Full wwPDB X-ray Structure Validation Report ⓘ

Nov 13, 2023 – 10:15 AM JST

PDB ID	:	5WY1
Title	:	Crystal structure of mouse DNA methyltransferase 1 (T1505A mutant)
Authors	:	Kanada, K.; Takeshita, K.; Suetake, I.; Tajima, S.; Nakagawa, A.
Deposited on	:	2017-01-10
Resolution	:	3.27 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 ⓘ 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:

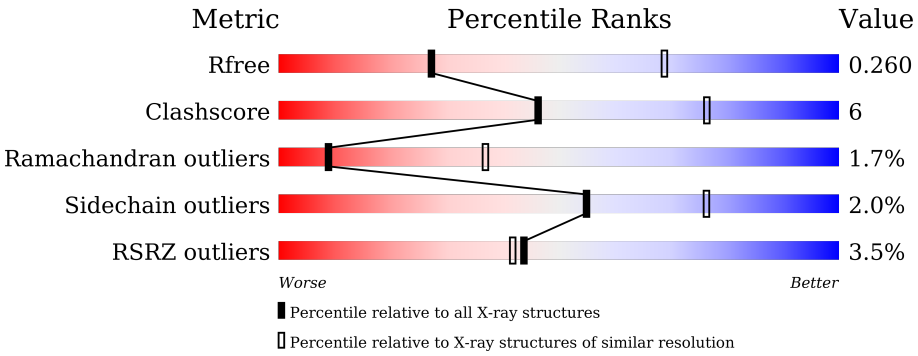
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.27 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	1177 (3.32-3.24)
Clashscore	141614	1044 (3.30-3.26)
Ramachandran outliers	138981	1026 (3.30-3.26)
Sidechain outliers	138945	1025 (3.30-3.26)
RSRZ outliers	127900	1141 (3.32-3.24)

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	1330	3% 71% 14% 14%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9135 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 DNA (cytosine-5)-methyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1140	Total	C	N	O	S	0	0	0
			9108	5765	1586	1698	59			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1505	ALA	THR	engineered mutation	UNP P13864

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

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	23	Total	O	0	0
			23	23		

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: DNA (cytosine-5)-methyltransferase 1

Chain A:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	135.00Å 97.80Å 130.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.27 45.98 – 3.27	Depositor EDS
% Data completeness (in resolution range)	90.4 (50.00-3.27) 90.4 (45.98-3.27)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.32 (at 3.25Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.203 , 0.263 0.205 , 0.260	Depositor DCC
R_{free} test set	1255 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	73.5	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 45.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.014 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	9135	wwPDB-VP
Average B, all atoms (Å^2)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% 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: ZN

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	0/9328	0.72	0/12612

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	9108	0	8888	114	0
2	A	4	0	0	0	0
3	A	23	0	0	5	0
All	All	9135	0	8888	114	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 6.

All (114) 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:446:LYS:NZ	1:A:454:SER:O	1.57	1.34

