Full wwPDB X-ray Structure Validation Report (i)

Oct 28, 2024 – 03:02 am GMT

PDB ID	:	2VPV
Title	:	Dimerization Domain of Mif2p
Authors	:	Cohen, R.L.; Espelin, C.W.; Sorger, P.K.; Harrison, S.C.; Simons, K.T.
Deposited on	:	2008-03-05
Resolution	:	2.70 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

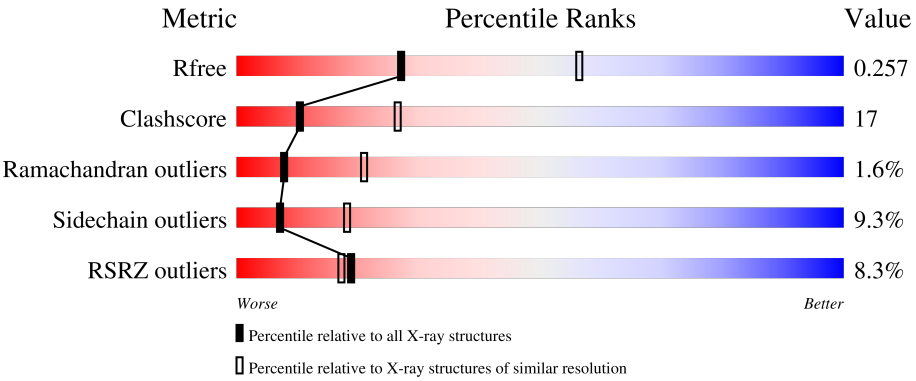
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

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 <=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	166	6% 42% 11% 43%
1	B	166	3% 30% 20% 45%

2 Entry composition

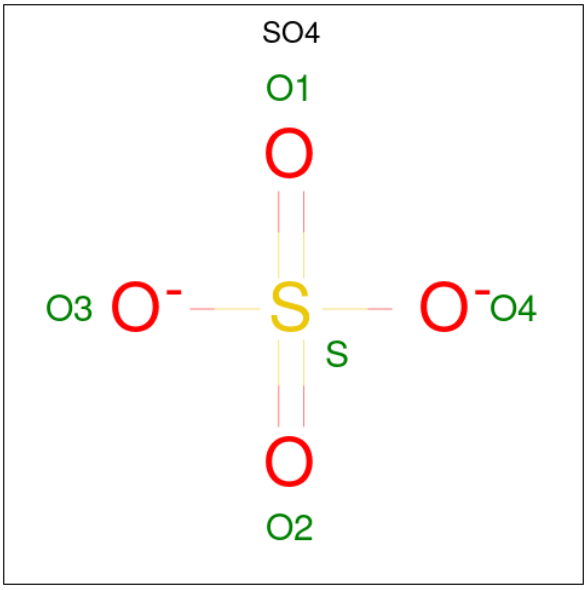
There are 3 unique types of molecules in this entry. The entry contains 1519 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 PROTEIN MIF2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	94	Total	C	N	O	S	13	0	0
			752	490	123	136	3			
1	B	92	Total	C	N	O	S	19	0	0
			735	481	120	131	3			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total	O	0	0
			18	18		

Continued on next page...

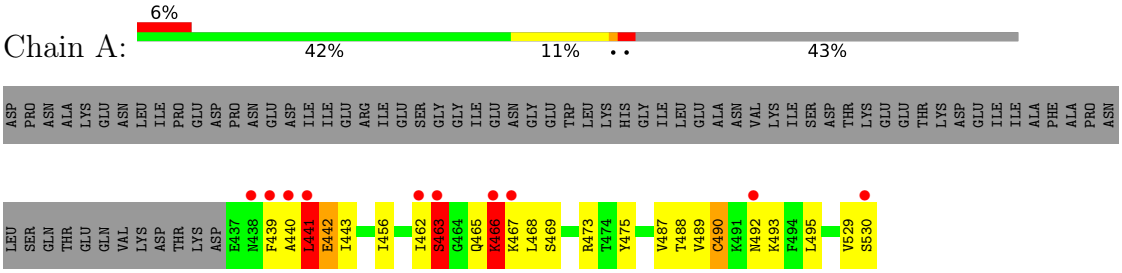
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	9	Total O 9 9	0	0

