



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 03:38 am GMT

PDB ID : 2PM7
Title : Crystal structure of yeast Sec13/31 edge element of the COPII vesicular coat, selenomethionine version
Authors : Goldberg, J.; Fath, S.; Mancias, J.D.; Bi, X.
Deposited on : 2007-04-20
Resolution : 2.35 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

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:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

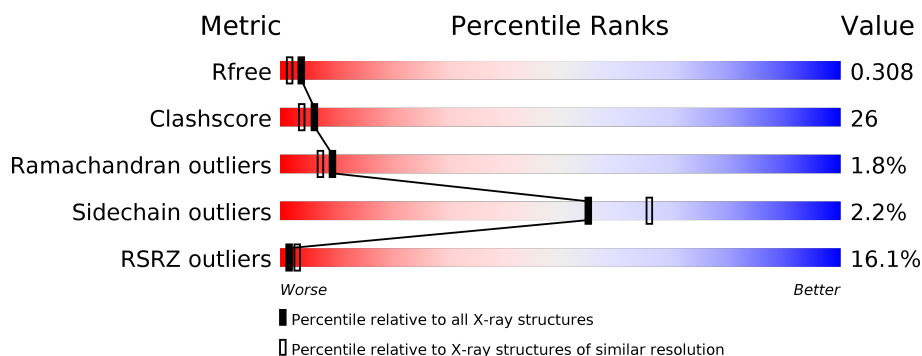
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	1015 (2.36-2.36)
Clashscore	122126	1081 (2.36-2.36)
Ramachandran outliers	120053	1066 (2.36-2.36)
Sidechain outliers	120020	1067 (2.36-2.36)
RSRZ outliers	108989	1002 (2.36-2.36)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	399	
1	C	399	
2	B	297	
2	D	297	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10195 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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						ZeroOcc	AltConf	Trace
1	A	345	Total	C	N	O	S	Se	0	0	0
			2729	1723	449	547	1	9			
1	C	347	Total	C	N	O	S	Se	0	0	0
			2746	1734	452	550	1	9			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	365	GLY	-	CLONING ARTIFACT	UNP P38968
A	366	ALA	-	CLONING ARTIFACT	UNP P38968
A	367	MSE	-	CLONING ARTIFACT	UNP P38968
A	368	GLY	-	CLONING ARTIFACT	UNP P38968
A	369	SER	-	CLONING ARTIFACT	UNP P38968
A	449	MSE	LEU	ENGINEERED	UNP P38968
A	455	MSE	MET	MODIFIED RESIDUE	UNP P38968
A	536	MSE	LEU	ENGINEERED	UNP P38968
A	537	MSE	MET	MODIFIED RESIDUE	UNP P38968
A	540	MSE	MET	MODIFIED RESIDUE	UNP P38968
A	614	MSE	MET	MODIFIED RESIDUE	UNP P38968
A	615	MSE	LEU	ENGINEERED	UNP P38968
A	622	MSE	LEU	ENGINEERED	UNP P38968
A	674	MSE	LEU	ENGINEERED	UNP P38968
C	365	GLY	-	CLONING ARTIFACT	UNP P38968
C	366	ALA	-	CLONING ARTIFACT	UNP P38968
C	367	MSE	-	CLONING ARTIFACT	UNP P38968
C	368	GLY	-	CLONING ARTIFACT	UNP P38968
C	369	SER	-	CLONING ARTIFACT	UNP P38968
C	449	MSE	LEU	ENGINEERED	UNP P38968
C	455	MSE	MET	MODIFIED RESIDUE	UNP P38968
C	536	MSE	LEU	ENGINEERED	UNP P38968
C	537	MSE	MET	MODIFIED RESIDUE	UNP P38968
C	540	MSE	MET	MODIFIED RESIDUE	UNP P38968
C	614	MSE	MET	MODIFIED RESIDUE	UNP P38968

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	615	MSE	LEU	ENGINEERED	UNP P38968
C	622	MSE	LEU	ENGINEERED	UNP P38968
C	674	MSE	LEU	ENGINEERED	UNP P38968

- Molecule 2 is a protein called Protein transport protein SEC13.

