



# wwPDB EM Validation Summary Report ⓘ

Apr 29, 2024 – 07:10 am BST

PDB ID : 4BZI  
EMDB ID : EMD-2428  
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 : 2013-07-26  
Resolution : 23.00 Å(reported)  
Based on initial model : 1M2O

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : 0.0.1.dev92  
Mogul : 1.8.4, CSD as541be (2020)  
MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36.2

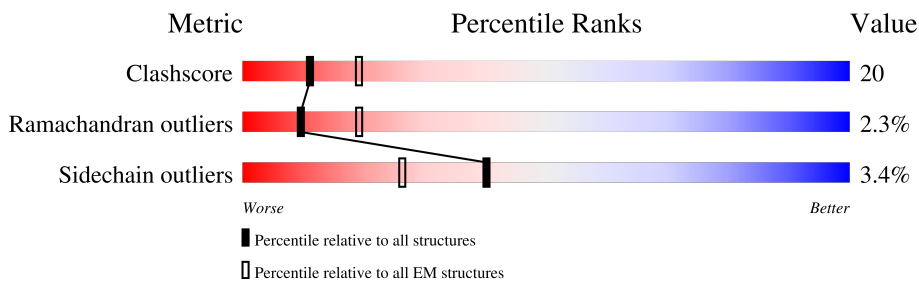
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 23.00 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.



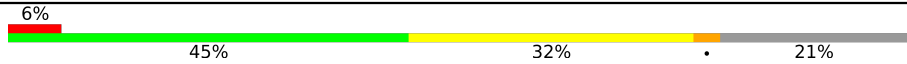
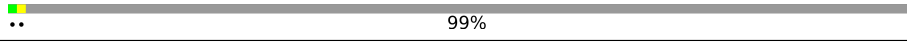
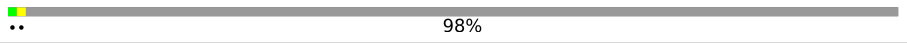
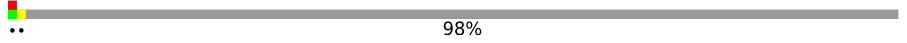
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion  $< 40\%$ ). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	768	
1	D	768	
1	G	768	
2	B	190	
2	J	190	
2	K	190	
3	E	926	
3	L	926	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	M	926	
4	F	926	
4	N	926	
4	O	926	

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 39154 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 SEC23P.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	733	5780	3679	969	1109	23	0	0
1	D	733	5780	3679	969	1109	23	0	0
1	G	733	5780	3679	969	1109	23	0	0

- Molecule 2 is a protein called SAR1P.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	164	1299	836	220	239	4	0	0
2	J	164	1299	836	220	239	4	0	0
2	K	164	1299	836	220	239	4	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	139	ALA	GLY	conflict	UNP C8ZIG2
J	139	ALA	GLY	conflict	UNP C8ZIG2
K	139	ALA	GLY	conflict	UNP C8ZIG2

- Molecule 3 is a protein called SEC24P.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	E	735	5823	3702	999	1084	38	0	0
3	L	735	5823	3702	999	1084	38	0	0
3	M	735	5823	3702	999	1084	38	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	408	THR	ALA	conflict	UNP C8ZAD6
E	865	VAL	ALA	conflict	UNP C8ZAD6
L	408	THR	ALA	conflict	UNP C8ZAD6
L	865	VAL	ALA	conflict	UNP C8ZAD6
M	408	THR	ALA	conflict	UNP C8ZAD6
M	865	VAL	ALA	conflict	UNP C8ZAD6

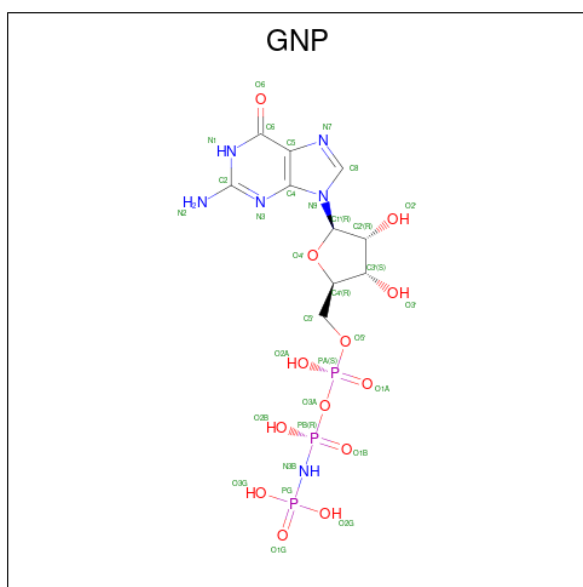
- Molecule 4 is a protein called SEC24P.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	F	13	Total	C	N	O	0	0
			109	70	19	20		
4	N	14	Total	C	N	O	0	0
			117	74	21	22		
4	O	14	Total	C	N	O	0	0
			117	74	21	22		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
5	A	1	Total	Zn	0
			1	1	
5	D	1	Total	Zn	0
			1	1	
5	E	1	Total	Zn	0
			1	1	
5	G	1	Total	Zn	0
			1	1	
5	L	1	Total	Zn	0
			1	1	
5	M	1	Total	Zn	0
			1	1	

- Molecule 6 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter code: GNP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
6	B	1	32	10	6	13	3	0
6	J	1	32	10	6	13	3	0
6	K	1	32	10	6	13	3	0

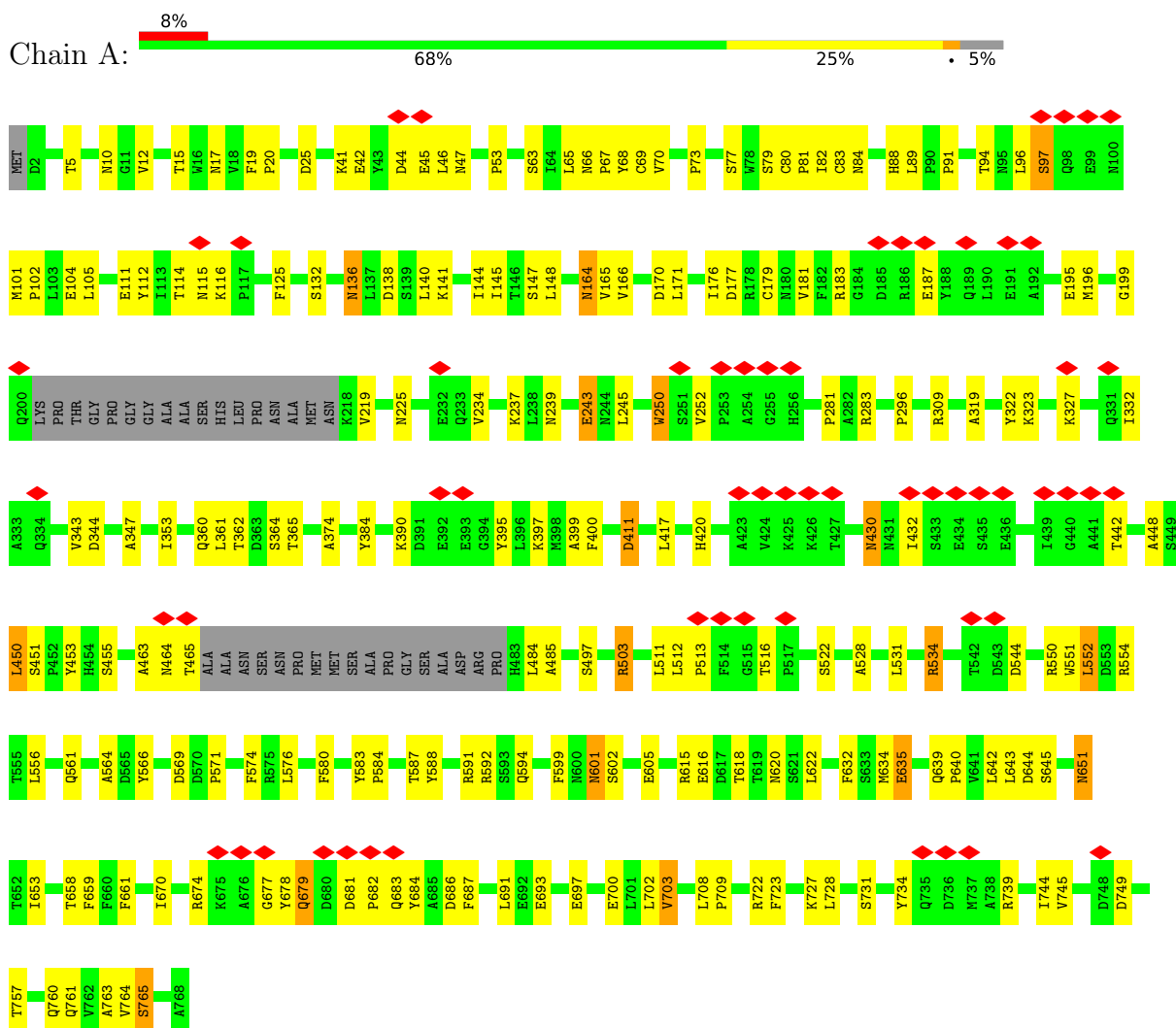
- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
7	B	1	1	1	0
7	J	1	1	1	0
7	K	1	1	1	0

