



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 3, 2024 – 07:27 PM EST

PDB ID : 1MVH  
Title : structure of the SET domain histone lysine methyltransferase Clr4  
Authors : Min, J.R.; Zhang, X.; Cheng, X.D.; Grewal, S.I.S.; Xu, R.-M.  
Deposited on : 2002-09-25  
Resolution : 2.30 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

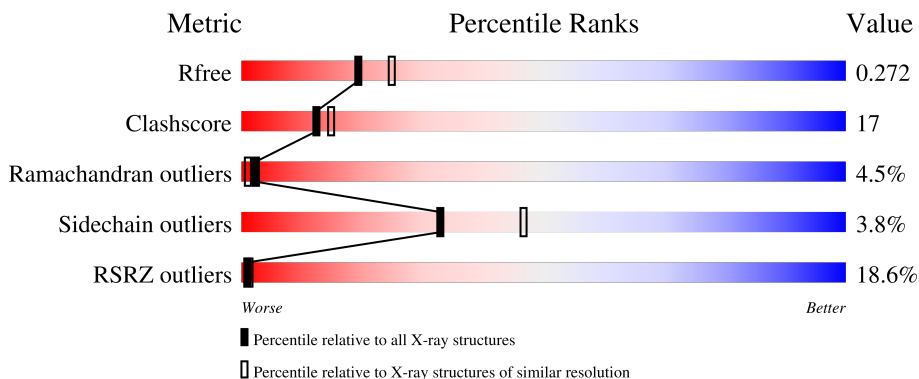
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	A	299	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 2240 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 Cryptic loci regulator 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	269	2136	1345	358	419	14	0	0	0

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



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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	3	3	3	0	0

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

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

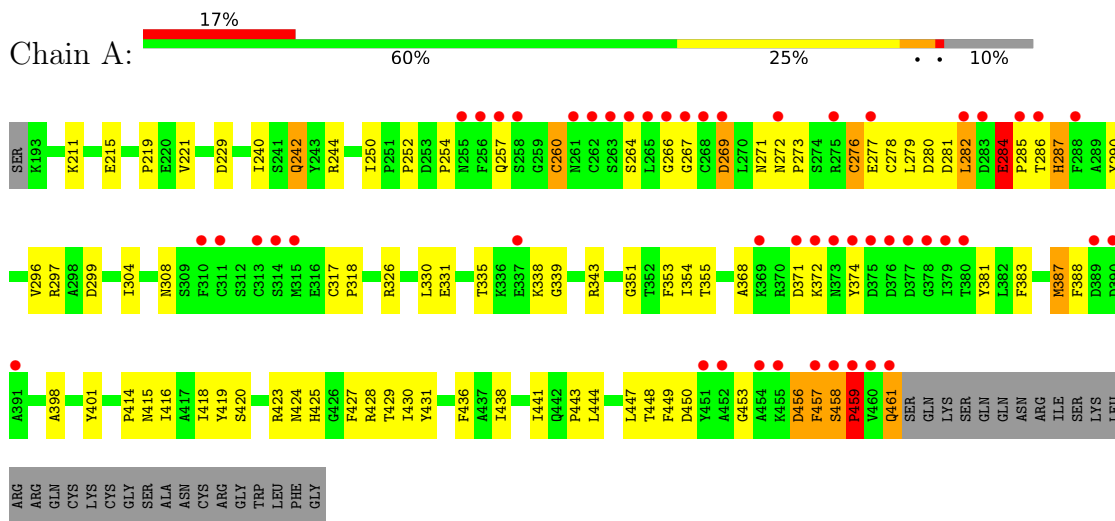
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	95	Total O 95 95	0	0

### 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: Cryptic loci regulator 4



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.07Å 75.07Å 190.59Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.30 38.43 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-2.30) 98.2 (38.43-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.33 (at 2.20Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.225 , 0.269 0.226 , 0.272	Depositor DCC
$R_{free}$ test set	1312 reflections (7.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtrriage
Anisotropy	0.133	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 54.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2240	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% 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: 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	0.38	0/2187	0.60	1/2964 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	284	GLU	C-N-CD	5.21	139.35	128.40

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	A	2136	0	1999	70	1
2	A	5	0	0	0	0
3	A	3	0	0	0	0
4	A	1	0	0	0	0
5	A	95	0	0	2	0
All	All	2240	0	1999	70	1

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 17.