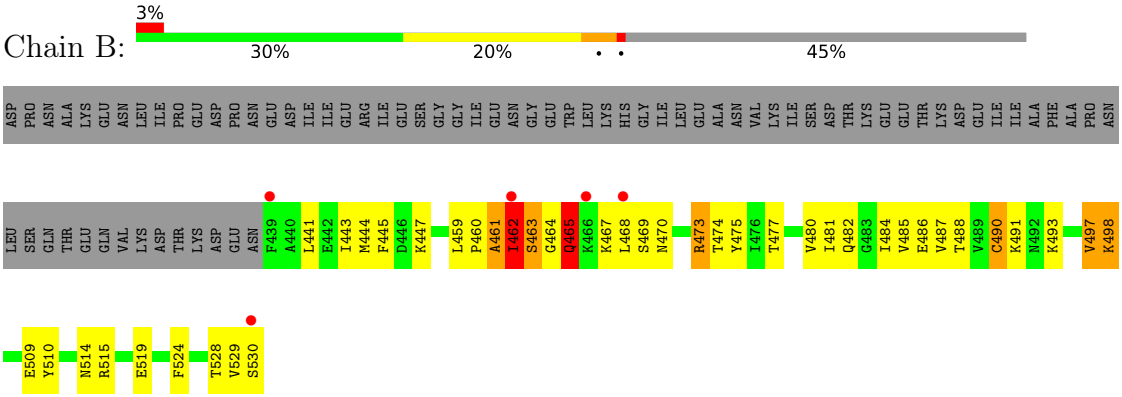
3 Residue-property plots (i)

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: PROTEIN MIF2



• Molecule 1: PROTEIN MIF2



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	106.02Å 106.02Å 106.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.70 30.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-2.70) 86.5 (30.00-2.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.90 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.4.0066	Depositor
R, R_{free}	0.208 , 0.242 0.220 , 0.257	Depositor DCC
R_{free} test set	468 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å^2)	56.1	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 18.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.054 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	1519	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.71	1/768 (0.1%)	0.78	1/1031 (0.1%)
1	B	0.74	1/751 (0.1%)	1.00	4/1008 (0.4%)
All	All	0.72	2/1519 (0.1%)	0.90	5/2039 (0.2%)

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	3
1	B	1	7
All	All	1	10

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	465	GLN	C-N	-11.21	1.08	1.34
1	A	463	SER	C-N	9.92	1.50	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	465	GLN	C-N-CA	13.69	155.92	121.70
1	B	465	GLN	O-C-N	-12.73	102.33	122.70
1	B	465	GLN	CA-C-N	8.93	136.84	117.20
1	B	462	ILE	O-C-N	-7.22	111.16	122.70
1	A	442	GLU	N-CA-C	-6.43	93.65	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	465	GLN	CA

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	441	LEU	Peptide
1	A	463	SER	Mainchain
1	A	465	GLN	Peptide
1	B	460	PRO	Peptide
1	B	461	ALA	Peptide
1	B	462	ILE	Peptide,Mainchain
1	B	463	SER	Peptide
1	B	464	GLY	Peptide
1	B	465	GLN	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	752	0	757	22	1
1	B	735	0	744	27	1
2	A	5	0	0	0	0
3	A	18	0	0	2	0
3	B	9	0	0	1	0
All	All	1519	0	1501	49	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 (49) 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:441:LEU:CG	1:A:442:GLU:HB2	1.69	1.20
1:A:441:LEU:HG	1:A:442:GLU:CB	1.77	1.13
1:B:461:ALA:CB	1:B:514:ASN:ND2	2.12	1.12
1:A:441:LEU:HD11	3:A:2006:HOH:O	1.54	1.08
1:B:461:ALA:HB2	1:B:514:ASN:HD22	1.26	0.98

