



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 20, 2017 – 07:15 AM EDT

PDB ID : 4BZK
EMDB ID: : EMD-2431
Title : The structure of the COPII coat assembled on membranes
Authors : Zanetti, G.; Prinz, S.; Daum, S.; Meister, A.; Schekman, R.; Bacia, K.; Briggs, J.A.G.
Deposited on : unknown
Resolution : 40.00 Å(reported)
Based on PDB ID : 2PM9, 2PM6

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

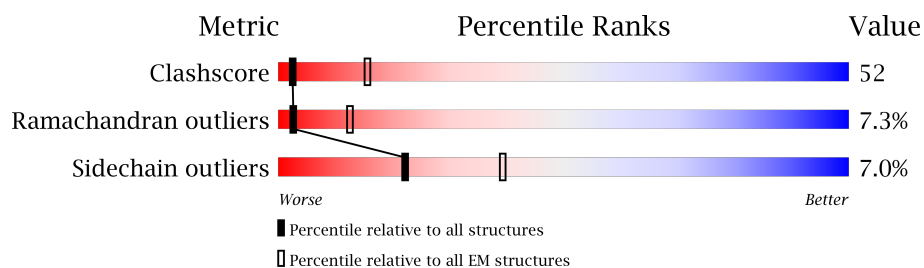
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 40.00 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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	1273	
1	C	1273	
2	B	297	
3	F	297	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 15423 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 PROTEIN TRANSPORT PROTEIN SEC31.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	691	Total	C	N	O	S	0	0
			5410	3406	908	1084	12		
1	C	693	Total	C	N	O	S	0	0
			5427	3417	911	1087	12		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	367	SER	THR	conflict	UNP P38968
C	367	SER	THR	conflict	UNP P38968

- Molecule 2 is a protein called PROTEIN TRANSPORT PROTEIN SEC13.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	279	Total	C	N	O	S	0	0
			2196	1403	375	415	3		

- Molecule 3 is a protein called PROTEIN TRANSPORT PROTEIN SEC13.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	F	280	Total	C	N	O	S	0	0
			2205	1402	376	418	9		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	11	MET	LEU	conflict	UNP Q04491
F	17	MET	LEU	conflict	UNP Q04491
F	24	MET	LEU	conflict	UNP Q04491
F	80	MET	LEU	conflict	UNP Q04491
F	115	MET	LEU	conflict	UNP Q04491
F	222	MET	LEU	conflict	UNP Q04491

