wwPDB EM Validation Summary Report (i)

Jul 3, 2024 – 01:22 am BST

PDB ID	:	7O2W
EMDB ID	:	EMD-12700
Title	:	Structure of the C9orf72-SMCR8 complex
Authors	:	Noerpel, J.; Cavadini, S.; Schenk, A.D.; Graff-Meyer, A.; Chao, J.; Bhaskar, V.
Deposited on	:	2021-03-31
Resolution	:	Not provided

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

We welcome your comments at *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 (i) 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 (i)) were used in the production of this report:

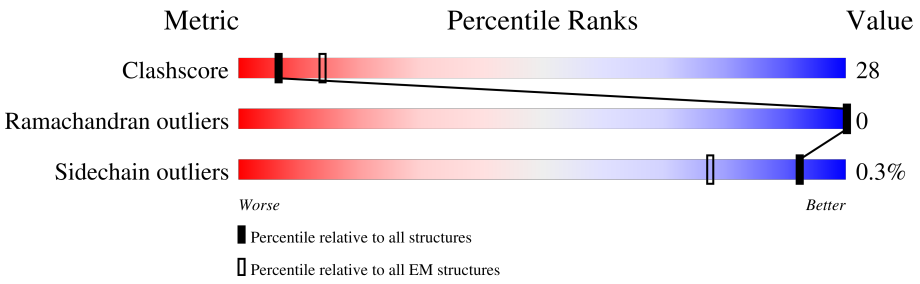
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 unknown.

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	593	36% 34% • 30%
2	B	1206	18% 17% 65%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6431 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 Ubiquitin-like protein SMT3,Guanine nucleotide exchange C9orf72.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	418	3151	2037	533	565	16	0	0

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-111	MET	-	initiating methionine	UNP Q12306
A	-110	ALA	-	expression tag	UNP Q12306
A	-109	HIS	-	expression tag	UNP Q12306
A	-108	HIS	-	expression tag	UNP Q12306
A	-107	HIS	-	expression tag	UNP Q12306
A	-106	HIS	-	expression tag	UNP Q12306
A	-105	HIS	-	expression tag	UNP Q12306
A	-104	HIS	-	expression tag	UNP Q12306
A	-103	GLY	-	expression tag	UNP Q12306
A	-7	SER	-	linker	UNP Q12306
A	-6	SER	-	linker	UNP Q12306
A	-5	GLY	-	linker	UNP Q12306
A	-4	LEU	-	linker	UNP Q12306
A	-3	ASP	-	linker	UNP Q12306
A	-2	ALA	-	linker	UNP Q12306
A	-1	ALA	-	linker	UNP Q12306
A	0	ALA	-	linker	UNP Q12306

- Molecule 2 is a protein called Guanine nucleotide exchange protein SMCR8,Guanine nucleotide exchange protein SMCR8,Maltose/maltodextrin-binding periplasmic protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	422	3280	2112	561	592	15	0	0

