



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 22, 2023 – 06:15 pm GMT

PDB ID : 7QUM  
Title : ATAD2 in complex with FragLite2  
Authors : Turberville, S.; Martin, M.P.; Hope, I.; Noble, M.E.M.  
Deposited on : 2022-01-18  
Resolution : 1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.32.1  
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) : 2.32.1

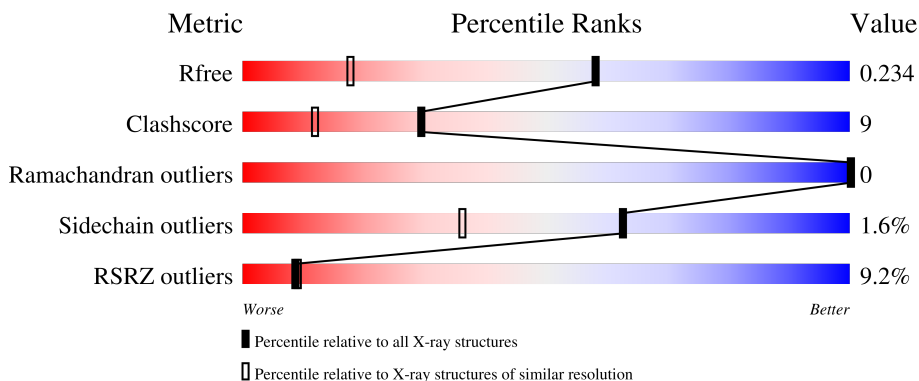
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	130	

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
5	PYZ	AAA	1210	-	-	X	-

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 2501 atoms, of which 1167 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.

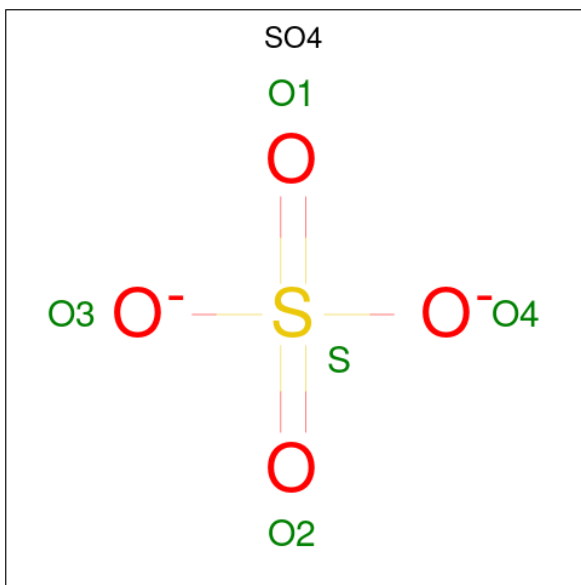
- Molecule 1 is a protein called ATPase family AAA domain-containing protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	AAA	130	2258	705	1128	201	219	5	43	5	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	979	SER	-	expression tag	UNP Q6PL18
AAA	980	MET	-	expression tag	UNP Q6PL18

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).

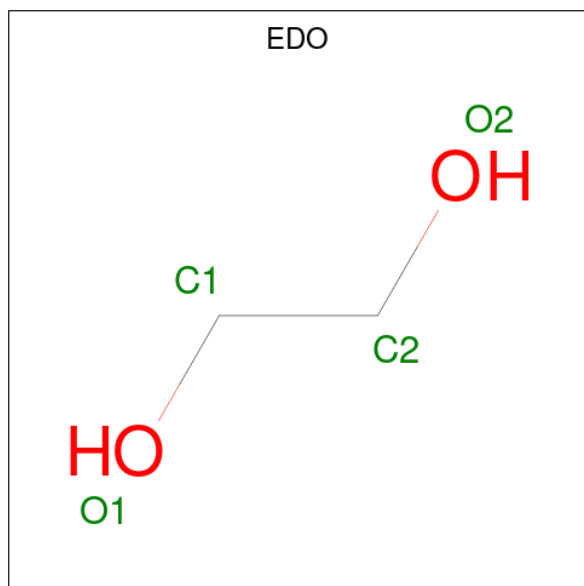


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total	O	S	0	0
			5	4	1		
2	AAA	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