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:918:ILE:HG22	1:A:919:GLU:H	1.44	0.81
1:A:828:LEU:O	1:A:831:ILE:HG22	1.84	0.78
1:A:1456:HIS:HB2	3:A:2103:HOH:O	1.87	0.74
1:A:950:PHE:O	1:A:996:ARG:NH1	2.22	0.71
1:A:766:VAL:HG11	1:A:831:ILE:HD12	1.73	0.70
1:A:775:LYS:N	1:A:776:PRO:HD3	2.07	0.70
1:A:444:CYS:SG	1:A:458:GLY:HA3	2.32	0.69
1:A:1512:TRP:HB3	1:A:1515:LEU:HD12	1.75	0.68
1:A:412:ARG:HD2	1:A:495:MET:SD	2.35	0.67
1:A:918:ILE:HG22	1:A:919:GLU:N	2.12	0.65
1:A:595:GLY:HA3	1:A:1504:HIS:CE1	2.34	0.63
1:A:744:GLU:O	1:A:747:ARG:N	2.33	0.62
1:A:1157:GLY:HA3	1:A:1587:ALA:HB3	1.82	0.61
1:A:1489:PRO:O	1:A:1490:GLU:HB2	2.00	0.61
1:A:1518:ARG:HA	1:A:1542:LEU:HB2	1.84	0.60
1:A:414:THR:HG21	1:A:453:PRO:O	2.01	0.60
1:A:839:TYR:CE2	1:A:841:ALA:HB2	2.36	0.60
1:A:1466:LYS:NZ	3:A:2101:HOH:O	2.35	0.58
1:A:1458:LEU:HD12	1:A:1496:THR:HG21	1.85	0.58
1:A:379:ASP:O	1:A:380:ALA:C	2.41	0.57
1:A:479:GLU:HA	1:A:516:LYS:HE2	1.86	0.57
1:A:766:VAL:CG1	1:A:831:ILE:HD12	2.35	0.56
1:A:1027:PRO:O	1:A:1030:THR:HB	2.05	0.56
1:A:1266:PHE:HB3	1:A:1320:ALA:HB3	1.87	0.55
1:A:522:LEU:HD21	1:A:586:CYS:SG	2.48	0.54
1:A:1074:SER:HB3	1:A:1077:ASP:OD2	2.07	0.54
1:A:1412:LEU:HD22	1:A:1415:HIS:HB2	1.89	0.53
1:A:416:PHE:HA	1:A:442:SER:O	2.09	0.52
1:A:1456:HIS:CG	1:A:1457:LYS:H	2.27	0.52
1:A:453:PRO:HA	1:A:495:MET:HE2	1.91	0.52
1:A:999:ARG:HH11	1:A:1047:TRP:HZ3	1.57	0.52
1:A:430:THR:HG22	1:A:430:THR:O	2.10	0.51
1:A:726:LYS:HA	1:A:770:PRO:HA	1.91	0.51
1:A:1030:THR:HG22	1:A:1032:ARG:H	1.76	0.51
1:A:1449:LEU:HD11	1:A:1455:ALA:HB2	1.93	0.51
1:A:932:LYS:NZ	1:A:1056:PHE:O	2.40	0.50
1:A:1360:LEU:O	1:A:1361:SER:CB	2.59	0.50
1:A:1456:HIS:CB	3:A:2103:HOH:O	2.54	0.50
1:A:362:CYS:SG	1:A:364:GLN:HG3	2.52	0.50
1:A:689:CYS:SG	1:A:690:LEU:N	2.84	0.50
1:A:691:LYS:HA	1:A:693:ARG:NH1	2.26	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:873:TYR:CE2	1:A:875:GLN:HA	2.47	0.50
1:A:766:VAL:HB	1:A:831:ILE:HG13	1.94	0.49
1:A:1449:LEU:HD12	1:A:1453:VAL:HG12	1.95	0.48
1:A:1332:GLU:HB3	1:A:1359:ARG:HD2	1.96	0.48
1:A:930:ILE:HD12	1:A:1056:PHE:CE2	2.49	0.48
1:A:639:ILE:O	1:A:640:GLU:HG3	2.13	0.47
1:A:989:LEU:O	1:A:1340:ARG:NH1	2.47	0.47
1:A:1489:PRO:C	1:A:1491:SER:H	2.18	0.47
1:A:1005:CYS:SG	1:A:1006:GLY:N	2.86	0.47
1:A:472:LEU:H	1:A:603:ALA:HB1	1.79	0.47
1:A:1456:HIS:CG	1:A:1457:LYS:N	2.82	0.47
1:A:1515:LEU:HD22	1:A:1535:MET:HA	1.97	0.47
1:A:551:PHE:HB2	3:A:2112:HOH:O	2.14	0.46
1:A:386:MET:HG3	1:A:387:LEU:N	2.31	0.46
1:A:734:ILE:HD11	1:A:828:LEU:HD13	1.98	0.46