There are 83 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-44	MET	-	initiating methionine	UNP Q8TEV9
B	-43	ASP	-	expression tag	UNP Q8TEV9
B	-42	SER	-	expression tag	UNP Q8TEV9
B	-41	ALA	-	expression tag	UNP Q8TEV9
B	-40	TRP	-	expression tag	UNP Q8TEV9
B	-39	SER	-	expression tag	UNP Q8TEV9
B	-38	HIS	-	expression tag	UNP Q8TEV9
B	-37	PRO	-	expression tag	UNP Q8TEV9
B	-36	GLN	-	expression tag	UNP Q8TEV9
B	-35	PHE	-	expression tag	UNP Q8TEV9
B	-34	GLU	-	expression tag	UNP Q8TEV9
B	-33	LYS	-	expression tag	UNP Q8TEV9
B	-32	GLY	-	expression tag	UNP Q8TEV9
B	-31	GLY	-	expression tag	UNP Q8TEV9
B	-30	GLY	-	expression tag	UNP Q8TEV9
B	-29	SER	-	expression tag	UNP Q8TEV9
B	-28	GLY	-	expression tag	UNP Q8TEV9
B	-27	GLY	-	expression tag	UNP Q8TEV9
B	-26	GLY	-	expression tag	UNP Q8TEV9
B	-25	SER	-	expression tag	UNP Q8TEV9
B	-24	GLY	-	expression tag	UNP Q8TEV9
B	-23	GLY	-	expression tag	UNP Q8TEV9
B	-22	SER	-	expression tag	UNP Q8TEV9
B	-21	ALA	-	expression tag	UNP Q8TEV9
B	-20	TRP	-	expression tag	UNP Q8TEV9
B	-19	SER	-	expression tag	UNP Q8TEV9
B	-18	HIS	-	expression tag	UNP Q8TEV9
B	-17	PRO	-	expression tag	UNP Q8TEV9
B	-16	GLN	-	expression tag	UNP Q8TEV9
B	-15	PHE	-	expression tag	UNP Q8TEV9
B	-14	GLU	-	expression tag	UNP Q8TEV9
B	-13	LYS	-	expression tag	UNP Q8TEV9
B	-12	SER	-	expression tag	UNP Q8TEV9
B	-11	ALA	-	expression tag	UNP Q8TEV9
B	-10	VAL	-	expression tag	UNP Q8TEV9
B	-9	ASP	-	expression tag	UNP Q8TEV9
B	-8	LEU	-	expression tag	UNP Q8TEV9
B	-7	GLU	-	expression tag	UNP Q8TEV9
B	-6	VAL	-	expression tag	UNP Q8TEV9
B	-5	LEU	-	expression tag	UNP Q8TEV9
B	-4	PHE	-	expression tag	UNP Q8TEV9
B	-3	GLN	-	expression tag	UNP Q8TEV9
B	-2	GLY	-	expression tag	UNP Q8TEV9

Continued on next page...

Continued from previous page...