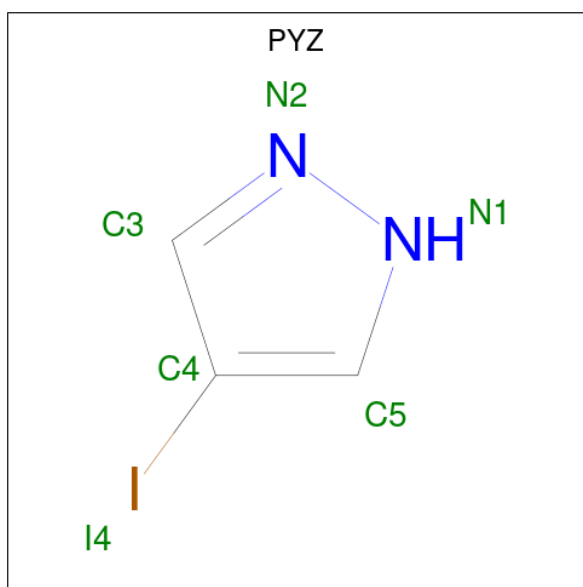
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total Cl 1 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total C H O 10 2 6 2	1	0
4	AAA	1	Total C H O 10 2 6 2	1	0
4	AAA	1	Total C H O 10 2 6 2	1	0
4	AAA	1	Total C H O 10 2 6 2	1	0
4	AAA	1	Total C H O 10 2 6 2	1	0
4	AAA	1	Total C H O 10 2 6 2	1	0

- Molecule 5 is 4-IODOPYRAZOLE (three-letter code: PYZ) (formula: C<sub>3</sub>H<sub>3</sub>IN<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	I	N		
5	AAA	1	9	3	3	1	2	0	0

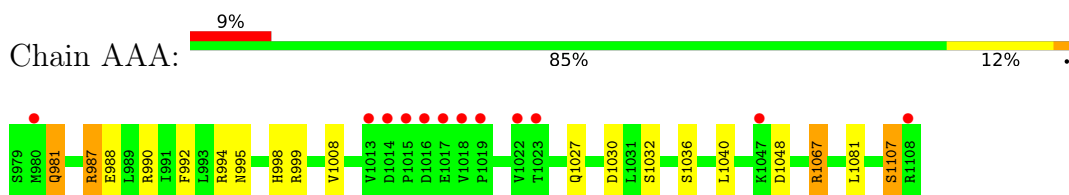
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	163	Total	O	0	0
			163	163		

### 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: ATPase family AAA domain-containing protein 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.16Å 79.16Å 138.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	61.42 – 1.50 61.42 – 1.50	Depositor EDS
% Data completeness (in resolution range)	96.5 (61.42-1.50) 96.4 (61.42-1.50)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 1.50Å)	Xtrriage
Refinement program	REFMAC 5.8.0258, REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.197 , 0.223 0.207 , 0.234	Depositor DCC
$R_{free}$ test set	2055 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.5	Xtrriage
Anisotropy	0.030	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 52.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2501	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.82% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: SO4, PYZ, CL, 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	0.95	0/1147	1.13	7/1544 (0.5%)

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
1	AAA	0	1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	AAA	1067	ARG	NE-CZ-NH1	-8.67	115.96	120.30
1	AAA	987[A]	ARG	NE-CZ-NH2	-6.42	117.09	120.30
1	AAA	987[B]	ARG	NE-CZ-NH2	-6.42	117.09	120.30
1	AAA	990	ARG	NE-CZ-NH2	-6.40	117.10	120.30
1	AAA	987[A]	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	AAA	987[B]	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	AAA	994	ARG	CG-CD-NE	-5.18	100.93	111.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	1107[B]	SER	Mainchain



## 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	1130	1128	1119	19	0
2	AAA	10	0	0	0	0
3	AAA	1	0	0	1	0
4	AAA	24	36	30	7	0
5	AAA	6	3	3	5	2
6	AAA	163	0	0	9	0
All	All	1334	1167	1152	22	2

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 9.

All (22) 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)
4:AAA:1206:EDO:H21	6:AAA:1391:HOH:O	1.71	0.88
1:AAA:1027[B]:GLN:OE1	6:AAA:1302:HOH:O	1.92	0.86
3:AAA:1202:CL:CL	6:AAA:1383:HOH:O	2.30	0.85
1:AAA:995:ASN:CB	5:AAA:1210:PYZ:C3	2.48	0.83
4:AAA:1206:EDO:C2	6:AAA:1391:HOH:O	2.32	0.70
1:AAA:987[B]:ARG:HD2	6:AAA:1332:HOH:O	1.92	0.69
1:AAA:1030:ASP:OD2	1:AAA:1032:SER:OG	2.12	0.63
1:AAA:995:ASN:ND2	6:AAA:1306:HOH:O	2.32	0.61
1:AAA:998:HIS:HD2	4:AAA:1204:EDO:H11	1.66	0.59
1:AAA:988:GLU:OE2	4:AAA:1206:EDO:H12	2.03	0.59
1:AAA:995:ASN:HB3	5:AAA:1210:PYZ:C3	2.31	0.58
1:AAA:1107[B]:SER:OG	6:AAA:1303:HOH:O	2.18	0.53
1:AAA:1081[B]:LEU:C	1:AAA:1081[B]:LEU:HD23	2.29	0.52
1:AAA:998:HIS:CD2	4:AAA:1204:EDO:H11	2.47	0.49
1:AAA:1048:ASP:OD2	6:AAA:1304:HOH:O	2.21	0.45
1:AAA:1036:SER:HB3	4:AAA:1208:EDO:H12	1.99	0.44
1:AAA:999:ARG:NH1	5:AAA:1210:PYZ:N2	2.65	0.44
1:AAA:992:PHE:CD1	5:AAA:1210:PYZ:H5	2.52	0.44
1:AAA:1008:VAL:HG11	6:AAA:1342:HOH:O	2.19	0.41
1:AAA:981:GLN:NE2	1:AAA:981:GLN:H	2.16	0.41
1:AAA:995:ASN:CB	5:AAA:1210:PYZ:H3	2.13	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:1040:LEU:HD12	4:AAA:1208:EDO:H11	2.02	0.40