### 3 Residue-property plots [i](#)

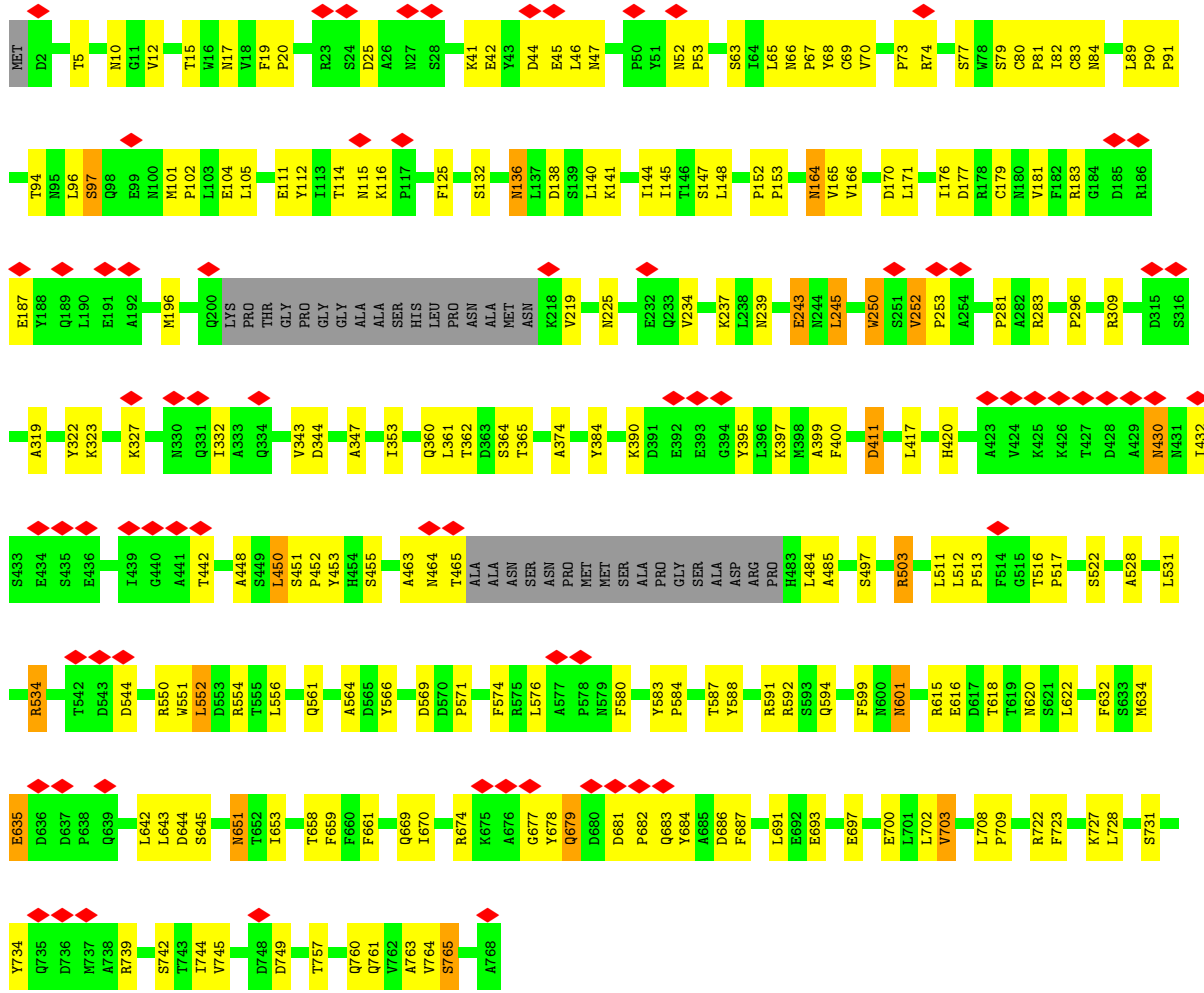
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. 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. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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: SEC23P

