



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:50 PM BST

PDB ID : 3J40
EMDB ID: : EMD-5678
Title : Validated Near-Atomic Resolution Structure of Bacteriophage Epsilon15 Derived from Cryo-EM and Modeling
Authors : Baker, M.L.; Hryc, C.F.; Zhang, Q.; Wu, W.; Jakana, J.; Haase-Pettingell, C.; Afonine, P.V.; Adams, P.D.; King, J.A.; Jiang, W.; Chiu, W.
Deposited on : 2013-05-30
Resolution : 4.50 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

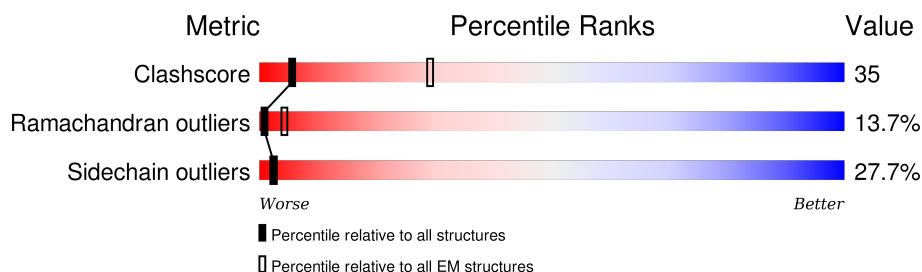
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.50 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	H	111	25% 49% 23% .
1	I	111	25% 42% 28% 5%
1	J	111	25% 41% 29% 5%
1	K	111	23% 52% 22% .
1	L	111	23% 41% 28% 7%
1	M	111	21% 46% 29% 5%
1	N	111	27% 43% 26% .
2	A	335	29% 50% 19% .
2	B	335	38% 38% 22% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	335	<div><div></div><div>37%</div><div>39%</div><div>21%</div><div></div></div>
2	D	335	<div><div></div><div>28%</div><div>49%</div><div>20%</div><div></div></div>
2	E	335	<div><div></div><div>37%</div><div>39%</div><div>22%</div><div></div></div>
2	F	335	<div><div></div><div>37%</div><div>42%</div><div>19%</div><div></div></div>
2	G	335	<div><div></div><div>38%</div><div>44%</div><div>15%</div><div></div></div>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 24066 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 gp10.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	N	111	Total	C	N	O	S	0	0
			850	521	143	181	5		
1	M	111	Total	C	N	O	S	0	0
			850	521	143	181	5		
1	H	111	Total	C	N	O	S	0	0
			850	521	143	181	5		
1	K	111	Total	C	N	O	S	0	0
			850	521	143	181	5		
1	I	111	Total	C	N	O	S	0	0
			850	521	143	181	5		
1	J	111	Total	C	N	O	S	0	0
			850	521	143	181	5		
1	L	111	Total	C	N	O	S	0	0
			850	521	143	181	5		

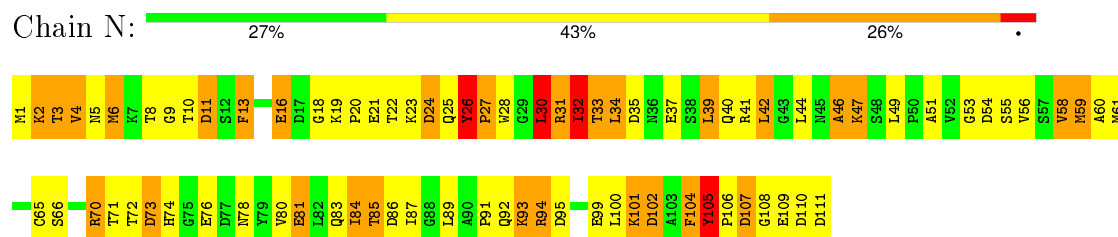
- Molecule 2 is a protein called gp7.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		
2	F	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		
2	B	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		
2	G	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		
2	D	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		
2	C	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		
2	E	335	Total	C	N	O	S	0	0
			2588	1620	453	502	13		

