



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 4, 2019 – 05:37 PM EDT

PDB ID : 6H4T  
Title : Crystal structure of human KDM4A in complex with compound 19a  
Authors : Le Bihan, Y.V.; van Montfort, R.L.M.  
Deposited on : 2018-07-23  
Resolution : 2.38 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : rb-20031633  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

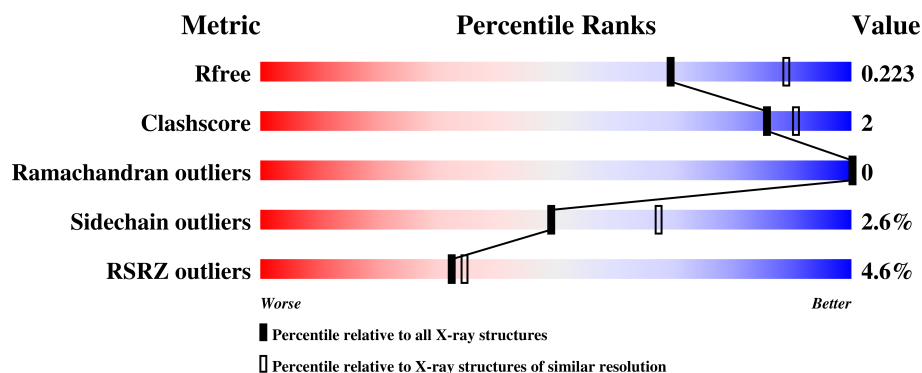
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.38 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))
$R_{free}$	111664	4850 (2.40-2.36)
Clashscore	122126	5405 (2.40-2.36)
Ramachandran outliers	120053	5324 (2.40-2.36)
Sidechain outliers	120020	5326 (2.40-2.36)
RSRZ outliers	108989	4741 (2.40-2.36)

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	360	<div> <div>7%</div> <div> <div></div> <div>86%</div> <div>8%</div> <div>6%</div> </div> </div>
1	B	360	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>6%</div> <div>6%</div> </div> </div>
1	C	360	<div> <div>6%</div> <div> <div></div> <div>88%</div> <div>5%</div> <div>• 6%</div> </div> </div>
1	D	360	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>6%</div> <div>• 5%</div> </div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 11699 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 Lysine-specific demethylase 4A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	340	Total	C	N	O	S	0	0	0
			2674	1749	429	482	14			
1	B	340	Total	C	N	O	S	0	0	0
			2712	1765	444	489	14			
1	C	338	Total	C	N	O	S	0	3	0
			2672	1741	434	482	15			
1	D	342	Total	C	N	O	S	0	1	0
			2712	1767	439	492	14			

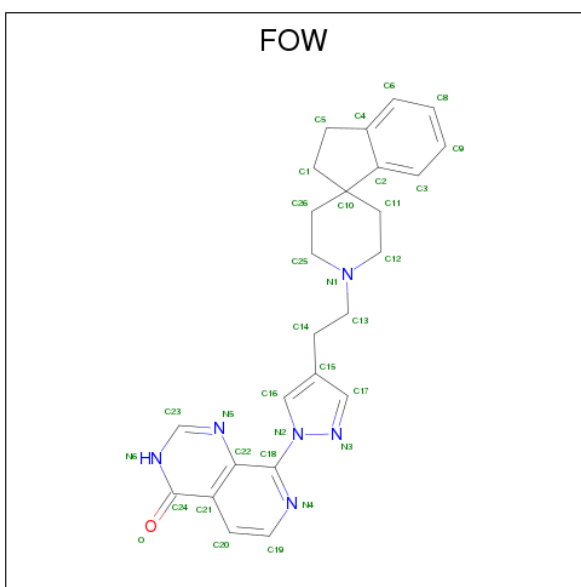
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP O75164
B	0	SER	-	expression tag	UNP O75164
C	0	SER	-	expression tag	UNP O75164
D	0	SER	-	expression tag	UNP O75164

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Zn	0	0
			2	2		
2	A	2	Total	Zn	0	0
			2	2		
2	D	2	Total	Zn	0	0
			2	2		
2	C	2	Total	Zn	0	0
			2	2		

- Molecule 3 is 8-[4-(2-spiro[1,2-dihydroindene-3,4'-piperidine]-1'-ylethyl)pyrazol-1-yl]-3 {H}-pyrido[3,4-d]pyrimidin-4-one (three-letter code: FOW) (formula: C<sub>25</sub>H<sub>26</sub>N<sub>6</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			32	25	6	1		
3	B	1	Total	C	N	O	0	0
			32	25	6	1		
3	C	1	Total	C	N	O	0	0
			32	25	6	1		
3	D	1	Total	C	N	O	0	0
			32	25	6	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

