wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2024 – 11:05 PM EST

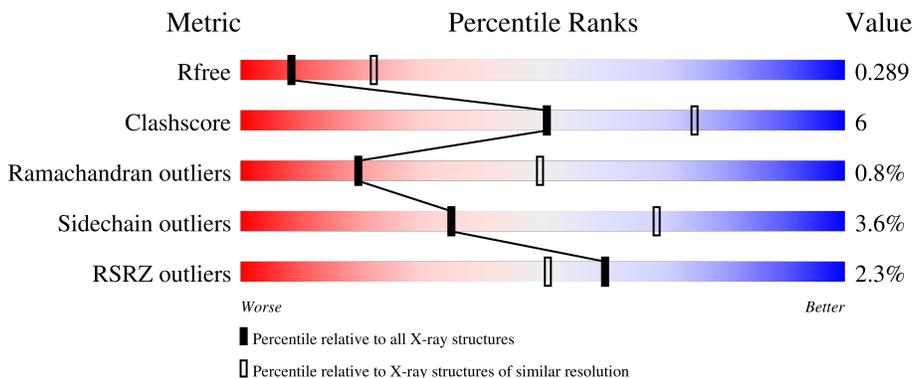|		
PDB ID	:	3MZL
Title	:	Sec13/Sec31 edge element, loop deletion mutant
Authors	:	Whittle, J.R.; Schwartz, T.U.
Deposited on	:	2010-05-12
Resolution	:	2.80 Å(reported)

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 2.80 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		0.289
Clashscore		6
Ramachandran outliers		0.8%
Sidechain outliers		3.6%
RSRZ outliers		2.3%

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}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 of the lower 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% 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	297	75% 18% • 6%
1	C	297	78% 15% • 6%
1	E	297	% 72% 21% • 6%
1	G	297	77% 17% 6%
2	B	345	6% 77% 18% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	345	3% 81% 14% 5%
2	F	345	2% 78% 18% 5%
2	H	345	4% 80% 15% 5%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 19212 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 SEC13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	279	Total	C	N	O	S	0	0	0
			2196	1403	375	415	3			
1	C	279	Total	C	N	O	S	0	0	0
			2196	1403	375	415	3			
1	E	279	Total	C	N	O	S	0	0	0
			2196	1403	375	415	3			
1	G	279	Total	C	N	O	S	0	0	0
			2196	1403	375	415	3			

- Molecule 2 is a protein called Protein transport protein SEC31.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	329	Total	C	N	O	S	0	0	0
			2607	1652	431	519	5			
2	D	329	Total	C	N	O	S	0	0	0
			2607	1652	431	519	5			
2	F	329	Total	C	N	O	S	0	0	0
			2607	1652	431	519	5			
2	H	329	Total	C	N	O	S	0	0	0
			2607	1652	431	519	5			

There are 8 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	368	GLY	-	expression tag	UNP P38968
B	369	PRO	-	expression tag	UNP P38968
D	368	GLY	-	expression tag	UNP P38968
D	369	PRO	-	expression tag	UNP P38968
F	368	GLY	-	expression tag	UNP P38968
F	369	PRO	-	expression tag	UNP P38968
H	368	GLY	-	expression tag	UNP P38968
H	369	PRO	-	expression tag	UNP P38968

WORLDWIDE PDB PROTEIN DATA BANK

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 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 SEC13
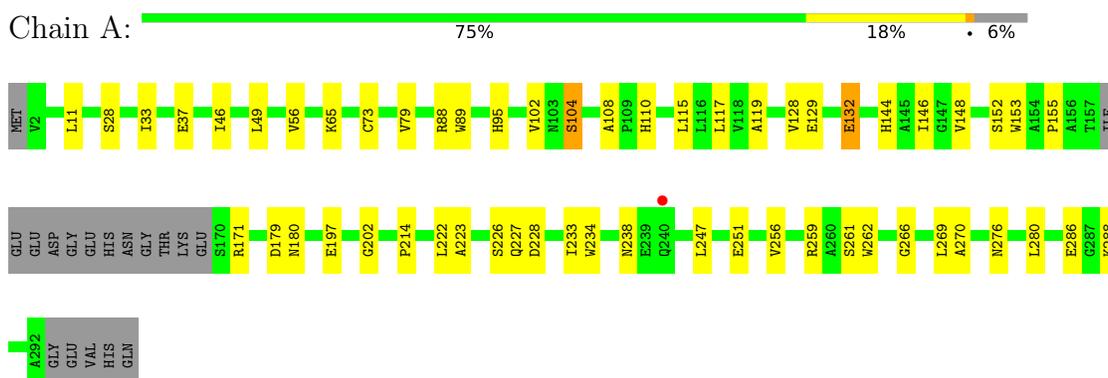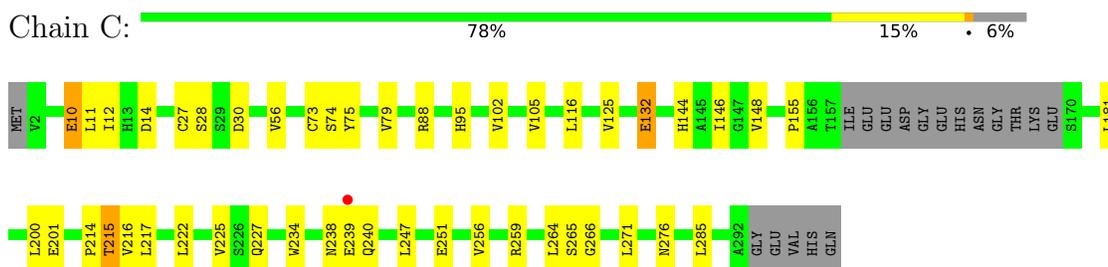
Chain A:

• Molecule 1: Protein transport protein SEC13

Chain C:

• Molecule 1: Protein transport protein SEC13

Chain E:

• Molecule 1: Protein transport protein SEC13
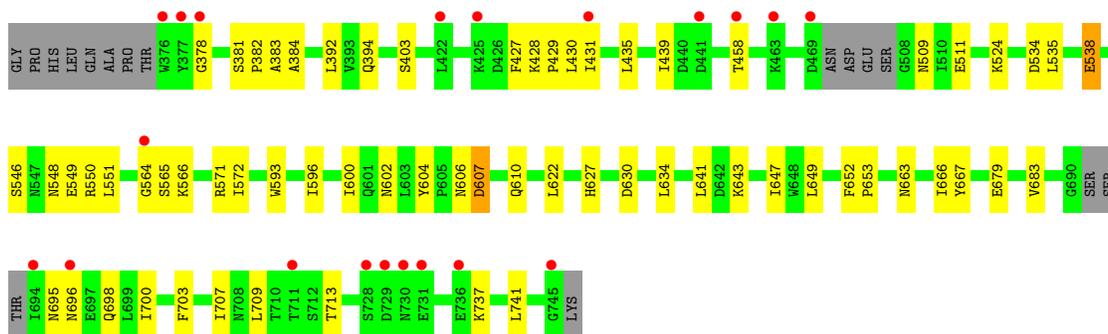
Chain G:

• Molecule 2: Protein transport protein SEC31

Chain B:

• Molecule 2: Protein transport protein SEC31

Chain D:
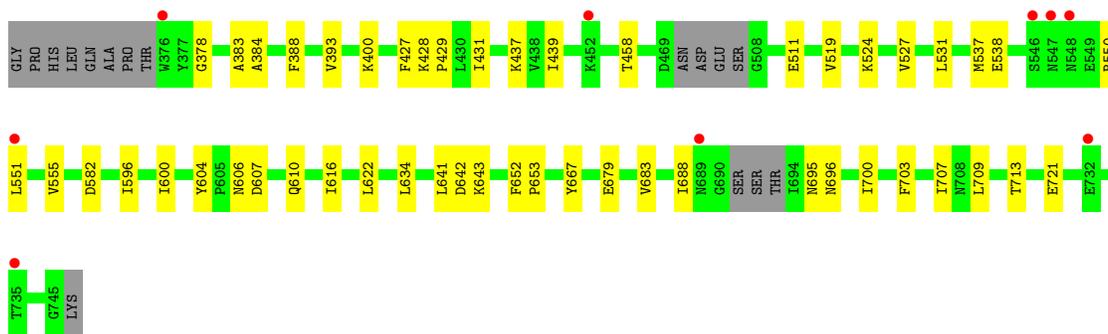
• Molecule 2: Protein transport protein SEC31

Chain F:

I694 N695 N696 E697 Q698 L699 I700 F703 I707 N708 L709 T713 G745 LYS

• Molecule 2: Protein transport protein SEC31

Chain H:

GLY PRO HIS LEU GLN ALA PRO THR W376 Y377 G378 E379 P380 A384 V393 L412 K423 T424 K425 D426 F427 K428 P429 L430 I431 N432 Q433 K437 V438 I439 N443 L449 T458 A467 F468 D469 ASN ASP GLU SER G508 Q512 K516 D534 L544 E549

R550 Y559 S565 R571 V590 S591 Q592 W593 I596 I600 N606 I616 L622 K623 L634 K643 I647 W648 L649 F652 P653 Y667 F687 I688 N689 G690 SER SER THR I694 N695 Q698 K702 F703 I707 N708 L709 T713 F716

