



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 21, 2024 – 06:04 PM EST

PDB ID : 8UL4  
Title : Structure of rsKiirro using SSX after illumination with 6.74 mJ/mm<sup>2</sup> of 405 nm light  
Authors : Baxter, J.M.; van Thor, J.J.  
Deposited on : 2023-10-16  
Resolution : 1.75 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](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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

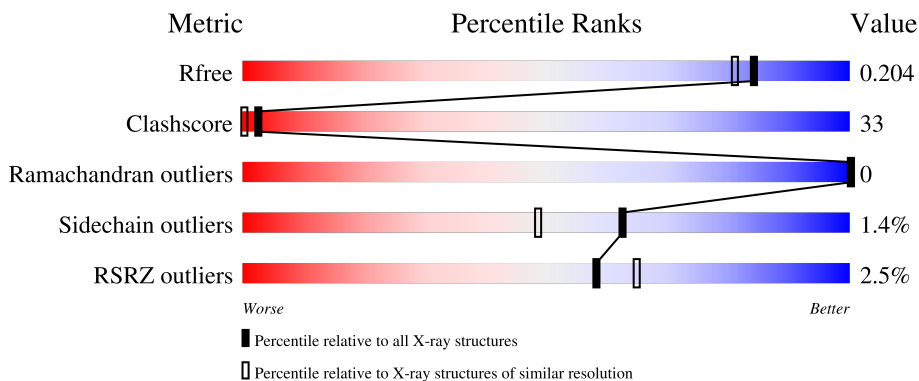
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


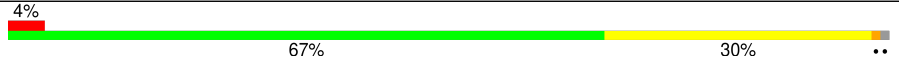
The reported resolution of this entry is 1.75 Å.

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))
$R_{free}$	164625	2888 (1.76-1.76)
Clashscore	180529	3097 (1.76-1.76)
Ramachandran outliers	177936	3072 (1.76-1.76)
Sidechain outliers	177891	3072 (1.76-1.76)
RSRZ outliers	164620	2887 (1.76-1.76)

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	1	220	 73% 24% ..
2	2	220	 67% 30% ..

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
2	A1BE5	2	64	-	-	X	-
4	GOL	2	301	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7286 atoms, of which 3406 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 rsKiiro cis structure.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	1	218	3453	1121	1699	298	326	9	49	0	0

- Molecule 2 is a protein called rsKiiro cis structure.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	2	218	3453	1121	1699	298	326	9	49	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



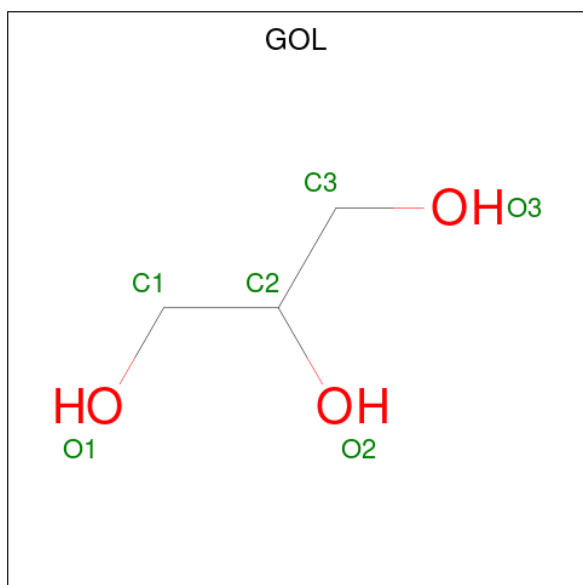
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
3	1	1	5	4	1	0	0
3	1	1	5	4	1	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	2	1	Total	O	S	0	0
			5	4	1		
3	2	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	2	1	Total	C	H	O	2	0
			14	3	8	3		