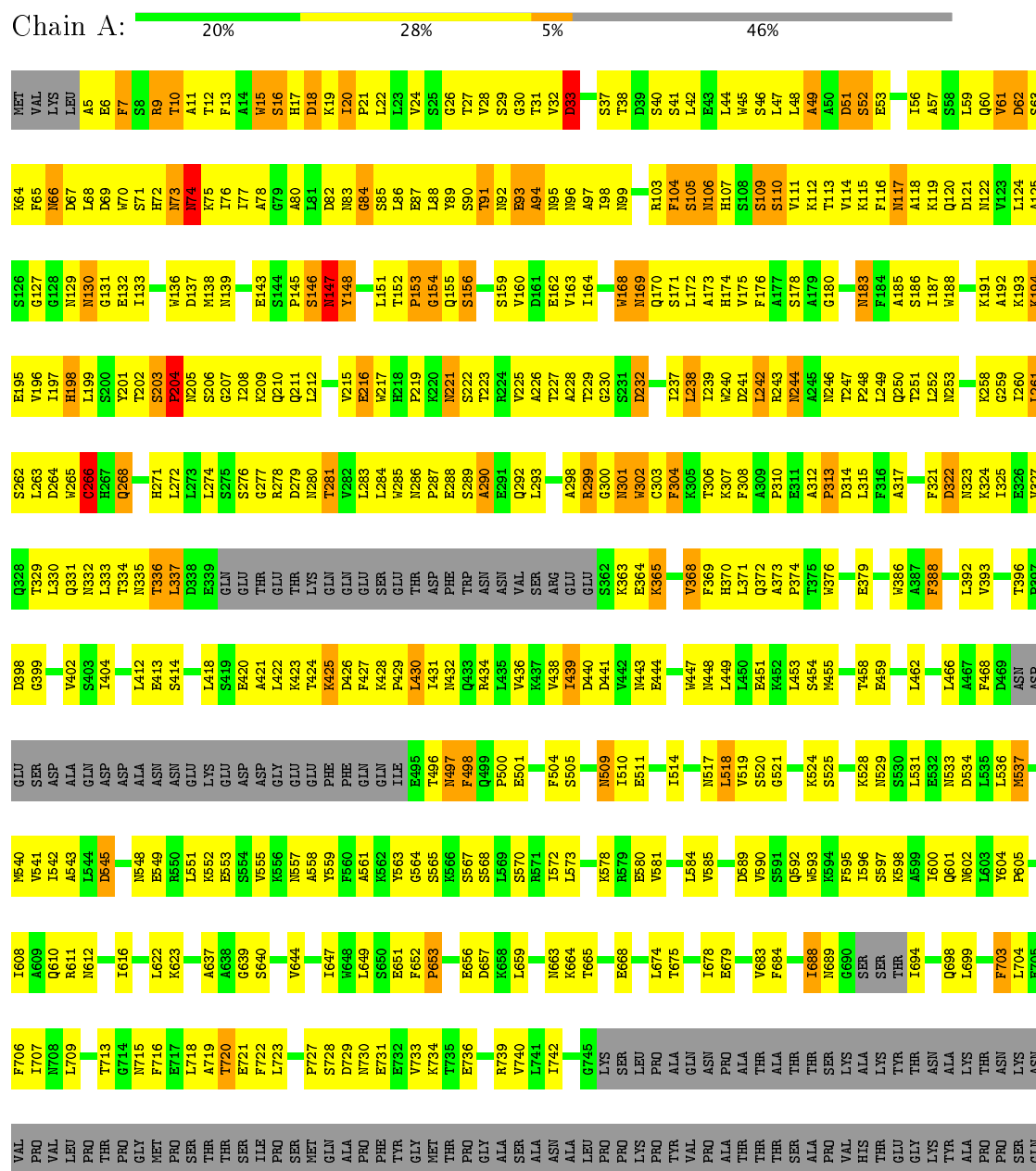
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		AltConf
4	A	45	Total 45	O 45	0
4	B	61	Total 61	O 61	0
4	C	68	Total 68	O 68	0
4	F	11	Total 11	O 11	0

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 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.

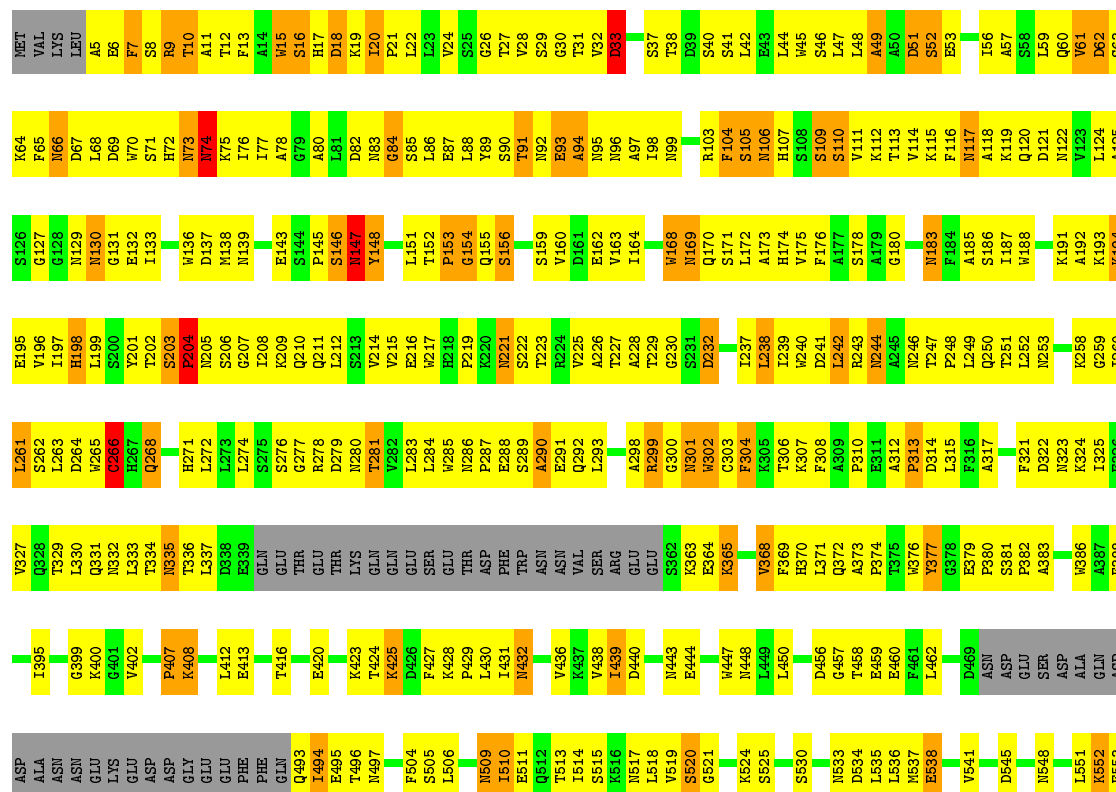
• Molecule 1: PROTEIN TRANSPORT PROTEIN SEC31

