wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:42 PM BST

PDB ID	:	3IZO
EMDB ID:	:	EMD-5172
Title	:	Model of the fiber tail and its interactions with the penton base of human adenovirus by cryo-electron microscopy
Authors	:	Liu, H.
Deposited on	:	2010-11-05
Resolution	:	3.60 Å(reported)

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

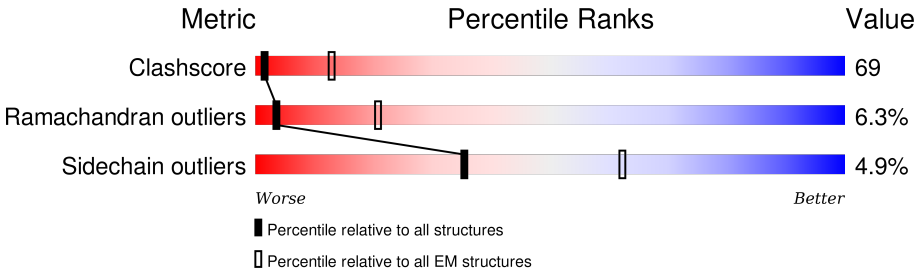
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 3.60 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	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 <=5%

Mol	Chain	Length	Quality of chain
1	A	571	38% 36% 6% 20%
1	B	571	37% 37% 5% 20%
1	C	571	37% 37% 5% 20%
1	D	571	39% 36% 5% 20%
1	E	571	38% 36% 5% 20%
2	F	581	98%
2	G	581	98%
2	H	581	98%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18537 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 Penton protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	455	Total	C	N	O	S	0	0
			3642	2308	632	690	12		
1	B	455	Total	C	N	O	S	0	0
			3642	2308	632	690	12		
1	C	455	Total	C	N	O	S	0	0
			3642	2308	632	690	12		
1	D	455	Total	C	N	O	S	0	0
			3642	2308	632	690	12		
1	E	455	Total	C	N	O	S	0	0
			3642	2308	632	690	12		

- Molecule 2 is a protein called Fiber.

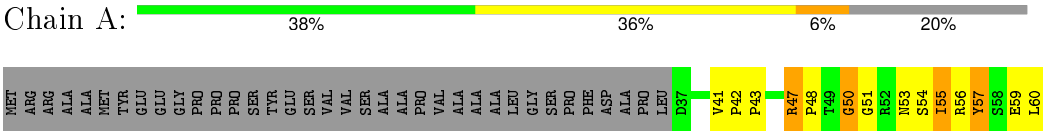
Mol	Chain	Residues	Atoms				AltConf	Trace
2	F	13	Total	C	N	O	0	0
			109	70	14	25		
2	G	13	Total	C	N	O	0	0
			109	70	14	25		
2	H	13	Total	C	N	O	0	0
			109	70	14	25		

