



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 18, 2021 – 10:01 pm BST

PDB ID : 7NW1
Title : Crystal structure of UFC1 in complex with UBA5
Authors : Manoj Kumar, P.; Padala, P.; Isupov, M.N.; Wiener, R.
Deposited on : 2021-03-16
Resolution : 1.95 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 : 2.23.1
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) : 2.23.1

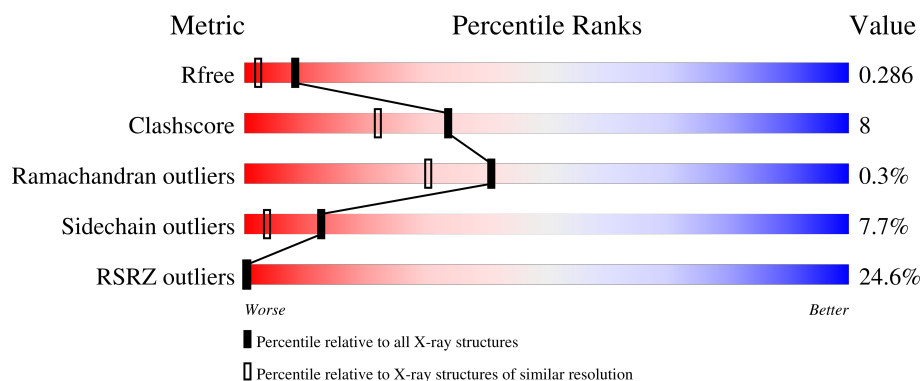
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



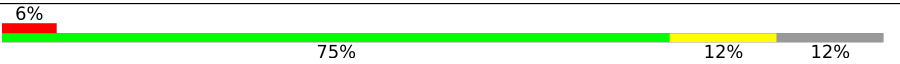
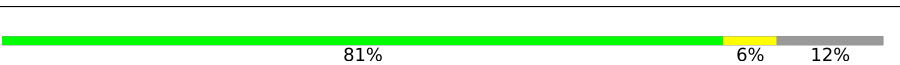
The reported resolution of this entry is 1.95 Å.

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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 $\leq 5\%$. 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	AAA	168	
1	BBB	168	
2	CCC	16	
2	FFF	16	

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	EDO	AAA	208	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 3051 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 Ubiquitin-fold modifier-conjugating enzyme 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	163	Total	C	N	O	S	0	0	0
			1343	867	232	239	5			
1	BBB	163	Total	C	N	O	S	0	0	0
			1343	867	232	239	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	GLY	-	expression tag	UNP Q9Y3C8
BBB	0	GLY	-	expression tag	UNP Q9Y3C8

- Molecule 2 is a protein called Ubiquitin-like modifier-activating enzyme 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	CCC	14	Total	C	N	O	S	0	0	0
			108	65	17	23	3			
2	FFF	14	Total	C	N	O	S	0	0	0
			108	65	17	23	3			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



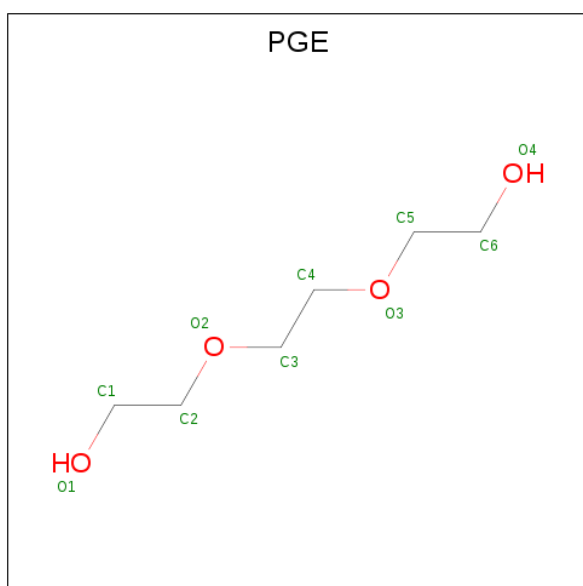
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	BBB	1	Total	C	O	0	0
			4	2	2		
3	BBB	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	C	O	0	0
			7	4	3		
4	AAA	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total	C	O	0	0
			10	6	4		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

