



Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 07:28 am BST

PDB ID : 5KPJ
Title : Mouse pgp methylated protein
Authors : Xia, D.; Esser, L.; Zhou, F.
Deposited on : 2016-07-04
Resolution : 3.50 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 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)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

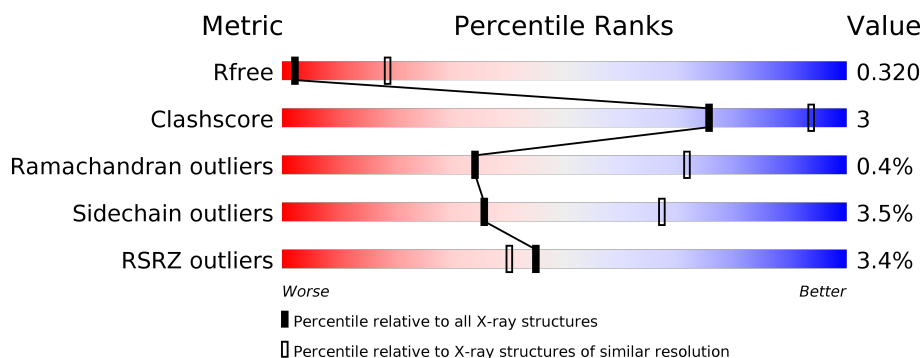
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 ≥ 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	1282	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9139 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 Multidrug resistance protein 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1177	Total	C	N	O	S	0	0	0
			9137	5879	1547	1673	38			


There are 6 discrepancies between the modelled and reference sequences:

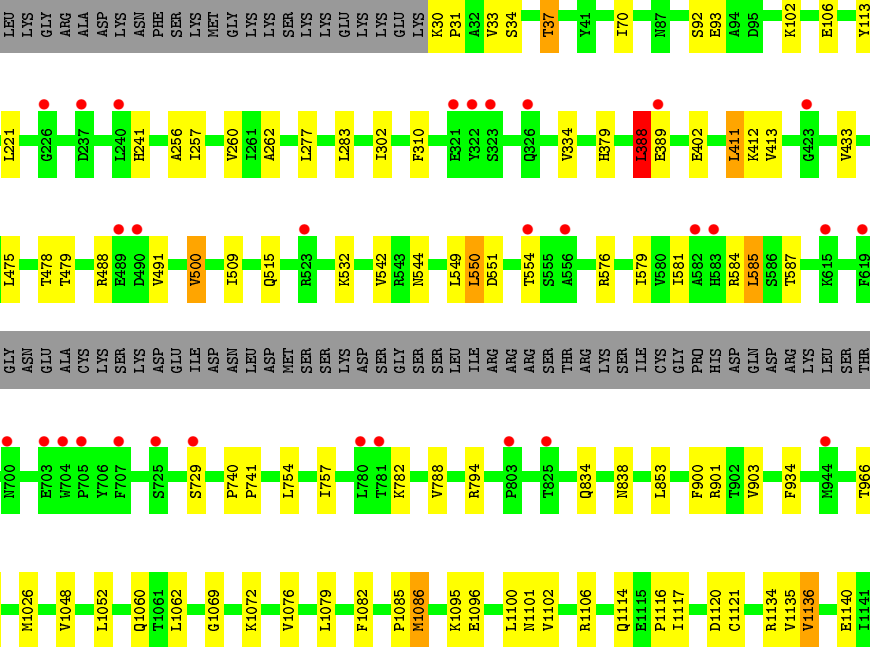
Chain	Residue	Modelled	Actual	Comment	Reference
A	1277	HIS	-	expression tag	UNP P21447
A	1278	HIS	-	expression tag	UNP P21447
A	1279	HIS	-	expression tag	UNP P21447
A	1280	HIS	-	expression tag	UNP P21447
A	1281	HIS	-	expression tag	UNP P21447
A	1282	HIS	-	expression tag	UNP P21447

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	O	0	0
			2	2		

These plots are drawn for all protein, RNA and DNA 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 ($\text{RSRZ} > 2$). 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.

- Chain A: 



The visualization displays the evolutionary paths of Chain A, categorized by generation (1 to 100). The chart is structured into three main horizontal panels, each representing a different phase of the chain's development. The top panel (generations 1-33) shows a high degree of diversity, with many different lineages represented by various colored squares. The middle panel (generations 34-66) shows a significant reduction in diversity, with only a few lineages remaining. The bottom panel (generations 67-100) shows a further reduction in diversity, with only one lineage remaining. The colors of the squares represent different lineages, with blue, green, and red being the most prominent. The grid is organized into columns, each representing a different lineage. The top section shows a high degree of diversity, with many different lineages represented. The middle section shows a significant reduction in diversity, with only a few lineages remaining. The bottom section shows a further reduction in diversity, with only one lineage remaining.