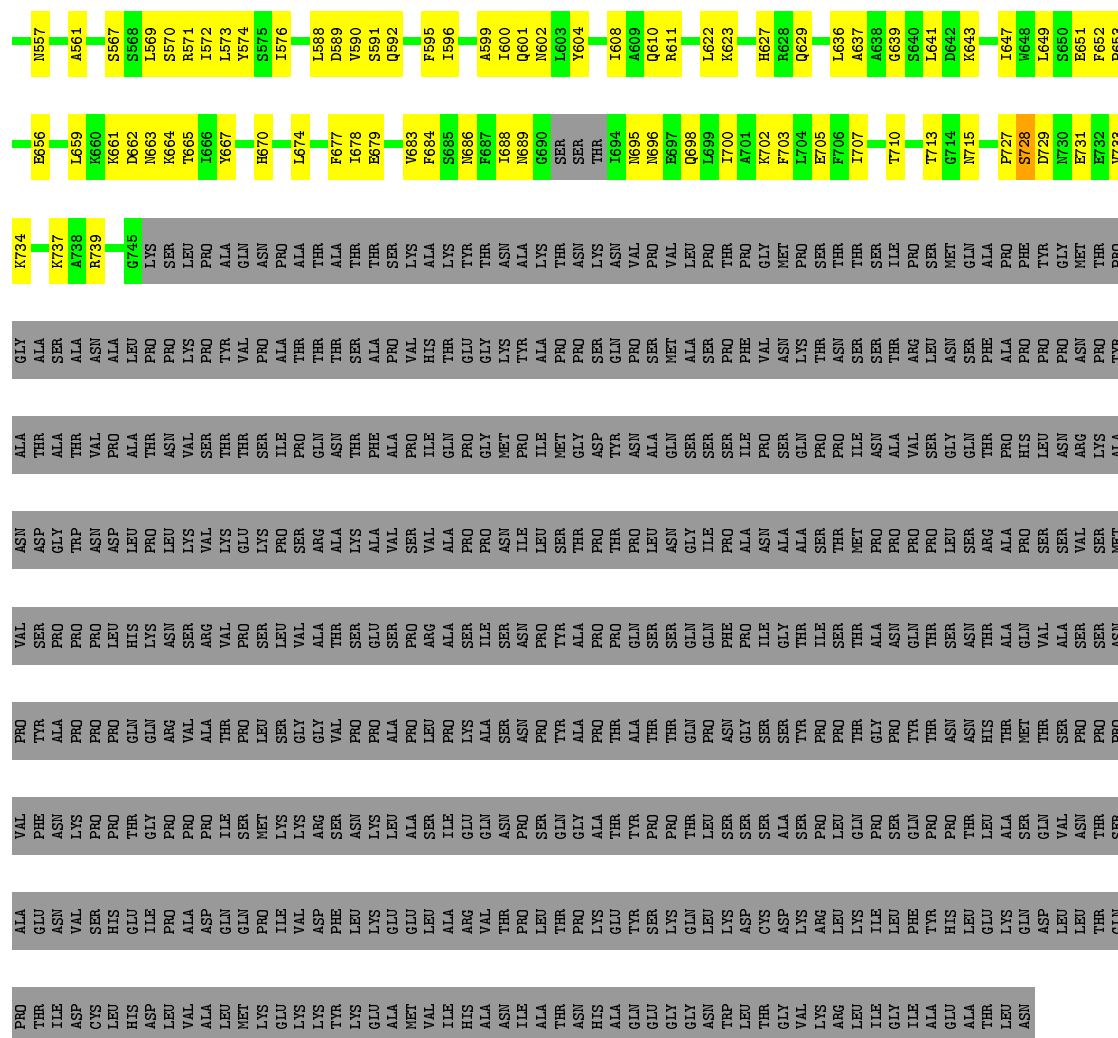


GLN	GLN	TYR	TYR	ALA	GLN	PRO	ASN	PRO	ASN	PRO
GLU	GLU	SER	SER	THR	SER	LEU	LEU	SER	THR	THR
GLY	GLY	LYS	PRO	THR	SER	ASN	ASN	ASN	GLN	GLN
ASN	ASN	GLN	THR	PRO	GLN	GLY	SER	ILE	SER	SER
TRP	TRP	LEU	SER	ASN	PHE	PRO	PRO	PRO	SER	SER
LEU	LEU	SER	SER	GLY	PRO	ALA	ALA	ALA	ILE	VAL
THR	THR	CYS	CYS	SER	ILE	ASN	ASN	PRO	PRO	ASN
GLY	GLY	ASP	ASP	ALA	SER	GLY	ALA	ALA	GLN	LYS
VAL	VAL	LYS	LYS	TYR	TYR	PRO	SER	ALA	PRO	THR
LYS	LYS	ARG	PRO	PRO	ILE	SER	THR	SER	PRO	ASN
ARG	ARG	LEU	LEU	PRO	SER	THR	THR	THR	PRO	SER
LEU	LEU	LYS	GLN	THR	THR	THR	ASN	SER	ILE	SER
ILE	ILE	PHE	GLN	TYR	GLN	PRO	VAL	GLY	GLN	ASN
ALA	ALA	TYR	PRO	ASN	SER	LEU	SER	LEU	SER	LEU
ALA	ALA	HIS	THR	ASN	ASN	THR	ASN	THR	GLN	ASN
THR	THR	GLU	LEU	HIS	THR	ALA	ALA	ARG	PRO	SER
LEU	LEU	LYS	ALA	THR	ALA	ALA	ALA	THR	ILE	THR
LEU	LEU	GLN	SER	MET	GLN	PRO	HIS	LEU	PRO	PRO
ASN	ASN	ASP	GLN	THR	VAL	SER	LEU	PRO	ASN	THR
