



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 01:27 am BST

PDB ID : 5KNL
Title : Crystal structure of S. pombe ubiquitin E1 (Uba1) in complex with Ubc15 and ubiquitin
Authors : Olsen, S.K.; Lv, Z.; Yuan, L.; Williams, K.
Deposited on : 2016-06-28
Resolution : 2.50 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

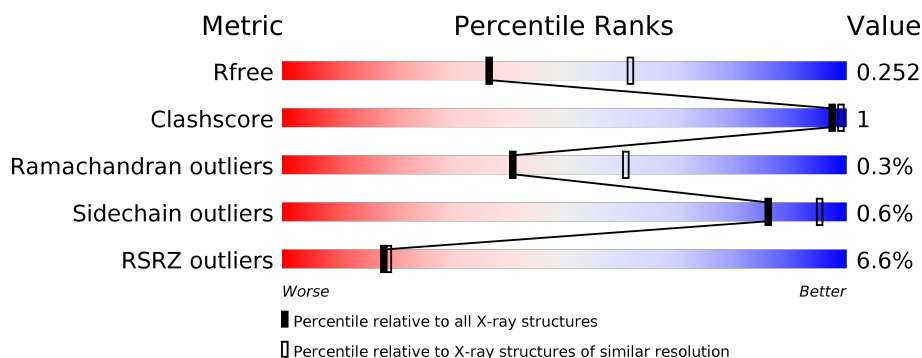
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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 $\leq 5\%$. 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	1001	
1	D	1001	
2	B	96	
2	E	96	
3	C	175	
3	F	175	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 38561 atoms, of which 19138 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 Ubiquitin-activating enzyme E1 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	981	Total	C	H	N	O	S	0	0	0
			15342	4926	7637	1259	1481	39			
1	D	977	Total	C	H	N	O	S	0	0	0
			15280	4907	7607	1254	1473	39			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	SER	-	expression tag	UNP O94609
D	12	SER	-	expression tag	UNP O94609

- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	76	Total	C	H	N	O	S	0	0	0
			1236	375	623	119	118	1			
2	E	76	Total	C	H	N	O	S	0	0	0
			1236	375	623	119	118	1			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	initiating methionine	UNP A0A081CFF0
B	-18	GLY	-	expression tag	UNP A0A081CFF0
B	-17	SER	-	expression tag	UNP A0A081CFF0
B	-16	SER	-	expression tag	UNP A0A081CFF0
B	-15	HIS	-	expression tag	UNP A0A081CFF0
B	-14	HIS	-	expression tag	UNP A0A081CFF0
B	-13	HIS	-	expression tag	UNP A0A081CFF0
B	-12	HIS	-	expression tag	UNP A0A081CFF0
B	-11	HIS	-	expression tag	UNP A0A081CFF0
B	-10	HIS	-	expression tag	UNP A0A081CFF0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	SER	-	expression tag	UNP A0A081CFF0
B	-8	SER	-	expression tag	UNP A0A081CFF0
B	-7	GLY	-	expression tag	UNP A0A081CFF0
B	-6	LEU	-	expression tag	UNP A0A081CFF0
B	-5	VAL	-	expression tag	UNP A0A081CFF0
B	-4	PRO	-	expression tag	UNP A0A081CFF0
B	-3	ARG	-	expression tag	UNP A0A081CFF0
B	-2	GLY	-	expression tag	UNP A0A081CFF0
B	-1	SER	-	expression tag	UNP A0A081CFF0
B	0	HIS	-	expression tag	UNP A0A081CFF0
B	6	ARG	LYS	engineered mutation	UNP A0A081CFF0
B	11	ARG	LYS	engineered mutation	UNP A0A081CFF0
B	27	ARG	LYS	engineered mutation	UNP A0A081CFF0
B	29	ARG	LYS	engineered mutation	UNP A0A081CFF0
B	33	ARG	LYS	engineered mutation	UNP A0A081CFF0
B	48	ARG	LYS	engineered mutation	UNP A0A081CFF0
B	57	ALA	SER	engineered mutation	UNP A0A081CFF0
B	63	ARG	LYS	engineered mutation	UNP A0A081CFF0
E	-19	MET	-	initiating methionine	UNP A0A081CFF0
E	-18	GLY	-	expression tag	UNP A0A081CFF0
E	-17	SER	-	expression tag	UNP A0A081CFF0
E	-16	SER	-	expression tag	UNP A0A081CFF0
E	-15	HIS	-	expression tag	UNP A0A081CFF0
E	-14	HIS	-	expression tag	UNP A0A081CFF0
E	-13	HIS	-	expression tag	UNP A0A081CFF0
E	-12	HIS	-	expression tag	UNP A0A081CFF0
E	-11	HIS	-	expression tag	UNP A0A081CFF0
E	-10	HIS	-	expression tag	UNP A0A081CFF0
E	-9	SER	-	expression tag	UNP A0A081CFF0
E	-8	SER	-	expression tag	UNP A0A081CFF0
E	-7	GLY	-	expression tag	UNP A0A081CFF0
E	-6	LEU	-	expression tag	UNP A0A081CFF0
E	-5	VAL	-	expression tag	UNP A0A081CFF0
E	-4	PRO	-	expression tag	UNP A0A081CFF0
E	-3	ARG	-	expression tag	UNP A0A081CFF0
E	-2	GLY	-	expression tag	UNP A0A081CFF0
E	-1	SER	-	expression tag	UNP A0A081CFF0
E	0	HIS	-	expression tag	UNP A0A081CFF0
E	6	ARG	LYS	engineered mutation	UNP A0A081CFF0
E	11	ARG	LYS	engineered mutation	UNP A0A081CFF0
E	27	ARG	LYS	engineered mutation	UNP A0A081CFF0
E	29	ARG	LYS	engineered mutation	UNP A0A081CFF0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	33	ARG	LYS	engineered mutation	UNP A0A081CFF0
E	48	ARG	LYS	engineered mutation	UNP A0A081CFF0
E	57	ALA	SER	engineered mutation	UNP A0A081CFF0
E	63	ARG	LYS	engineered mutation	UNP A0A081CFF0

- Molecule 3 is a protein called Ubiquitin-conjugating enzyme E2 15.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	171	Total	C	H	N	O	S	0	0	0
			2718	877	1346	230	258	7			
3	F	162	Total	C	H	N	O	S	0	1	0
			2610	840	1302	222	239	7			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	168	LEU	-	expression tag	UNP Q9Y818
C	169	GLU	-	expression tag	UNP Q9Y818
C	170	HIS	-	expression tag	UNP Q9Y818
C	171	HIS	-	expression tag	UNP Q9Y818
C	172	HIS	-	expression tag	UNP Q9Y818
C	173	HIS	-	expression tag	UNP Q9Y818
C	174	HIS	-	expression tag	UNP Q9Y818
C	175	HIS	-	expression tag	UNP Q9Y818
F	168	LEU	-	expression tag	UNP Q9Y818
F	169	GLU	-	expression tag	UNP Q9Y818
F	170	HIS	-	expression tag	UNP Q9Y818
F	171	HIS	-	expression tag	UNP Q9Y818
F	172	HIS	-	expression tag	UNP Q9Y818
F	173	HIS	-	expression tag	UNP Q9Y818
F	174	HIS	-	expression tag	UNP Q9Y818
F	175	HIS	-	expression tag	UNP Q9Y818

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		
4	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is water.

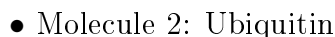
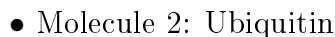
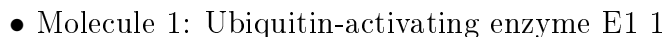
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	32	Total	O	0	0
			32	32		
5	B	7	Total	O	0	0
			7	7		
5	C	6	Total	O	0	0
			6	6		

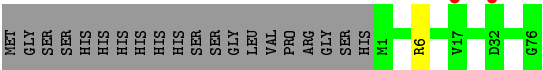
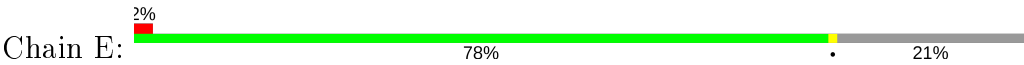
Continued on next page...