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	88.65Å 137.17Å 186.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	21.45 – 3.50 39.70 – 3.50	Depositor EDS
% Data completeness (in resolution range)	93.6 (21.45-3.50) 93.5 (39.70-3.50)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.98 (at 3.48Å)	Xtriage
Refinement program	PHENIX (dev_2443: ???)	Depositor
R, R_{free}	0.285 , 0.315 0.295 , 0.320	Depositor DCC
R_{free} test set	1359 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	109.6	Xtriage
Anisotropy	0.473	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.17 , -8.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.31$, $\langle L^2 \rangle = 0.14$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	9139	wwPDB-VP
Average B, all atoms (Å ²)	127.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.42% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: M3L

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	0.30	0/9282	0.59	7/12550 (0.1%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	443	LEU	CA-CB-CG	7.83	133.31	115.30
1	A	550	LEU	CA-CB-CG	6.50	130.26	115.30
1	A	388	LEU	CB-CG-CD1	6.17	121.48	111.00
1	A	411	LEU	CA-CB-CG	5.56	128.10	115.30
1	A	439	LEU	CA-CB-CG	5.51	127.96	115.30
1	A	453	ASP	CB-CG-OD1	5.39	123.16	118.30
1	A	585	LEU	CB-CG-CD1	5.24	119.91	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	9137	0	9326	57	0
2	A	2	0	0	0	0
All	All	9139	0	9326	57	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 3.

All (57) 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:1134:ARG:HG2	1:A:1136:VAL:HG23	1.74	0.68
1:A:379:HIS:HB3	1:A:457:ILE:HG22	1.80	0.62
1:A:277:LEU:HB3	1:A:782:LYS:HG3	1.85	0.59
1:A:1022:LEU:HB2	1:A:1100:LEU:HD23	1.84	0.59
1:A:544:ASN:OD1	1:A:576:ARG:NH2	2.38	0.56
1:A:788:VAL:HG21	1:A:1004:ILE:HD12	1.88	0.56
1:A:1069:GLY:H	1:A:1072:M3L:HM32	1.72	0.55
1:A:1229:ARG:O	1:A:1232:THR:OG1	2.26	0.53
1:A:1026:MET:SD	1:A:1095:LYS:NZ	2.78	0.52
1:A:102:LYS:NZ	1:A:106:GLU:OE2	2.40	0.52
1:A:1079:LEU:HD23	1:A:1194:LEU:HD21	1.91	0.52
1:A:70:ILE:HD13	1:A:113:TYR:HB3	1.90	0.52
1:A:1176:GLN:HG2	1:A:1179:ARG:HH21	1.74	0.52
1:A:473:PRO:HG2	1:A:532:LYS:HB3	1.91	0.52
1:A:1106:ARG:O	1:A:1188:ARG:NH1	2.44	0.51
1:A:491:VAL:HG21	1:A:542:VAL:HG13	1.95	0.49
1:A:433:VAL:HG13	1:A:549:LEU:HD23	1.95	0.49
1:A:256:ALA:HB2	1:A:1117:ILE:HG21	1.95	0.48
1:A:388:LEU:HB2	1:A:413:VAL:HB	1.96	0.48
1:A:1136:VAL:HG13	1:A:1140:GLU:HB2	1.94	0.48
1:A:1086:MET:N	1:A:1086:MET:SD	2.87	0.48
1:A:200:PHE:HE1	1:A:334:VAL:HG13	1.79	0.48
1:A:156:ILE:HD13	1:A:439:LEU:HG	1.96	0.47
1:A:1244:ASN:HB2	1:A:1246:M3L:HM13	1.97	0.47
1:A:221:LEU:HD22	1:A:302:ILE:HD13	1.97	0.46
1:A:257:ILE:HA	1:A:260:VAL:HG12	1.98	0.46
1:A:1116:PRO:HG3	1:A:1178:GLN:HA	1.98	0.45
1:A:554:THR:HG21	1:A:587:THR:HG21	1.99	0.45
1:A:729:SER:OG	1:A:968:GLU:O	2.34	0.45
1:A:389:GLU:HG2	1:A:412:LYS:HG2	2.00	0.44
1:A:92:SER:OG	1:A:93:GLU:N	2.50	0.44
1:A:1217:ALA:O	1:A:1221:ARG:NH1	2.51	0.44
1:A:34:SER:OG	1:A:37:THR:OG1	2.35	0.44
1:A:262:ALA:HB2	1:A:1082:PHE:HZ	1.83	0.43
1:A:500:VAL:HG11	1:A:509:ILE:HD12	2.01	0.43
1:A:129:VAL:HG23	1:A:184:ASP:HB3	2.00	0.43
1:A:1072:M3L:H	1:A:1072:M3L:HG2	1.59	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1076:VAL:HG13	1:A:1194:LEU:HB3	1.99	0.43
1:A:478:THR:HG22	1:A:479:THR:H	1.82	0.43
1:A:834:GLN:O	1:A:838:ASN:N	2.51	0.43
1:A:160:ASP:OD1	1:A:901:ARG:NH2	2.52	0.43
1:A:900:PHE:HA	1:A:903:VAL:HG12	2.01	0.42
1:A:754:LEU:HD12	1:A:757:ILE:HD12	2.00	0.42
1:A:584:ARG:O	1:A:587:THR:OG1	2.31	0.42
1:A:740:PRO:HA	1:A:741:PRO:HD3	1.91	0.42
1:A:30:LYS:HA	1:A:31:PRO:HD3	1.91	0.42
1:A:1142:VAL:HG22	1:A:1146:LYS:HE3	2.00	0.42
1:A:488:ARG:HG2	1:A:491:VAL:HG23	2.01	0.41
1:A:551:ASP:HA	1:A:581:ILE:HB	2.01	0.41
1:A:853:LEU:HB3	1:A:973:VAL:HG11	2.01	0.41
1:A:585:LEU:HD23	1:A:622:VAL:HG13	2.03	0.41
1:A:1021:GLY:HA3	1:A:1101:ASN:HB2	2.01	0.41
1:A:475:LEU:HA	1:A:475:LEU:HD23	1.89	0.41
1:A:1230:LEU:HD23	1:A:1230:LEU:HA	1.88	0.40
1:A:388:LEU:HD23	1:A:579:ILE:HD11	2.03	0.40
1:A:1062:LEU:HD11	1:A:1240:VAL:HG23	2.03	0.40
1:A:1230:LEU:O	1:A:1233:ILE:HG22	2.21	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	1171/1282 (91%)	1128 (96%)	38 (3%)	5 (0%)	34 72