E721 N730 E731 K734 K737 V740 G745 LYS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	156.22Å 46.62Å 192.00Å 90.00° 93.48° 90.00°	Depositor
Resolution (Å)	29.38 – 2.80 29.38 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.6 (29.38-2.80) 83.4 (29.38-2.80)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.12 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.267 , 0.300 0.261 , 0.289	Depositor DCC
R_{free} test set	1792 reflections (2.58%)	wwPDB-VP
Wilson B-factor (Å^2)	48.6	Xtriage
Anisotropy	0.591	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 8.8	EDS
L-test for twinning[2]	$<|L|> = 0.34, <L^2> = 0.17$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Reported twinning fraction	0.772 for H, K, L 0.228 for -h,-k,l	Depositor
Outliers	0 of 69354 reflections	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	19212	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.31 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.2494e-03.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.35	0/2256	0.53	0/3079
1	C	0.34	0/2256	0.52	0/3079
1	E	0.35	0/2256	0.52	0/3079
1	G	0.34	0/2256	0.51	0/3079
2	B	0.34	0/2647	0.49	0/3571
2	D	0.33	0/2647	0.48	0/3571
2	F	0.34	0/2647	0.48	0/3571
2	H	0.34	0/2647	0.48	0/3571
All	All	0.34	0/19612	0.50	0/26600

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 ⓘ

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	2196	0	2138	29	0
1	C	2196	0	2138	24	0
1	E	2196	0	2138	38	0
1	G	2196	0	2138	29	0
2	B	2607	0	2601	33	0
2	D	2607	0	2601	25	0
2	F	2607	0	2601	32	0
2	H	2607	0	2601	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	19212	0	18956	217	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 6.

The worst 5 of 217 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)
1:G:222:LEU:HB2	1:G:234:TRP:HB2	1.71	0.72
1:E:95:HIS:CE1	1:E:128:VAL:HG21	2.25	0.71
1:C:10:GLU:HG3	1:C:30:ASP:HB3	1.71	0.71
1:G:33:ILE:HD11	1:G:56:VAL:HG11	1.72	0.70
2:H:623:LYS:HD3	2:H:647:ILE:HD12	1.73	0.69

There are no symmetry-related clashes.

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	275/297 (93%)	259 (94%)	14 (5%)	2 (1%)	22	53
1	C	275/297 (93%)	259 (94%)	15 (6%)	1 (0%)	34	66
1	E	275/297 (93%)	258 (94%)	13 (5%)	4 (2%)	10	33
1	G	275/297 (93%)	260 (94%)	13 (5%)	2 (1%)	22	53
2	B	323/345 (94%)	304 (94%)	15 (5%)	4 (1%)	13	39
2	D	323/345 (94%)	300 (93%)	21 (6%)	2 (1%)	25	56
2	F	323/345 (94%)	299 (93%)	22 (7%)	2 (1%)	25	56
2	H	323/345 (94%)	303 (94%)	18 (6%)	2 (1%)	25	56
All	All	2392/2568 (93%)	2242 (94%)	131 (6%)	19 (1%)	19	49

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	146	ILE
2	D	378	GLY
2	F	380	PRO
2	H	378	GLY
2	H	380	PRO

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	237/252 (94%)	228 (96%)	9 (4%)	33	67
1	C	237/252 (94%)	225 (95%)	12 (5%)	24	55
1	E	237/252 (94%)	227 (96%)	10 (4%)	30	63
1	G	237/252 (94%)	230 (97%)	7 (3%)	41	75
2	B	292/306 (95%)	282 (97%)	10 (3%)	37	71
2	D	292/306 (95%)	282 (97%)	10 (3%)	37	71
2	F	292/306 (95%)	282 (97%)	10 (3%)	37	71
2	H	292/306 (95%)	284 (97%)	8 (3%)	44	78
All	All	2116/2232 (95%)	2040 (96%)	76 (4%)	35	69

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	572	ILE
2	H	590	VAL
2	F	630	ASP
1	G	240	GLN
2	H	721	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	509	ASN
2	H	695	ASN
2	F	695	ASN
2	H	509	ASN
2	F	625	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 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 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	279/297 (93%)	-0.05	1 (0%)	92	91	25, 39, 59, 73	0
1	C	279/297 (93%)	0.01	1 (0%)	92	91	25, 38, 58, 73	0
1	E	279/297 (93%)	-0.04	3 (1%)	80	75	25, 38, 59, 73	0
1	G	279/297 (93%)	-0.08	1 (0%)	92	91	25, 38, 58, 73	0
2	B	329/345 (95%)	0.50	20 (6%)	21	13	34, 59, 80, 98	0
2	D	329/345 (95%)	0.32	9 (2%)	54	44	34, 59, 80, 98	0
2	F	329/345 (95%)	0.36	8 (2%)	59	49	34, 59, 80, 98	0
2	H	329/345 (95%)	0.41	14 (4%)	35	25	34, 59, 80, 98	0
All	All	2432/2568 (94%)	0.20	57 (2%)	60	51	25, 51, 74, 98	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	730	ASN	6.2
2	B	745	GLY	5.4
2	B	729	ASP	5.1
2	B	731	GLU	4.8
2	B	376	TRP	4.8

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 monosaccharides in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.4　Ligands ⓘ

There are no ligands in this entry.

6.5　Other polymers ⓘ

There are no such residues in this entry.