Continued from previous page...

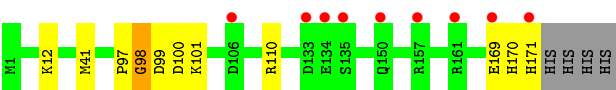
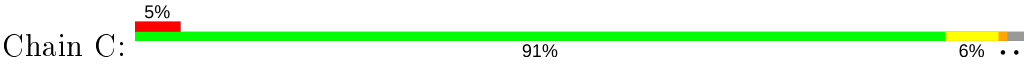
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	46	Total 46	O 46	0	0
5	E	2	Total 2	O 2	0	0
5	F	1	Total 1	O 1	0	0

- Molecule 1: Ubiquitin-activating enzyme E1 1

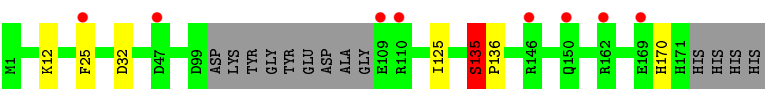
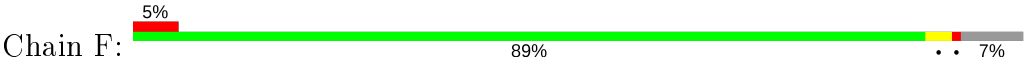




• Molecule 3: Ubiquitin-conjugating enzyme E2 15



• Molecule 3: Ubiquitin-conjugating enzyme E2 15



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	77.11Å 82.24Å 135.38Å 102.10° 95.78° 90.86°	Depositor
Resolution (Å)	28.31 – 2.50 28.41 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.9 (28.31-2.50) 96.9 (28.41-2.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.08 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.216 , 0.251 0.219 , 0.252	Depositor DCC
R_{free} test set	10829 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å ²)	29.5	Xtriage
Anisotropy	0.260	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 35.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	38561	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.22	0/7876	0.37	0/10660
1	D	0.23	0/7844	0.37	0/10618
2	B	0.22	0/618	0.43	0/831
2	E	0.22	0/618	0.40	0/831
3	C	0.23	0/1411	0.40	0/1915
3	F	0.24	0/1348	0.46	1/1829 (0.1%)
All	All	0.23	0/19715	0.39	1/26684 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	F	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	135	SER	C-N-CD	-5.85	107.73	120.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	F	135	SER	Peptide

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	7705	7637	7634	10	0
1	D	7673	7607	7607	11	0
2	B	613	623	625	3	0
2	E	613	623	625	1	0
3	C	1372	1346	1341	5	0
3	F	1308	1302	1298	4	0
4	A	20	0	0	0	0
4	B	5	0	0	1	0
4	D	15	0	0	0	0
4	F	5	0	0	0	0
5	A	32	0	0	0	0
5	B	7	0	0	0	0
5	C	6	0	0	0	0
5	D	46	0	0	0	0
5	E	2	0	0	1	0
5	F	1	0	0	0	0
All	All	19423	19138	19130	31	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 1.