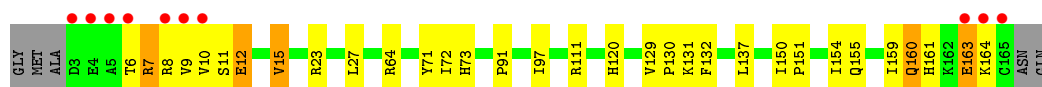
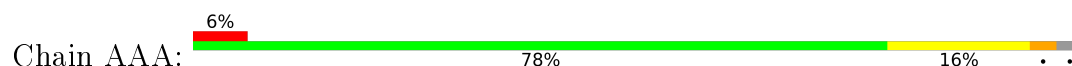
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	AAA	32	Total	O	0	0
			32	32		
7	BBB	15	Total	O	0	0
			15	15		
7	CCC	2	Total	O	0	0
			2	2		
7	FFF	2	Total	O	0	0
			2	2		

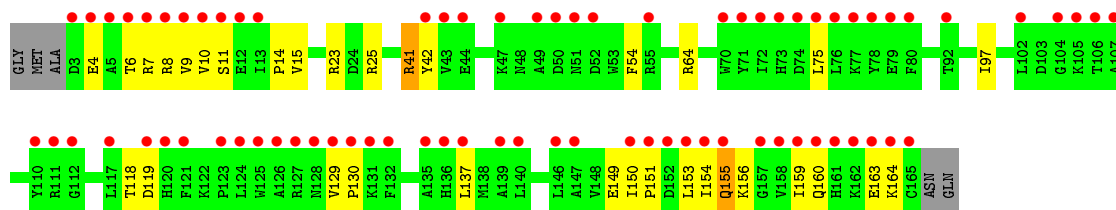
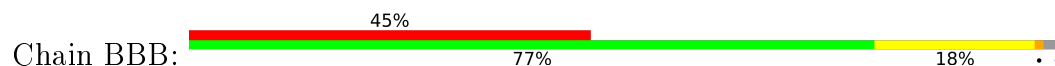
3 Residue-property plots [i](#)

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.

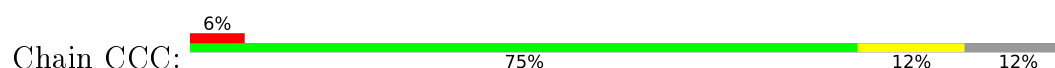
- Molecule 1: Ubiquitin-fold modifier-conjugating enzyme 1



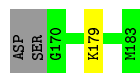
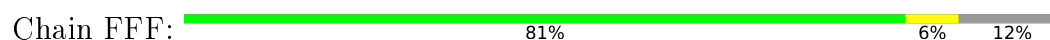
- Molecule 1: Ubiquitin-fold modifier-conjugating enzyme 1



- Molecule 2: Ubiquitin-like modifier-activating enzyme 5



- Molecule 2: Ubiquitin-like modifier-activating enzyme 5



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	47.20 Å 67.42 Å 133.98 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.76 – 1.95 33.71 – 1.95	Depositor EDS
% Data completeness (in resolution range)	47.1 (23.76-1.95) 47.0 (33.71-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.03 (at 1.95 Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.215 , 0.280 0.215 , 0.286	Depositor DCC
R_{free} test set	818 reflections (5.41%)	wwPDB-VP
Wilson B-factor (Å ²)	37.3	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 65.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3051	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.51% 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: PGE, PEG, GOL, EDO

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	AAA	0.47	0/1380	0.82	0/1871
1	BBB	0.32	0/1380	0.71	0/1871
2	CCC	0.43	0/107	0.70	0/137
2	FFF	0.61	0/107	0.93	0/137
All	All	0.41	0/2974	0.77	0/4016

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	AAA	1343	0	1337	34	0
1	BBB	1343	0	1337	16	0
2	CCC	108	0	109	2	0
2	FFF	108	0	109	0	0
3	AAA	36	0	54	15	0
3	BBB	8	0	12	0	0
4	AAA	14	0	20	0	0
5	AAA	10	0	14	0	0
6	AAA	30	0	40	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	AAA	32	0	0	0	0
7	BBB	15	0	0	1	0
7	CCC	2	0	0	0	0
7	FFF	2	0	0	0	0
All	All	3051	0	3032	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 8.