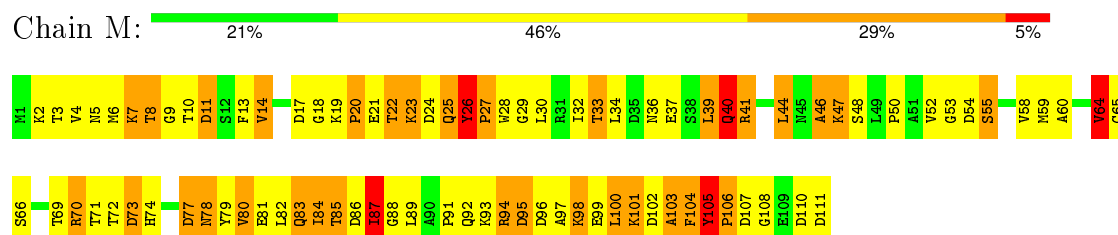
3 Residue-property plots

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.

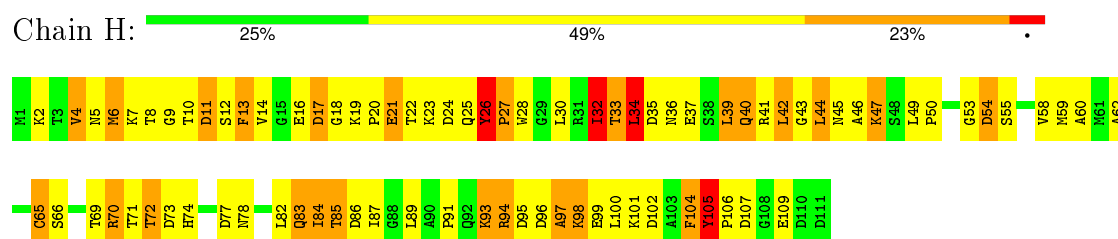
- Molecule 1: gp10



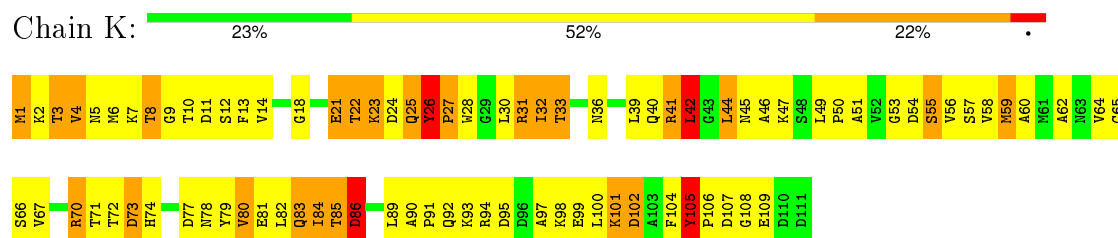
- Molecule 1: gp10



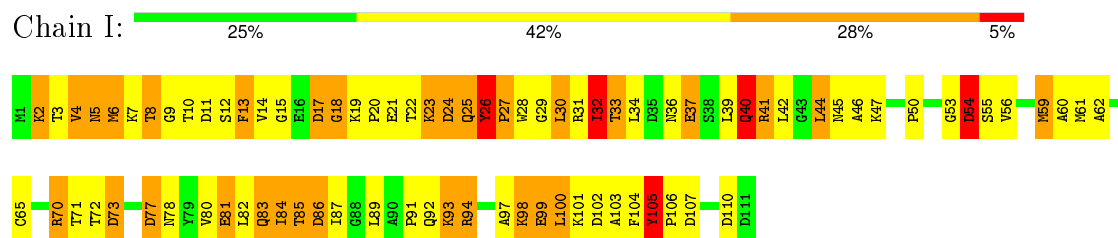
- Molecule 1: gp10



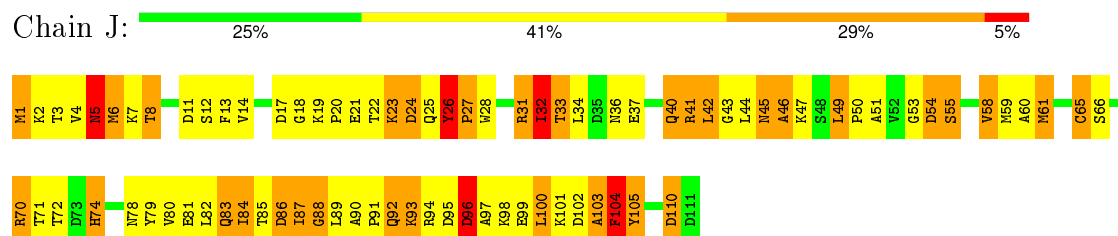
- Molecule 1: gp10



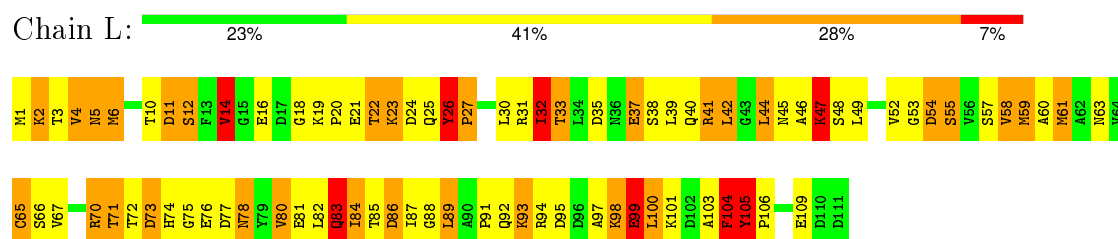
- Molecule 1: gp10



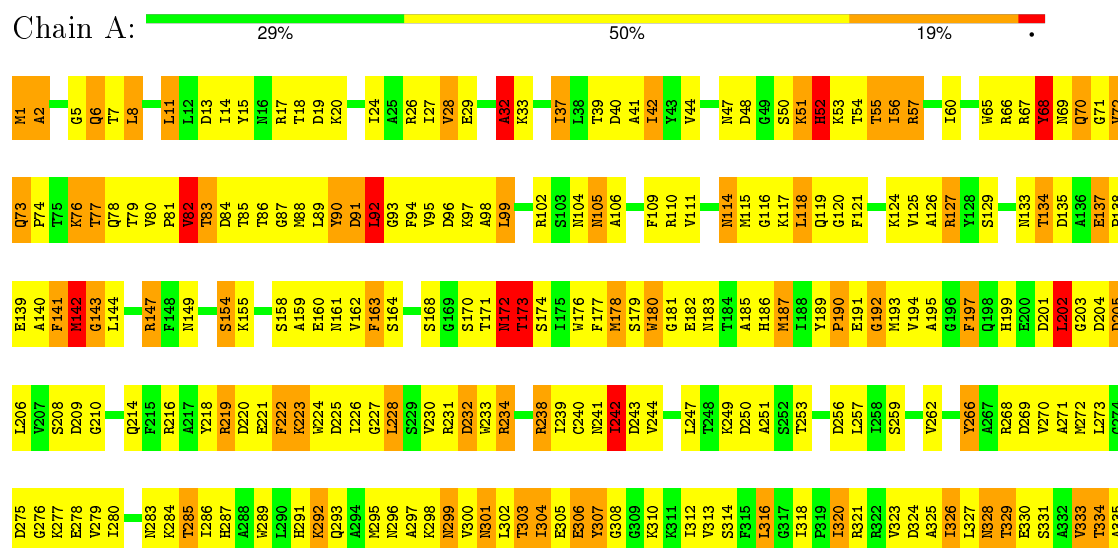
• Molecule 1: gp10



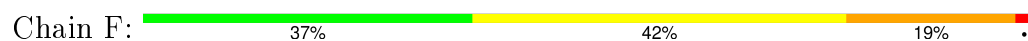
• Molecule 1: gp10

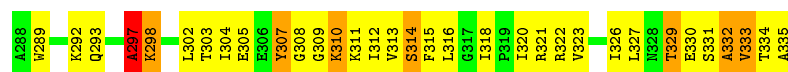


• Molecule 2: gp7



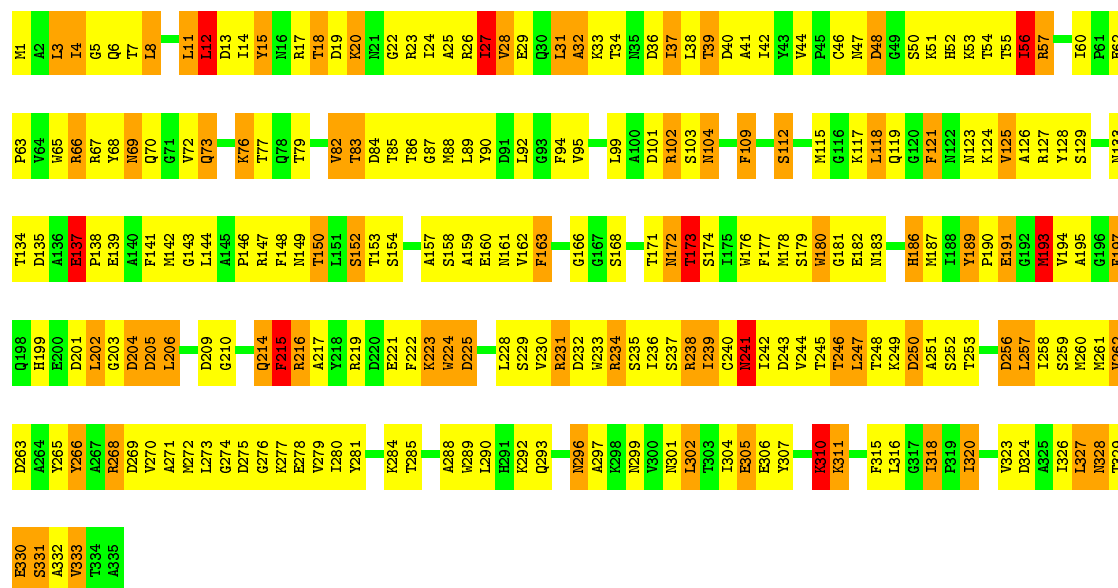
• Molecule 2: gp7





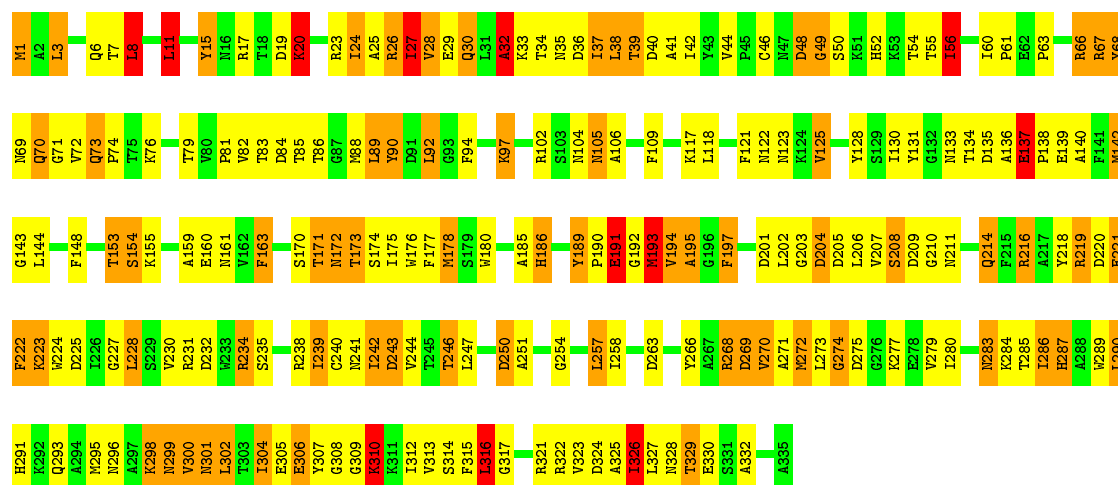
• Molecule 2: gp7

Chain D: 28% 49% 20%



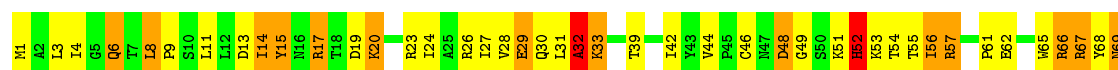
• Molecule 2: gp7

Chain C: 37% 39% 21%