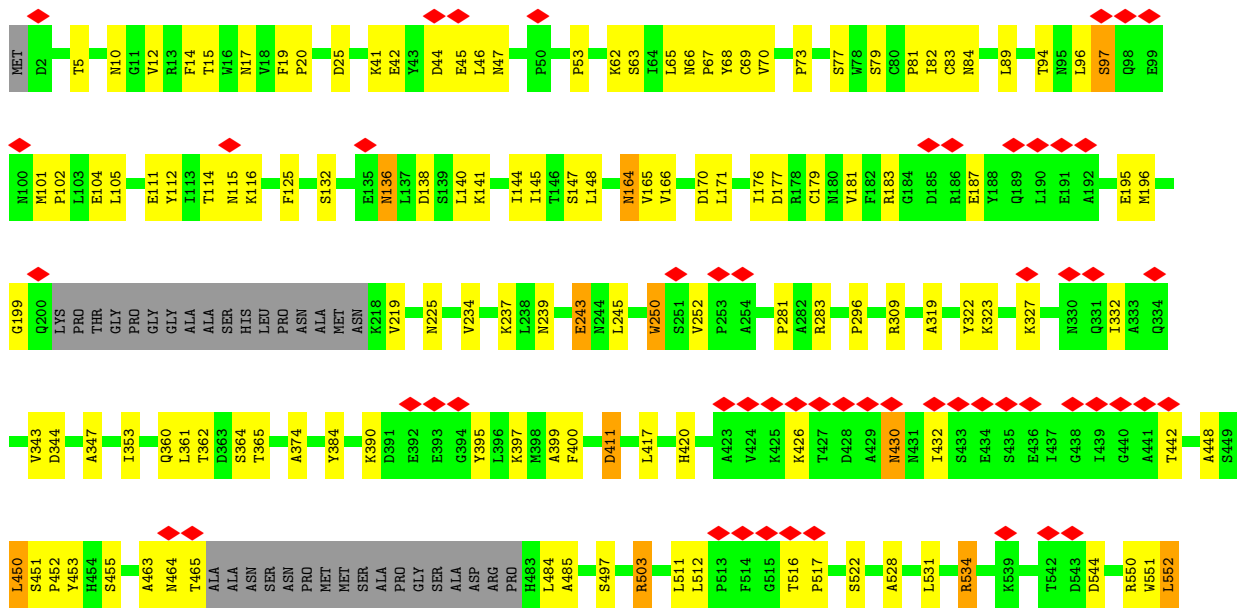


- Molecule 1: SEC23P

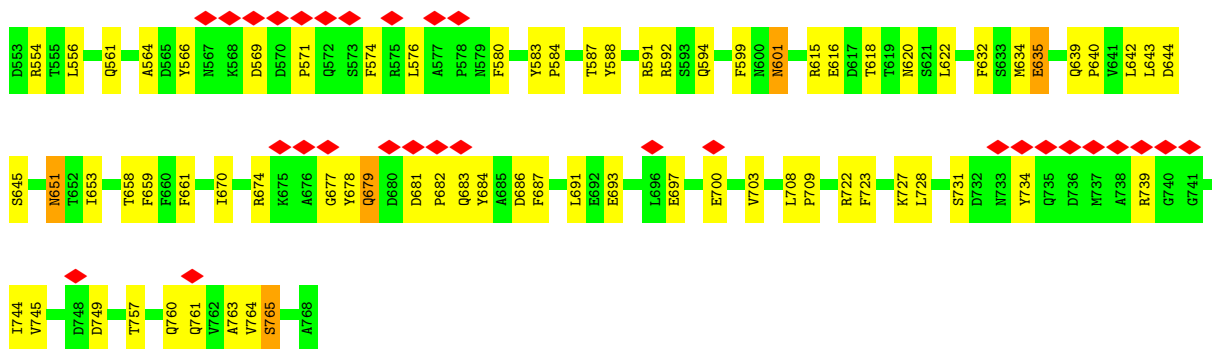




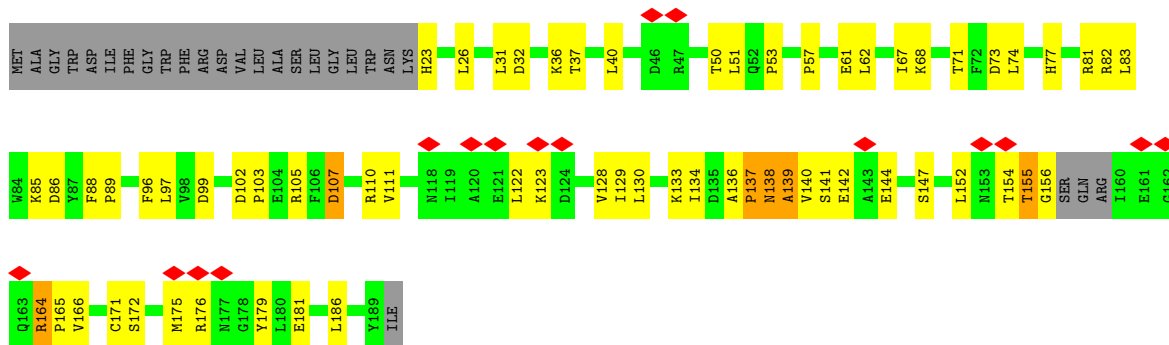
• Molecule 1: SEC23P



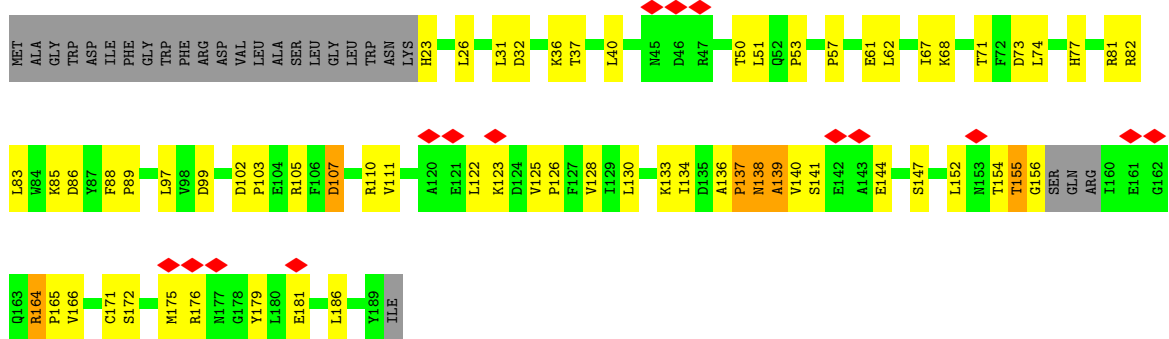




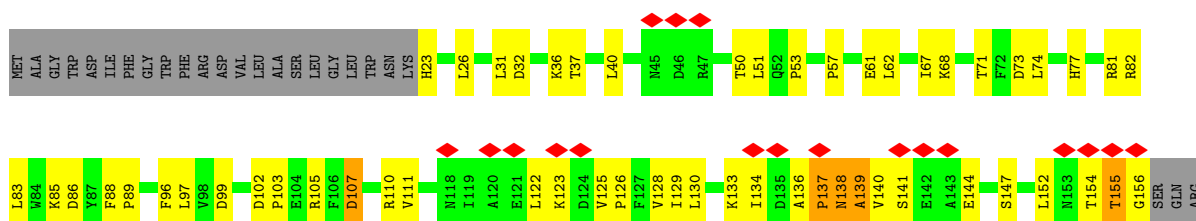
• Molecule 2: SAR1P



• Molecule 2: SAR1P

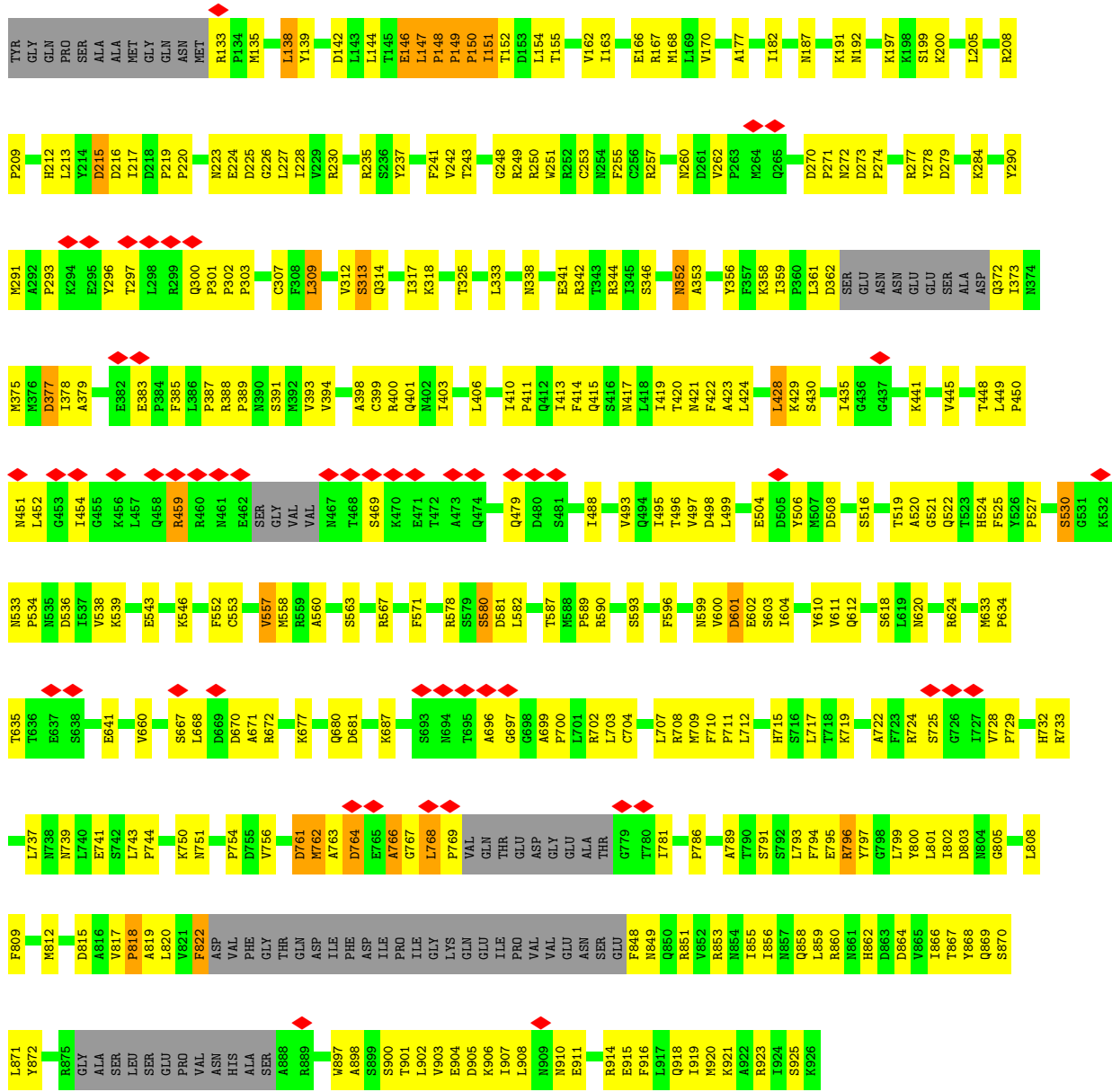


• Molecule 2: SAR1P





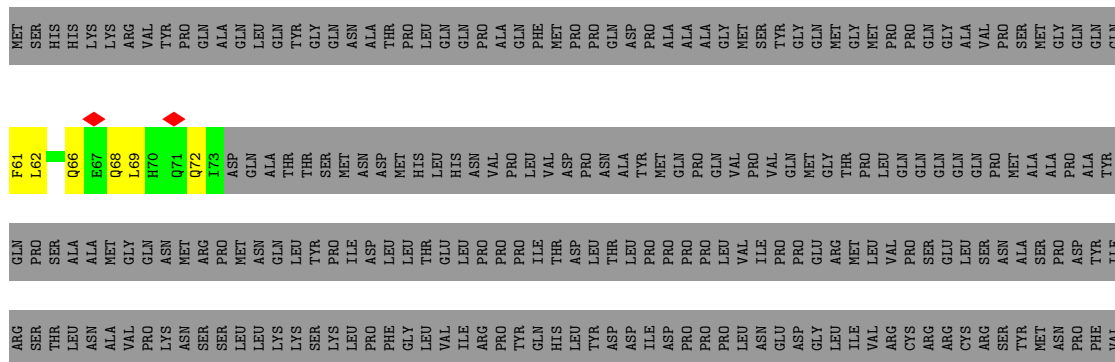




● Molecule 4: SEC24P

Chain F: ..

99%









## 4 Experimental information

Property	Value	Source
EM reconstruction method	TOMOGRAPHY	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of tilted images used	15000	Depositor
Resolution determination method	Not provided	
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
Maximum voxel value	3.686	Depositor
Minimum voxel value	-2.502	Depositor
Average voxel value	0.000	Depositor
Voxel value standard deviation	1.000	Depositor
Recommended contour level	1.4	Depositor
Tomogram size (Å)	275.2, 275.2, 275.2	wwPDB
Tomogram dimensions	64, 64, 64	wwPDB
Tomogram angles (°)	90.0, 90.0, 90.0	wwPDB
Grid spacing (Å)	4.3, 4.3, 4.3	Depositor



## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GNP, ZN, MG

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	A	0.38	0/5915	0.62	0/8052
1	D	0.38	0/5915	0.62	0/8052
1	G	0.38	0/5915	0.62	0/8052
2	B	0.34	0/1327	0.61	0/1800
2	J	0.34	0/1327	0.61	0/1800
2	K	0.34	0/1327	0.61	0/1800
3	E	0.39	0/5943	0.68	1/8064 (0.0%)
3	L	0.39	0/5943	0.68	1/8064 (0.0%)
3	M	0.39	0/5943	0.68	1/8064 (0.0%)
4	F	0.35	0/111	0.47	0/150
4	N	0.35	0/118	0.47	0/158
4	O	0.35	0/118	0.48	0/158
All	All	0.38	0/39902	0.64	3/54214 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	768	LEU	CA-CB-CG	5.33	127.56	115.30
