



## wwPDB EM Validation Summary Report ⓘ

Jan 25, 2022 – 03:06 pm GMT

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

EMDB validation analysis : 0.0.0.dev97  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.26

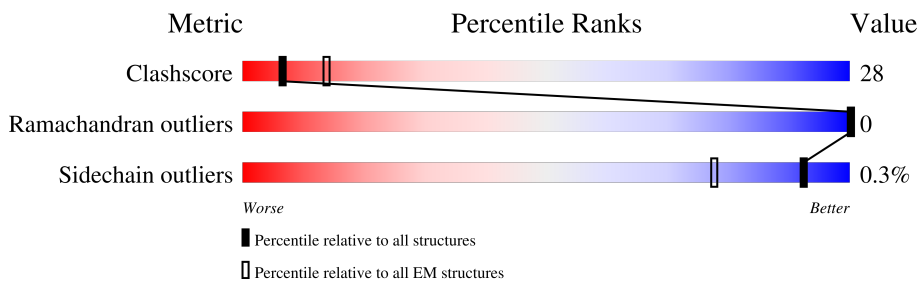
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	593	
2	B	1206	

## 2 Entry composition

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...*

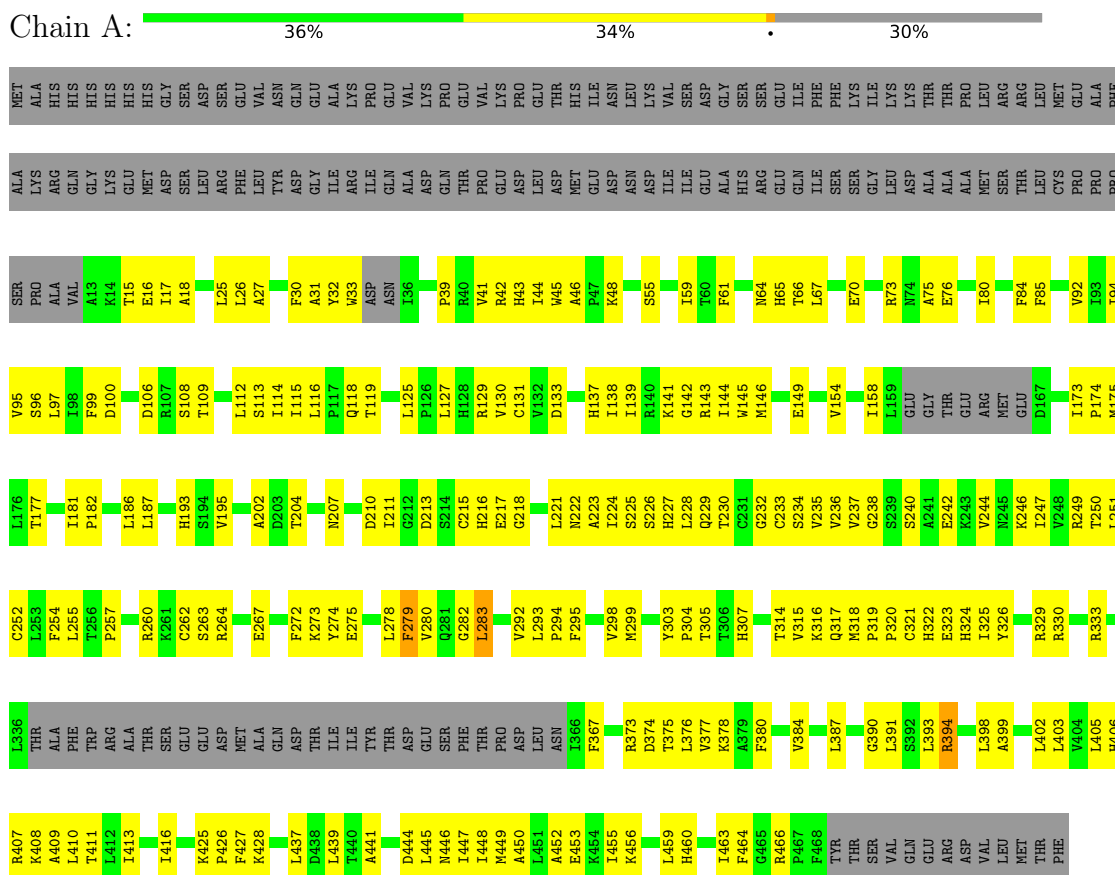
*Continued from previous page...*

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

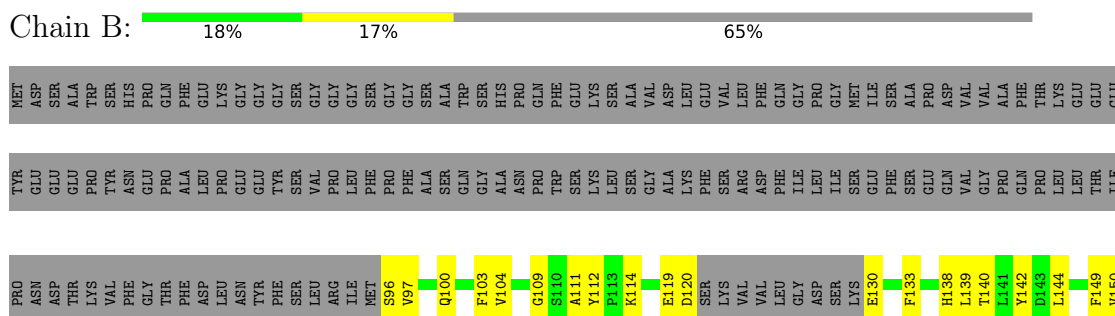
### 3 Residue-property plots

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



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





## 4 Experimental information

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 ( $e^-/\text{\AA}^2$ )	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( $^{\circ}$ )	Ideal( $^{\circ}$ )
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

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

This section was not generated.

### 6.2 Central slices

This section was not generated.

### 6.3 Largest variance slices

This section was not generated.

### 6.4 Orthogonal surface views

This section was not generated.

### 6.5 Mask visualisation

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

## 7 Map analysis

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

### 7.1 Map-value distribution

This section was not generated.

### 7.2 Volume estimate versus contour level

This section was not generated.

### 7.3 Rotationally averaged power spectrum

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

## 8 Fourier-Shell correlation

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

## 9 Map-model fit

This section was not generated.