Full wwPDB X-ray Structure Validation Report ⓘ

Oct 6, 2022 – 07:08 PM EDT

PDB ID	:	5SJH
Title	:	Crystal Structure of human phosphodiesterase 10
Authors	:	Joseph, C.; Benz, J.; Flohr, A.; Rudolph, M.G.
Deposited on	:	2022-02-01
Resolution	:	2.10 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:

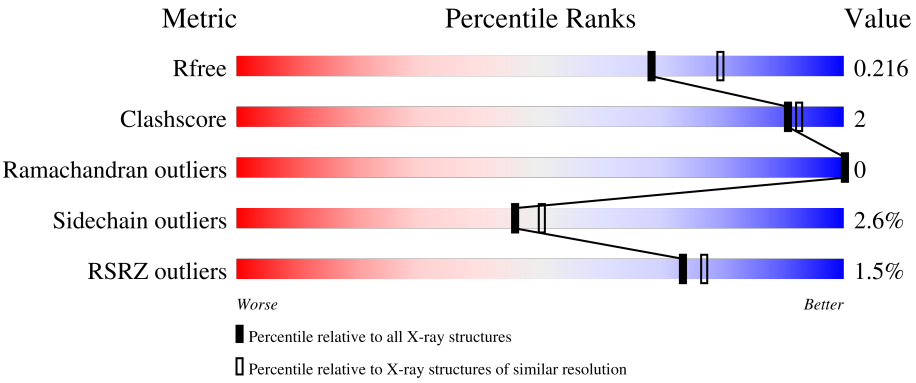
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	343	84% 5% • 10%
1	B	343	86% • • 9%
1	C	343	87% • 9%
1	D	343	84% 6% • 10%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 11045 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 cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	310	Total	C	N	O	S	0	12	0
			2586	1658	441	462	25			
1	B	311	Total	C	N	O	S	0	6	0
			2558	1638	435	461	24			
1	C	312	Total	C	N	O	S	0	10	0
			2583	1654	443	462	24			
1	D	310	Total	C	N	O	S	0	10	0
			2578	1650	444	460	24			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	447	GLY	-	expression tag	UNP Q9Y233
A	448	SER	-	expression tag	UNP Q9Y233
B	447	GLY	-	expression tag	UNP Q9Y233
B	448	SER	-	expression tag	UNP Q9Y233
C	447	GLY	-	expression tag	UNP Q9Y233
C	448	SER	-	expression tag	UNP Q9Y233
D	447	GLY	-	expression tag	UNP Q9Y233
D	448	SER	-	expression tag	UNP Q9Y233

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0
3	B	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0
3	D	1	Total Mg 1 1	0	0

- Molecule 4 is water.

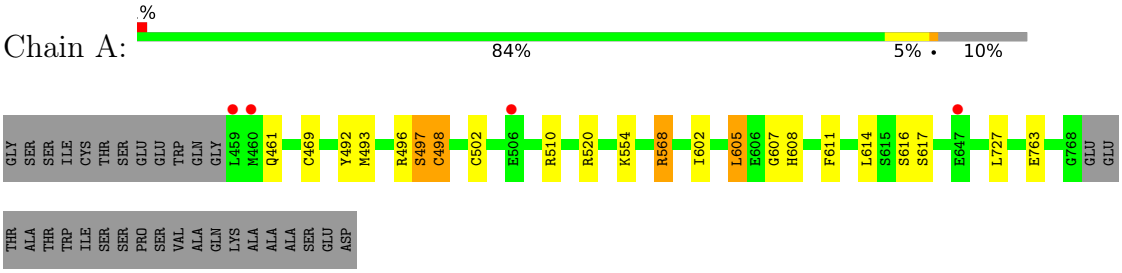
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	187	Total O 187 187	0	0
4	B	202	Total O 202 202	0	0
4	C	202	Total O 202 202	0	0
4	D	141	Total O 141 141	0	0

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.