3	M	768	LEU	CA-CB-CG	5.33	127.56	115.30
3	L	768	LEU	CA-CB-CG	5.32	127.55	115.30

There are no chirality outliers.

There are no planarity outliers.

### 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	5780	0	5663	201	0
1	D	5780	0	5663	205	0
1	G	5780	0	5663	193	0
2	B	1299	0	1288	65	0
2	J	1299	0	1288	59	0
2	K	1299	0	1288	61	0
3	E	5823	0	5853	291	0
3	L	5823	0	5853	293	0
3	M	5823	0	5853	301	0
4	F	109	0	105	14	0
4	N	117	0	110	14	0
4	O	117	0	110	12	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
5	L	1	0	0	0	0
5	M	1	0	0	0	0
6	B	32	0	13	5	0
6	J	32	0	13	4	0
6	K	32	0	13	5	0
7	B	1	0	0	0	0
7	J	1	0	0	0	0
7	K	1	0	0	0	0
All	All	39154	0	38776	1589	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 20.

The worst 5 of 1589 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:91:PRO:HB3	1:D:669:GLN:NE2	1.18	1.43
1:A:91:PRO:CB	1:D:669:GLN:HE21	1.35	1.38
1:D:183:ARG:NH1	3:L:383:GLU:HB2	1.53	1.22
1:G:183:ARG:NH1	3:M:383:GLU:HB2	1.53	1.22
1:A:183:ARG:NH1	3:E:383:GLU:HB2	1.53	1.21

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	727/768 (95%)	687 (94%)	35 (5%)	5 (1%)	22	63
1	D	727/768 (95%)	688 (95%)	34 (5%)	5 (1%)	22	63
1	G	727/768 (95%)	688 (95%)	34 (5%)	5 (1%)	22	63
2	B	160/190 (84%)	142 (89%)	13 (8%)	5 (3%)	4	27
2	J	160/190 (84%)	142 (89%)	13 (8%)	5 (3%)	4	27
2	K	160/190 (84%)	142 (89%)	13 (8%)	5 (3%)	4	27
3	E	723/926 (78%)	640 (88%)	55 (8%)	28 (4%)	3	23
3	L	723/926 (78%)	641 (89%)	54 (8%)	28 (4%)	3	23
3	M	723/926 (78%)	641 (89%)	54 (8%)	28 (4%)	3	23
4	F	11/926 (1%)	10 (91%)	1 (9%)	0	100	100
4	N	11/926 (1%)	10 (91%)	1 (9%)	0	100	100
4	O	11/926 (1%)	10 (91%)	1 (9%)	0	100	100
All	All	4863/8430 (58%)	4441 (91%)	308 (6%)	114 (2%)	9	34