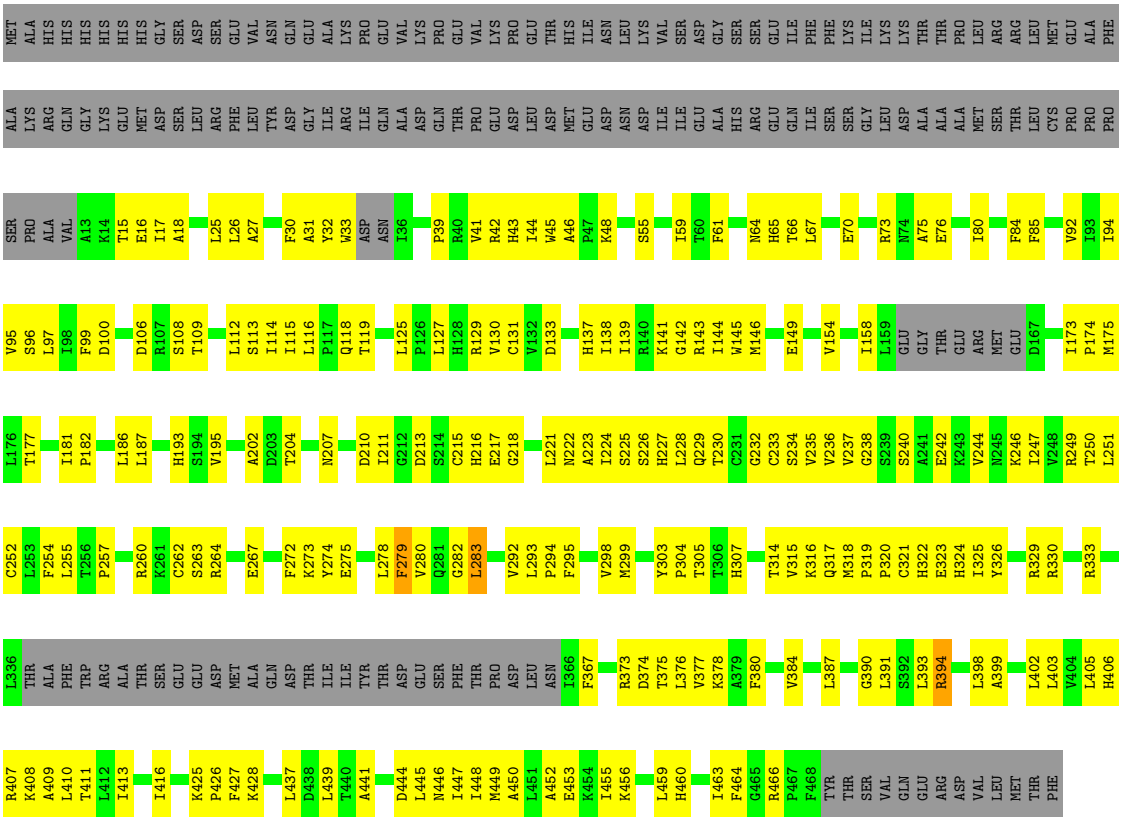
Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	PRO	-	expression tag	UNP Q8TEV9
B	0	GLY	-	expression tag	UNP Q8TEV9
B	600	ASP	-	linker	UNP Q8TEV9
B	601	THR	-	linker	UNP Q8TEV9
B	602	GLY	-	linker	UNP Q8TEV9
B	603	SER	-	linker	UNP Q8TEV9
B	604	THR	-	linker	UNP Q8TEV9
B	605	GLY	-	linker	UNP Q8TEV9
B	606	SER	-	linker	UNP Q8TEV9
B	607	THR	-	linker	UNP Q8TEV9
B	608	SER	-	linker	UNP Q8TEV9
B	609	GLY	-	linker	UNP Q8TEV9
B	610	THR	-	linker	UNP Q8TEV9
B	611	LEU	-	linker	UNP Q8TEV9
B	612	GLU	-	linker	UNP Q8TEV9
B	613	VAL	-	linker	UNP Q8TEV9
B	614	LEU	-	linker	UNP Q8TEV9
B	615	PHE	-	linker	UNP Q8TEV9
B	616	GLN	-	linker	UNP Q8TEV9
B	617	GLY	-	linker	UNP Q8TEV9
B	618	PRO	-	linker	UNP Q8TEV9
B	619	GLY	-	linker	UNP Q8TEV9
B	938	GLY	-	linker	UNP Q8TEV9
B	939	GLY	-	linker	UNP Q8TEV9
B	940	SER	-	linker	UNP Q8TEV9
B	941	GLY	-	linker	UNP Q8TEV9
B	942	SER	-	linker	UNP Q8TEV9
B	943	GLU	-	linker	UNP Q8TEV9
B	944	ASN	-	linker	UNP Q8TEV9
B	945	LEU	-	linker	UNP Q8TEV9
B	946	TYR	-	linker	UNP Q8TEV9
B	947	PHE	-	linker	UNP Q8TEV9
B	948	GLN	-	linker	UNP Q8TEV9
B	949	GLY	-	linker	UNP Q8TEV9
B	950	GLY	-	linker	UNP Q8TEV9
B	951	THR	-	linker	UNP Q8TEV9
B	952	SER	-	linker	UNP Q8TEV9
B	953	SER	-	linker	UNP Q8TEV9
B	954	GLY	-	linker	UNP Q8TEV9
B	955	MET	-	linker	UNP Q8TEV9

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. 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.

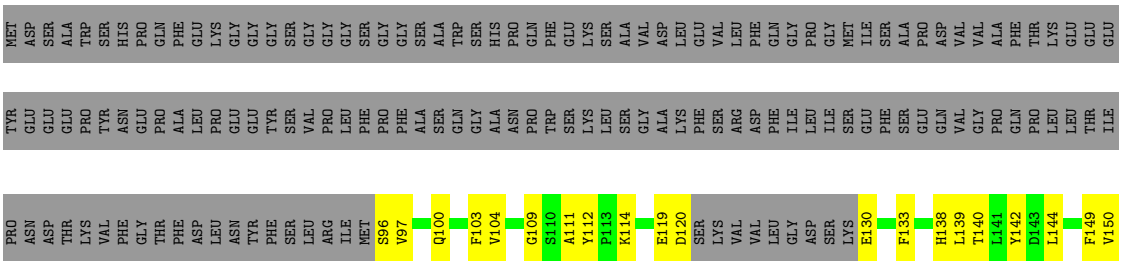
• Molecule 1: Ubiquitin-like protein SMT3,Guanine nucleotide exchange C9orf72

Chain A: 36% 34% • 30%



• Molecule 2: Guanine nucleotide exchange protein SMCR8,Guanine nucleotide exchange protein SMCR8,Maltose/maltodextrin-binding periplasmic protein

