



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 30, 2017 – 08:30 PM EST

PDB ID : 5M4Y
Title : Crystal structure of the Sec3/Sso2 complex at 2.20 angstrom resolution
Authors : Zhang, Y.B.; Dong, G.
Deposited on : 2016-10-19
Resolution : 2.20 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

<http://wwpdb.org/validation/2016/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
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028442
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028442

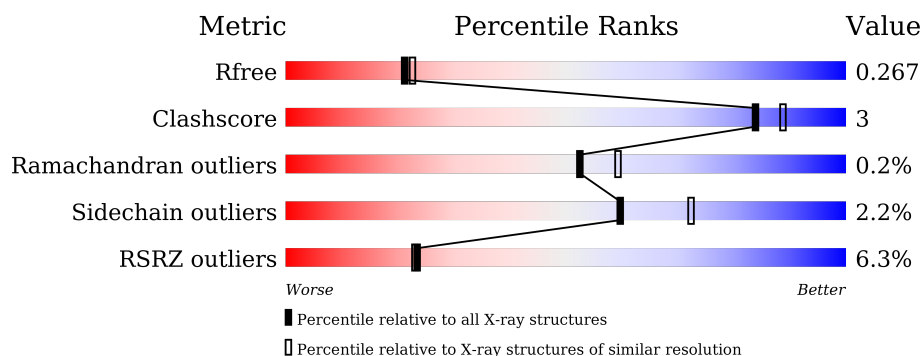
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	210	<div> <div>9%</div> <div>74% 10% • 16%</div> </div>
1	C	210	<div> <div>6%</div> <div>79% 7% 14%</div> </div>
1	E	210	<div> <div>14%</div> <div>75% 10% • 14%</div> </div>
2	B	250	<div> <div>2%</div> <div>65% • 31%</div> </div>
2	D	250	<div> <div>%</div> <div>62% 7% 31%</div> </div>
2	F	250	<div> <div></div> <div>65% • 32%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	D	401	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9010 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 SSO2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	177	Total	C	N	O	S	0	0	0
			1445	888	253	298	6			
1	C	181	Total	C	N	O	S	0	0	0
			1476	904	261	305	6			
1	E	181	Total	C	N	O	S	0	0	0
			1476	907	258	305	6			

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	GLY	-	expression tag	UNP P39926
A	33	SER	-	expression tag	UNP P39926
A	34	HIS	-	expression tag	UNP P39926
A	35	MET	-	expression tag	UNP P39926
A	228	LEU	-	expression tag	UNP P39926
A	229	THR	-	expression tag	UNP P39926
A	230	GLN	-	expression tag	UNP P39926
A	231	LEU	-	expression tag	UNP P39926
A	232	PHE	-	expression tag	UNP P39926
A	233	ASN	-	expression tag	UNP P39926
A	234	ASP	-	expression tag	UNP P39926
A	235	MET	-	expression tag	UNP P39926
A	236	GLU	-	expression tag	UNP P39926
A	237	GLU	-	expression tag	UNP P39926
A	238	LEU	-	expression tag	UNP P39926
A	239	VAL	-	expression tag	UNP P39926
A	240	ILE	-	expression tag	UNP P39926
A	241	GLU	-	expression tag	UNP P39926
C	32	GLY	-	expression tag	UNP P39926
C	33	SER	-	expression tag	UNP P39926
C	34	HIS	-	expression tag	UNP P39926
C	35	MET	-	expression tag	UNP P39926
C	228	LEU	-	expression tag	UNP P39926

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	229	THR	-	expression tag	UNP P39926
C	230	GLN	-	expression tag	UNP P39926
C	231	LEU	-	expression tag	UNP P39926
C	232	PHE	-	expression tag	UNP P39926
C	233	ASN	-	expression tag	UNP P39926
C	234	ASP	-	expression tag	UNP P39926
C	235	MET	-	expression tag	UNP P39926
C	236	GLU	-	expression tag	UNP P39926
C	237	GLU	-	expression tag	UNP P39926
C	238	LEU	-	expression tag	UNP P39926
C	239	VAL	-	expression tag	UNP P39926
C	240	ILE	-	expression tag	UNP P39926
C	241	GLU	-	expression tag	UNP P39926
E	32	GLY	-	expression tag	UNP P39926
E	33	SER	-	expression tag	UNP P39926
E	34	HIS	-	expression tag	UNP P39926
E	35	MET	-	expression tag	UNP P39926
E	228	LEU	-	expression tag	UNP P39926
E	229	THR	-	expression tag	UNP P39926
E	230	GLN	-	expression tag	UNP P39926
E	231	LEU	-	expression tag	UNP P39926
E	232	PHE	-	expression tag	UNP P39926
E	233	ASN	-	expression tag	UNP P39926
E	234	ASP	-	expression tag	UNP P39926
E	235	MET	-	expression tag	UNP P39926
E	236	GLU	-	expression tag	UNP P39926
E	237	GLU	-	expression tag	UNP P39926
E	238	LEU	-	expression tag	UNP P39926
E	239	VAL	-	expression tag	UNP P39926
E	240	ILE	-	expression tag	UNP P39926
E	241	GLU	-	expression tag	UNP P39926