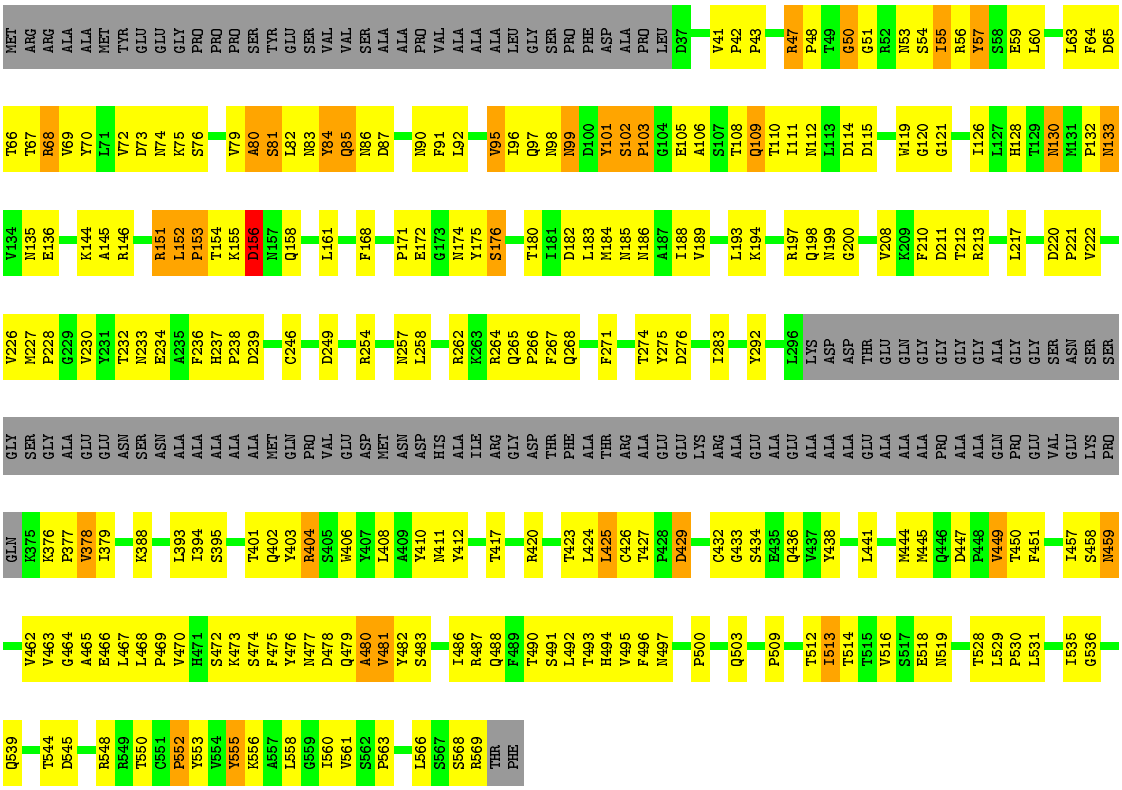
3 Residue-property plots (i)

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: Penton protein

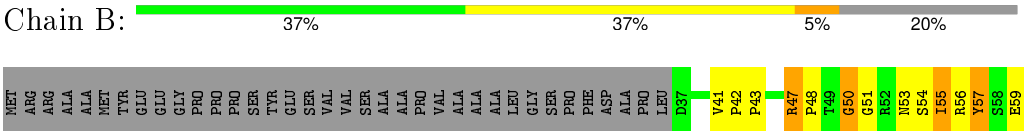
Chain A:

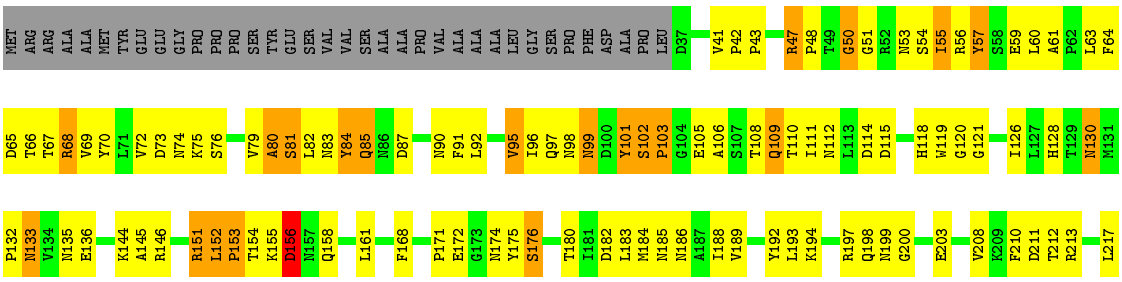


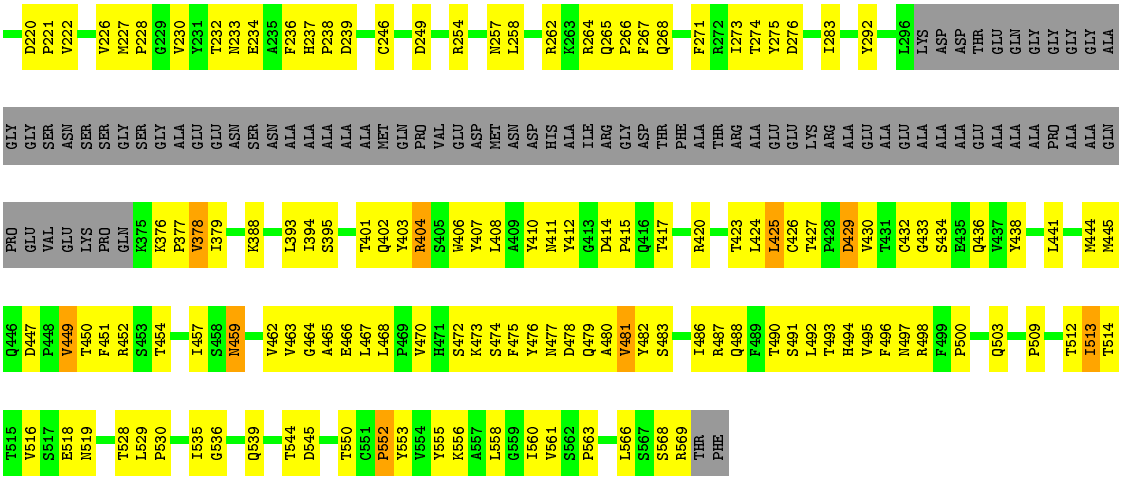


- Molecule 1: Penton protein

Chain B:

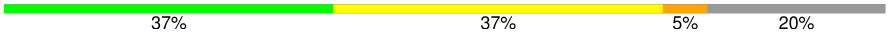


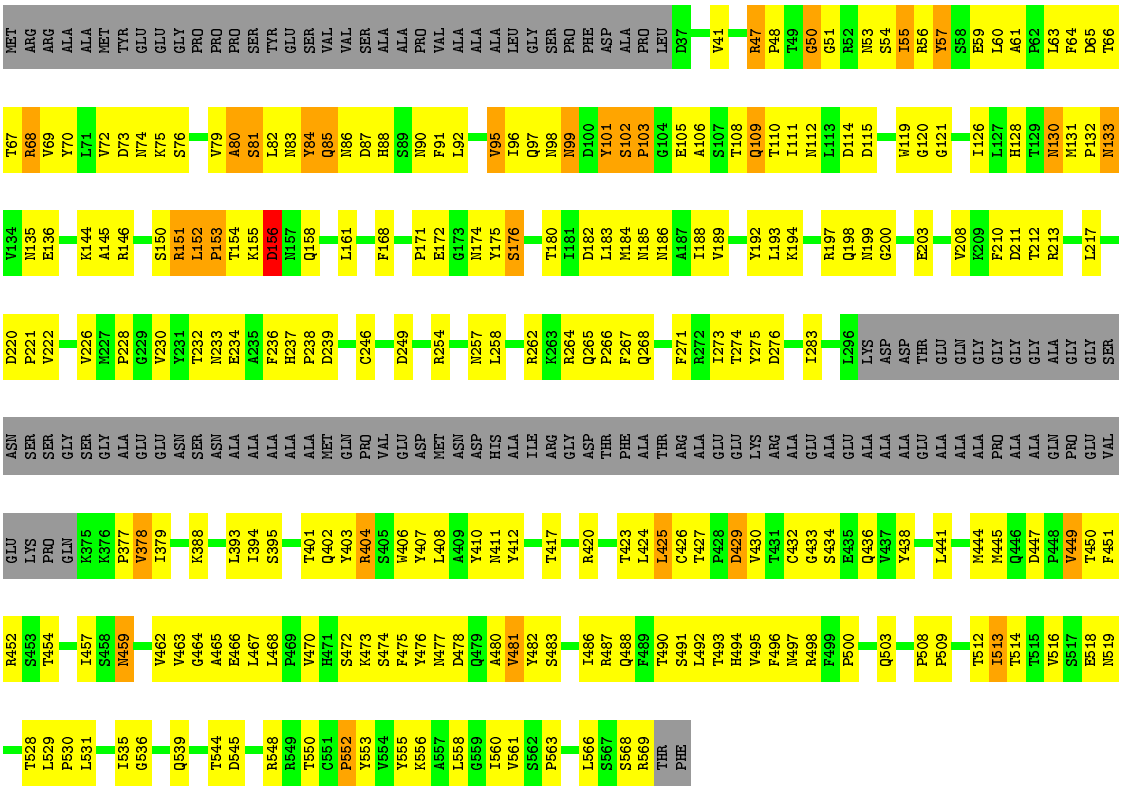




• Molecule 1: Penton protein

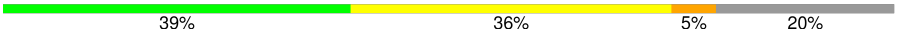
Chain C:

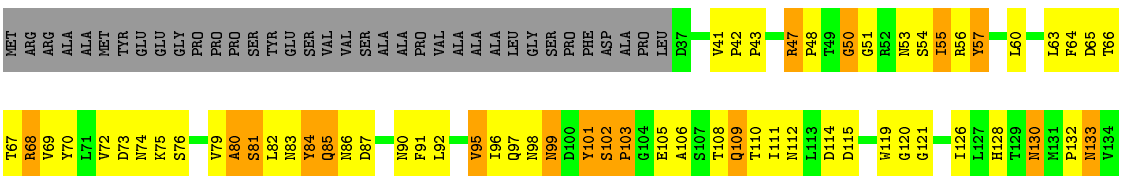


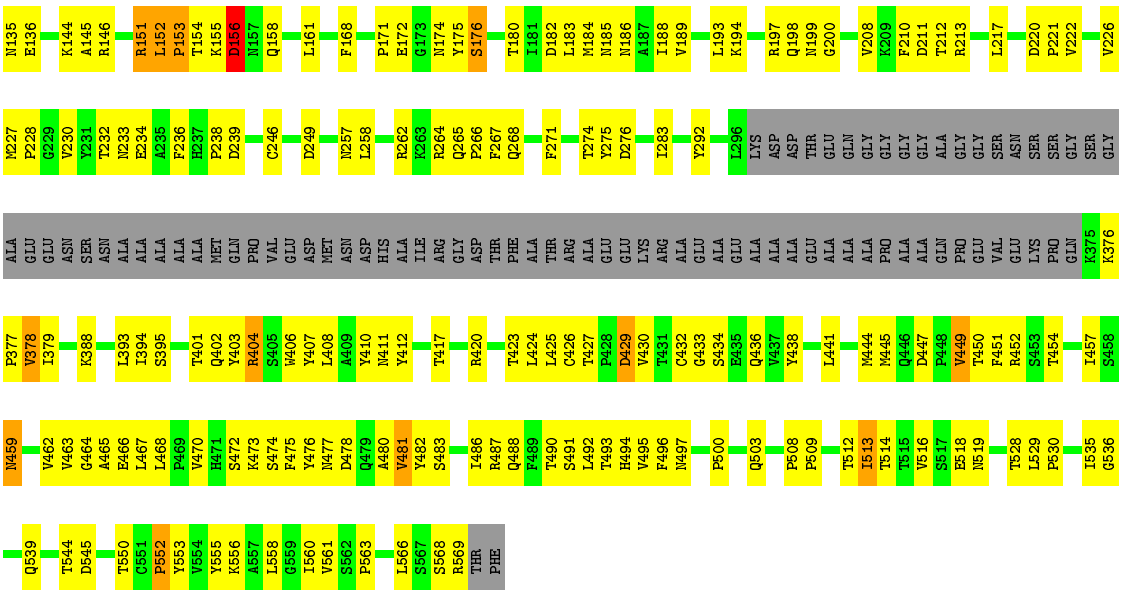


• Molecule 1: Penton protein

Chain D:

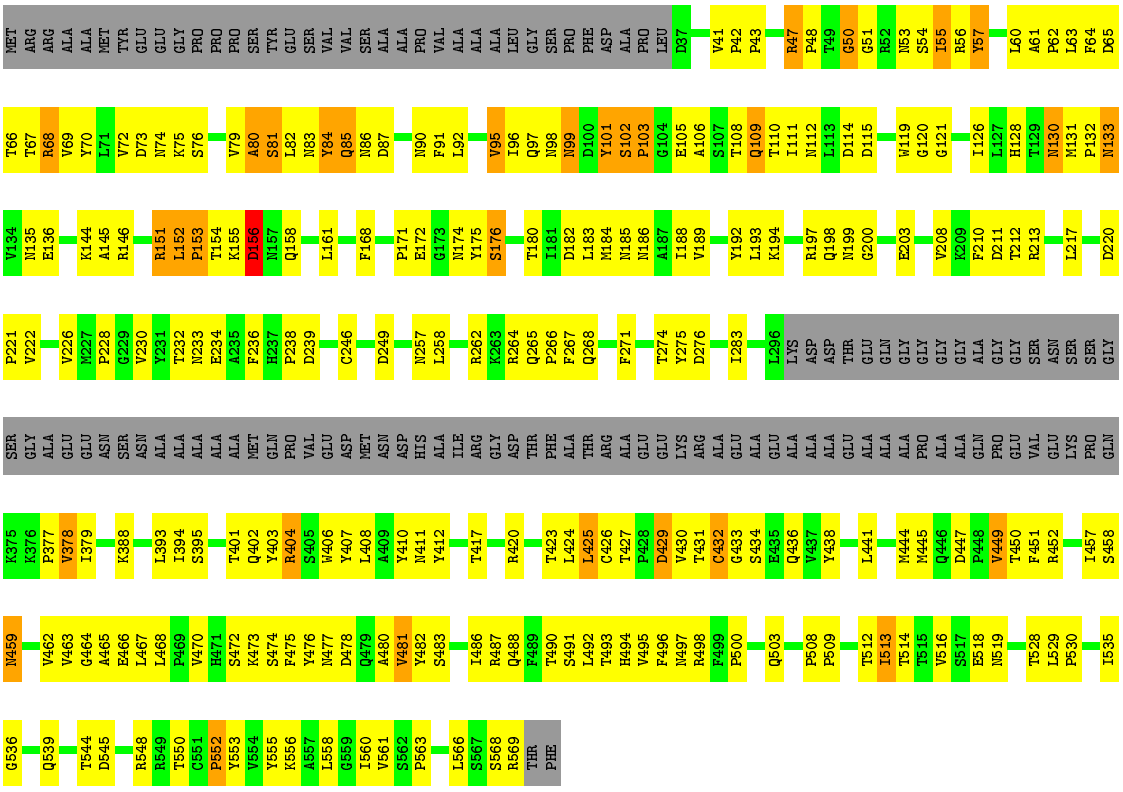






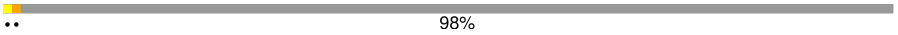
- Molecule 1: Penton protein

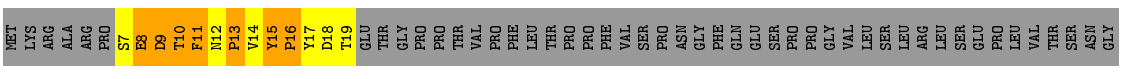
Chain E:

