



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 5, 2017 – 02:59 PM EDT

PDB ID : 5V9P  
Title : Crystal structure of pyrrolidine amide inhibitor [(3S)-3-(4-bromo-1H-pyrazol-1-yl)pyrrolidin-1-yl][3-(propan-2-yl)-1H-pyrazol-5-yl]methanone (compound 35) in complex with KDM5A  
Authors : Kiefer, J.R.; Liang, J.; Vinogradova, M.  
Deposited on : 2017-03-23  
Resolution : 3.00 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029077  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20029077

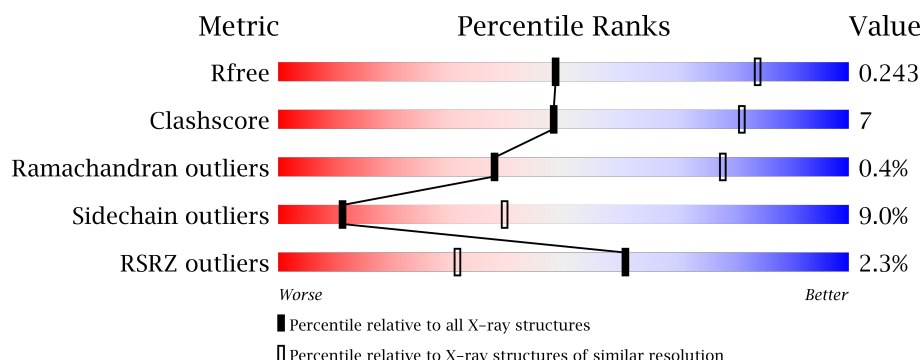
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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  $\geq 3$ , 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	790	 2% 58% 13% 27%
2	B	10	 60% 40%

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	90S	A	904	-	-	-	X
6	SO4	A	905	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9103 atoms, of which 4411 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 Lysine-specific demethylase 5A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	577	Total	C	H	N	O	S	0	2	0
			8998	2956	4397	759	848	38			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	11	SER	-	expression tag	UNP P29375
A	798	GLY	-	expression tag	UNP P29375
A	799	ASN	-	expression tag	UNP P29375
A	800	SER	-	expression tag	UNP P29375

- Molecule 2 is a protein called Lysine-specific demethylase 5A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	10	Total	C	H	N	O	0	0	0
			63	30	13	10	10			

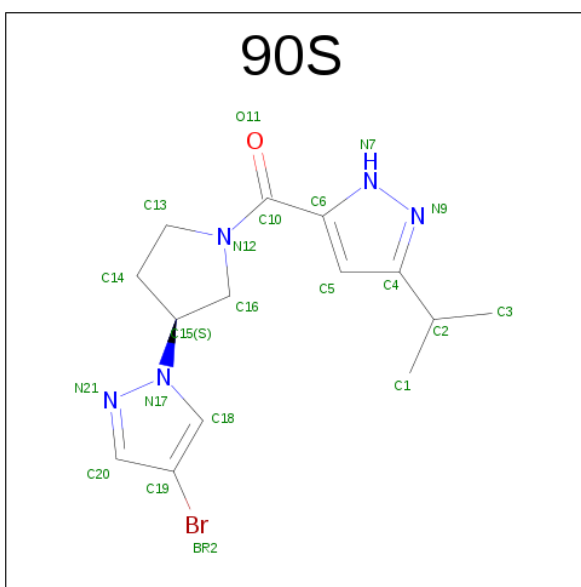
- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ni	0	0
			1	1		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Zn	0	0
			2	2		

- Molecule 5 is [(3S)-3-(4-bromo-1H-pyrazol-1-yl)pyrrolidin-1-yl][3-(propan-2-yl)-1H-pyrazol-5-yl]methanone (three-letter code: 90S) (formula: C<sub>14</sub>H<sub>18</sub>BrN<sub>5</sub>O).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	Br	C	H	N	O	0	0
			22	1	14	1	5	1		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



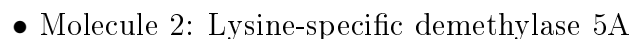
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	O S	0	0
			5	4 1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	12	Total	O	0	0
			12	12		



- Molecule 1: Lysine-specific demethylase 5A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	158.95Å 158.95Å 90.59Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.16 – 3.00 34.16 – 2.99	Depositor EDS
% Data completeness (in resolution range)	87.7 (34.16-3.00) 87.7 (34.16-2.99)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.00 (at 3.00Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.216 , 0.248 0.211 , 0.243	Depositor DCC
$R_{free}$ test set	1153 reflections (4.92%)	DCC
Wilson B-factor (Å <sup>2</sup> )	51.9	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	9103	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% 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: 90S, ZN, SO4, NI

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	0.43	0/4738	0.56	0/6440

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	A	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	176	VAL	Peptide
1	A	466	MET	Peptide

### 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	4601	4397	4383	61	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	50	13	13	2	0
3	A	1	0	0	0	0
4	A	2	0	0	0	0
5	A	21	1	0	2	0
6	A	5	0	0	0	0
7	A	12	0	0	0	0
All	All	4692	4411	4396	63	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 7.