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
5:AAA:1210:PYZ:C3	5:AAA:1210:PYZ:C4[12_564]	1.90	0.30
5:AAA:1210:PYZ:C3	5:AAA:1210:PYZ:C3[12_564]	1.99	0.21

## 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	AAA	132/130 (102%)	128 (97%)	4 (3%)	0	<b>100</b> <b>100</b>

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	128/123 (104%)	126 (98%)	2 (2%)	<b>62</b> <b>36</b>

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	981	GLN
1	AAA	1067	ARG

Sometimes 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 1 is monoatomic - leaving 9 for Mogul analysis.

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	EDO	AAA	1207	-	3,3,3	0.42	0	2,2,2	0.66	0
2	SO4	AAA	1201	-	4,4,4	0.61	0	6,6,6	0.61	0
4	EDO	AAA	1208	-	3,3,3	0.11	0	2,2,2	0.26	0
5	PYZ	AAA	1210	-	6,6,6	0.73	0	2,7,7	2.76	2 (100%)
4	EDO	AAA	1204	-	3,3,3	0.13	0	2,2,2	0.28	0
2	SO4	AAA	1209	-	4,4,4	0.37	0	6,6,6	0.05	0
4	EDO	AAA	1206	-	3,3,3	0.30	0	2,2,2	0.22	0
4	EDO	AAA	1203	-	3,3,3	0.42	0	2,2,2	0.32	0
4	EDO	AAA	1205	-	3,3,3	0.54	0	2,2,2	0.41	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
4	EDO	AAA	1207	-	-	0/1/1/1	-
4	EDO	AAA	1208	-	-	1/1/1/1	-
5	PYZ	AAA	1210	-	-	-	0/1/1/1
4	EDO	AAA	1204	-	-	1/1/1/1	-
4	EDO	AAA	1206	-	-	0/1/1/1	-
4	EDO	AAA	1203	-	-	0/1/1/1	-
4	EDO	AAA	1205	-	-	1/1/1/1	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	AAA	1210	PYZ	C3-C4-I4	-2.95	119.61	127.53
5	AAA	1210	PYZ	C5-C4-I4	2.57	134.42	127.53

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AAA	1204	EDO	O1-C1-C2-O2
4	AAA	1205	EDO	O1-C1-C2-O2
4	AAA	1208	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	AAA	1208	EDO	2	0
5	AAA	1210	PYZ	5	2
4	AAA	1204	EDO	2	0
4	AAA	1206	EDO	3	0

## 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	130/130 (100%)	0.81	12 (9%) <b>9</b> <b>9</b>	18, 28, 61, 91	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	1018	VAL	7.1
1	AAA	1019	PRO	6.5
1	AAA	1016	ASP	4.7
1	AAA	1015	PRO	4.7
1	AAA	1014	ASP	4.5
1	AAA	1017	GLU	4.5
1	AAA	1013	VAL	3.9
1	AAA	980	MET	3.8
1	AAA	1022	VAL	3.0
1	AAA	1108[A]	ARG	2.5
1	AAA	1023	THR	2.4
1	AAA	1047	LYS	2.2

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	EDO	AAA	1204	4/4	0.62	0.39	60,70,72,72	1
4	EDO	AAA	1207	4/4	0.80	0.20	20,48,52,54	1
3	CL	AAA	1202	1/1	0.87	0.15	70,70,70,70	0
4	EDO	AAA	1208	4/4	0.87	0.23	20,57,64,64	1
4	EDO	AAA	1206	4/4	0.88	0.55	20,42,45,49	1
2	SO4	AAA	1209	5/5	0.89	0.14	73,74,76,79	5
5	PYZ	AAA	1210	6/6	0.93	0.49	33,71,103,113	2
4	EDO	AAA	1203	4/4	0.94	0.12	28,30,37,38	1
4	EDO	AAA	1205	4/4	0.95	0.20	20,34,36,39	1
2	SO4	AAA	1201	5/5	0.98	0.11	30,32,36,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.