All (70) 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:254:PRO:HA	1:A:257:GLN:HE21	1.29	0.98
1:A:252:PRO:HG3	1:A:429:THR:HG22	1.52	0.90
1:A:371:ASP:HA	1:A:374:TYR:HD1	1.44	0.83
1:A:277:GLU:HA	1:A:280:ASP:OD2	1.87	0.74
1:A:371:ASP:HA	1:A:374:TYR:CD1	2.22	0.73
1:A:414:PRO:HG3	1:A:450:ASP:OD1	1.94	0.68
1:A:254:PRO:CA	1:A:257:GLN:HE21	2.07	0.65
1:A:457:PHE:C	1:A:459:PRO:HD2	2.17	0.65
1:A:273:PRO:HB3	1:A:286:THR:HG22	1.79	0.64
1:A:250:ILE:O	1:A:250:ILE:HG13	1.98	0.64
1:A:338:LYS:HD2	5:A:16:HOH:O	1.99	0.62
1:A:271:ASN:C	1:A:273:PRO:HD3	2.20	0.62
1:A:276:CYS:HB3	1:A:279:LEU:HG	1.83	0.61
1:A:282:LEU:HD11	1:A:284:GLU:OE2	2.01	0.60
1:A:211:LYS:O	1:A:215:GLU:HG3	2.02	0.59
1:A:287:HIS:HB2	1:A:297:ARG:NH1	2.18	0.58
1:A:308:ASN:OD1	1:A:423:ARG:HB3	2.04	0.57
1:A:296:VAL:HG12	1:A:351:GLY:O	2.04	0.57
1:A:387:MET:HG3	1:A:420:SER:OG	2.04	0.57
1:A:267:GLY:HA3	1:A:318:PRO:HD2	1.85	0.57
1:A:368:ALA:O	1:A:372:LYS:HG3	2.05	0.56
1:A:284:GLU:CB	1:A:285:PRO:CD	2.86	0.54
1:A:388:PHE:CE2	1:A:428:ARG:HD3	2.43	0.53
1:A:244:ARG:NH2	5:A:15:HOH:O	2.43	0.52
1:A:414:PRO:O	1:A:438:ILE:HD11	2.09	0.52
1:A:278:CYS:HA	1:A:425:HIS:CE1	2.45	0.52
1:A:415:ASN:C	1:A:438:ILE:HG12	2.30	0.51
1:A:424:ASN:OD1	1:A:427:PHE:HD1	1.94	0.51
1:A:221:VAL:HA	1:A:330:LEU:O	2.11	0.51
1:A:429:THR:HG23	1:A:430:ILE:HG13	1.94	0.50
1:A:286:THR:O	1:A:287:HIS:HB3	2.10	0.50
1:A:368:ALA:HA	1:A:371:ASP:HB2	1.94	0.49
1:A:284:GLU:HB2	1:A:285:PRO:HD2	1.94	0.49
1:A:387:MET:HB2	1:A:431:TYR:OH	2.13	0.49
1:A:450:ASP:O	1:A:450:ASP:CG	2.51	0.49
1:A:416:ILE:HG12	1:A:448:THR:O	2.14	0.48
1:A:290:TYR:OH	1:A:304:ILE:HG21	2.14	0.48
1:A:287:HIS:HB2	1:A:297:ARG:HH12	1.79	0.47
1:A:296:VAL:HG12	1:A:351:GLY:HA2	1.96	0.47
1:A:353:PHE:HB2	1:A:436:PHE:CE1	2.49	0.47

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:335:THR:OG1	1:A:338:LYS:O	2.33	0.47
1:A:296:VAL:HG13	1:A:296:VAL:O	2.16	0.46
1:A:457:PHE:C	1:A:459:PRO:CD	2.83	0.46
1:A:443:PRO:O	1:A:444:LEU:HB2	2.15	0.46
1:A:281:ASP:O	1:A:282:LEU:HB2	2.16	0.46
1:A:260:CYS:HA	1:A:277:GLU:HG3	1.98	0.45
1:A:354:ILE:HG22	1:A:355:THR:HG22	1.99	0.44
1:A:441:ILE:HD13	1:A:447:LEU:CD2	2.48	0.44
1:A:387:MET:HB2	1:A:431:TYR:CZ	2.53	0.43
1:A:457:PHE:N	1:A:459:PRO:HD2	2.34	0.43
1:A:272:ASN:N	1:A:273:PRO:HD3	2.33	0.42
1:A:331:GLU:OE2	1:A:343:ARG:NH2	2.51	0.42
1:A:252:PRO:CG	1:A:429:THR:HG22	2.36	0.42
1:A:296:VAL:HG12	1:A:351:GLY:CA	2.50	0.42
1:A:326:ARG:HB3	1:A:326:ARG:NH1	2.34	0.42
1:A:419:TYR:CE1	1:A:457:PHE:HE2	2.37	0.42
1:A:418:ILE:HD12	1:A:418:ILE:N	2.35	0.42
1:A:458:SER:N	1:A:459:PRO:CD	2.83	0.42
1:A:453:GLY:O	1:A:456:ASP:OD1	2.37	0.41
1:A:240:ILE:HD12	1:A:242:GLN:O	2.20	0.41
1:A:273:PRO:CB	1:A:286:THR:HG22	2.47	0.41
1:A:398:ALA:HA	1:A:401:TYR:O	2.21	0.41
1:A:299:ASP:OD1	1:A:299:ASP:N	2.53	0.41
1:A:219:PRO:O	1:A:326:ARG:NH2	2.54	0.41
1:A:284:GLU:HB3	1:A:285:PRO:CD	2.49	0.41
1:A:326:ARG:HB3	1:A:326:ARG:HH11	1.86	0.41
1:A:308:ASN:OD1	1:A:423:ARG:O	2.39	0.40
1:A:449:PHE:CD1	1:A:449:PHE:C	2.94	0.40
1:A:267:GLY:HA3	1:A:317:CYS:HA	2.03	0.40
1:A:461:GLN:OE1	1:A:461:GLN:C	2.60	0.40