All (63) 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:724:ASP:OD1	1:A:725:LEU:N	2.19	0.75
1:A:646:GLU:OE2	1:A:686:SER:OG	2.07	0.71
1:A:174:SER:OG	1:A:175:GLY:N	2.24	0.69
1:A:391:GLU:OE2	1:A:554:ARG:NH1	2.25	0.69
1:A:531:ASP:OD1	1:A:607:ARG:NH2	2.31	0.64
1:A:176:VAL:HG23	1:A:177:SER:H	1.64	0.61
1:A:415:SER:OG	1:A:469:PRO:O	2.19	0.60
1:A:174:SER:O	1:A:175:GLY:O	2.21	0.58
1:A:369:MET:HG3	1:A:370:ALA:N	2.20	0.54
1:A:610:VAL:HG23	1:A:611:PHE:HD2	1.73	0.54
2:B:227:UNK:O	2:B:228:UNK:C	2.58	0.51
1:A:610:VAL:HG23	1:A:611:PHE:CD2	2.47	0.50
1:A:714:LYS:CB	1:A:715:CYS:HA	2.40	0.50
1:A:676:CYS:SG	1:A:679:CYS:HB2	2.51	0.50
1:A:705:ASP:OD1	1:A:705:ASP:N	2.45	0.49
1:A:676:CYS:HB3	1:A:679:CYS:HB2	1.94	0.49
1:A:714:LYS:CB	1:A:715:CYS:CA	2.90	0.49
1:A:72:PRO:HG3	1:A:361:TYR:HD2	1.77	0.49
1:A:589:ALA:HB2	1:A:631:LEU:HD12	1.94	0.48
1:A:651:GLU:O	1:A:654:VAL:HG22	2.13	0.48
1:A:32:PRO:O	1:A:36:ILE:HG13	2.14	0.48
1:A:425:LYS:NZ	1:A:436:GLU:OE1	2.43	0.48
2:B:221:UNK:O	2:B:224:UNK:N	2.47	0.48
1:A:724:ASP:HB2	1:A:728:LEU:HD12	1.97	0.47
1:A:485:GLU:OE1	5:A:904:90S:N9	2.48	0.47
1:A:450:VAL:HA	1:A:457:ALA:HA	1.96	0.46
1:A:114:ILE:HD13	1:A:114:ILE:H	1.78	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:29:PHE:CZ	1:A:53:PRO:HB3	2.51	0.46
1:A:87:ARG:HG3	1:A:118:TYR:HB3	1.98	0.46
1:A:689:THR:O	1:A:715:CYS:N	2.49	0.46
1:A:777:GLU:CB	1:A:781:PHE:HB2	2.46	0.45
1:A:487:HIS:CD2	1:A:610:VAL:HG22	2.51	0.45
1:A:487:HIS:NE2	1:A:610:VAL:HG22	2.32	0.44
1:A:679:CYS:H	1:A:680:ARG:HA	1.83	0.44
1:A:711:MET:O	1:A:712:GLN:C	2.55	0.44
1:A:542:PRO:HB3	1:A:552:VAL:HG11	1.98	0.44
1:A:374:LYS:HE2	1:A:385:VAL:CG1	2.48	0.43
1:A:75:GLN:HG3	1:A:409:TYR:HE1	1.83	0.43
1:A:635:VAL:O	1:A:639:LEU:N	2.42	0.43
1:A:646:GLU:O	1:A:650:ARG:N	2.44	0.43
1:A:451:LEU:O	1:A:457:ALA:HB2	2.17	0.43
1:A:679:CYS:HB3	1:A:681:THR:H	1.83	0.43
1:A:667:GLU:OE2	1:A:719:ARG:NH2	2.38	0.43
1:A:171:LEU:HB3	1:A:177:SER:O	2.19	0.43
1:A:170:GLU:O	1:A:174:SER:CB	2.67	0.43
1:A:541:ASN:HB3	1:A:544:VAL:HG23	2.00	0.43
1:A:657:GLY:O	1:A:711:MET:O	2.36	0.43
1:A:711:MET:O	1:A:713:LYS:O	2.37	0.42
1:A:459:ILE:O	1:A:459:ILE:CG1	2.67	0.42
1:A:182:GLN:O	1:A:183:MET:HB3	2.19	0.42
1:A:726:PRO:O	1:A:730:TYR:HB2	2.19	0.42
1:A:104:THR:HG22	1:A:104:THR:O	2.18	0.41
1:A:14:VAL:H	1:A:543:ASN:HD21	1.68	0.41
1:A:491:SER:HB3	1:A:565:THR:HB	2.02	0.41
1:A:714:LYS:CB	1:A:715:CYS:HB2	2.50	0.41
1:A:456:LEU:HD13	1:A:456:LEU:N	2.36	0.41
1:A:606:ARG:NH1	1:A:678:ALA:HB1	2.36	0.41
1:A:69:ARG:HA	1:A:361:TYR:O	2.21	0.41
1:A:707:CYS:HB3	1:A:709:CYS:SG	2.61	0.41
1:A:86:VAL:HG23	1:A:87:ARG:N	2.36	0.41
1:A:472:TYR:OH	5:A:904:90S:C1	2.69	0.40
1:A:387:THR:HA	1:A:576:GLN:OE1	2.22	0.40
1:A:780:LEU:HD12	1:A:780:LEU:H	1.86	0.40