• Molecule 2: gp7

Chain E: 37% 39% 22%



K284	T285	I286	H287	A288	W289	L290	H291	R292	Q293	N296	A297	K298	N299	V300	N301	L302	T303	I304	E305	E306	Y307	G308	G309	K310	F315	I316	G317	I318	F319	I320	R321	R322	I326	L327	N328	T329	E330	S331	A332	V333	T334	A335														
T215	L151	S154	K155	A156	A157	S158	A159	E160	N161	V162	F163	S164	G169	S170	T171	N172	T173	S174	W175	W176	F177	W178	S179	W180	G181	E182	N183	T184	A185	H186	M187	I188	Y189	P190	E191	G192	M193	V194	F197	Q198	H199	E200	D201	L202	G203	D204	D205	L206	V207	S208	D209	G210	N211	G212	G213	Q214
R215	R216	A217	Y218	R219	F222	K223	W224	D225	L226	G227	L228	S229	W230	R231	D232	W233	R234	S235	I236	S237	R238	I239	C240	R241	D242	D243	T246	L247	T248	K249	D250	A251	S252	T253	D256	L257	I258	S259	W260	D263	A264	D269	Y270	A271	W272	L273	G274	D275	G276	K277	E278	V279	I280			
Q70	G71	W72	Q73	P74	T77	Q78	T79	V82	T83	M88	L89	Y90	F94	V95	D96	R97	L99	F100	D101	R102	S103	N104	N105	A106	E113	K117	L118	Q119	G120	F121	N122	N123	K124	R127	Y128	S129	I130	Y131	T134	D135	A136	E137	P138	E139	A140	F141	R147	F148	N149							

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, 1	Depositor
Number of images	14000	Depositor
Resolution determination method	Gold Standard Definition (FSC at 0.143 cut-off) with two independent maps	Depositor
CTF correction method	per particle	Depositor
Microscope	JEOL 3200FSC	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	17	Depositor
Minimum defocus (nm)	400	Depositor
Maximum defocus (nm)	2700	Depositor
Magnification	50000	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	H	0.29	0/862	0.81	1/1166 (0.1%)
1	I	0.31	0/862	0.79	0/1166
1	J	0.31	0/862	0.81	1/1166 (0.1%)
1	K	0.31	0/862	0.81	1/1166 (0.1%)
1	L	0.32	0/862	0.76	0/1166