1:A:879:ARG:NH2	1:A:1328:PRO:O	2.48	0.46
1:A:1543:HIS:HD2	1:A:1546:GLN:H	1.63	0.46
1:A:1569:GLY:O	1:A:1574:ARG:NE	2.39	0.46
1:A:416:PHE:CD2	1:A:494:LEU:HD22	2.50	0.46
1:A:625:LYS:HE3	1:A:851:GLY:HA2	1.98	0.46
1:A:766:VAL:CG1	1:A:831:ILE:CD1	2.94	0.46
1:A:1434:PHE:CG	1:A:1435:PRO:HD2	2.50	0.46
1:A:381:VAL:HG23	1:A:386:MET:HB3	1.98	0.46
1:A:1309:VAL:CG1	1:A:1311:GLN:HB3	2.46	0.46
1:A:772:ASP:C	1:A:774:SER:H	2.18	0.45
1:A:1418:LYS:HE2	1:A:1420:MET:HE1	1.99	0.45
1:A:992:PRO:HD2	1:A:1336:VAL:HG22	1.99	0.45
1:A:783:THR:HG21	1:A:897:CYS:HB2	1.98	0.45
1:A:932:LYS:HB2	1:A:1056:PHE:CE2	2.51	0.44
1:A:655:ARG:HD3	1:A:662:CYS:SG	2.57	0.44
1:A:1500:TRP:O	1:A:1503:PRO:HD2	2.17	0.44
1:A:522:LEU:HD13	1:A:581:ILE:CG2	2.47	0.44
1:A:1377:LEU:HD22	1:A:1412:LEU:HD11	2.00	0.44
1:A:1000:ILE:HG12	1:A:1019:LEU:HD23	2.00	0.43
1:A:640:GLU:HB3	1:A:641:LYS:H	1.53	0.43
1:A:966:ASP:HB2	1:A:967:PRO:HD2	1.99	0.43
1:A:1413:ARG:O	1:A:1414:ASP:HB2	2.19	0.43
1:A:1330:PHE:H	1:A:1356:ASN:HD21	1.65	0.43
1:A:995:TYR:OH	1:A:1359:ARG:HG2	2.19	0.43
1:A:1181:ARG:NH1	1:A:1189:VAL:HB	2.34	0.42
1:A:357:PRO:HB2	1:A:366:LEU:HB2	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:444:CYS:HG	1:A:458:GLY:HA3	1.85	0.42
1:A:918:ILE:HD11	1:A:928:SER:HB3	2.01	0.42
1:A:1480:SER:OG	1:A:1488:ASP:N	2.49	0.42
1:A:526:PRO:O	1:A:527:ASP:HB2	2.19	0.41
1:A:1139:LEU:HD12	1:A:1140:PRO:HD2	2.02	0.41
1:A:826:MET:HE1	1:A:830:TYR:HB3	2.02	0.41
1:A:472:LEU:N	1:A:603:ALA:HB1	2.35	0.41
1:A:617:ALA:N	3:A:2107:HOH:O	2.53	0.41
1:A:1502:LEU:HB2	1:A:1503:PRO:HD3	2.01	0.41
1:A:416:PHE:CG	1:A:494:LEU:HD22	2.56	0.41
1:A:635:PHE:C	1:A:637:GLU:H	2.24	0.41
1:A:1481:CYS:HA	1:A:1485:LYS:O	2.20	0.41
1:A:1584:PRO:HB2	1:A:1585:PRO:HD3	2.03	0.41
1:A:517:ILE:CD1	1:A:540:THR:HG21	2.51	0.41
1:A:939:LEU:HD22	1:A:1001:LYS:O	2.20	0.41
1:A:1231:GLY:HA3	1:A:1245:LYS:HG2	2.01	0.41
1:A:1429:ARG:HA	1:A:1547:HIS:CG	2.56	0.41
1:A:383:GLU:HB3	1:A:384:PRO:CD	2.51	0.41
1:A:772:ASP:O	1:A:774:SER:N	2.43	0.41
1:A:1511:HIS:O	1:A:1512:TRP:C	2.59	0.41
1:A:1526:SER:HB3	1:A:1539:GLY:HA3	2.02	0.41
1:A:626:LEU:HB3	1:A:1259:ASP:OD2	2.21	0.40
1:A:694:CYS:HA	1:A:695:PRO:HD3	1.95	0.40
1:A:1370:VAL:HG23	1:A:1524:PHE:HA	2.03	0.40
1:A:715:MET:HG3	1:A:716:PRO:HD2	2.03	0.40
1:A:879:ARG:NH2	1:A:1329:LEU:HD23	2.37	0.40
1:A:999:ARG:NH1	1:A:1047:TRP:CZ3	2.89	0.40
1:A:626:LEU:C	1:A:626:LEU:HD23	2.42	0.40
1:A:931:THR:CG2	1:A:932:LYS:N	2.84	0.40
1:A:1179:ALA:HA	1:A:1182:LEU:HD12	2.04	0.40
1:A:1268:LEU:C	1:A:1268:LEU:HD23	2.42	0.40