5 of 114 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	703	VAL
2	B	139	ALA
2	B	155	THR
1	D	703	VAL
3	E	147	LEU

### 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	644/668 (96%)	616 (96%)	28 (4%)	29	53
1	D	644/668 (96%)	616 (96%)	28 (4%)	29	53
1	G	644/668 (96%)	616 (96%)	28 (4%)	29	53
2	B	138/159 (87%)	136 (99%)	2 (1%)	67	80
2	J	138/159 (87%)	136 (99%)	2 (1%)	67	80
2	K	138/159 (87%)	136 (99%)	2 (1%)	67	80
3	E	660/819 (81%)	640 (97%)	20 (3%)	41	63
3	L	660/819 (81%)	640 (97%)	20 (3%)	41	63
3	M	660/819 (81%)	640 (97%)	20 (3%)	41	63
4	F	12/817 (2%)	12 (100%)	0	100	100
4	N	13/817 (2%)	13 (100%)	0	100	100
4	O	13/817 (2%)	13 (100%)	0	100	100
All	All	4364/7389 (59%)	4214 (97%)	150 (3%)	40	60

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	213	LEU
3	M	739	ASN
3	L	428	LEU
3	M	138	LEU
1	D	592	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 97 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	464	ASN
3	L	374	ASN
1	G	601	ASN
2	J	66	ASN
3	L	524	HIS

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 9 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GNP	J	1190	7	29,34,34	1.85	7 (24%)	33,54,54	2.31	6 (18%)
6	GNP	K	1190	7	29,34,34	1.85	7 (24%)	33,54,54	2.31	6 (18%)
6	GNP	B	1190	7	29,34,34	1.85	7 (24%)	33,54,54	2.31	6 (18%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GNP	J	1190	7	-	4/14/38/38	0/3/3/3
6	GNP	K	1190	7	-	4/14/38/38	0/3/3/3
6	GNP	B	1190	7	-	4/14/38/38	0/3/3/3

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1190	GNP	C6-N1	4.60	1.41	1.33
6	K	1190	GNP	C6-N1	4.59	1.41	1.33
6	J	1190	GNP	C6-N1	4.56	1.41	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1190	GNP	PB-O2B	-4.16	1.45	1.56
6	J	1190	GNP	PB-O2B	-4.14	1.45	1.56

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	1190	GNP	C5-C6-N1	-8.39	111.95	123.43
6	K	1190	GNP	C5-C6-N1	-8.38	111.98	123.43
6	J	1190	GNP	C5-C6-N1	-8.36	111.99	123.43
6	B	1190	GNP	C2-N1-C6	6.19	125.77	115.93
6	K	1190	GNP	C2-N1-C6	6.18	125.75	115.93

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

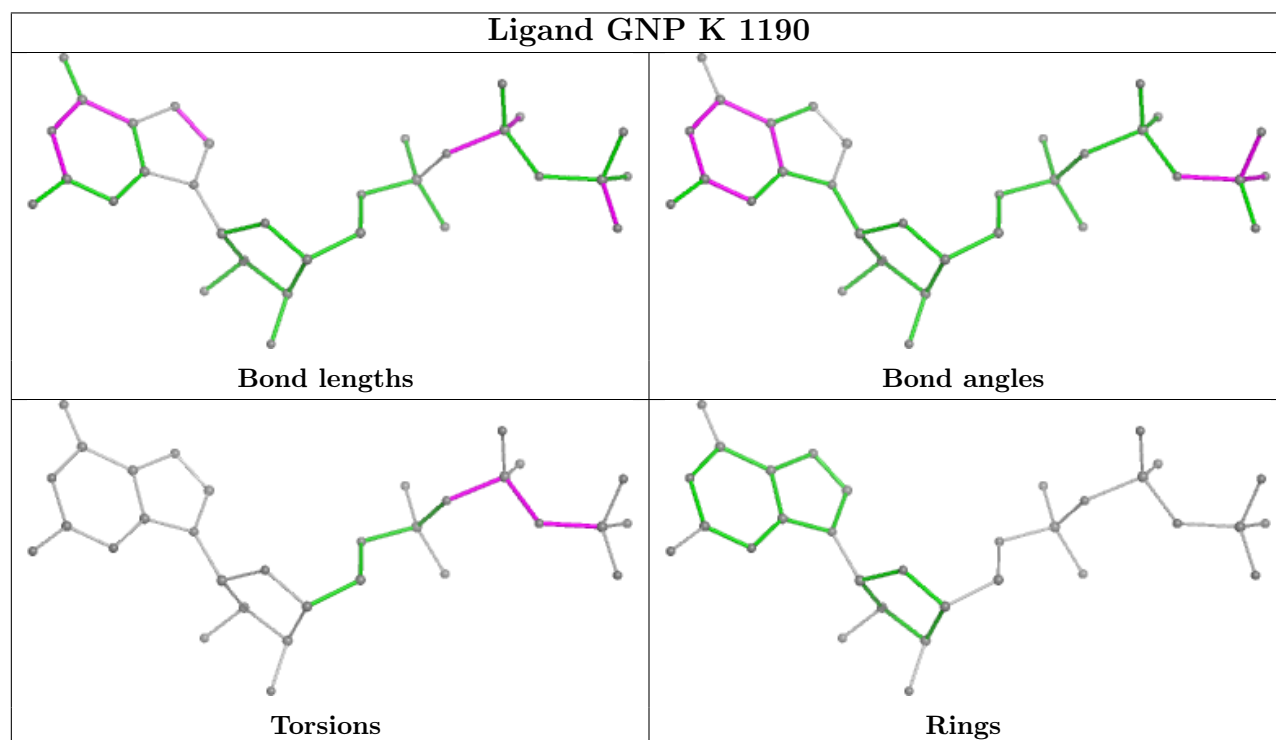
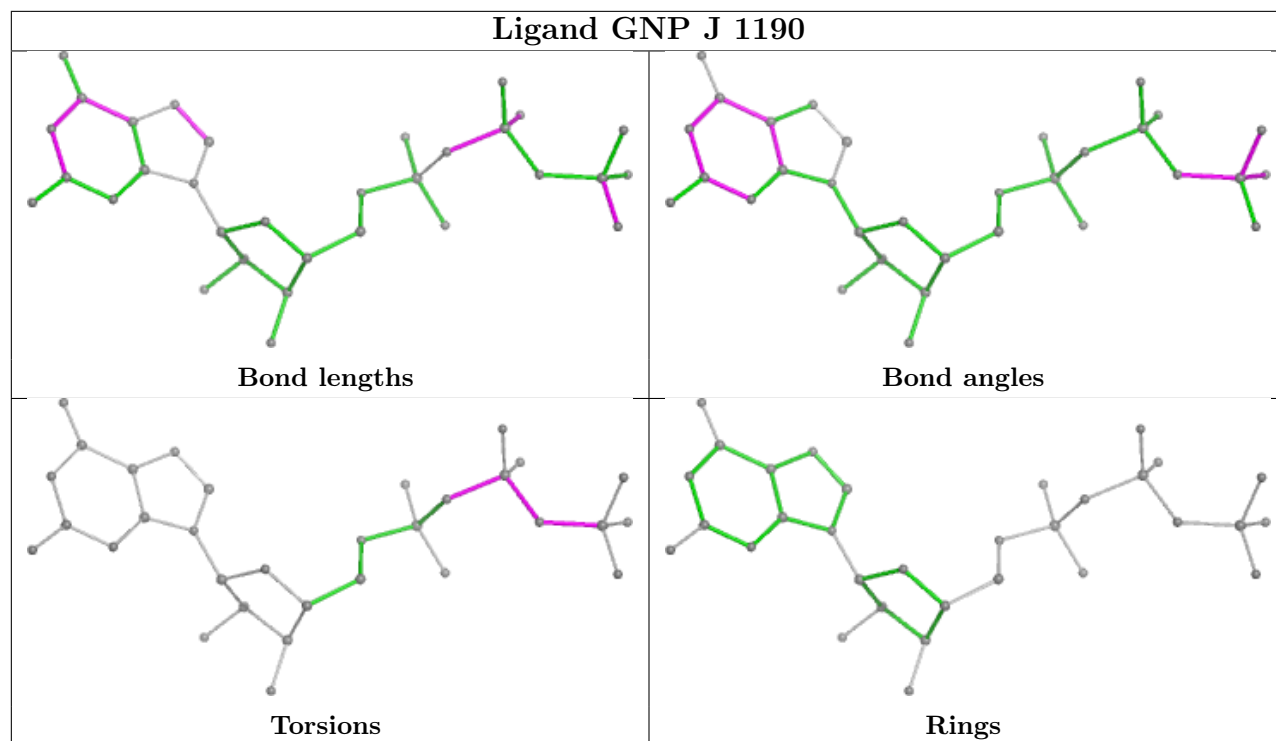
Mol	Chain	Res	Type	Atoms
6	B	1190	GNP	PB-N3B-PG-O1G
6	B	1190	GNP	PG-N3B-PB-O1B
6	B	1190	GNP	PA-O3A-PB-O1B
6	J	1190	GNP	PB-N3B-PG-O1G
6	J	1190	GNP	PG-N3B-PB-O1B