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	573/790 (72%)	521 (91%)	50 (9%)	2 (0%)	44 81

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	175	GLY
1	A	174	SER

### 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	500/710 (70%)	454 (91%)	46 (9%)	11 38

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	GLU
1	A	18	GLU
1	A	38	ARG
1	A	55	LYS
1	A	65	VAL
1	A	105	LEU
1	A	107	ILE
1	A	113[A]	LYS
1	A	113[B]	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	114	ILE
1	A	122	LYS
1	A	137	LYS
1	A	138	LYS
1	A	361	TYR
1	A	363	LEU
1	A	369	MET
1	A	401	ILE
1	A	415	SER
1	A	428	ARG
1	A	448	MET
1	A	456	LEU
1	A	464	SER
1	A	466	MET
1	A	508	SER
1	A	516	GLU
1	A	517	VAL
1	A	576	GLN
1	A	607	ARG
1	A	616	LEU
1	A	641	LEU
1	A	655	GLN
1	A	663	GLU
1	A	671	ASP
1	A	675	GLN
1	A	679	CYS
1	A	680	ARG
1	A	689	THR
1	A	693	ASN
1	A	695	GLU
1	A	698	VAL
1	A	699	CYS
1	A	705	ASP
1	A	722	LEU
1	A	725	LEU
1	A	769	ASP
1	A	780	LEU

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 ⓘ

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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	$\# Z  > 2$
5	90S	A	904	3	19,23,23	1.25	2 (10%)	18,33,33	1.91	6 (33%)
6	SO4	A	905	-	4,4,4	0.16	0	6,6,6	0.26	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
5	90S	A	904	3	-	0/8/25/25	0/3/3/3
6	SO4	A	905	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	904	90S	C4-N9	-2.46	1.30	1.34
5	A	904	90S	N21-N17	3.58	1.40	1.35

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	904	90S	O11-C10-N12	-4.01	115.42	122.36
5	A	904	90S	O11-C10-C6	-2.67	113.75	118.04
5	A	904	90S	C13-C14-C15	-2.30	99.02	104.61
5	A	904	90S	C16-N12-C10	2.48	133.10	123.75
5	A	904	90S	C14-C13-N12	2.98	106.90	103.23
5	A	904	90S	C18-N17-C15	3.22	128.84	125.56

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	904	90S	2	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, 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	A	577/790 (73%)	-0.28	13 (2%) 61 31	13, 48, 126, 158	0
2	B	0/10	-	-	-	-
All	All	577/800 (72%)	-0.28	13 (2%) 61 31	13, 48, 126, 158	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	765	VAL	4.6
1	A	767	LEU	3.4
1	A	656	MET	2.8
1	A	772	ASP	2.7
1	A	766	MET	2.7
1	A	739	TYR	2.5
1	A	730	TYR	2.5
1	A	626	CYS	2.5
1	A	781	PHE	2.5
1	A	742	TRP	2.4
1	A	743	VAL	2.2
1	A	783	LYS	2.2
1	A	779	ASP	2.1

### 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, 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	90S	A	904	21/21	0.94	0.33	4.10	10,33,63,221	0
6	SO4	A	905	5/5	0.95	0.26	2.22	41,84,104,117	0
3	NI	A	901	1/1	1.00	0.23	0.64	13,13,13,13	0
4	ZN	A	902	1/1	0.99	0.14	-0.85	70,70,70,70	0
4	ZN	A	903	1/1	0.95	0.04	-2.49	93,93,93,93	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.