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:AAA:159:ILE:HG22	6:AAA:212:GOL:H11	1.51	0.93
1:AAA:132:PHE:H	3:AAA:208:EDO:H11	1.35	0.89
1:AAA:132:PHE:H	3:AAA:208:EDO:C1	1.88	0.86
1:AAA:132:PHE:N	3:AAA:208:EDO:H11	1.94	0.81
1:AAA:154:ILE:HG13	1:AAA:159:ILE:HD11	1.60	0.81

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	AAA	161/168 (96%)	155 (96%)	6 (4%)	0	100	100
1	BBB	161/168 (96%)	153 (95%)	7 (4%)	1 (1%)	25	14
2	CCC	12/16 (75%)	12 (100%)	0	0	100	100
2	FFF	12/16 (75%)	11 (92%)	1 (8%)	0	100	100
All	All	346/368 (94%)	331 (96%)	14 (4%)	1 (0%)	41	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	25	ARG

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	AAA	143/146 (98%)	129 (90%)	14 (10%)	8	2
1	BBB	143/146 (98%)	134 (94%)	9 (6%)	18	7
2	CCC	12/14 (86%)	12 (100%)	0	100	100
2	FFF	12/14 (86%)	11 (92%)	1 (8%)	11	3
All	All	310/320 (97%)	286 (92%)	24 (8%)	13	4

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

Mol	Chain	Res	Type
1	BBB	4	GLU
1	BBB	15	VAL
1	BBB	9	VAL
1	BBB	23	ARG
1	AAA	27	LEU

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry

19 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	$\# Z > 2$
6	GOL	AAA	209	-	5,5,5	0.12	0	5,5,5	0.34	0
3	EDO	BBB	201	-	3,3,3	0.20	0	2,2,2	0.45	0
3	EDO	AAA	204	-	3,3,3	0.19	0	2,2,2	0.03	0
6	GOL	AAA	210	-	5,5,5	0.15	0	5,5,5	0.44	0
4	PEG	AAA	205	-	6,6,6	0.15	0	5,5,5	0.14	0
6	GOL	AAA	211	-	5,5,5	0.10	0	5,5,5	0.28	0
3	EDO	AAA	202	-	3,3,3	0.04	0	2,2,2	0.09	0
4	PEG	AAA	203	-	6,6,6	0.31	0	5,5,5	0.28	0
3	EDO	AAA	217	-	3,3,3	0.24	0	2,2,2	0.42	0
3	EDO	AAA	201	-	3,3,3	0.08	0	2,2,2	0.48	0
3	EDO	AAA	207	-	3,3,3	0.10	0	2,2,2	0.25	0
6	GOL	AAA	212	-	5,5,5	0.18	0	5,5,5	0.51	0
3	EDO	AAA	214	-	3,3,3	0.12	0	2,2,2	0.18	0
3	EDO	AAA	215	-	3,3,3	0.09	0	2,2,2	0.14	0
6	GOL	AAA	213	-	5,5,5	0.14	0	5,5,5	0.35	0
3	EDO	AAA	216	-	3,3,3	0.10	0	2,2,2	0.20	0
5	PGE	AAA	206	-	9,9,9	0.40	0	8,8,8	0.37	0
3	EDO	BBB	202	-	3,3,3	0.21	0	2,2,2	0.17	0
3	EDO	AAA	208	-	3,3,3	0.27	0	2,2,2	0.70	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
6	GOL	AAA	209	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	BBB	201	-	-	1/1/1/1	-
3	EDO	AAA	204	-	-	1/1/1/1	-
6	GOL	AAA	210	-	-	4/4/4/4	-
4	PEG	AAA	205	-	-	1/4/4/4	-
6	GOL	AAA	211	-	-	2/4/4/4	-
3	EDO	AAA	202	-	-	1/1/1/1	-
4	PEG	AAA	203	-	-	3/4/4/4	-
3	EDO	AAA	217	-	-	1/1/1/1	-
3	EDO	AAA	201	-	-	0/1/1/1	-
3	EDO	AAA	207	-	-	1/1/1/1	-
6	GOL	AAA	212	-	-	0/4/4/4	-
3	EDO	AAA	214	-	-	0/1/1/1	-
3	EDO	AAA	215	-	-	1/1/1/1	-
6	GOL	AAA	213	-	-	0/4/4/4	-
3	EDO	AAA	216	-	-	0/1/1/1	-
5	PGE	AAA	206	-	-	3/7/7/7	-
3	EDO	BBB	202	-	-	1/1/1/1	-
3	EDO	AAA	208	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	AAA	210	GOL	O1-C1-C2-O2
6	AAA	210	GOL	O1-C1-C2-C3
6	AAA	210	GOL	C1-C2-C3-O3
6	AAA	211	GOL	C1-C2-C3-O3
5	AAA	206	PGE	O2-C3-C4-O3