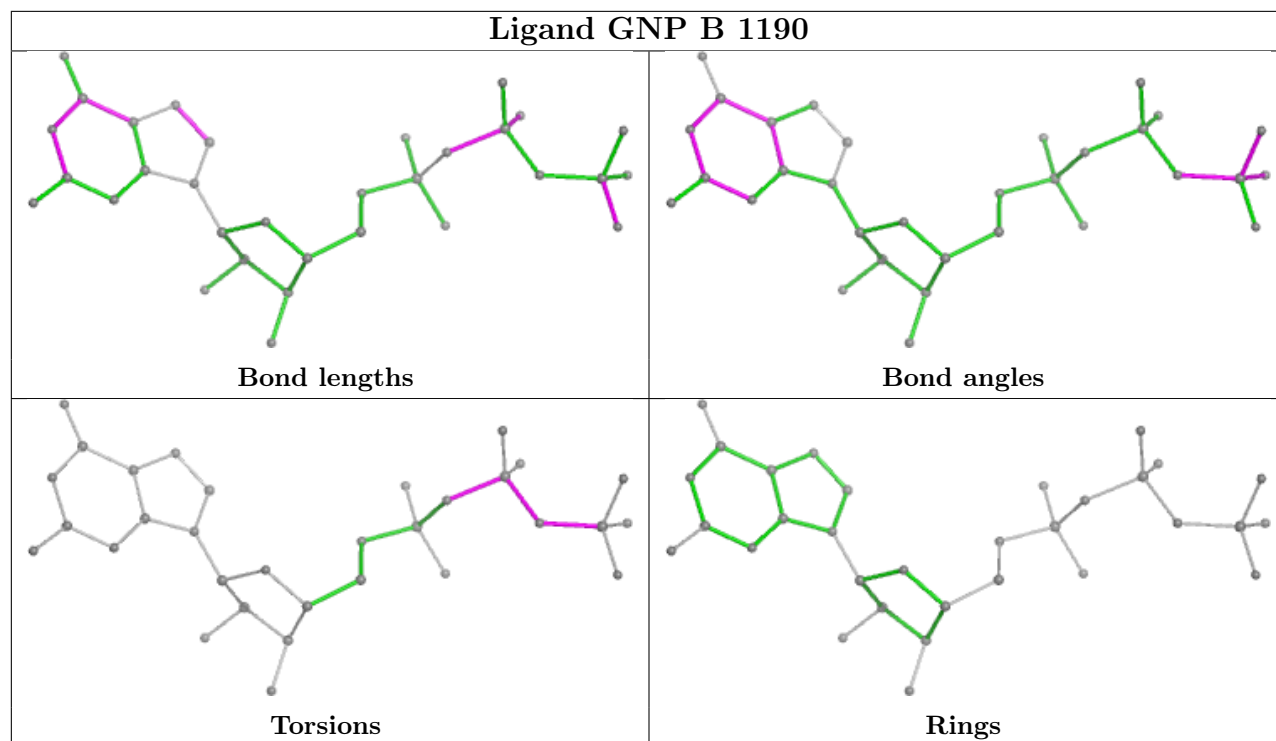
There are no ring outliers.

3 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	J	1190	GNP	4	0
6	K	1190	GNP	5	0
6	B	1190	GNP	5	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

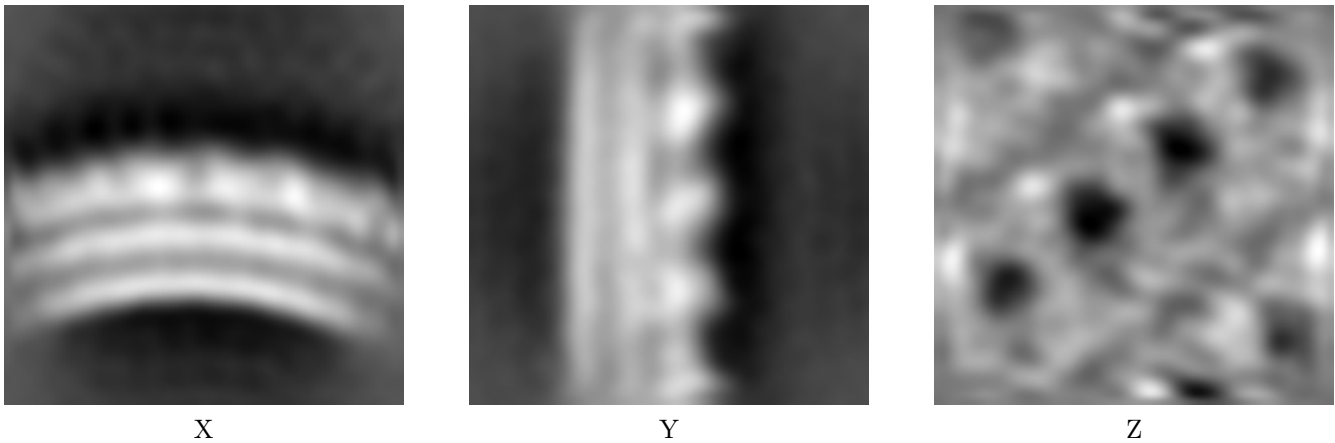
There are no chain breaks in this entry.