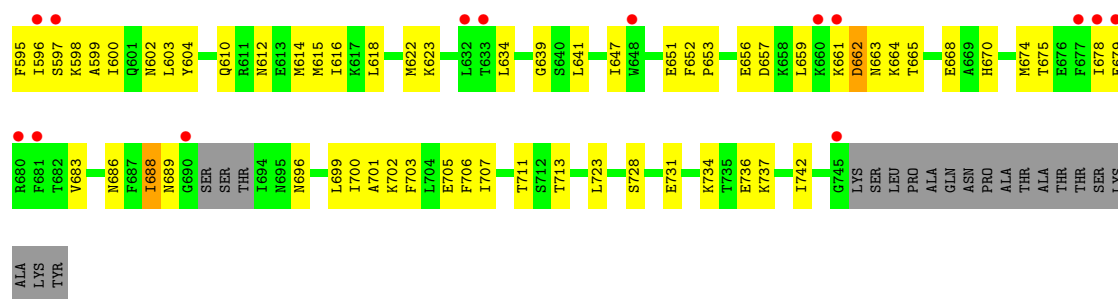
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	279	Total	C	N	O	S	Se	0	0	0
			2196	1397	375	415	3	6			
2	D	288	Total	C	N	O	S	Se	0	0	0
			2263	1438	387	429	3	6			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q04491
B	11	MSE	LEU	ENGINEERED	UNP Q04491
B	17	MSE	LEU	ENGINEERED	UNP Q04491
B	24	MSE	LEU	ENGINEERED	UNP Q04491
B	80	MSE	LEU	ENGINEERED	UNP Q04491
B	115	MSE	LEU	ENGINEERED	UNP Q04491
B	222	MSE	LEU	ENGINEERED	UNP Q04491
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q04491
D	11	MSE	LEU	ENGINEERED	UNP Q04491
D	17	MSE	LEU	ENGINEERED	UNP Q04491
D	24	MSE	LEU	ENGINEERED	UNP Q04491
D	80	MSE	LEU	ENGINEERED	UNP Q04491
D	115	MSE	LEU	ENGINEERED	UNP Q04491
D	222	MSE	LEU	ENGINEERED	UNP Q04491

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	29	Total	O	0	0
			29	29		
3	B	60	Total	O	0	0
			60	60		
3	C	50	Total	O	0	0
			50	50		
3	D	122	Total	O	0	0
			122	122		



• Molecule 2: Protein transport protein SEC13



• Molecule 2: Protein transport protein SEC13



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	128.19Å 52.50Å 133.09Å 90.00° 108.34° 90.00°	Depositor
Resolution (Å)	30.00 – 2.35 24.83 – 2.35	Depositor EDS
% Data completeness (in resolution range)	96.7 (30.00-2.35) 96.9 (24.83-2.35)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.47 (at 2.36Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.242 , 0.298 0.257 , 0.308	Depositor DCC
R_{free} test set	3487 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	47.4	Xtriage
Anisotropy	0.883	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 63.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.009 for l,-k,h	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	10195	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.31	0/2764	0.52	0/3712
1	C	0.35	0/2781	0.55	0/3735
2	B	0.36	0/2250	0.63	0/3055
2	D	0.43	0/2318	0.71	0/3146
All	All	0.36	0/10113	0.60	0/13648

There are no bond length outliers.

There are no bond angle outliers.

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	2729	0	2696	176	0
1	C	2746	0	2715	155	0
2	B	2196	0	2126	138	0
2	D	2263	0	2191	89	0
3	A	29	0	0	18	1
3	B	60	0	0	25	0
3	C	50	0	0	9	0
3	D	122	0	0	22	1
All	All	10195	0	9728	517	1

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 26.

The worst 5 of 517 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:D:55:PRO:HA	3:D:417:HOH:O	1.51	1.10
1:A:496:THR:HG23	1:C:557:ASN:HB3	1.33	1.09
1:A:431:ILE:HD11	1:A:683:VAL:HG11	1.35	1.08
1:C:623:LYS:HE3	1:C:647:ILE:HD12	1.40	1.04
1:C:408:LYS:H	1:C:408:LYS:HD3	1.21	1.04

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:118:HOH:O	3:D:311:HOH:O[1_644]	2.14	0.06

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	339/399 (85%)	292 (86%)	39 (12%)	8 (2%)	6	4
1	C	341/399 (86%)	315 (92%)	22 (6%)	4 (1%)	14	13
2	B	275/297 (93%)	246 (90%)	26 (10%)	3 (1%)	16	15
2	D	284/297 (96%)	258 (91%)	19 (7%)	7 (2%)	6	4
All	All	1239/1392 (89%)	1111 (90%)	106 (9%)	22 (2%)	9	7

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	439	ILE
1	A	497	ASN
2	D	135	THR
2	D	294	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	439	ILE

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 X-ray entries followed by that with respect to entries of similar resolution.

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	306/340 (90%)	300 (98%)	6 (2%)	58	69
1	C	308/340 (91%)	299 (97%)	9 (3%)	45	55
2	B	237/245 (97%)	234 (99%)	3 (1%)	71	81
2	D	244/245 (100%)	238 (98%)	6 (2%)	50	60
All	All	1095/1170 (94%)	1071 (98%)	24 (2%)	55	66

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	432	ASN
1	C	510	ILE
2	D	204	SER
1	C	440	ASP
1	C	509	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	240	GLN
1	C	433	GLN
2	D	149	ASN
2	B	276	ASN
1	C	443	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](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	336/399 (84%)	1.30	86 (25%) 0 1	52, 87, 110, 119	0
1	C	338/399 (84%)	0.81	44 (13%) 3 6	35, 67, 99, 117	0
2	B	273/297 (91%)	0.90	36 (13%) 3 5	28, 64, 105, 120	0
2	D	282/297 (94%)	0.74	32 (11%) 5 8	27, 49, 90, 113	0
All	All	1229/1392 (88%)	0.95	198 (16%) 1 3	27, 70, 106, 120	0

The worst 5 of 198 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	216	VAL	8.3
2	B	216	VAL	6.6
1	C	494	ILE	6.5
1	A	439	ILE	6.3
1	A	550	ARG	6.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.