There are no ring outliers.

3 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	AAA	212	GOL	1	0
6	AAA	213	GOL	2	0
3	AAA	208	EDO	15	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	AAA	163/168 (97%)	0.63	10 (6%) 21 29	18, 29, 98, 158	0
1	BBB	163/168 (97%)	2.72	76 (46%) 0 0	34, 82, 138, 173	0
2	CCC	14/16 (87%)	0.73	1 (7%) 16 24	41, 54, 85, 85	0
2	FFF	14/16 (87%)	0.38	0 100 100	26, 43, 59, 69	0
All	All	354/368 (96%)	1.58	87 (24%) 0 0	18, 51, 122, 173	0

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	165	CYS	16.6
1	AAA	165	CYS	11.6
1	BBB	6	THR	11.0
1	BBB	158	VAL	9.9
1	BBB	121	PHE	9.8

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 monosaccharides 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
6	GOL	AAA	209	6/6	0.79	0.15	53,58,61,61	0
4	PEG	AAA	203	7/7	0.80	0.23	34,47,55,57	7
3	EDO	AAA	217	4/4	0.80	0.15	44,45,50,54	4
3	EDO	AAA	216	4/4	0.81	0.16	53,58,59,69	0
5	PGE	AAA	206	10/10	0.82	0.17	45,48,55,59	0
3	EDO	BBB	202	4/4	0.83	0.20	48,53,59,61	4
3	EDO	AAA	214	4/4	0.86	0.20	62,65,67,71	0
3	EDO	AAA	202	4/4	0.87	0.20	55,58,61,61	0
3	EDO	AAA	207	4/4	0.89	0.13	47,48,50,53	4
3	EDO	AAA	215	4/4	0.89	0.17	65,69,73,77	0
6	GOL	AAA	210	6/6	0.89	0.13	36,44,50,54	0
4	PEG	AAA	205	7/7	0.92	0.19	48,54,57,58	7
6	GOL	AAA	212	6/6	0.92	0.15	36,46,49,54	0
6	GOL	AAA	213	6/6	0.92	0.17	36,47,50,54	0
3	EDO	BBB	201	4/4	0.93	0.12	60,61,66,69	0
6	GOL	AAA	211	6/6	0.94	0.15	37,48,55,55	0
3	EDO	AAA	201	4/4	0.94	0.12	37,39,43,47	0
3	EDO	AAA	204	4/4	0.94	0.14	41,47,51,55	0
3	EDO	AAA	208	4/4	0.95	0.24	38,39,40,43	0

6.5 Other polymers [i](#)

There are no such residues in this entry.