



wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 01:21 AM EST

PDB ID : 7KAP
EMDB ID : EMD-22779
Title : Cryo-EM structure of the Sec complex from *S. cerevisiae*, Sec61 pore mutant, class with Sec62, conformation 1 (C1)
Authors : Itskanov, S.; Park, E.
Deposited on : 2020-10-01
Resolution : 4.10 Å(reported)

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 ⓘ 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.dev43
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)	:	2.31.2

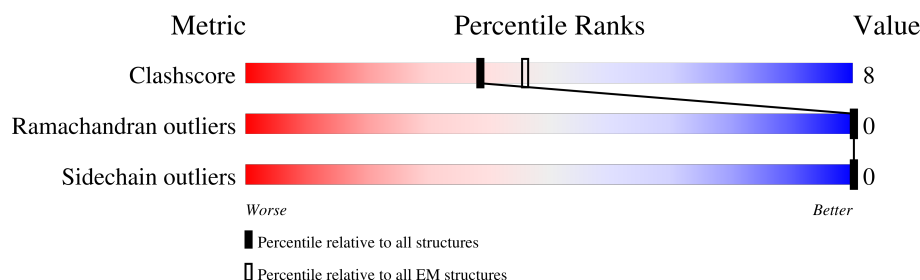
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.10 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 ≥ 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	480	69% 21% 9%
2	C	80	51% 18% 31%
3	B	82	29% 10% 61%
4	D	694	58% 13% 29%
5	E	206	48% 19% 33%
6	F	193	78% 21% .
7	G	56	100%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10715 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 SEC61.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	435	Total	C	N	O	S	0	0
			3307	2188	532	576	11		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	90	LEU	MET	variant	UNP P32915
A	185	ILE	THR	variant	UNP P32915
A	294	ILE	MET	variant	UNP P32915
A	450	LEU	MET	variant	UNP P32915

- Molecule 2 is a protein called Protein transport protein SSS1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	55	Total	C	N	O	S	0	0
			435	294	71	69	1		

- Molecule 3 is a protein called Protein transport protein SBH1.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	B	32	Total	C	N	O	0	0
			229	157	38	34		

- Molecule 4 is a protein called Protein translocation protein SEC63.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	492	Total	C	N	O	S	0	0
			3864	2517	654	679	14		

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-13	GLY	-	expression tag	UNP P14906

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-12	GLY	-	expression tag	UNP P14906
D	-11	SER	-	expression tag	UNP P14906
D	-10	GLY	-	expression tag	UNP P14906
D	-9	GLY	-	expression tag	UNP P14906
D	-8	SER	-	expression tag	UNP P14906
D	-7	GLY	-	expression tag	UNP P14906
D	-6	GLY	-	expression tag	UNP P14906
D	-5	SER	-	expression tag	UNP P14906
D	-4	GLY	-	expression tag	UNP P14906
D	-3	GLY	-	expression tag	UNP P14906
D	-2	SER	-	expression tag	UNP P14906
D	-1	GLY	-	expression tag	UNP P14906
D	0	GLY	-	expression tag	UNP P14906
D	1	SER	-	expression tag	UNP P14906
D	664	ALA	-	expression tag	UNP P14906
D	665	GLY	-	expression tag	UNP P14906
D	666	GLY	-	expression tag	UNP P14906
D	667	ALA	-	expression tag	UNP P14906
D	668	THR	-	expression tag	UNP P14906
D	669	THR	-	expression tag	UNP P14906
D	670	ALA	-	expression tag	UNP P14906
D	671	SER	-	expression tag	UNP P14906
D	672	GLY	-	expression tag	UNP P14906
D	673	THR	-	expression tag	UNP P14906
D	674	GLY	-	expression tag	UNP P14906
D	675	GLU	-	expression tag	UNP P14906
D	676	ASN	-	expression tag	UNP P14906
D	677	LEU	-	expression tag	UNP P14906
D	678	TYR	-	expression tag	UNP P14906
D	679	PHE	-	expression tag	UNP P14906
D	680	GLN	-	expression tag	UNP P14906

- Molecule 5 is a protein called Translocation protein SEC66.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	138	Total	C	N	O	S	0	0
			1129	722	200	202	5		

- Molecule 6 is a protein called Translocation protein SEC72.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	190	Total	C	N	O	S	0	0
			1471	934	256	274	7		