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:461:ALA:CB	1:B:514:ASN:HD22	1.80	0.92
1:A:490:CYS:HB3	3:A:2009:HOH:O	1.78	0.84
1:A:439:PHE:CE2	1:A:467:LYS:HB3	2.19	0.77
1:A:466:LYS:HD3	1:A:466:LYS:C	2.10	0.72
1:A:439:PHE:CD2	1:A:467:LYS:HB3	2.26	0.70
1:B:461:ALA:HB3	1:B:514:ASN:ND2	2.04	0.69
1:B:490:CYS:HB3	3:B:2004:HOH:O	1.94	0.68
1:B:461:ALA:HB1	1:B:514:ASN:ND2	2.10	0.67
1:A:441:LEU:HG	1:A:442:GLU:HB2	0.80	0.66
1:B:484:ILE:HG12	1:B:497:VAL:HG23	1.78	0.64
1:B:468:LEU:HD13	1:B:469:SER:N	2.13	0.64
1:B:482:GLN:NE2	1:B:519:GLU:O	2.32	0.61
1:A:473:ARG:HG2	1:A:529:VAL:CG1	2.32	0.59
1:B:445:PHE:CE2	1:B:447:LYS:HB3	2.38	0.59
1:A:468:LEU:HD13	1:A:469:SER:N	2.19	0.57
1:B:461:ALA:HB1	1:B:514:ASN:CG	2.25	0.57
1:B:441:LEU:HD21	1:B:443:ILE:HG13	1.85	0.57
1:A:441:LEU:CD2	1:A:442:GLU:HB2	2.35	0.56
1:B:459:LEU:O	1:B:461:ALA:HB2	2.07	0.55
1:A:439:PHE:CG	1:A:467:LYS:HD3	2.44	0.53
1:B:461:ALA:HB2	1:B:514:ASN:ND2	1.92	0.53
1:B:529:VAL:HG22	1:B:530:SER:N	2.26	0.50
1:B:529:VAL:HG22	1:B:530:SER:H	1.77	0.50
1:A:441:LEU:HD21	1:A:456:ILE:O	2.12	0.50
1:B:486:GLU:HB2	1:B:515:ARG:HG3	1.96	0.47
1:A:489:VAL:O	1:A:490:CYS:C	2.53	0.46
1:A:466:LYS:HD3	1:A:466:LYS:O	2.16	0.45
1:B:474:THR:OG1	1:B:528:THR:HG22	2.15	0.45
1:B:488:THR:O	1:B:510:TYR:HA	2.17	0.45
1:A:487:VAL:O	1:A:493:LYS:HA	2.17	0.45
1:B:480:VAL:HG12	1:B:498:LYS:HA	1.99	0.43
1:A:529:VAL:HG22	1:A:530:SER:N	2.34	0.43
1:A:439:PHE:CE2	1:A:469:SER:HB2	2.54	0.42
1:B:487:VAL:O	1:B:493:LYS:HA	2.18	0.42
1:B:490:CYS:HB3	1:B:491:LYS:H	1.45	0.42
1:B:467:LYS:NZ	1:B:469:SER:OG	2.36	0.42
1:A:441:LEU:CD2	1:A:442:GLU:CB	2.98	0.41
1:A:440:ALA:O	1:A:441:LEU:HB2	2.20	0.41
1:B:473:ARG:CG	1:B:473:ARG:HH11	2.33	0.41
1:A:488:THR:HA	1:A:492:ASN:O	2.19	0.41
1:A:441:LEU:CG	1:A:442:GLU:CB	2.63	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:461:ALA:CB	1:B:514:ASN:CG	2.76	0.41
1:B:470:ASN:HA	1:B:509:GLU:HA	2.02	0.40
1:B:477:THR:O	1:B:524:PHE:HA	2.21	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:475:TYR:OH	1:B:475:TYR:OH[9_555]	2.11	0.09

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	92/166 (55%)	83 (90%)	7 (8%)	2 (2%)	5 15
1	B	90/166 (54%)	83 (92%)	6 (7%)	1 (1%)	12 30
All	All	182/332 (55%)	166 (91%)	13 (7%)	3 (2%)	8 21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	462	ILE
1	A	463	SER
1	A	466	LYS

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	82/146 (56%)	76 (93%)	6 (7%)	11 29
1	B	80/146 (55%)	71 (89%)	9 (11%)	4 11
All	All	162/292 (56%)	147 (91%)	15 (9%)	7 18

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	441	LEU
1	A	443	ILE
1	A	462	ILE
1	A	466	LYS
1	A	490	CYS
1	A	495	LEU
1	B	444	MET
1	B	463	SER
1	B	465	GLN
1	B	473	ARG
1	B	481	ILE
1	B	485	VAL
1	B	490	CYS
1	B	497	VAL
1	B	498	LYS

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	SO4	A	1531	-	4,4,4	1.05	1 (25%)	6,6,6	0.61	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1531	SO4	O4-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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	465:GLN	C	466:LYS	N	1.08

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^{th} 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	A	92/166 (55%)	0.24	10 (10%) 12 11	39, 48, 81, 102	0
1	B	89/166 (53%)	0.39	5 (5%) 31 29	40, 57, 74, 79	0
All	All	181/332 (54%)	0.32	15 (8%) 19 17	39, 53, 78, 102	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	462	ILE	4.2
1	A	530	SER	4.1
1	A	467	LYS	3.9
1	B	530	SER	3.9
1	A	466	LYS	3.9
1	A	463	SER	3.9
1	A	439	PHE	3.8
1	A	440	ALA	3.8
1	B	468	LEU	3.7
1	B	439	PHE	3.6
1	A	441	LEU	3.5
1	A	462	ILE	3.4
1	B	466	LYS	2.9
1	A	438	ASN	2.7
1	A	492	ASN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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^{th} 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^2)	Q<0.9
2	SO4	A	1531	5/5	0.96	0.06	54,55,55,55	5

6.5 Other polymers

There are no such residues in this entry.