Chain B: 18% 17% 65%



R151 P152 F153 C154 Y157 I158 S159 A160 D161 Q162 H163 K164 I165 M166 Q167 Q170 E171 A174 E175 F176 S177 R178 N187 F191 E194 L195 E196 K197 K198 L199 K200 D201 L202 D203 Y204 T205 R206 T207 V208 L209 T213 E214 K217 K218 A219 N220 D221 K222 G223 F224 Y225 S226
S227 Q228 E231 K232 V239 E240 K241 S242 I243 I244 E245 H246 Q247 L250 LYS GLN ILE ARG SER TYR PRO HIS ARG LYS LEU LYS GLY HIS ASP LEU CYS PRO GLY GLU MET GLU HIS ILE GLN ASP GLN ALA SER GLN ALA SER THR THR SER ASN PRO ASP GLU SER ALA ASP THR
ASP LEU TYR THR CYS ARG PRO ALA TYR THR PRO LYS LEU ILE LYS ALA LYS SER THR LYS C314 F315 D316 K317 K320 T321 L322 E323 E324 L325 C326 D327 T328 E329 Y330 F331 T332 Q333 T334 L335 A336 Q337 L338 S339 H340 I341 M344 F345 L349 C350 Y351 L352 L353 T354 S355 Q356
R359 A360 L361 L362 K363 N369 F370 L371 F372 E373 D374 PHE VAL GLU VAL ASP ASP ARG MET VAL GLU LYS GLN GLU SER ILE PRO SER LYS PRO SER GLN ASP ARG PRO PRO SER SER LEU GLU GLU CYS PRO ILE PRO LYS VAL LEU ILE SER VAL GLY SER TYR LYS SER SER VAL
GLU SER VAL LEU ILE LYS MET GLU GLN GLU LEU GLY ASP GLU GLU TYR LYS ASP THR GLY SER THR GLY SER THR SER GLY THR LEU GLU VAL LEU PHE GLN GLY PRO GLY ARG GLN LYS ASP GLN GLY PHE ARG VAL ASP PHE SER VAL GLU ASN ALA ASN PRO SER SER ARG ASP ASN
SER CYS GLU GLY PHE PRO ALA TYR GLU LEU ASP PRO SER HIS LEU LEU ALA SER ARG ASP ILE SER LYS THR SER LEU ASP ASN TYR SER ASP THR THR SER TYR VAL SER SER VAL ALA SER THR SER SER ASP ARG ILE PRO SER ALA TYR PRO ALA GLY LEU SER SER ASP R701 H702
K703 A706 A710 F713 Y717 P718 F719 A720 H721 P722 A723 L727 G730 R731 T732 L733 V734 V735 E740 R744 F753 V754 Y757 G758 K765 H766 W767 A768 I774 MET ASP PHE Q778 K779 W780 K781 L782 I783 G784 R787 VAL ALA SER PRO ALA GLY GLY
L797 L800 S801 R802 Y803 S804 T807 S808 I809 L810 R818 C819 P820 L821 Y822 R823 G824 T825 L826 V827 P828 A831 D832 H833 R834 T835 Q836 I837 G840 Y843 Y844 L845 H846 V847 Q848 S849 M850 L851 T852 C855 S856 F859 L860 Y861 T862 F863 C864 H865 H866 L867 H868
L869 P870 T871 T876 E877 E878 R883 Q884 M885 S886 F887 L888 K889 L890 T891 L892 G893 L894 V895 N896 E897 R900 L905 L908 L909 K910 L911 E916 SER PRO GLY THR SER HIS PRO MET LEU ARG PHE ASP Y929 F933 K936 ILE GLY GLY SER GLY SER GLU GLU ASN LEU
TYR PHE GLN GLY GLY THR SER SER GLY MET LYS ILE GLU GLU GLY LYS LEU VAL ILE TRP ILE ASN GLY ASP LYS GLY TYR ASN GLY LEU ALA GLU VAL GLY LYS LYS PHE GLU LYS ASP THR GLY ILE LYS VAL THR VAL GLU HIS PRO ASP LYS LEU PHE GLU VAL
ALA ALA THR GLY ASP GLY PRO ASP ILE ILE PHE TRP ALA HIS ASP ARG PHE GLY GLY TYR ALA GLN SER GLY LEU LEU ALA GLU ILE THR PRO ASP LYS ALA PHE GLN ASP LYS LEU TYR PRO PHE THR TRP ASP ALA VAL ARG TYR ASN GLY LYS LEU ILE ALA PHE VAL
GLU ALA LEU SER LEU ILE TYR ASN LYS ASP LEU PRO ASN PRO PRO LYS THR TRP ALA LEU ASP LYS GLU LEU LYS ALA LYS GLY LYS SER ALA LEU MET PHE ASN LEU GLN GLU PRO TYR PHE THR TRP PRO LEU GLU ILE ALA ASP LEU ALA ALA ASP MET GLY GLY ALA TYR PRO ILE ALA PHE LYS
TYR GLU ASN GLY LYS TYR ASP ILE LYS ASP VAL GLY VAL ASP ASN ALA GLY ALA LYS ALA GLY LEU THR PHE LEU VAL ASP LEU ILE LYS ASN LYS HIS MET ASN ALA ASP THR ASP TYR SER ILE ALA GLU ALA ALA PHE ASN PHE LYS GLY GLU THR ALA MET THR ILE ASN GLY GLY PRO TRP
ALA TRP SER ASN ILE ASP THR SER LYS VAL ASN TYR GLY VAL ASN LYS PRO PHE VAL GLY VAL LEU SER ALA GLY ILE ASN ALA ALA SER PRO ASN LYS GLU LEU ALA ALA LYS GLU PHE LEU GLU ASP TYR ASN TYR LEU LEU THR ASP LEU GLY GLU
GLU ALA VAL ASN LYS ASP PRO LEU GLY ALA VAL ALA LYS SER TYR GLU GLU GLU LEU ALA LYS ASP PRO ARG ILE ALA ALA THR MET GLU ASN ALA GLN LYS GLY GLU ILE MET PRO ASN ILE PRO GLN MET SER ALA PHE TRP TYR ALA VAL ARG THR ALA VAL ILE ASN ALA
ALA SER GLY ARG GLN THR VAL ASP GLU ALA ALA LEU LYS ASP ALA GLN THR