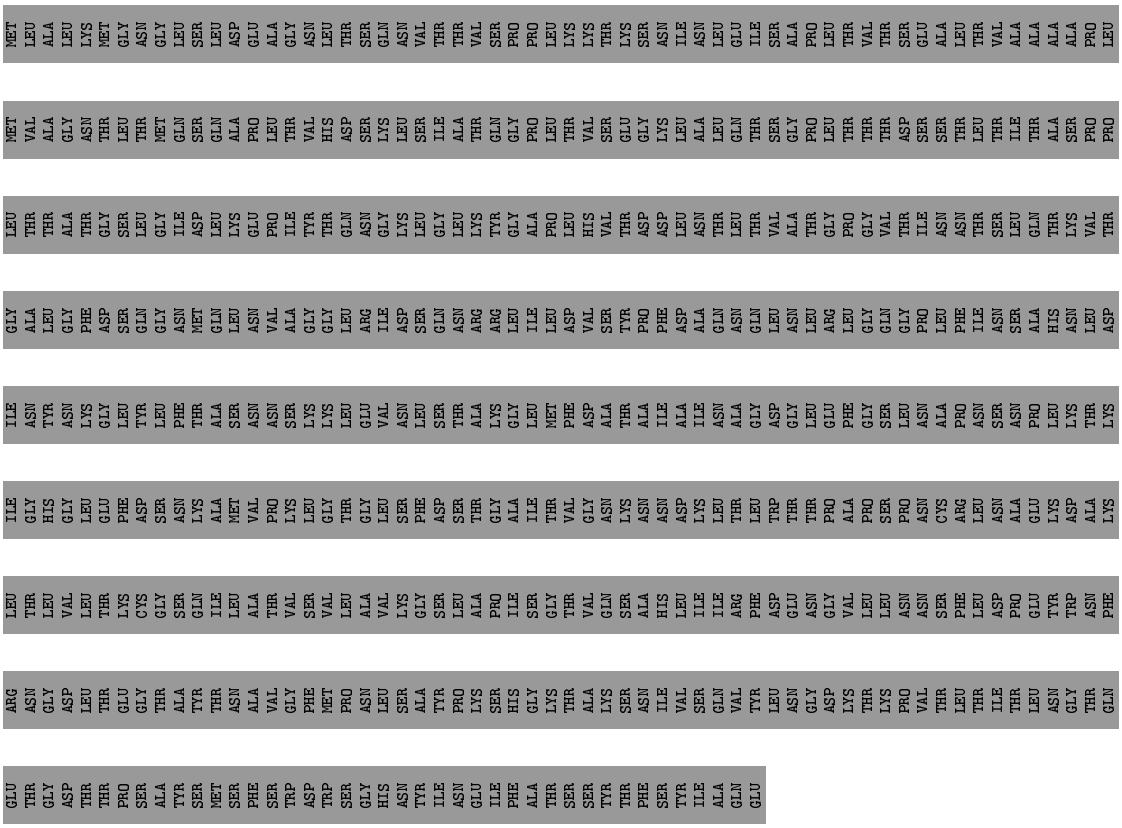


- Molecule 2: Fiber

Chain F:

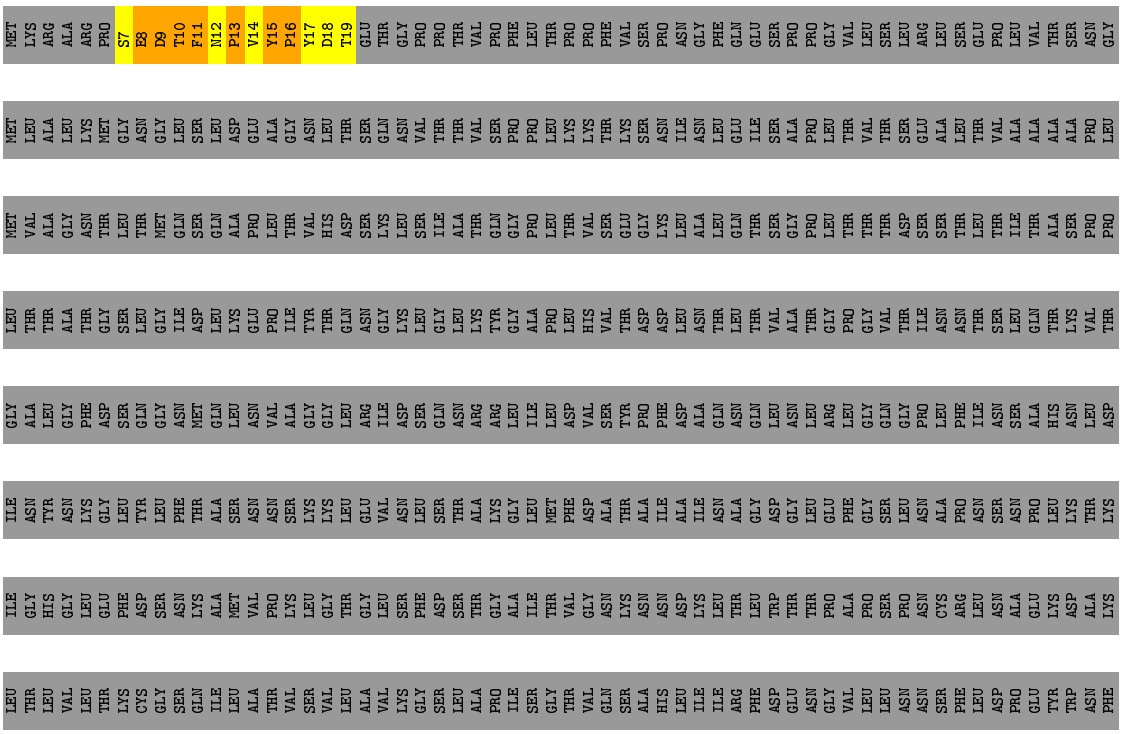


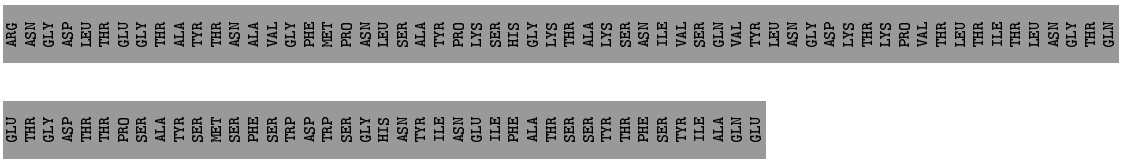




- Molecule 2: Fiber

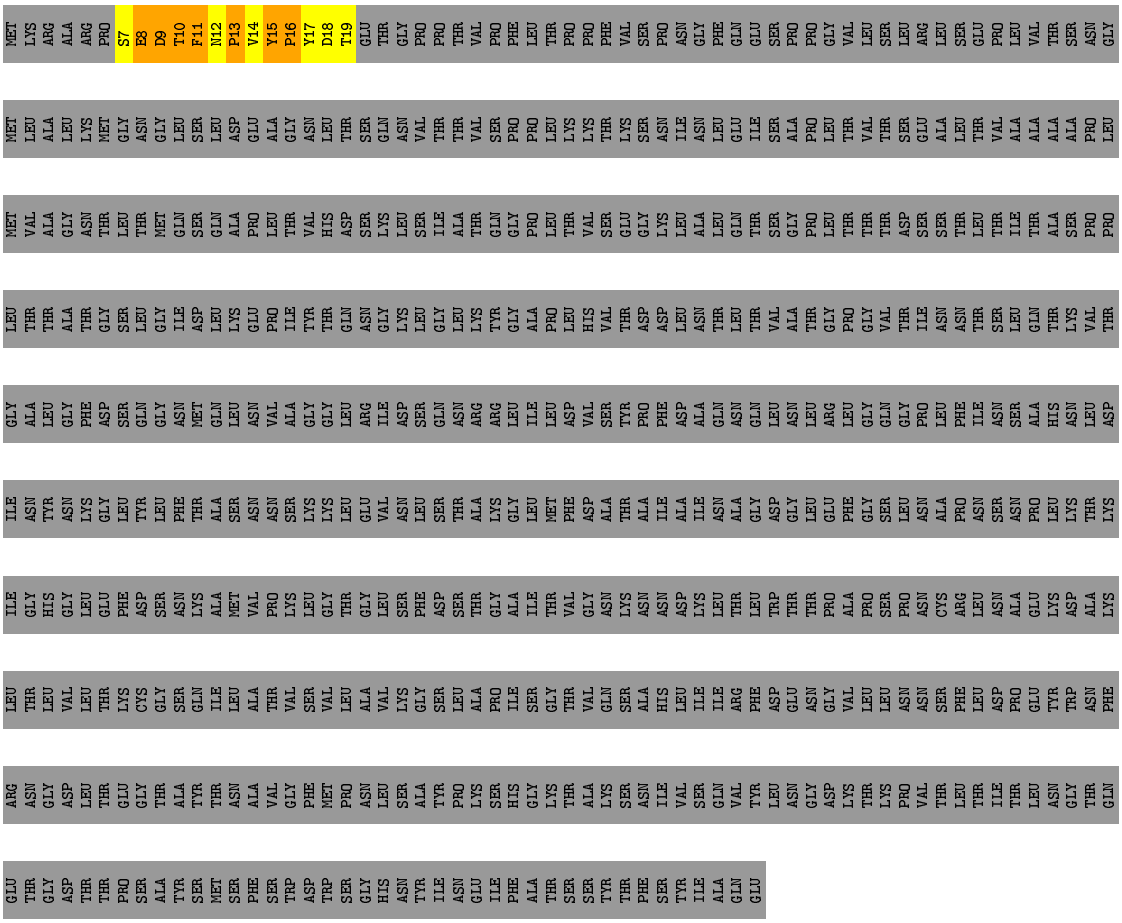
Chain G: 98%





- Molecule 2: Fiber

Chain H: 98%



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	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	Titan Krios	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	KODAK SO163 FILM	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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.33	0/3733	0.54	0/5088
1	B	0.33	0/3733	0.55	0/5088
1	C	0.33	0/3733	0.55	0/5088
1	D	0.33	0/3733	0.55	0/5088
1	E	0.33	0/3733	0.54	0/5088
2	F	0.22	0/113	0.39	0/156
2	G	0.22	0/113	0.39	0/156
2	H	0.22	0/113	0.39	0/156
All	All	0.33	0/19004	0.54	0/25908

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
All	All	0	10

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	47	ARG	Sidechain
1	A	68	ARG	Sidechain
1	B	47	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	68	ARG	Sidechain
1	C	47	ARG	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	A	3642	0	3571	777	0
1	B	3642	0	3571	780	0
1	C	3642	0	3571	757	0
1	D	3642	0	3571	735	0
1	E	3642	0	3571	742	0
2	F	109	0	88	55	0
2	G	109	0	88	52	0
2	H	109	0	88	55	0
All	All	18537	0	18119	2535	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 69.

The worst 5 of 2535 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:410:TYR:CE2	1:D:425:LEU:HD12	1.34	1.62
1:D:425:LEU:HD13	1:E:172:GLU:CB	1.27	1.59
1:A:267:PHE:CZ	1:B:80:ALA:HA	1.34	1.58
1:C:450:THR:HG23	1:D:57:TYR:CE1	1.41	1.54
1:A:450:THR:HG23	1:B:57:TYR:CE1	1.45	1.51