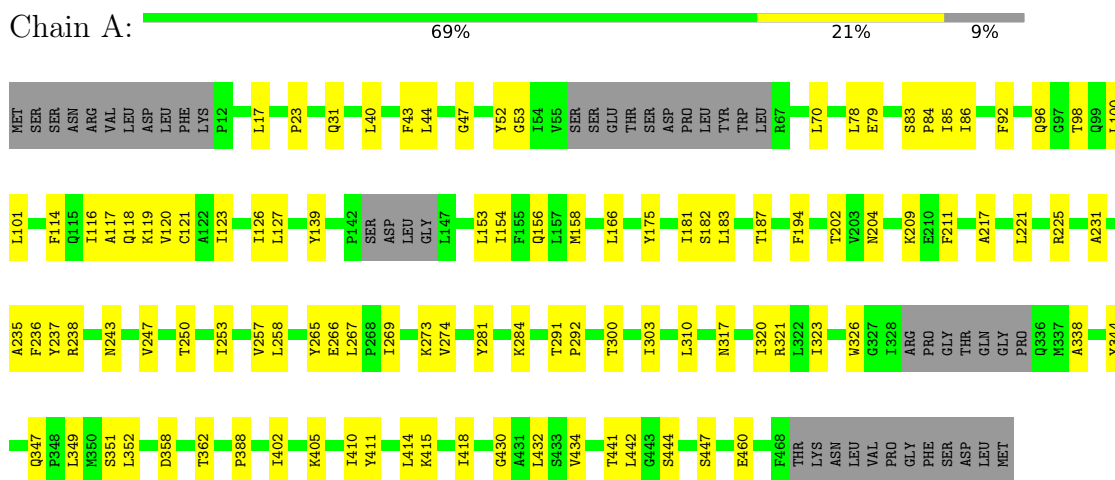
- Molecule 7 is a protein called Protein transport protein Sec62.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	G	56	Total	C	N	O	0	0
			280	168	56	56		

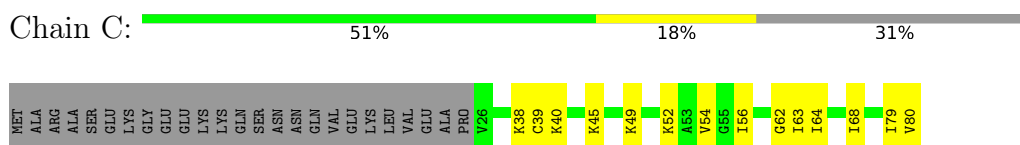
3 Residue-property plots

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.

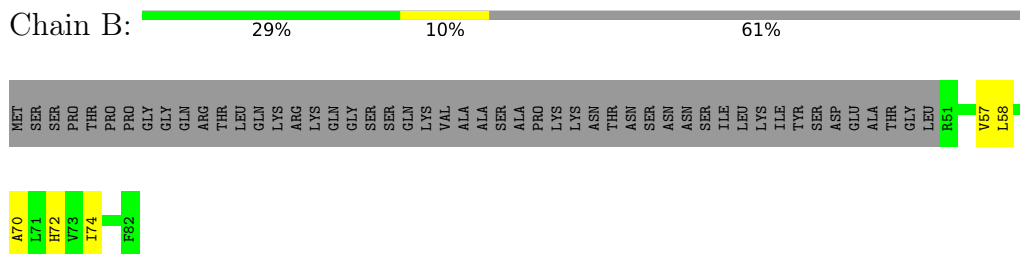
- Molecule 1: Protein transport protein SEC61



- Molecule 2: Protein transport protein SSS1

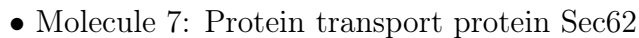


- Molecule 3: Protein transport protein SBH1



- Molecule 4: Protein translocation protein SEC63





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	17341	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	48.8	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	43478	Depositor
Image detector	GATAN K2 SUMMIT (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	0.26	0/3373	0.43	1/4582 (0.0%)
2	C	0.26	0/443	0.36	0/596
3	B	0.26	0/233	0.39	0/318
4	D	0.26	0/3955	0.42	0/5375
5	E	0.25	0/1147	0.38	0/1540
6	F	0.24	0/1491	0.39	0/2018
All	All	0.26	0/10642	0.41	1/14429 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	266	GLU	C-N-CA	-5.27	108.52	121.70

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	3307	0	3444	67	0
2	C	435	0	481	9	0
3	B	229	0	243	5	0
4	D	3864	0	3901	55	0
5	E	1129	0	1145	23	0
6	F	1471	0	1512	25	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	280	0	58	0	0
All	All	10715	0	10784	169	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 8.

The worst 5 of 169 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)
6:F:92:ARG:NH2	6:F:96:GLU:OE1	2.21	0.74
4:D:277:THR:HG23	4:D:355:THR:HG22	1.70	0.73
4:D:30:THR:HG21	4:D:100:TRP:HE1	1.54	0.70
4:D:409:ILE:HD11	4:D:415:LEU:HD13	1.74	0.69
1:A:40:LEU:HD11	1:A:187:THR:HG21	1.75	0.68

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	427/480 (89%)	415 (97%)	12 (3%)	0	100	100
2	C	53/80 (66%)	52 (98%)	1 (2%)	0	100	100
3	B	30/82 (37%)	30 (100%)	0	0	100	100
4	D	484/694 (70%)	471 (97%)	13 (3%)	0	100	100
5	E	136/206 (66%)	134 (98%)	2 (2%)	0	100	100
6	F	188/193 (97%)	186 (99%)	2 (1%)	0	100	100
All	All	1318/1735 (76%)	1288 (98%)	30 (2%)	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	347/406 (86%)	347 (100%)	0	100	100
2	C	46/67 (69%)	46 (100%)	0	100	100
3	B	23/69 (33%)	23 (100%)	0	100	100
4	D	422/615 (69%)	422 (100%)	0	100	100
5	E	120/191 (63%)	120 (100%)	0	100	100
6	F	157/165 (95%)	157 (100%)	0	100	100
All	All	1115/1513 (74%)	1115 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
5	E	181	GLN
6	F	155	GLN
6	F	106	HIS
4	D	334	HIS
5	E	123	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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-22779. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

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.