W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 26, 2014 – 11:26 PM GMT

PDB ID : 2PM6
Title : Crystal Structure of yeast Sec13/31 edge element of the COPII vesicular coat, native version
Authors : Goldberg, J.; Fath, S.; Mancias, J.D.; Bi, X.
Deposited on : 2007-04-20
Resolution : 2.45 Å(reported)

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

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.45 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	Percentile Ranks	Value
Rfree		0.248
Clashscore		31
Ramachandran outliers		2.7%
Sidechain outliers		3.0%
RSRZ outliers		4.8%

Worse ◄ ► Better

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	3566 (2.50-2.42)
Clashscore	79885	4471 (2.50-2.42)
Ramachandran outliers	78287	4383 (2.50-2.42)
Sidechain outliers	78261	4385 (2.50-2.42)
RSRZ outliers	66119	3568 (2.50-2.42)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

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 hydrogen and 0 are deuterium.

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	0	0	0
			2729	1728	449	547	5			
1	C	347	Total	C	N	O	S	0	0	0
			2746	1739	452	550	5			

There are 10 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	MET	-	CLONING ARTIFACT	UNP P38968
A	368	GLY	-	CLONING ARTIFACT	UNP P38968
A	369	SER	-	CLONING ARTIFACT	UNP P38968
C	365	GLY	-	CLONING ARTIFACT	UNP P38968
C	366	ALA	-	CLONING ARTIFACT	UNP P38968
C	367	MET	-	CLONING ARTIFACT	UNP P38968
C	368	GLY	-	CLONING ARTIFACT	UNP P38968
C	369	SER	-	CLONING ARTIFACT	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	0	0	0
			2196	1403	375	415	3			
2	D	288	Total	C	N	O	S	0	0	0
			2263	1444	387	429	3			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total	O	0	0
			28	28		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	61	Total O 61 61		0	0
3	C	51	Total O 51 51		0	0
3	D	121	Total O 121 121		0	0

3　Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron 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 dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.
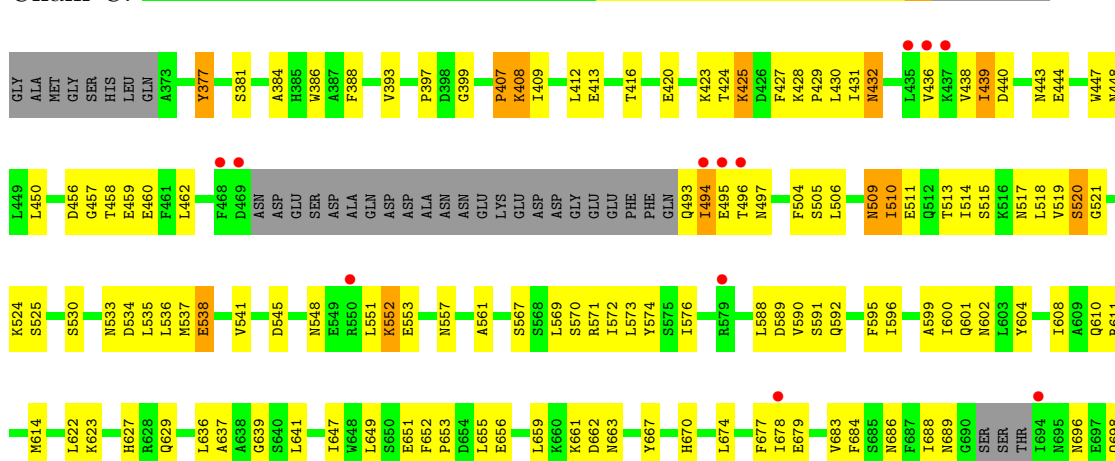
- Molecule 1: Protein transport protein SEC31
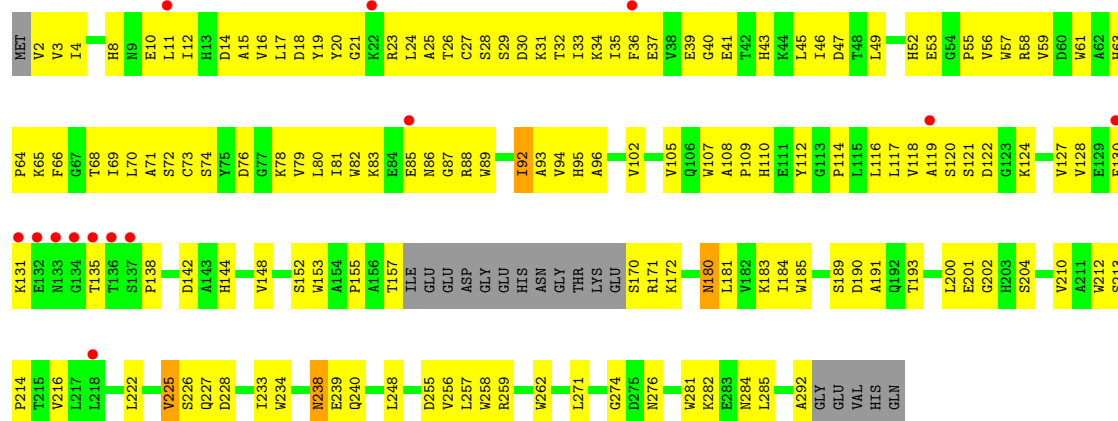
Chain A:
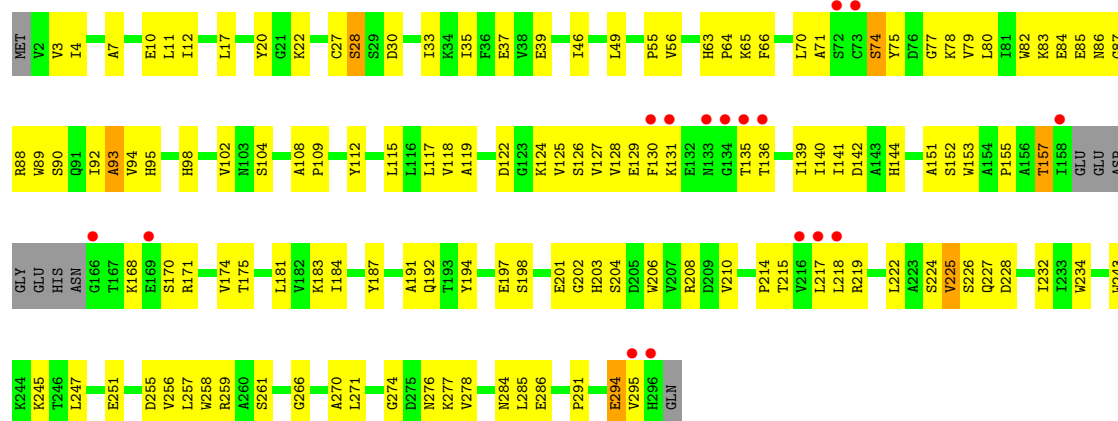
- Molecule 1: Protein transport protein SEC31