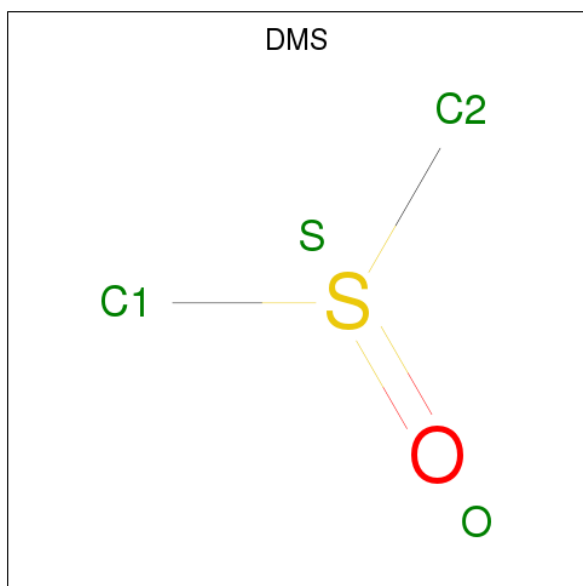
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Cl	0	0
			1	1		
4	A	1	Total	Cl	0	0
			1	1		
4	D	1	Total	Cl	0	0
			1	1		
4	C	1	Total	Cl	0	0
			1	1		

- Molecule 5 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
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula:  $C_2H_6OS$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	O	S	0	0
			4	2	1	1		
6	B	1	Total	C	O	S	0	0
			4	2	1	1		
6	C	1	Total	C	O	S	0	0
			4	2	1	1		

*Continued on next page...*

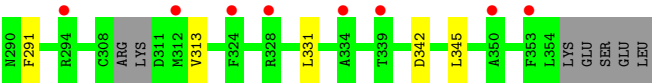
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	1	Total 4	C 2	O 1	S 1	0	0
6	D	1	Total 4	C 2	O 1	S 1	0	0
6	D	1	Total 4	C 2	O 1	S 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	178	Total 178	O 178	0	0
7	B	242	Total 244	O 244	0	2
7	C	129	Total 133	O 133	0	4
7	D	200	Total 204	O 204	0	4







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.86Å 101.81Å 142.68Å 90.00° 99.20° 90.00°	Depositor
Resolution (Å)	43.64 – 2.38 43.64 – 2.38	Depositor EDS
% Data completeness (in resolution range)	100.0 (43.64-2.38) 100.0 (43.64-2.38)	Depositor EDS
$R_{merge}$	0.26	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 2.37Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.173 , 0.216 0.184 , 0.223	Depositor DCC
$R_{free}$ test set	3335 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.1	Xtriage
Anisotropy	0.125	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 77.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.030 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11699	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.90 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.1926e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GOL, ZN, FOW, DMS, CL

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	0.52	0/2760	0.69	1/3760 (0.0%)
1	B	0.55	0/2798	0.69	1/3805 (0.0%)
1	C	0.48	0/2757	0.68	1/3757 (0.0%)
1	D	0.54	0/2797	0.69	0/3807
All	All	0.52	0/11112	0.69	3/15129 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	308	CYS	N-CA-C	5.58	126.08	111.00
1	A	308	CYS	N-CA-C	5.30	125.31	111.00
1	B	308	CYS	N-CA-C	5.18	124.97	111.00

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	2674	0	2455	12	0
1	B	2712	0	2518	10	0
1	C	2672	0	2417	9	0
1	D	2712	0	2495	16	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	32	0	0	1	0
3	B	32	0	0	1	0
3	C	32	0	0	1	0
3	D	32	0	0	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	B	6	0	8	1	0
6	B	8	0	12	0	0
6	C	4	0	6	0	0
6	D	12	0	18	0	0
7	A	178	0	0	1	0
7	B	244	0	0	1	0
7	C	133	0	0	1	0
7	D	204	0	0	4	0
All	All	11699	0	9929	50	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 2.

The worst 5 of 50 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:B:126:THR:HB	5:B:404:GOL:H2	1.34	1.07
1:D:9:ASN:HB2	1:D:15:MET:HE3	1.54	0.88
1:A:150:ILE:HG23	1:A:289:THR:HG22	1.80	0.64
1:B:158:ASP:HA	1:B:172:ASN:HD21	1.64	0.62
1:A:158:ASP:HA	1:A:172:ASN:HD21	1.65	0.62

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	A	336/360 (93%)	330 (98%)	6 (2%)	0	100	100
1	B	336/360 (93%)	329 (98%)	7 (2%)	0	100	100
1	C	335/360 (93%)	326 (97%)	9 (3%)	0	100	100
1	D	337/360 (94%)	332 (98%)	5 (2%)	0	100	100
All	All	1344/1440 (93%)	1317 (98%)	27 (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	A	259/316 (82%)	253 (98%)	6 (2%)	53	71
1	B	267/316 (84%)	263 (98%)	4 (2%)	67	82
1	C	254/316 (80%)	244 (96%)	10 (4%)	35	51
1	D	266/316 (84%)	257 (97%)	9 (3%)	40	57
All	All	1046/1264 (83%)	1017 (97%)	29 (3%)	49	65

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	132	TYR
1	C	163	GLU
1	D	172	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	156	ILE
1	C	182	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	172	ASN
1	B	172	ASN