## 6 Tomogram visualisation [i](#)

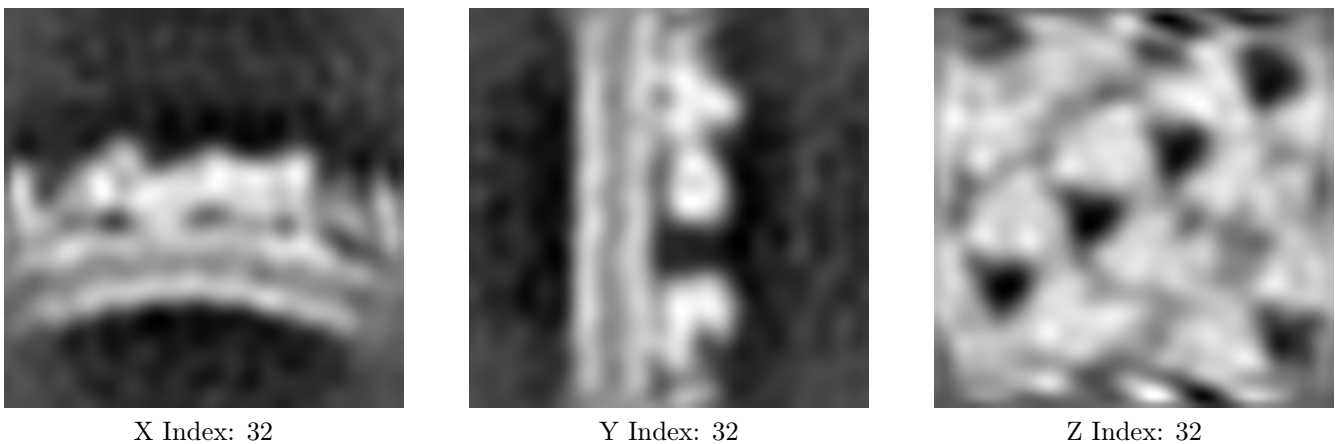
This section contains visualisations of the EMDB entry EMD-2428. These allow visual inspection of the internal detail of the tomogram and identification of artifacts.

### 6.1 Orthogonal projections [i](#)



The images above show the tomogram projected in three orthogonal directions.

### 6.2 Central slices [i](#)



The images above show central slices of the tomogram in three orthogonal directions.

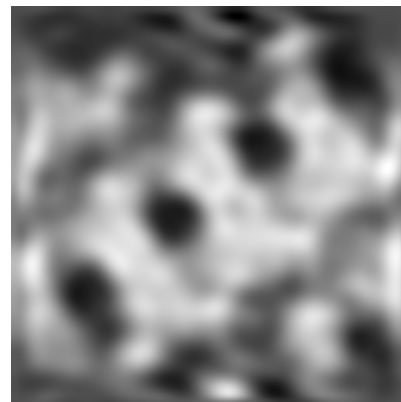
### 6.3 Largest variance slices [i](#)



X Index: 32



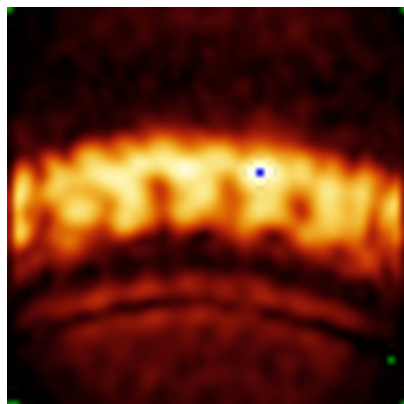
Y Index: 26



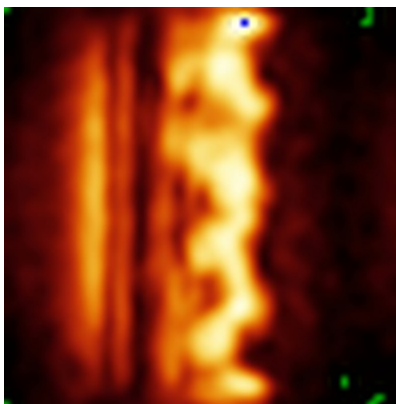
Z Index: 36

The images above show the largest variance slices of the tomogram in three orthogonal directions.

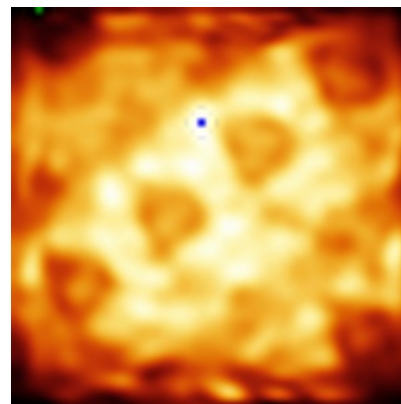
### 6.4 Orthogonal standard-deviation projections (False-color) [i](#)



X



Y



Z

The images above show the tomogram projected in three orthogonal directions.

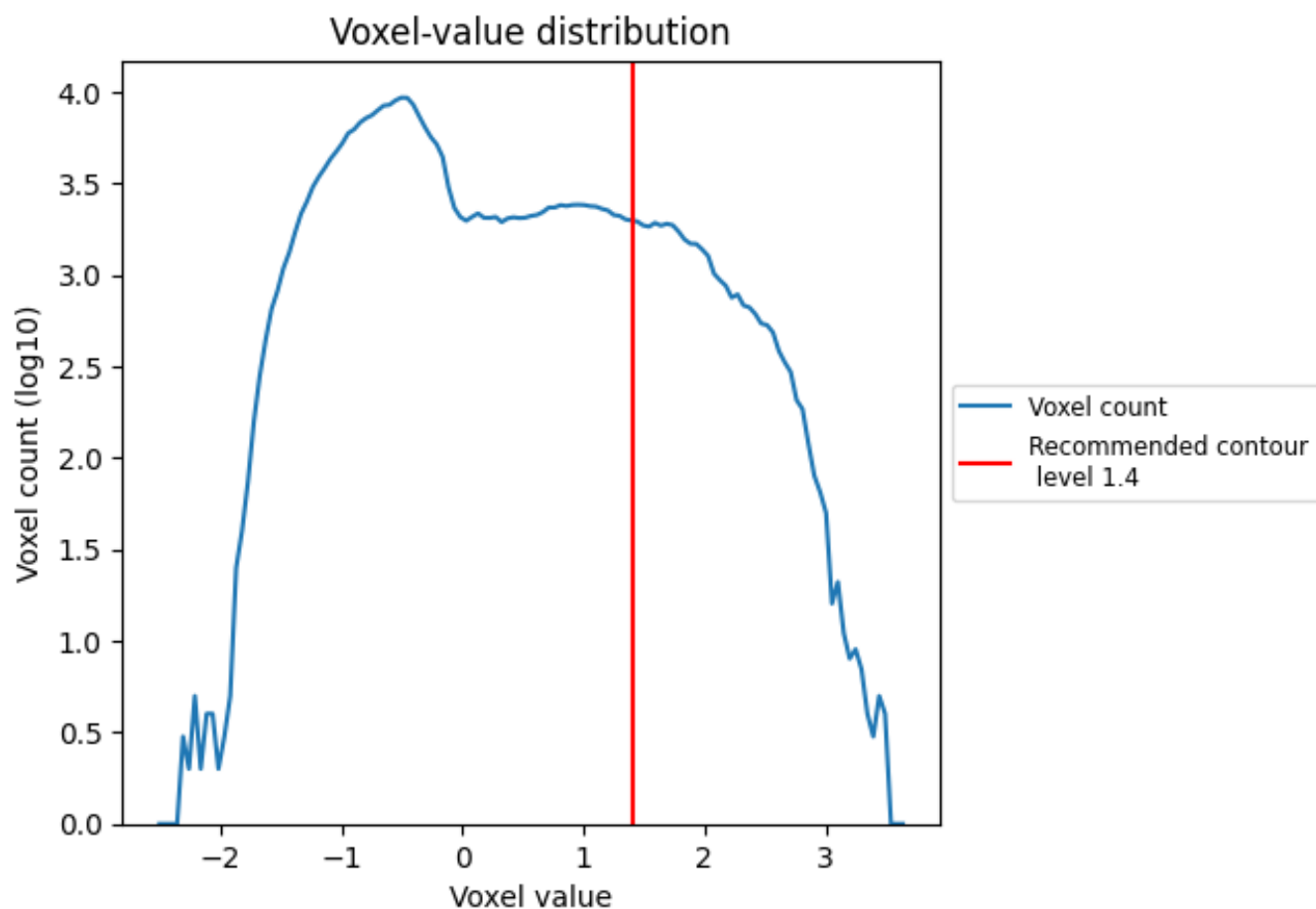
### 6.5 Mask visualisation [i](#)

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

## 7 Tomogram analysis [i](#)

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

### 7.1 Voxel-value distribution [i](#)



The voxel-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic.