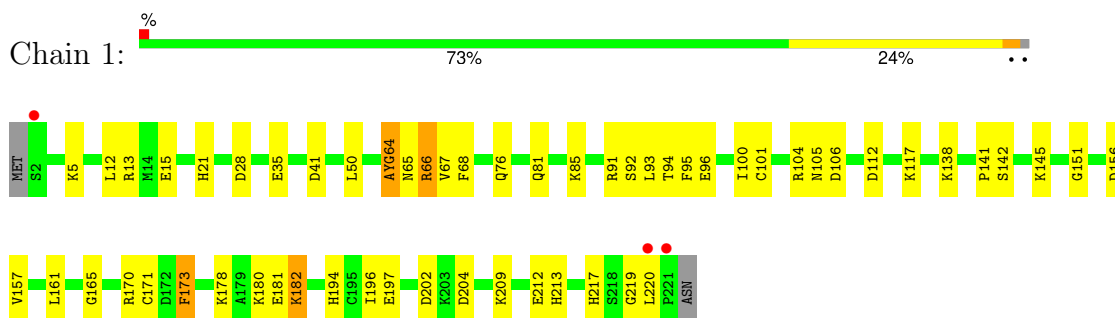
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	1	171	Total	O	0	0
			171	171		
5	2	175	Total	O	0	0
			175	175		

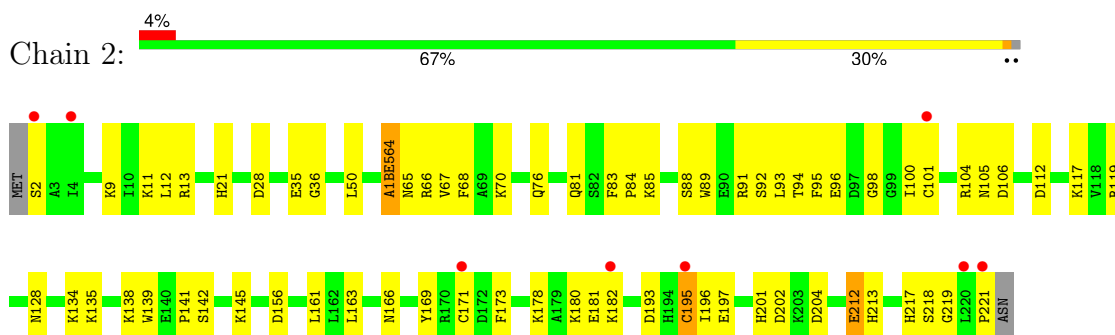
### 3 Residue-property plots [i](#)

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: rsKiiri cis structure



- Molecule 2: rsKiiri cis structure



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	39.60Å 74.50Å 78.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	54.23 – 1.75 54.17 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.6 (54.23-1.75) 99.6 (54.17-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.15 (at 1.75Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.169 , 0.203 0.169 , 0.204	Depositor DCC
$R_{free}$ test set	1160 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.7	Xtrriage
Anisotropy	0.020	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 53.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	7286	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.98% 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, GOL, PIA, A1BE5

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	1	0.65	0/1779	0.87	2/2397 (0.1%)
2	2	0.60	0/1779	0.71	0/2397
All	All	0.62	0/3558	0.80	2/4794 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	1	173	PHE	CB-CA-C	-5.98	98.45	110.40
1	1	173	PHE	N-CA-CB	5.03	119.65	110.60

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	1	1754	1699	1598	131	0
2	2	1754	1699	1584	177	0
3	1	10	0	0	0	0
3	2	10	0	0	0	0
4	2	6	8	8	4	0
5	1	171	0	0	50	0
5	2	175	0	0	57	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	3880	3406	3190	224	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 33.