### 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 23 ligands modelled in this entry, 12 are monoatomic - leaving 11 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	$\# Z  > 2$
3	FOW	A	403	2	32,37,37	0.49	0	41,54,54	1.45	7 (17%)
3	FOW	B	403	2	32,37,37	0.57	0	41,54,54	1.46	8 (19%)
5	GOL	B	404	-	5,5,5	0.15	0	5,5,5	0.60	0
6	DMS	B	406	-	3,3,3	0.24	0	3,3,3	0.12	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	DMS	B	407	-	3,3,3	0.23	0	3,3,3	0.39	0
3	FOW	C	403	2	32,37,37	0.49	0	41,54,54	1.36	7 (17%)
6	DMS	C	405	-	3,3,3	0.30	0	3,3,3	0.31	0
3	FOW	D	403	2	32,37,37	0.47	0	41,54,54	1.41	7 (17%)
6	DMS	D	405	-	3,3,3	0.23	0	3,3,3	0.69	0
6	DMS	D	406	-	3,3,3	0.29	0	3,3,3	0.23	0
6	DMS	D	407	-	3,3,3	0.33	0	3,3,3	0.43	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
3	FOW	A	403	2	-	0/5/33/33	0/6/6/6
3	FOW	B	403	2	-	0/5/33/33	0/6/6/6
5	GOL	B	404	-	-	0/4/4/4	0/0/0/0
6	DMS	B	406	-	-	0/0/0/0	0/0/0/0
6	DMS	B	407	-	-	0/0/0/0	0/0/0/0
3	FOW	C	403	2	-	0/5/33/33	0/6/6/6
6	DMS	C	405	-	-	0/0/0/0	0/0/0/0
3	FOW	D	403	2	-	0/5/33/33	0/6/6/6
6	DMS	D	405	-	-	0/0/0/0	0/0/0/0
6	DMS	D	406	-	-	0/0/0/0	0/0/0/0
6	DMS	D	407	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	403	FOW	C18-C22-C21	-4.62	114.74	119.63
3	B	403	FOW	C18-C22-C21	-4.52	114.83	119.63
3	D	403	FOW	C18-C22-C21	-4.22	115.15	119.63
3	A	403	FOW	C18-C22-C21	-4.19	115.19	119.63
3	A	403	FOW	C19-C20-C21	-2.78	117.16	119.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	403	FOW	1	0
3	B	403	FOW	1	0
5	B	404	GOL	1	0
3	C	403	FOW	1	0
3	D	403	FOW	1	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	340/360 (94%)	0.38	24 (7%) 16 17	27, 47, 92, 107	0
1	B	340/360 (94%)	0.02	5 (1%) 73 75	22, 38, 67, 105	1 (0%)
1	C	338/360 (93%)	0.41	21 (6%) 20 22	31, 55, 102, 127	0
1	D	342/360 (95%)	0.23	13 (3%) 40 43	22, 48, 89, 112	1 (0%)
All	All	1360/1440 (94%)	0.26	63 (4%) 32 34	22, 47, 92, 127	2 (0%)

The worst 5 of 63 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	329	TYR	5.7
1	C	332	TRP	4.9
1	C	326	PRO	4.4
1	C	309	ARG	4.3
1	A	324	PHE	4.3

### 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. 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	CL	A	404	1/1	0.82	0.12	73,73,73,73	0
6	DMS	C	405	4/4	0.84	0.26	127,127,128,128	0
6	DMS	B	406	4/4	0.86	0.26	61,62,63,63	4
2	ZN	C	402	1/1	0.87	0.13	68,68,68,68	1
6	DMS	D	406	4/4	0.88	0.22	60,61,62,63	4
4	CL	B	405	1/1	0.89	0.20	72,72,72,72	0
4	CL	D	404	1/1	0.90	0.16	90,90,90,90	0
6	DMS	B	407	4/4	0.91	0.24	64,64,64,65	4
4	CL	C	404	1/1	0.91	0.15	71,71,71,71	0
6	DMS	D	407	4/4	0.92	0.20	80,81,81,82	4
6	DMS	D	405	4/4	0.92	0.23	62,62,62,65	4
3	FOW	C	403	32/32	0.95	0.15	32,44,67,68	32
3	FOW	A	403	32/32	0.95	0.19	33,38,46,46	32
5	GOL	B	404	6/6	0.95	0.28	11,16,17,21	6
3	FOW	D	403	32/32	0.95	0.16	29,41,60,61	32
3	FOW	B	403	32/32	0.96	0.15	17,28,41,42	32
2	ZN	A	401	1/1	0.97	0.15	50,50,50,50	0
2	ZN	D	402	1/1	0.98	0.09	46,46,46,46	1
2	ZN	B	402	1/1	0.99	0.10	38,38,38,38	1
2	ZN	D	401	1/1	0.99	0.11	44,44,44,44	0
2	ZN	B	401	1/1	0.99	0.11	37,37,37,37	0
2	ZN	C	401	1/1	0.99	0.12	48,48,48,48	0
2	ZN	A	402	1/1	1.00	0.09	43,43,43,43	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.