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	1118/1330 (84%)	1010 (90%)	89 (8%)	19 (2%)	9 37

All (19) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	465	GLY
1	A	715	MET
1	A	1203	ALA
1	A	1350	ASP
1	A	1361	SER
1	A	380	ALA
1	A	477	GLY
1	A	602	ARG
1	A	772	ASP
1	A	1040	THR
1	A	636	SER
1	A	640	GLU
1	A	690	LEU
1	A	776	PRO
1	A	1034	TYR
1	A	774	SER
1	A	1009	LYS
1	A	654	ARG
1	A	918	ILE

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	1001/1161 (86%)	981 (98%)	20 (2%)	55 76

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	387	LEU
1	A	451	GLU
1	A	495	MET
1	A	498	SER
1	A	529	VAL
1	A	565	SER
1	A	566	GLN
1	A	575	ASP
1	A	741	MET
1	A	751	GLN
1	A	775	LYS
1	A	843	SER
1	A	849	GLU
1	A	981	SER
1	A	999	ARG
1	A	1095	LYS
1	A	1213	ARG
1	A	1272	ARG
1	A	1350	ASP
1	A	1480	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	523	GLN
1	A	525	ASN
1	A	1043	ASN
1	A	1159	HIS
1	A	1160	GLN
1	A	1296	GLN
1	A	1356	ASN
1	A	1407	HIS
1	A	1510	ASN
1	A	1543	HIS

5.3.3 RNA

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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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^{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	1140/1330 (85%)	0.05	40 (3%) 44 42	42, 84, 137, 168	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	409	PRO	3.7
1	A	924	ARG	3.6
1	A	1002	GLU	3.4
1	A	489	PHE	3.3
1	A	390	GLU	3.3
1	A	729	GLN	2.9
1	A	719	LYS	2.9
1	A	1406	SER	2.8
1	A	894	HIS	2.7
1	A	691	LYS	2.7
1	A	895	LYS	2.6
1	A	832	HIS	2.6
1	A	393	SER	2.5
1	A	715	MET	2.5
1	A	1022	TYR	2.4
1	A	1000	ILE	2.4
1	A	923	GLY	2.4
1	A	939	LEU	2.4
1	A	1004	HIS	2.4
1	A	1012	VAL	2.3
1	A	831	ILE	2.3
1	A	937	TYR	2.3
1	A	410	MET	2.3
1	A	453	PRO	2.3
1	A	1003	ILE	2.3
1	A	1450	GLY	2.3
1	A	919	GLU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1070	ASP	2.2
1	A	1059	VAL	2.2
1	A	1005	CYS	2.2
1	A	1013	ASN	2.2
1	A	1611	LYS	2.1
1	A	1001	LYS	2.1
1	A	407	ASP	2.1
1	A	760	LEU	2.1
1	A	601	ARG	2.0
1	A	1452	GLY	2.0
1	A	693	ARG	2.0
1	A	722	HIS	2.0
1	A	408	SER	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^{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	ZN	A	2001	1/1	0.99	0.14	73,73,73,73	0
2	ZN	A	2002	1/1	0.99	0.07	88,88,88,88	0
2	ZN	A	2003	1/1	0.99	0.09	82,82,82,82	0
2	ZN	A	2004	1/1	1.00	0.13	53,53,53,53	0

6.5 Other polymers (i)

There are no such residues in this entry.