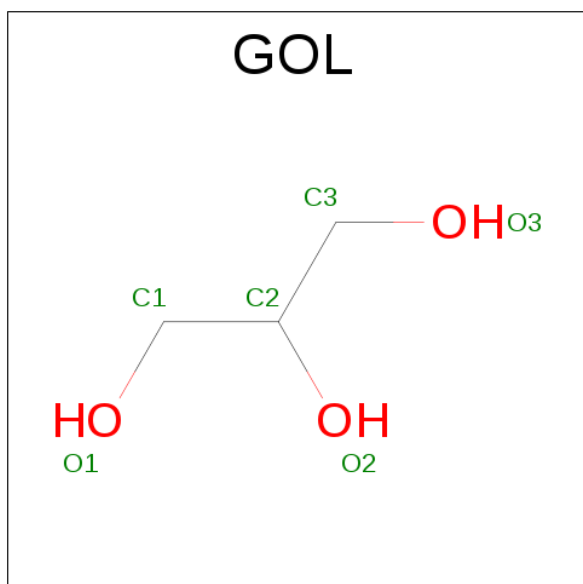
- Molecule 2 is a protein called Exocyst complex component SEC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	173	Total	C	N	O	S	0	0	0
			1433	914	250	266	3			
2	D	173	Total	C	N	O	S	0	0	0
			1433	914	250	266	3			
2	F	171	Total	C	N	O	S	0	0	0
			1416	904	246	263	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	71	GLN	-	expression tag	UNP P33332
B	72	GLY	-	expression tag	UNP P33332
B	73	HIS	-	expression tag	UNP P33332
B	74	MET	-	expression tag	UNP P33332
D	71	GLN	-	expression tag	UNP P33332
D	72	GLY	-	expression tag	UNP P33332
D	73	HIS	-	expression tag	UNP P33332
D	74	MET	-	expression tag	UNP P33332
F	71	GLN	-	expression tag	UNP P33332
F	72	GLY	-	expression tag	UNP P33332
F	73	HIS	-	expression tag	UNP P33332
F	74	MET	-	expression tag	UNP P33332

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	42	Total	O	0	0
			42	42		
4	B	80	Total	O	0	0
			80	80		
4	C	29	Total	O	0	0
			29	29		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	76	Total 76	O 76	0	0
4	E	20	Total 20	O 20	0	0
4	F	78	Total 78	O 78	0	0

PHE	SER	ALA	PRO	THR	GLU	ARG	THR	ARG	ARG	SER	GLU	THR	GLU	SER	VAL	ASN	PRO	VAL	SER	THR	THR	ALA	ALA	SER	VAL	GLU	TYR	HIS	ALA	GLY	MET	LYS	SER	LEU	ASN	LYS	ALA	PRO	TYR	SER	SER	ASN	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

● Molecule 2: Exocyst complex component SEC3



GLN	GLY	HIS	MET	S75	R86	R99	LYS	THR	G102	Y108	D117	K130	I131	E132	M133	K134	H146	W170	E179	K180	I184	S185	E186	M192	K211	E222	S234	Y237	Y243	V247	I248	T249	ASN	ARG	PRO	PRO	GLY	SER	ALA	VAL	TYR	SER	PRO	ILE	LYS	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PRO	THR	SER	ASN	PHE	THR	THR	ASN	THR	THR	GLN	SER	VAL	GLY	SER	VAL	PRO	PHE	SER	ALA	PRO	THR	GLU	ARG	THR	ARG	ARG	SER	GLU	THR	GLU	SER	VAL	ASN	PRO	VAL	SER	THR	PRO	ALA	ALA	VAL	VAL	GLU	TYR	HIS	ALA	GLY	MET	LYS	SER	LEU	ASN	PRO	GLY	SER	ALA	VAL	TYR	SER	PRO	SER	ASN	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

● Molecule 2: Exocyst complex component SEC3



GLN	GLY	HIS	MET	S75	E83	K87	D98	HIS	LYS	THR	G102	I114	S122	L139	V176	M192	S193	E222	I248	THR	ASN	ARG	ARG	PRO	GLY	SER	VAL	SER	SER	PRO	ILE	LYS	SER	PRO	THR	SER	ASN	PHE	THR	THR	GLN	SER	SER	VAL	GLY	SER	VAL	PRO	PRO	PHE
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