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	451/571 (79%)	341 (76%)	84 (19%)	26 (6%)	2 25
1	B	451/571 (79%)	341 (76%)	85 (19%)	25 (6%)	2 26
1	C	451/571 (79%)	341 (76%)	85 (19%)	25 (6%)	2 26
1	D	451/571 (79%)	341 (76%)	85 (19%)	25 (6%)	2 26
1	E	451/571 (79%)	340 (75%)	85 (19%)	26 (6%)	2 25
2	F	11/581 (2%)	3 (27%)	2 (18%)	6 (54%)	0 0
2	G	11/581 (2%)	3 (27%)	2 (18%)	6 (54%)	0 0
2	H	11/581 (2%)	3 (27%)	2 (18%)	6 (54%)	0 0
All	All	2288/4598 (50%)	1713 (75%)	430 (19%)	145 (6%)	3 23

5 of 145 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	81	SER
1	A	156	ASP
1	A	378	VAL
2	F	8	GLU
2	F	9	ASP

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 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	412/489 (84%)	391 (95%)	21 (5%)	29 70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	412/489 (84%)	392 (95%)	20 (5%)	31 71
1	C	412/489 (84%)	392 (95%)	20 (5%)	31 71
1	D	412/489 (84%)	393 (95%)	19 (5%)	33 73
1	E	412/489 (84%)	392 (95%)	20 (5%)	31 71
2	F	13/489 (3%)	12 (92%)	1 (8%)	16 56
2	G	13/489 (3%)	12 (92%)	1 (8%)	16 56
2	H	13/489 (3%)	12 (92%)	1 (8%)	16 56
All	All	2099/3912 (54%)	1996 (95%)	103 (5%)	35 71

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	101	TYR
1	C	276	ASP
1	E	211	ASP
1	C	108	THR
1	C	156	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 71 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	109	GLN
1	C	411	ASN
1	E	257	ASN
1	C	130	ASN
1	C	186	ASN

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

There are no carbohydrates 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.