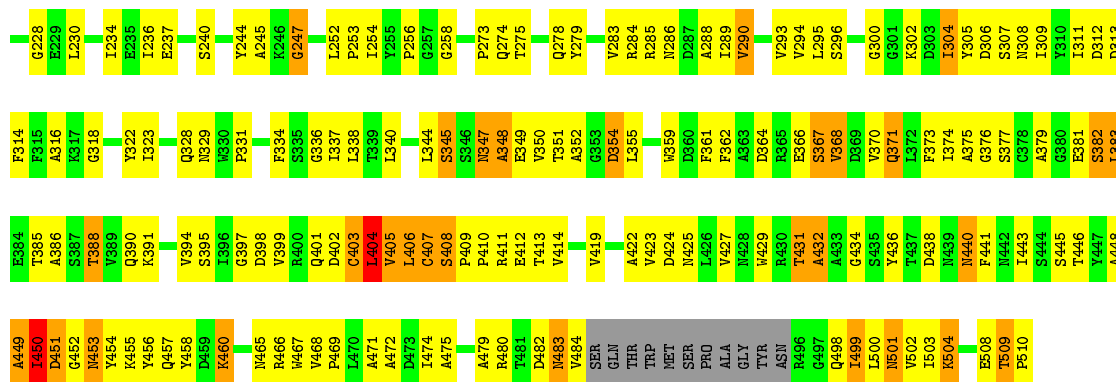
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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: Tail sheath protein Gp18