		LEU	VAL	SER	ALA	SER	ASN	LEU	ASN	GLY
		LEU	ASN	PRO	THR	VAL	ARG	LEU	GLY	THR
		THR	THR	PRO	THR	SER	SER	VAL	THR	THR
		GLN	SER	PRO	GLN	ASN	LYS	VAL	ASN	ASN
		THR	ALA	VAL	THR	VAL	ASN	THR	LYS	THR
		ILE	GLU	PHE	TYR	SER	ASP	GLY	ALA	PHE
		ASP	THR	THR	ALA	PRO	GLY	PRO	VAL	ALA
		ASP	CYS	LYS	LYS	PRO	LYS	SER	SER	ALA
		LEU	HIS	GLU	GLN	HIS	LEU	THR	ALA	THR
		ASP	ILE	GLY	GLN	THR	LYS	PRO	LYS	THR
		LEU	PRO	PRO	ARG	ASN	ASN	LEU	GLY	THR
		VAL	LYS	PRO	VAL	VAL	VAL	GLU	ALA	PHE
		LYS	LYS	LYS	THR	THR	THR	SER	VAL	ALA
		GLU	LYS	LYS	GLY	GLY	VAL	PRO	VAL	ALA
		ALA	VAL	LYS	THR	ALA	ARG	SER	SER	THR
		THR	THR	THR	THR	THR	THR	THR	THR	THR
		ASN	ASN	ASN	ALA	ILE	ALA	ALA	ALA	GLN
		THR	THR	THR	GLN	LYS	SER	PRO	PRO	ILE
		ASN	THR	THR	THR	ARG	PRO	PRO	THR	THR
		HIS	LYS	ALA	THR	VAL	ILE	ILE	PRO	GLY
		THR	THR	ASN	GLY	GLU	TYR	TYR	ASN	THR
		THR	THR	THR	THR	THR	THR	THR	THR	THR
		ASN	ASN	THR	THR	THR	THR	THR	THR	THR
		HIS	LYS	ALA	THR	VAL	ILE	ILE	PRO	ASP