Chain C:

L699 I700 A701 K702 F703 L704 E705 F706 I707 T710 T713 G714 N715 F716 E717 P727 S728 D729 N730 E731 E732 V733 K734 K737 A738 R739 I742 G745 LYS SER LEU PRO ALA GLN ASN PRO ALA THR ALA THR THR SER LYS ALA LYS TYR

• Molecule 2: Protein transport protein SEC13

Chain B:

MET V2 V3 I4 H8 N9 E10 L11 I12 H13 D14 A15 V16 L17 D18 Y19 Y20 G21 K22 R23 L24 A25 T26 C27 S28 S29 D30 K31 T32 I33 K34 I35 F36 E37 V38 E39 G40 E41 T42 H43 K44 L45 I46 D47 T48 L49 H52 E53 G54 P55 V56 W57 R58 V59 D60 W61 A62 H63

P64 K65 F66 G67 T68 I69 L70 A71 S72 C73 S74 Y75 D76 G77 K78 V79 L80 I81 W82 K83 E84 E85 N86 G87 R88 W89 I92 A93 V94 H95 A96 V102 V105 Q106 W107 A108 P109 H110 E111 Y112 G113 P114 L115 L116 L117 V118 A119 S120 S121 D122 G123 K124 V127 V128 E129 F130

K131 E132 N133 G134 T135 T136 S137 P138 D142 A143 H144 V148 S152 W153 A154 P155 A156 T157 ILE GLU GLU GLY GLU HIS ASN GLY THR LYS GLU S170 R171 K172 N180 L181 V182 K183 I184 W185 S189 D190 A191 Q192 T193 L200 E201 H203 S204 V210 A211 W212 S213

P214 T215 V216 L217 L218 L222 V225 S226 Q227 D228 I233 W234 N238 E239 Q240 L248 D255 V256 L257 W258 R259 W262 L271 G274 D275 N276 W281 K282 E283 N284 L285 A292 GLY GLU VAL HIS GLN

• Molecule 2: Protein transport protein SEC13

Chain D:

MET V2 V3 I4 A7 E10 L11 I12 L17 Y20 G21 K22 C27 S28 S29 D30 I33 K34 I35 F36 E37 V38 E39 I46 L49 P55 V56 H63 P64 K65 F66 L70 A71 S72 C73 S74 Y75 D76 G77 K78 V79 L80 I81 W82 K83 E84 E85 N86 G87