1	M	0.31	0/862	0.80	0/1166
1	N	0.30	0/862	0.75	1/1166 (0.1%)
2	A	0.31	0/2637	0.72	3/3573 (0.1%)
2	B	0.31	0/2637	0.69	1/3573 (0.0%)
2	C	0.30	0/2637	0.70	1/3573 (0.0%)
2	D	0.31	0/2637	0.71	1/3573 (0.0%)
2	E	0.31	0/2637	0.69	1/3573 (0.0%)
2	F	0.31	0/2637	0.69	0/3573
2	G	0.30	0/2637	0.69	1/3573 (0.0%)
All	All	0.31	0/24493	0.72	12/33173 (0.0%)

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	H	0	2
1	I	0	2
1	J	0	3
1	K	0	1
1	L	0	2
1	M	0	2
1	N	0	1
2	A	0	4
2	B	0	1
2	C	0	3
2	D	0	4

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
2	E	0	4
2	F	0	5
2	G	0	5
All	All	0	39

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	92	LEU	CA-CB-CG	6.04	129.19	115.30
2	G	99	LEU	CA-CB-CG	6.00	129.09	115.30
2	A	202	LEU	CA-CB-CG	5.63	128.25	115.30
2	B	316	LEU	CA-CB-CG	5.58	128.14	115.30
2	D	12	LEU	CA-CB-CG	5.55	128.06	115.30

There are no chirality outliers.