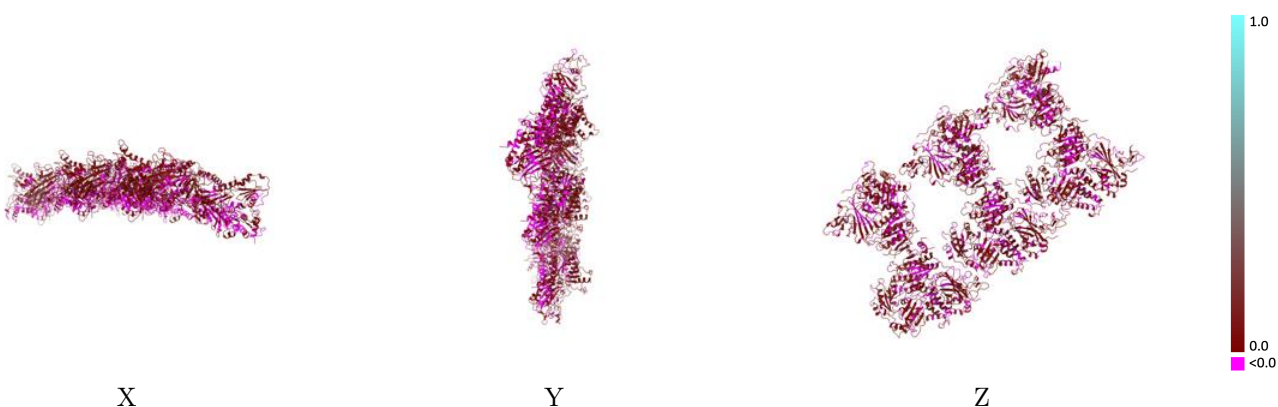
## 8 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-2428 and PDB model 4BZI. Per-residue inclusion information can be found in section 3 on page 7.

### 8.1 Map-model overlay [i](#)

This section was not generated.

### 8.2 Q-score mapped to coordinate model [i](#)

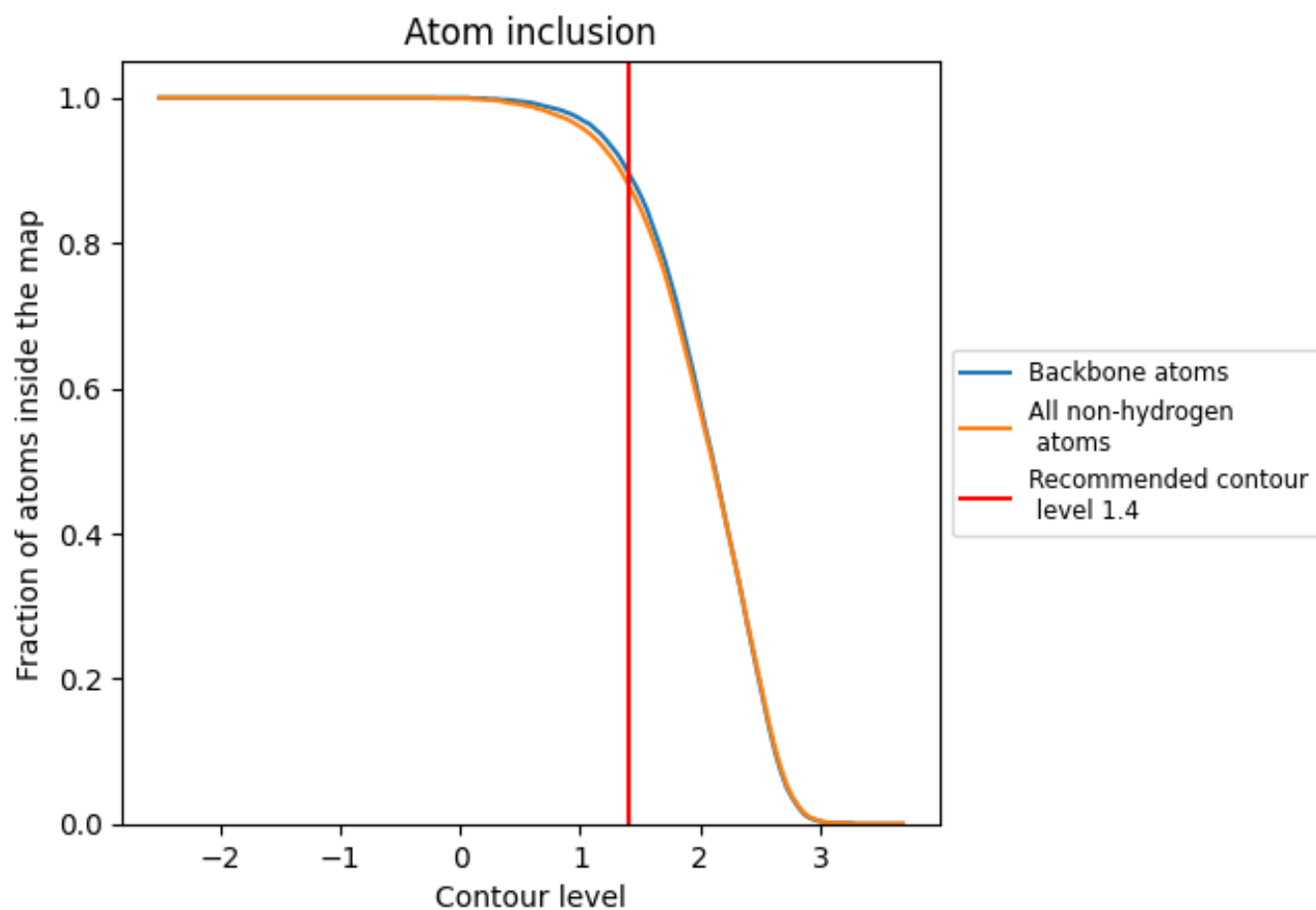


The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

### 8.3 Atom inclusion mapped to coordinate model [i](#)

This section was not generated.



























## 8.4 Atom inclusion [i](#)



At the recommended contour level, 90% of all backbone atoms, 88% of all non-hydrogen atoms, are inside the map.

## 8.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (1.4) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8810	 0.0470
A	 0.8970	 0.0440
B	 0.8910	 0.0620
D	 0.8800	 0.0450
E	 0.9120	 0.0470
F	 0.7160	 0.0410
G	 0.8640	 0.0520
J	 0.9000	 0.0480
K	 0.8270	 0.0500
L	 0.8420	 0.0410
M	 0.9070	 0.0450
N	 0.9230	 0.0770
O	 0.6150	 0.0390