SER	ALA	PRO	THR	ARG	THR	ARG	ARG	SER	GLU	THR	GLU	SER	VAL	ASN	PRO	VAL	SER	THR	PRO	ALA	ALA	VAL	GLU	TYR	HIS	ALA	GLY	MET	LYS	SER	LEU	ASN	LYS	ALA	PRO	TYR	SER	SER	SER	ASN	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	101.14Å 135.80Å 185.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.98 – 2.20 49.22 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.98-2.20) 99.7 (49.22-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.96 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.229 , 0.264 0.229 , 0.267	Depositor DCC
R_{free} test set	1735 reflections (2.67%)	DCC
Wilson B-factor (Å ²)	50.3	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 39.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9010	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.24% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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	0.23	0/1458	0.33	0/1958
1	C	0.22	0/1489	0.33	0/1998
1	E	0.23	0/1489	0.35	0/2000
2	B	0.24	0/1465	0.42	0/1978
2	D	0.24	0/1465	0.42	0/1978
2	F	0.24	0/1447	0.41	0/1953
All	All	0.23	0/8813	0.38	0/11865

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	1445	0	1408	10	0
1	C	1476	0	1434	7	0
1	E	1476	0	1438	11	0
2	B	1433	0	1433	7	0
2	D	1433	0	1433	13	0
2	F	1416	0	1419	5	0
3	D	6	0	8	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	42	0	0	1	0
4	B	80	0	0	2	0
4	C	29	0	0	0	0
4	D	76	0	0	3	0
4	E	20	0	0	1	0
4	F	78	0	0	2	0
All	All	9010	0	8573	51	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 3.

The worst 5 of 51 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:A:170:ASP:HB2	1:A:171:VAL:HG13	1.67	0.76
2:F:193:SER:OG	4:F:401:HOH:O	2.06	0.74
2:D:179:GLU:OE1	4:D:501:HOH:O	2.06	0.73
2:B:146:ASN:OD1	4:B:401:HOH:O	2.06	0.73
2:D:211:LYS:NZ	4:D:503:HOH:O	2.23	0.71

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 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	173/210 (82%)	169 (98%)	3 (2%)	1 (1%)	30	29
1	C	177/210 (84%)	174 (98%)	3 (2%)	0	100	100
1	E	177/210 (84%)	171 (97%)	6 (3%)	0	100	100
2	B	169/250 (68%)	165 (98%)	4 (2%)	0	100	100
2	D	169/250 (68%)	164 (97%)	5 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	167/250 (67%)	162 (97%)	4 (2%)	1 (1%)	30	29
All	All	1032/1380 (75%)	1005 (97%)	25 (2%)	2 (0%)	52	59

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	172	ASN
2	F	222	GLU

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	159/187 (85%)	152 (96%)	7 (4%)	35	42
1	C	161/187 (86%)	158 (98%)	3 (2%)	65	77
1	E	162/187 (87%)	155 (96%)	7 (4%)	35	43
2	B	161/230 (70%)	161 (100%)	0	100	100
2	D	161/230 (70%)	159 (99%)	2 (1%)	78	88
2	F	159/230 (69%)	157 (99%)	2 (1%)	76	87
All	All	963/1251 (77%)	942 (98%)	21 (2%)	60	72

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	174	GLN
2	D	247	VAL
1	E	205	LEU
1	C	172	ASN
1	E	218	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	GOL	D	401	-	5,5,5	0.35	0	5,5,5	0.18	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	D	401	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	401	GOL	1	0

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	177/210 (84%)	0.53	18 (10%) 9 8	35, 54, 99, 112	0
1	C	181/210 (86%)	0.35	12 (6%) 22 21	34, 52, 99, 112	0
1	E	181/210 (86%)	0.64	29 (16%) 3 2	37, 56, 120, 130	0
2	B	173/250 (69%)	0.12	5 (2%) 55 54	34, 43, 74, 117	0
2	D	173/250 (69%)	-0.07	2 (1%) 81 80	29, 40, 63, 116	0
2	F	171/250 (68%)	-0.10	1 (0%) 90 90	30, 40, 62, 81	0
All	All	1056/1380 (76%)	0.25	67 (6%) 23 23	29, 46, 100, 130	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	197	VAL	5.9
1	A	35	MET	5.6
1	E	191	LYS	5.6
1	E	156	GLN	5.4
1	A	171	VAL	5.3

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	GOL	D	401	6/6	0.54	0.42	6.15	61,64,67,75	0

6.5 Other polymers [i](#)

There are no such residues in this entry.