The worst 5 of 31 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:D:684:GLN:O	1:D:688:ASN:ND2	2.32	0.61
2:B:33:ARG:NH1	4:B:101:SO4:O3	2.36	0.59
1:D:1009:CYS:SG	1:D:1011:LYS:NZ	2.68	0.57
1:D:13:ASN:OD1	1:D:14:THR:N	2.37	0.57
1:A:997:ASP:O	1:A:999:ASN:N	2.40	0.55

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	977/1001 (98%)	923 (94%)	52 (5%)	2 (0%)	47	68
1	D	973/1001 (97%)	923 (95%)	49 (5%)	1 (0%)	51	73
2	B	74/96 (77%)	71 (96%)	3 (4%)	0	100	100
2	E	74/96 (77%)	71 (96%)	3 (4%)	0	100	100
3	C	169/175 (97%)	160 (95%)	6 (4%)	3 (2%)	8	14
3	F	159/175 (91%)	152 (96%)	5 (3%)	2 (1%)	12	21
All	All	2426/2544 (95%)	2300 (95%)	118 (5%)	8 (0%)	41	61

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	170	HIS
1	A	786	GLU
1	A	998	ALA
3	C	170	HIS
1	D	998	ALA

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	860/877 (98%)	853 (99%)	7 (1%)	81	93
1	D	857/877 (98%)	853 (100%)	4 (0%)	88	96
2	B	67/84 (80%)	66 (98%)	1 (2%)	65	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	67/84 (80%)	67 (100%)	0	100	100
3	C	154/160 (96%)	153 (99%)	1 (1%)	86	95
3	F	149/160 (93%)	149 (100%)	0	100	100
All	All	2154/2242 (96%)	2141 (99%)	13 (1%)	86	95

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	999	ASN
1	A	1002	ASP
1	D	999	ASN
1	A	841	ARG
1	D	80	GLN

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

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](#)

9 ligands are 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	$\# Z > 2$
4	SO4	D	1102	-	4,4,4	0.14	0	6,6,6	0.05	0
4	SO4	A	1101	-	4,4,4	0.15	0	6,6,6	0.06	0
4	SO4	B	101	-	4,4,4	0.14	0	6,6,6	0.05	0
4	SO4	A	1104	-	4,4,4	0.14	0	6,6,6	0.06	0
4	SO4	F	201	-	4,4,4	0.13	0	6,6,6	0.05	0
4	SO4	D	1103	-	4,4,4	0.14	0	6,6,6	0.06	0
4	SO4	D	1101	-	4,4,4	0.13	0	6,6,6	0.07	0
4	SO4	A	1103	-	4,4,4	0.14	0	6,6,6	0.07	0
4	SO4	A	1102	-	4,4,4	0.15	0	6,6,6	0.05	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
4	B	101	SO4	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	981/1001 (98%)	0.35	68 (6%) 16 17	11, 44, 103, 165	0
1	D	977/1001 (97%)	0.40	75 (7%) 13 13	12, 43, 111, 179	0
2	B	76/96 (79%)	0.05	0 100 100	22, 45, 88, 108	0
2	E	76/96 (79%)	0.35	2 (2%) 56 59	21, 61, 104, 146	0
3	C	171/175 (97%)	0.28	9 (5%) 26 28	22, 49, 98, 140	0
3	F	162/175 (92%)	0.41	8 (4%) 29 31	30, 57, 102, 146	1 (0%)
All	All	2443/2544 (96%)	0.36	162 (6%) 18 19	11, 45, 104, 179	1 (0%)

The worst 5 of 162 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	199	GLU	8.2
1	D	998	ALA	7.6
1	A	1000	GLY	7.2
1	A	198	GLU	6.2
1	D	911	LYS	6.0

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. 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	B-factors(\AA^2)	Q<0.9
4	SO4	F	201	5/5	0.86	0.16	129,129,130,131	0
4	SO4	B	101	5/5	0.90	0.17	99,100,101,101	0
4	SO4	D	1102	5/5	0.93	0.18	92,92,94,95	0
4	SO4	D	1101	5/5	0.95	0.13	55,56,60,61	0
4	SO4	A	1101	5/5	0.96	0.16	46,49,50,51	0
4	SO4	D	1103	5/5	0.98	0.09	37,39,44,44	0
4	SO4	A	1104	5/5	0.98	0.09	34,36,39,47	0
4	SO4	A	1103	5/5	0.98	0.09	49,51,51,53	0
4	SO4	A	1102	5/5	0.98	0.09	38,43,44,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.