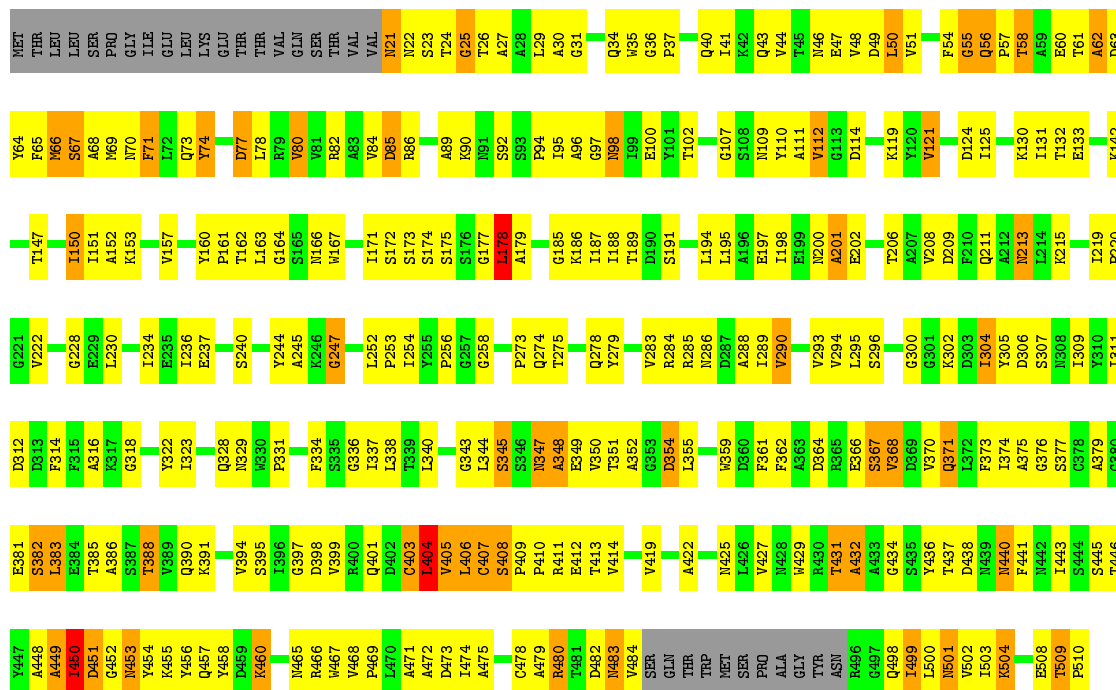
• Molecule 1: Tail sheath protein Gp18





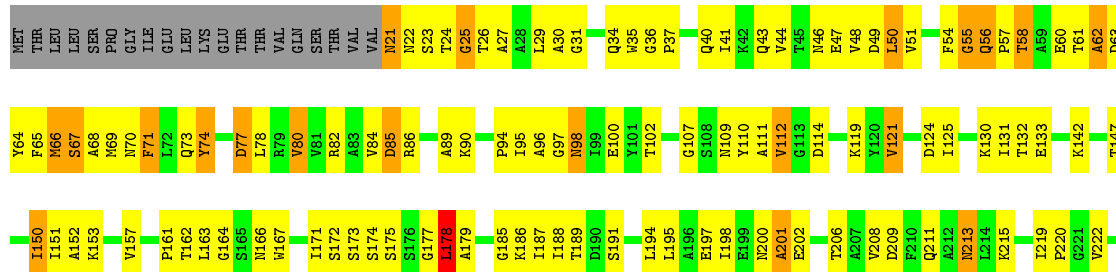
• Molecule 1: Tail sheath protein Gp18

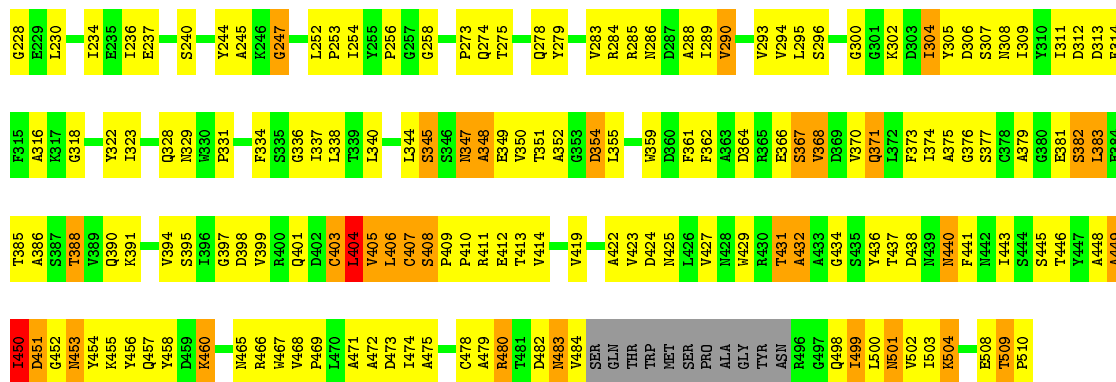
Chain C: 40% 44% 10% 6%



• Molecule 1: Tail sheath protein Gp18

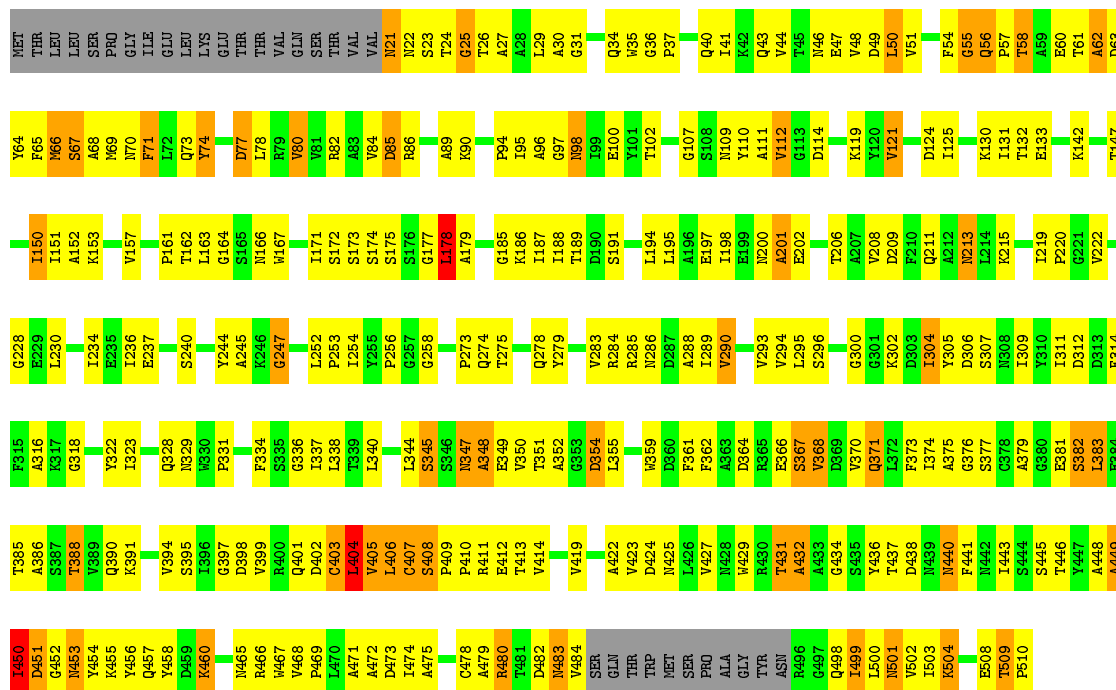
Chain D: 40% 44% 10% 6%





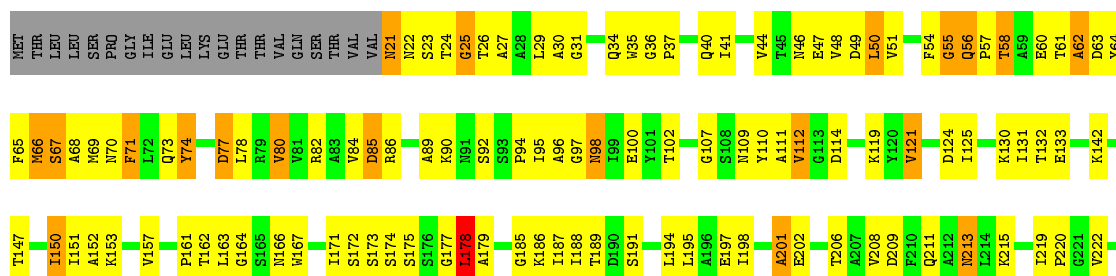
• Molecule 1: Tail sheath protein Gp18

Chain E: 40% 44% 10% 6%



• Molecule 1: Tail sheath protein Gp18

Chain F: 40% 43% 10% 6%



I450	I384	P315	Q228
D451	T385	A316	E229
G452	A386	K317	I230
N453	S387	G318	
Y454	T388		I234
K455	V389	Y322	E235
Y456	Q390	I323	I236
Q457	K391		E237
Y458		Q328	
D459	V394	N329	S240
K460	K395	W330	
	I396	P331	Y244
N465	G397		A245
R466	D398	F334	E246
W467	V399	S335	G247
V468	R400	G336	
P469	Q401	I337	I252
I470	D402	I338	P253
A471	C403	T339	I254
A472	L404	L340	I255
D473	V405		P256
I474	L406	G343	G257
A475	C407	I344	G258
	S408	S345	
C478	P409	S346	P273
A479	P410	N347	Q274
R480	R411	A348	I275
T481	E412	E349	
D482	T413	V350	Q278
N483	V414	T351	Y279
V484		A352	
SER	V419	G353	V283
GLN		D354	R284
THR	A422	L355	R285
TRP			I286
MET	N425	W359	D287
SER	L426	D360	A288
PRO	V427	F361	I289
ALA	R428	F362	
GLY	W429	A363	V290
TYR	R430	D364	
ASN	T431	E365	V293
	A432	E366	V294
R496	A433	S367	L295
Q497	G434	V368	S296
Q498	S435	D369	
I499	L500	V370	G300
N501	T437	Q371	G301
V502	D438	I372	K302
I503	W439	F373	D303
K504	N440	I374	I304
	F441	A375	I305
E508	N442	G376	D306
T509	I443	S377	S307
	S444	C378	N308
P510	S445	A379	I309
	T446	G380	V310
	T447	E381	I311
	A448	S382	D312
	A449	L383	D313
			F314