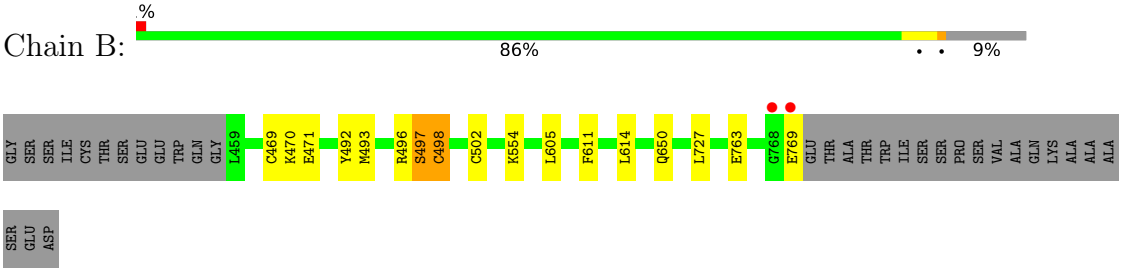
• Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

Chain A:



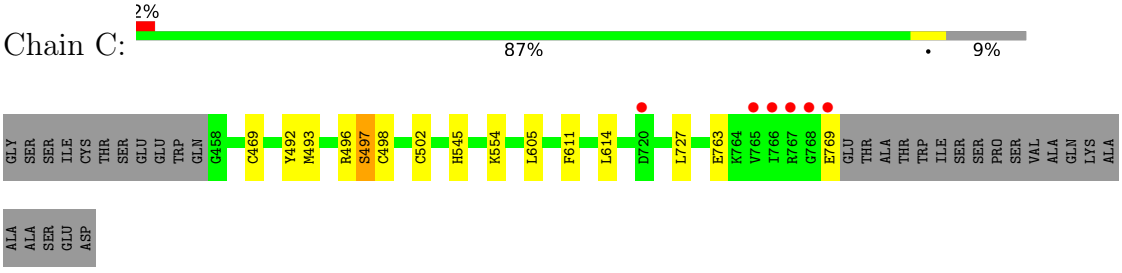
• Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

Chain B:



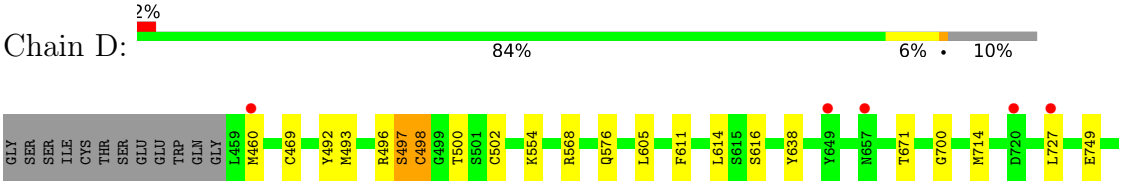
• Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

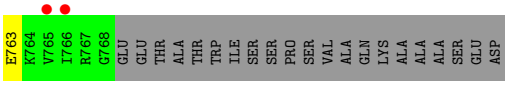
Chain C:



• Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

Chain D:




E763
K764
V765
I766
R767
G768
GLU
GLU
THR
ALA
THR
TRP
ILE
SER
SER
PRO
SER
VAL
ALA
GLN
LYS
ALA
ALA
ALA
SER
GLU
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	136.31Å 136.31Å 235.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.72 – 2.10 41.72 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.9 (41.72-2.10) 95.9 (41.72-2.10)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 2.10Å)	Xtriage
Refinement program	BUSTER 2.11.7 (19-MAR-2020), REFMAC 5.8.0258	Depositor
R, R_{free}	0.177 , 0.204 0.186 , 0.216	Depositor DCC
R_{free} test set	4586 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	35.5	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 48.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.032 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11045	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.62% 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: MG, 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	0.42	0/2684	0.53	0/3628
1	B	0.43	0/2638	0.54	0/3569
1	C	0.42	0/2675	0.53	0/3616
1	D	0.39	0/2670	0.52	0/3612
All	All	0.42	0/10667	0.53	0/14425

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	2586	0	2599	11	0
1	B	2558	0	2549	8	0
1	C	2583	0	2593	7	0
1	D	2578	0	2583	8	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	187	0	0	0	0
4	B	202	0	0	0	0
4	C	202	0	0	1	0
4	D	141	0	0	0	0
All	All	11045	0	10324	34	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.