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1160	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	515	GLN
1	A	1136	VAL
1	A	1120	ASP
1	A	411	LEU

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	969/1061 (91%)	935 (96%)	34 (4%)	36 67

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	33	VAL
1	A	37	THR
1	A	195	THR
1	A	241	HIS
1	A	283	LEU
1	A	310	PHE
1	A	388	LEU
1	A	402	GLU
1	A	443	LEU
1	A	458	ASN
1	A	470	SER
1	A	500	VAL
1	A	550	LEU
1	A	697	LEU
1	A	794	ARG
1	A	934	PHE
1	A	966	THR
1	A	1048	VAL
1	A	1052	LEU
1	A	1060	GLN
1	A	1085	PRO
1	A	1086	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1096	GLU
1	A	1102	VAL
1	A	1114	GLN
1	A	1121	CYS
1	A	1135	VAL
1	A	1150	ILE
1	A	1152	GLN
1	A	1155	ASP
1	A	1156	SER
1	A	1187	VAL
1	A	1200	SER
1	A	1206	SER

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	145	GLN
1	A	458	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are 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	$\# Z > 2$
1	M3L	A	1072	1	10,11,12	0.51	0	9,14,16	0.65	0
1	M3L	A	1246	1	10,11,12	0.48	0	9,14,16	0.46	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
1	M3L	A	1072	1	-	8/9/10/12	-
1	M3L	A	1246	1	-	3/9/10/12	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1072	M3L	N-CA-CB-CG
1	A	1246	M3L	N-CA-CB-CG
1	A	1246	M3L	C-CA-CB-CG
1	A	1072	M3L	CG-CD-CE-NZ
1	A	1072	M3L	CD-CE-NZ-CM1
1	A	1072	M3L	CD-CE-NZ-CM2
1	A	1072	M3L	CD-CE-NZ-CM3
1	A	1072	M3L	CE-CD-CG-CB
1	A	1072	M3L	CA-CB-CG-CD
1	A	1246	M3L	CG-CD-CE-NZ
1	A	1072	M3L	C-CA-CB-CG

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1072	M3L	2	0
1	A	1246	M3L	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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	1175/1282 (91%)	-0.07	40 (3%) 45 40	30, 124, 181, 225	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	704	TRP	6.7
1	A	326	GLN	4.6
1	A	690	PRO	4.2
1	A	705	PRO	4.0
1	A	582	ALA	3.8
1	A	321	GLU	3.8
1	A	489	GLU	3.7
1	A	725	SER	3.3
1	A	322	TYR	3.2
1	A	554	THR	3.1
1	A	523	ARG	3.1
1	A	781	THR	3.0
1	A	226	GLY	2.9
1	A	700	ASN	2.9
1	A	583	HIS	2.9
1	A	623	MET	2.8
1	A	944	MET	2.8
1	A	996	LYS	2.8
1	A	619	PHE	2.7
1	A	703	GLU	2.7
1	A	707	PHE	2.7
1	A	389	GLU	2.7
1	A	240	LEU	2.5
1	A	803	PRO	2.5
1	A	729	SER	2.5
1	A	556	ALA	2.4
1	A	825	THR	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1220	GLY	2.3
1	A	975	SER	2.3
1	A	41	TYR	2.3
1	A	1152	GLN	2.3
1	A	490	ASP	2.3
1	A	323	SER	2.3
1	A	87	ASN	2.2
1	A	237	ASP	2.2
1	A	780	LEU	2.2
1	A	1219	GLU	2.1
1	A	615	LYS	2.1
1	A	423	GLY	2.1
1	A	95	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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, 95th 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
1	M3L	A	1246	12/13	0.84	0.31	75,91,101,107	0
1	M3L	A	1072	12/13	0.92	0.50	122,127,134,140	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.