All (1) 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)
1:A:457:PHE:O	1:A:457:PHE:O[12_555]	2.03	0.17

## 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	A	267/299 (89%)	232 (87%)	23 (9%)	12 (4%)	<b>2</b> <b>1</b>

All (12) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	284	GLU
1	A	264	SER
1	A	269	ASP
1	A	381	TYR
1	A	266	GLY
1	A	339	GLY
1	A	282	LEU
1	A	287	HIS
1	A	456	ASP
1	A	457	PHE
1	A	458	SER
1	A	459	PRO

### 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	A	234/262 (89%)	225 (96%)	9 (4%)	<b>33</b> <b>47</b>

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	229	ASP
1	A	242	GLN
1	A	260	CYS
1	A	269	ASP
1	A	276	CYS
1	A	383	PHE
1	A	387	MET
1	A	459	PRO
1	A	461	GLN

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

Mol	Chain	Res	Type
1	A	257	GLN
1	A	272	ASN
1	A	293	GLN
1	A	373	ASN
1	A	399	GLN

### 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 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 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	$\# Z  > 2$
2	SO4	A	601	-	4,4,4	1.88	2 (50%)	6,6,6	0.91	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	SO4	O1-S	3.11	1.62	1.46
2	A	601	SO4	O3-S	-2.05	1.31	1.47

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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

### 6.1 Protein, DNA and RNA chains

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	269/299 (89%)	0.78	50 (18%) <b>1</b> <b>1</b>	20, 49, 100, 101	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	374	TYR	13.3
1	A	265	LEU	10.1
1	A	379	ILE	7.7
1	A	378	GLY	7.0
1	A	264	SER	6.0
1	A	263	SER	6.0
1	A	375	ASP	5.9
1	A	377	ASP	5.7
1	A	258	SER	5.5
1	A	262	CYS	5.2
1	A	266	GLY	5.2
1	A	286	THR	5.1
1	A	455	LYS	5.0
1	A	373	ASN	4.9
1	A	275	ARG	4.8
1	A	458	SER	4.8
1	A	267	GLY	4.5
1	A	371	ASP	4.5
1	A	285	PRO	4.1
1	A	269	ASP	3.9
1	A	372	LYS	3.8
1	A	457	PHE	3.8
1	A	261	ASN	3.7
1	A	376	ASP	3.6
1	A	460	VAL	3.6
1	A	380	THR	3.5
1	A	257	GLN	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	459	PRO	3.3
1	A	288	PHE	3.2
1	A	390	ASP	3.0
1	A	277	GLU	2.9
1	A	311	CYS	2.8
1	A	454	ALA	2.6
1	A	337	GLU	2.6
1	A	282	LEU	2.6
1	A	268	CYS	2.6
1	A	391	ALA	2.5
1	A	461	GLN	2.5
1	A	315	MET	2.4
1	A	314	SER	2.4
1	A	256	PHE	2.3
1	A	313	CYS	2.3
1	A	452	ALA	2.3
1	A	389	ASP	2.2
1	A	310	PHE	2.2
1	A	451	TYR	2.1
1	A	369	LYS	2.1
1	A	283	ASP	2.0
1	A	255	ASN	2.0
1	A	272	ASN	2.0

## 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 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.

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
-----	------	-------	-----	-------	------	-----	-----------------------------	-------

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	ZN	A	501	1/1	0.91	0.04	62,62,62,62	0
3	ZN	A	502	1/1	0.94	0.07	64,64,64,64	0
2	SO4	A	601	5/5	0.95	0.12	75,78,81,81	0
3	ZN	A	503	1/1	0.96	0.07	80,80,80,80	0
4	NI	A	504	1/1	0.99	0.17	35,35,35,35	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.