- Molecule 1: PROTEIN TRANSPORT PROTEIN SEC31

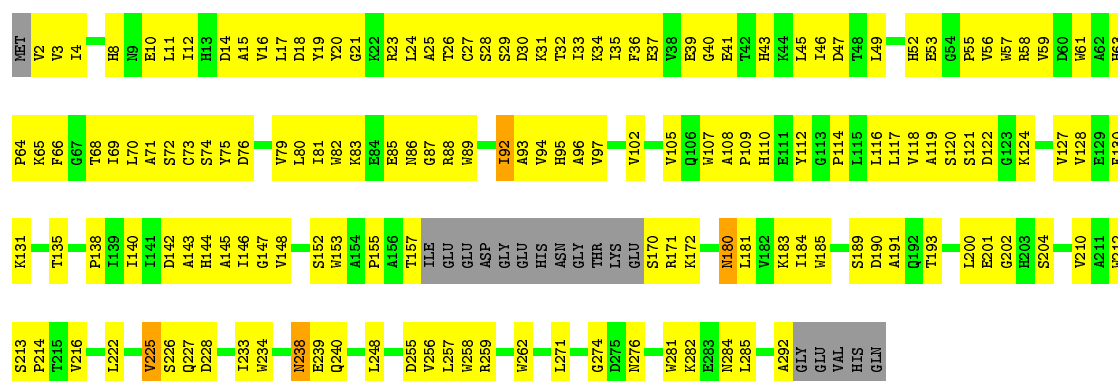
Chain C:





- Molecule 2: PROTEIN TRANSPORT PROTEIN SEC13

Chain B: 42% 51% 6%



- Molecule 3: PROTEIN TRANSPORT PROTEIN SEC13

Chain F:  23% 60% 11% 6%

G266	L196	V128	H63	Met
G267	L197	E129	F64	V2
	L198	F130	F65	V3
	L199	K131	F66	I4
A270	L200	E132	G67	A5
L271	E201	H133	T68	N6
S272		G134	L69	A7
G273	D205	T135	L70	
D274	V206	T136		E10
D275	R207		S71	M11
N276	R208	I141	A72	L12
G277	D209	D142	C73	H13
V278	V210	A143	S74	D14
T279	A211	H144	D76	A15
L280	N212	A145	G77	V16
K281	S213	I146	K78	M17
K282	P214	G147	V79	D18
E283	T215	V148	H80	I19
N284	V216	M149	I81	Y20
L285	L217	S150	H82	C21
E286	L218	A151	K83	C22
G287	R219	S152	E84	R23
	D220	H153	E85	N24
K288	S220	A154	H86	A25
V289	Y221	P155		T26
E290	N222	A156	H89	C27
P291	A223	T157		S28
A292	S224	I158	I92	S29
GLY	V225	I159	A93	D30
GLU	S226	GLU	V94	K31
VAL	Q227	GLU	H95	T32
HIS	D228	ASP	A96	I33
GLN	R229	GLY		K34
	D230	GLU	S99	I35
	C231	HIS		F36
	T232	ASN	A100	E37
	T233	GLY	S101	V38
	N234	THR	V102	E39
	T235	LVS	V105	G40
	Q236	E169	Q106	E41
N238	D237	S170	P107	T42
	E239	R171	A108	H43
	Q240	K172	P109	K44
	G241	F173	H110	L45
			E111	I46
		A178	Y112	D47
	L247	D179		T48
	L248	N180		L49
	K249	K183	H115	T50
	E250	I184	L116	G51
	E251	H185	V118	H52
		H186	A119	
P254	D255	K187	S120	P55
V256	T257	N188	S121	V56
N258	K258	A191	D122	H57
R259	S259	Q192	G123	R58
A260	A260	T193	K124	V59
S261	S261	Y194	V125	D60
D262	D262	H195	S126	H61
		V196	V127	A62