4 Experimental information

Property	Value	Source
Reconstruction method	Not provided	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	Not provided	Depositor
Voltage (kV)	Not provided	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.35	0/3677	0.78	21/5001 (0.4%)
1	B	0.35	0/3677	0.78	21/5001 (0.4%)
1	C	0.35	0/3677	0.78	21/5001 (0.4%)
1	D	0.35	0/3677	0.78	21/5001 (0.4%)
1	E	0.35	0/3677	0.78	21/5001 (0.4%)
1	F	0.35	0/3677	0.78	21/5001 (0.4%)
All	All	0.35	0/22062	0.78	126/30006 (0.4%)

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	4
1	B	0	4
1	C	0	4
1	D	0	4
1	E	0	4
1	F	0	4
All	All	0	24

There are no bond length outliers.

The worst 5 of 126 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	E	55	GLY	N-CA-C	18.03	158.17	113.10
1	A	55	GLY	N-CA-C	18.00	158.10	113.10
1	C	55	GLY	N-CA-C	18.00	158.10	113.10
1	D	55	GLY	N-CA-C	17.99	158.08	113.10
1	B	55	GLY	N-CA-C	17.99	158.07	113.10

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	21	ASN	Peptide
1	A	450	ILE	Peptide
1	A	451	ASP	Peptide
1	A	452	GLY	Peptide
1	B	21	ASN	Peptide

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	3613	0	3554	294	0
1	B	3613	0	3554	297	0
1	C	3613	0	3554	297	0
1	D	3613	0	3554	295	0
1	E	3613	0	3554	296	0
1	F	3613	0	3554	296	0
All	All	21678	0	21324	1775	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 41.

The worst 5 of 1775 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:A:409:PRO:O	1:A:454:TYR:CE1	1.65	1.49
1:B:409:PRO:O	1:B:454:TYR:CE1	1.65	1.49
1:F:409:PRO:O	1:F:454:TYR:CE1	1.65	1.49
1:C:409:PRO:O	1:C:454:TYR:CE1	1.65	1.48
1:D:409:PRO:O	1:D:454:TYR:CE1	1.65	1.47

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	A	475/510 (93%)	371 (78%)	87 (18%)	17 (4%)	4	38
1	B	475/510 (93%)	371 (78%)	87 (18%)	17 (4%)	4	38
1	C	475/510 (93%)	371 (78%)	87 (18%)	17 (4%)	4	38
1	D	475/510 (93%)	371 (78%)	87 (18%)	17 (4%)	4	38
1	E	475/510 (93%)	371 (78%)	87 (18%)	17 (4%)	4	38
1	F	475/510 (93%)	370 (78%)	88 (18%)	17 (4%)	4	38
All	All	2850/3060 (93%)	2225 (78%)	523 (18%)	102 (4%)	7	38

5 of 102 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	98	ASN
1	A	201	ALA
1	A	432	ALA
1	B	98	ASN
1	B	201	ALA

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	A	383/411 (93%)	341 (89%)	42 (11%)	8	34
1	B	383/411 (93%)	341 (89%)	42 (11%)	8	34
1	C	383/411 (93%)	340 (89%)	43 (11%)	7	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	383/411 (93%)	340 (89%)	43 (11%)	7	33
1	E	383/411 (93%)	340 (89%)	43 (11%)	7	33
1	F	383/411 (93%)	340 (89%)	43 (11%)	7	33
All	All	2298/2466 (93%)	2042 (89%)	256 (11%)	12	34

5 of 256 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	431	THR
1	D	202	GLU
1	F	367	SER
1	C	440	ASN
1	D	67	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	440	ASN
1	D	213	ASN
1	F	390	GLN
1	C	453	ASN
1	D	43	GLN

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 carbohydrates 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.