5 of 39 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	137	GLU	Peptide
2	A	32	ALA	Peptide
1	M	103	ALA	Peptide
1	M	77	ASP	Peptide
1	N	105	TYR	Peptide

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	H	850	0	820	78	0
1	I	850	0	820	83	0
1	J	850	0	820	74	0
1	K	850	0	820	84	0
1	L	850	0	820	68	0
1	M	850	0	820	74	0
1	N	850	0	820	82	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	2588	0	2542	194	0
2	B	2588	0	2542	181	0
2	C	2588	0	2542	169	0
2	D	2588	0	2542	211	0
2	E	2588	0	2542	180	0
2	F	2588	0	2542	195	0
2	G	2588	0	2542	146	0
All	All	24066	0	23534	1682	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 35.

The worst 5 of 1682 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:L:89:LEU:HG	1:L:91:PRO:HD3	1.48	0.94
2:D:101:ASP:HB2	2:D:104:ASN:HA	1.52	0.91
2:G:200:GLU:HG2	2:G:219:ARG:H	1.34	0.90
2:D:147:ARG:HH22	2:D:162:VAL:HG13	1.34	0.90
2:A:312:ILE:O	2:A:314:SER:N	2.06	0.89

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	H	109/111 (98%)	56 (51%)	38 (35%)	15 (14%)	0 6
1	I	109/111 (98%)	54 (50%)	34 (31%)	21 (19%)	0 3
1	J	109/111 (98%)	54 (50%)	33 (30%)	22 (20%)	0 2
1	K	109/111 (98%)	63 (58%)	27 (25%)	19 (17%)	0 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	109/111 (98%)	54 (50%)	31 (28%)	24 (22%)	0	1
1	M	109/111 (98%)	49 (45%)	34 (31%)	26 (24%)	0	1
1	N	109/111 (98%)	60 (55%)	34 (31%)	15 (14%)	0	6
2	A	333/335 (99%)	211 (63%)	82 (25%)	40 (12%)	0	9
2	B	333/335 (99%)	207 (62%)	85 (26%)	41 (12%)	0	8
2	C	333/335 (99%)	205 (62%)	80 (24%)	48 (14%)	0	6
2	D	333/335 (99%)	200 (60%)	92 (28%)	41 (12%)	0	8
2	E	333/335 (99%)	215 (65%)	79 (24%)	39 (12%)	0	9
2	F	333/335 (99%)	199 (60%)	93 (28%)	41 (12%)	0	8
2	G	333/335 (99%)	215 (65%)	87 (26%)	31 (9%)	1	16
All	All	3094/3122 (99%)	1842 (60%)	829 (27%)	423 (14%)	1	6

5 of 423 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	N	27	PRO
1	N	83	GLN
1	N	105	TYR
1	M	23	LYS
1	M	27	PRO

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	H	95/95 (100%)	70 (74%)	25 (26%)	0	6
1	I	95/95 (100%)	64 (67%)	31 (33%)	0	3
1	J	95/95 (100%)	64 (67%)	31 (33%)	0	3
1	K	95/95 (100%)	70 (74%)	25 (26%)	0	6
1	L	95/95 (100%)	62 (65%)	33 (35%)	0	2
1	M	95/95 (100%)	67 (70%)	28 (30%)	0	4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	N	95/95 (100%)	72 (76%)	23 (24%)	1	7
2	A	274/274 (100%)	195 (71%)	79 (29%)	0	4
2	B	274/274 (100%)	193 (70%)	81 (30%)	0	4
2	C	274/274 (100%)	199 (73%)	75 (27%)	0	5
2	D	274/274 (100%)	200 (73%)	74 (27%)	0	5
2	E	274/274 (100%)	198 (72%)	76 (28%)	0	4
2	F	274/274 (100%)	205 (75%)	69 (25%)	1	7
2	G	274/274 (100%)	209 (76%)	65 (24%)	1	8
All	All	2583/2583 (100%)	1868 (72%)	715 (28%)	2	4

5 of 715 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	90	TYR
2	D	150	THR
1	J	26	TYR
2	G	139	GLU
2	G	287	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	293	GLN
2	G	293	GLN
1	I	25	GLN
2	B	301	ASN
1	H	40	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 [i](#)

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.