R88 W89 S90 Q91 I92 A93 V94 H95 H98 V102 N103 S104 A108 P109 Y112 L115 L116 L117 V118 A119 D122 G123 K124 V125 S126 V127 V128 E129 F130 K131 E132 N133 G134 T135 T136 I139 I140 I141 D142 A143 H144 A151 S152 W153 A154 P155 A156 T157 I158 GLU GLU ASP

GLY GLU HIS ASN G166 T167 K168 E169 S170 R171 V174 T175 L181 V182 K183 I184 Y187 A191 Q192 T193 Y194 E197 S198 E201 G202 H203 S204 D205 W206 V207 R208 D209 V210 P214 T215 V216 L217 L218 R219 L222 A223 S224 V225 S226 Q227 D228 I232 I233 W234 W243

K244 K245 T246 L247 E251 D255 V256 L257 W258 R259 A260 S261 G266 A270 L271 G274 D275 N276 V277 V278 N284 L285 E286 P291 E294 V295 H296 GLN

4   Data and refinement statistics   (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	128.39Å 52.34Å 133.13Å 90.00° 108.62° 90.00°	Depositor
Resolution (Å)	30.00 – 2.45 26.00 – 2.45	Depositor EDS
% Data completeness (in resolution range)	88.0 (30.00-2.45) 88.0 (26.00-2.45)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	3.05 (at 2.44Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.251 , 0.301 0.256 , 0.248	Depositor DCC
R_{free} test set	2783 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	48.5	Xtriage
Anisotropy	0.612	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 45.8	EDS
Estimated twinning fraction	0.013 for l,-k,h	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	0 of 59182 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10195	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.33	0/2773	0.55	0/3744
1	C	0.38	0/2790	0.58	0/3767
2	B	0.39	0/2256	0.68	0/3079
2	D	0.44	0/2324	0.74	0/3170
All	All	0.39	0/10143	0.64	0/13760

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2729	0	2706	206	0
1	C	2746	0	2725	175	0
2	B	2196	0	2138	169	0
2	D	2263	0	2203	114	0
3	A	28	0	0	15	1
3	B	61	0	0	24	0
3	C	51	0	0	8	0
3	D	121	0	0	14	1
All	All	10195	0	9772	612	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 31.

The worst 5 of 612 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:D:129:GLU:HG2	2:D:130:PHE:H	1.14	1.08
1:C:425:LYS:HE2	1:C:686:ASN:HD21	1.21	1.06
1:C:688:ILE:HG13	1:C:689:ASN:H	1.21	1.04
1:C:408:LYS:H	1:C:408:LYS:HD3	1.22	1.04
2:D:55:PRO:HA	3:D:416:HOH:O	1.55	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	Distance(Å)	Clash(Å)
3:A:118:HOH:O	3:D:311:HOH:O[1_644]	2.02	0.18

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%)	285 (84%)	43 (13%)	11 (3%)	6	7
1	C	341/399 (86%)	298 (87%)	34 (10%)	9 (3%)	8	10
2	B	275/297 (93%)	241 (88%)	30 (11%)	4 (2%)	15	23
2	D	284/297 (96%)	255 (90%)	20 (7%)	9 (3%)	6	7
All	All	1239/1392 (89%)	1079 (87%)	127 (10%)	33 (3%)	8	9

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	497	ASN
2	B	131	LYS
1	C	715	ASN
2	D	131	LYS
2	D	218	LEU

5.3.2 Protein sidechains ⓘ

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	306/350 (87%)	297 (97%)	9 (3%)	55	79
1	C	308/350 (88%)	297 (96%)	11 (4%)	47	71
2	B	237/252 (94%)	232 (98%)	5 (2%)	66	87
2	D	244/252 (97%)	236 (97%)	8 (3%)	50	74
All	All	1095/1204 (91%)	1062 (97%)	33 (3%)	53	78

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	377	TYR
1	C	432	ASN
2	D	204	SER
1	C	381	SER
1	C	408	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	276	ASN
1	C	493	GLN
2	D	149	ASN
1	C	443	ASN
1	A	517	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates 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 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	345/399 (86%)	0.41	18 (5%)	26	26	49, 83, 114, 130	0
1	C	347/399 (86%)	0.16	12 (3%)	42	42	30, 66, 103, 130	0
2	B	279/297 (93%)	0.35	14 (5%)	28	28	15, 62, 110, 142	0
2	D	288/297 (96%)	0.13	16 (5%)	24	23	16, 47, 96, 135	0
All	All	1259/1392 (90%)	0.27	60 (4%)	29	30	15, 68, 109, 142	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	496	THR	7.8
2	B	133	ASN	7.0
2	B	136	THR	7.0
2	D	295	VAL	6.7
2	B	134	GLY	6.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.