The worst 5 of 224 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)
1:1:12:LEU:CB	5:1:552:HOH:O	1.69	1.36
2:2:81:GLN:CG	5:2:406:HOH:O	1.66	1.35
5:1:408:HOH:O	2:2:13:ARG:HD3	1.28	1.30
2:2:218:SER:HB2	5:2:535:HOH:O	1.31	1.27
1:1:85:LYS:NZ	2:2:181:GLU:OE2	1.64	1.26

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	1	213/220 (97%)	211 (99%)	2 (1%)	0	100	100
2	2	213/220 (97%)	207 (97%)	6 (3%)	0	100	100
All	All	426/440 (97%)	418 (98%)	8 (2%)	0	100	100

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	1	185/187 (99%)	182 (98%)	3 (2%)	58	42
2	2	185/187 (99%)	183 (99%)	2 (1%)	70	58
All	All	370/374 (99%)	365 (99%)	5 (1%)	62	49

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	1	66	ARG
1	1	67	VAL
1	1	182	LYS
2	2	195	CYS
2	2	212	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	1	76	GLN
1	1	206	ASN
2	2	65	ASN
2	2	116	ASN
2	2	201	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues 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
2	A1BE5	2	64	2	20,21,22	0.58	0	28,29,31	1.15	3 (10%)
1	PIA	1	64	1	20,21,22	0.73	0	28,29,31	1.02	3 (10%)

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
2	A1BE5	2	64	2	-	2/8/27/28	0/2/2/2
1	PIA	1	64	1	-	2/8/27/28	0/2/2/2

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	2	64	A1BE5	CG2-CB2-CA2	2.74	133.12	129.87
2	2	64	A1BE5	C2-CA2-N2	-2.40	107.23	108.95
1	1	64	PIA	N3-C1-N2	2.30	113.29	111.48
2	2	64	A1BE5	CB2-CA2-C2	2.20	125.02	122.36
1	1	64	PIA	CB1-CA1-C1	2.06	114.33	111.43

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	1	64	PIA	C3-CA3-N3-C2
2	2	64	A1BE5	N2-CA2-CB2-CG2
2	2	64	A1BE5	C2-CA2-CB2-CG2
1	1	64	PIA	C3-CA3-N3-C1

There are no ring outliers.

2 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	2	64	A1BE5	10	0
1	1	64	PIA	6	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

5 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	$\# Z  > 2$
3	SO4	1	302	-	4,4,4	0.32	0	6,6,6	0.08	0
3	SO4	1	301	-	4,4,4	0.31	0	6,6,6	0.10	0
4	GOL	2	301	-	5,5,5	0.13	0	5,5,5	0.42	0
3	SO4	2	303	-	4,4,4	0.34	0	6,6,6	0.07	0
3	SO4	2	302	-	4,4,4	0.33	0	6,6,6	0.09	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	GOL	2	301	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	2	301	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	2	301	GOL	4	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	1	2
2	2	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	1	64:PIA	C3	65:ASN	N	1.69
1	1	61:PHE	C	64:PIA	N1	1.63
1	2	64:A1BE5	C3	65:ASN	N	0.98
1	2	61:PHE	C	64:A1BE5	N1	0.83

## 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	1	217/220 (98%)	-0.35	3 (1%) 73 79	10, 15, 33, 103	217 (100%)
2	2	217/220 (98%)	-0.00	8 (3%) 45 52	5, 6, 9, 18	217 (100%)
All	All	434/440 (98%)	-0.18	11 (2%) 58 65	5, 11, 26, 103	434 (100%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	2	101	CYS	7.3
1	1	221	PRO	7.2
2	2	171	CYS	6.0
2	2	221	PRO	5.3
2	2	182	LYS	5.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
2	A1BE5	2	64	20/21	0.95	0.10	0,18,25,28	32
1	PIA	1	64	20/21	0.97	0.04	0,14,15,17	32

### 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(Å <sup>2</sup> )	Q<0.9
3	SO4	1	302	5/5	0.51	0.19	106,107,113,115	5
3	SO4	1	301	5/5	0.58	0.19	100,102,108,112	5
4	GOL	2	301	6/6	0.71	0.26	59,65,69,69	14
3	SO4	2	302	5/5	0.73	0.16	103,105,109,113	5
3	SO4	2	303	5/5	0.82	0.18	309,309,311,313	5

## 6.5 Other polymers [i](#)

There are no such residues in this entry.