4 Experimental information

Property	Value	Source
Reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of subtomograms used	192	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH TILTED IMAGE WITHIN TOMOGRAM	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	80	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3200	Depositor
Magnification	19500	Depositor
Image detector	GATAN MULTISCAN	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	0.41	0/5516	0.70	2/7479 (0.0%)
1	C	0.44	0/5533	0.71	2/7502 (0.0%)
2	B	0.39	0/2256	0.68	0/3079
3	F	0.41	0/2265	0.67	0/3085
All	All	0.42	0/15570	0.69	4/21145 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	154	GLY	N-CA-C	5.12	125.89	113.10
1	C	154	GLY	N-CA-C	5.10	125.84	113.10
1	C	156	SER	N-CA-C	5.08	124.72	111.00
1	A	156	SER	N-CA-C	5.08	124.72	111.00

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	A	5410	0	5272	604	0
1	C	5427	0	5289	604	0
2	B	2196	0	2138	186	0
3	F	2205	0	2131	304	0
4	A	45	0	0	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	61	0	0	25	0
4	C	68	0	0	25	0
4	F	11	0	0	4	0
All	All	15423	0	14830	1578	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 52.

The worst 5 of 1578 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:C:739:ARG:HA	3:F:20:TYR:CE1	1.28	1.64
1:C:664:LYS:C	3:F:285:LEU:HD21	1.30	1.51
1:C:739:ARG:CA	3:F:20:TYR:HE1	1.28	1.46
1:A:314:ASP:OD2	1:A:376:TRP:CD2	1.64	1.46
1:A:314:ASP:OD2	1:A:376:TRP:CE2	1.76	1.36

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	681/1273 (54%)	519 (76%)	106 (16%)	56 (8%)	1	16
1	C	683/1273 (54%)	532 (78%)	97 (14%)	54 (8%)	1	17
2	B	275/297 (93%)	241 (88%)	30 (11%)	4 (2%)	12	53
3	F	276/297 (93%)	214 (78%)	37 (13%)	25 (9%)	1	15
All	All	1915/3140 (61%)	1506 (79%)	270 (14%)	139 (7%)	2	19

5 of 139 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	33	ASP
1	A	49	ALA
1	A	74	ASN
1	A	110	SER
1	A	143	GLU

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	608/1113 (55%)	562 (92%)	46 (8%)	15	47
1	C	610/1113 (55%)	562 (92%)	48 (8%)	14	45
2	B	237/252 (94%)	232 (98%)	5 (2%)	59	80
3	F	238/252 (94%)	218 (92%)	20 (8%)	13	43
All	All	1693/2730 (62%)	1574 (93%)	119 (7%)	22	50

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	18	ASP
1	C	130	ASN
3	F	200	LEU
1	C	33	ASP
1	C	73	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	180	ASN
1	C	74	ASN
1	C	606	ASN
2	B	238	ASN
1	C	66	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 [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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	372:GLN	C	373:ALA	N	16.58
1	C	372:GLN	C	373:ALA	N	16.33