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	284568	Depositor
Resolution determination method	Not provided	
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	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	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.34	0/3216	0.57	1/4378 (0.0%)
2	B	0.34	0/3353	0.55	0/4541
All	All	0.34	0/6569	0.56	1/8919 (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	A	0	2
2	B	0	1
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	283	LEU	CA-CB-CG	5.95	128.99	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	279	PHE	Peptide
1	A	75	ALA	Peptide
2	B	766	HIS	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	3151	0	3085	176	0
2	B	3280	0	3150	185	0
All	All	6431	0	6235	354	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 28.

The worst 5 of 354 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)
2:B:774:ILE:HD13	2:B:779:LYS:NZ	1.44	1.31
2:B:774:ILE:HG21	2:B:779:LYS:CE	1.84	1.07
2:B:774:ILE:CG2	2:B:779:LYS:HE2	1.89	1.02
1:A:235:VAL:HG21	1:A:283:LEU:HD23	1.39	1.02
2:B:774:ILE:HD13	2:B:779:LYS:HZ3	1.29	0.94

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	410/593 (69%)	347 (85%)	63 (15%)	0	100	100
2	B	408/1206 (34%)	341 (84%)	67 (16%)	0	100	100
All	All	818/1799 (46%)	688 (84%)	130 (16%)	0	100	100

There are no Ramachandran outliers to report.

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	328/529 (62%)	327 (100%)	1 (0%)	92	92
2	B	334/1027 (32%)	333 (100%)	1 (0%)	92	92
All	All	662/1556 (42%)	660 (100%)	2 (0%)	92	92

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	394	ARG
2	B	744	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	333	GLN
2	B	853	GLN
2	B	865	HIS
1	A	322	HIS
1	A	222	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 monosaccharides 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.

6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-12700. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

This section was not generated.

6.2 Central slices (i)

This section was not generated.

6.3 Largest variance slices (i)

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color) (i)

This section was not generated.

6.5 Orthogonal surface views (i)

This section was not generated.

6.6 Mask visualisation (i)

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

7 Map analysis (i)

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

7.1 Map-value distribution (i)

This section was not generated.

7.2 Volume estimate versus contour level (i)

This section was not generated.

7.3 Rotationally averaged power spectrum (i)

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit (i)

This section was not generated.