All (34) 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:C:502:CYS:HB2	1:C:554:LYS:HE2	1.69	0.73
1:C:545:HIS:HD2	4:C:1003:HOH:O	1.75	0.68
1:B:727:LEU:HD21	1:B:763:GLU:HG3	1.77	0.67
1:A:727:LEU:HD21	1:A:763:GLU:HG3	1.78	0.65
1:D:727:LEU:HD21	1:D:763:GLU:HG3	1.79	0.64
1:C:727:LEU:HD21	1:C:763:GLU:HG3	1.81	0.61
1:C:493:MET:O	1:C:497:SER:HB2	2.04	0.57
1:D:493:MET:O	1:D:497:SER:HB2	2.06	0.56
1:B:493:MET:O	1:B:497:SER:HB2	2.06	0.55
1:A:493:MET:O	1:A:497:SER:HB2	2.08	0.53
1:A:510[A]:ARG:HD2	1:A:607:GLY:O	2.09	0.53
1:D:502[A]:CYS:HB2	1:D:554:LYS:HE2	1.92	0.52
1:A:502[A]:CYS:HB2	1:A:554:LYS:HE2	1.90	0.52
1:B:502[A]:CYS:HB2	1:B:554:LYS:HE2	1.92	0.51
1:A:498:CYS:O	1:A:502[B]:CYS:SG	2.71	0.48
1:B:498:CYS:O	1:B:502[B]:CYS:SG	2.72	0.48
1:C:502:CYS:CB	1:C:554:LYS:HE2	2.43	0.47
1:D:498:CYS:O	1:D:502[B]:CYS:SG	2.72	0.47
1:D:492:TYR:CZ	1:D:496:ARG:HD2	2.50	0.46
1:B:611:PHE:HB3	1:B:614:LEU:HD22	1.98	0.45
1:C:611:PHE:HB3	1:C:614:LEU:HD22	1.98	0.45
1:A:520:ARG:NH2	1:A:568[B]:ARG:NH2	2.65	0.44
1:B:492:TYR:CZ	1:B:496:ARG:HD2	2.53	0.44
1:D:611:PHE:HB3	1:D:614:LEU:HD22	1.99	0.43
1:D:700:GLY:HA3	1:D:714:MET:O	2.19	0.43
1:B:470:LYS:HG3	1:B:471:GLU:N	2.34	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:492:TYR:CZ	1:A:496:ARG:HD2	2.55	0.42
1:D:638:TYR:CD1	1:D:671:THR:HG21	2.55	0.42
1:A:510[B]:ARG:HG2	1:A:608:HIS:CE1	2.55	0.41
1:B:498:CYS:HB2	1:B:502[B]:CYS:SG	2.61	0.41
1:A:611:PHE:HB3	1:A:614:LEU:HD22	2.02	0.41
1:C:492:TYR:CZ	1:C:496:ARG:HD2	2.56	0.41
1:A:498:CYS:HB2	1:A:502[B]:CYS:SG	2.61	0.40
1:A:602:ILE:HA	1:A:605:LEU:HD22	2.04	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	320/343 (93%)	314 (98%)	6 (2%)	0	100	100
1	B	315/343 (92%)	309 (98%)	6 (2%)	0	100	100
1	C	320/343 (93%)	314 (98%)	6 (2%)	0	100	100
1	D	318/343 (93%)	313 (98%)	5 (2%)	0	100	100
All	All	1273/1372 (93%)	1250 (98%)	23 (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	292/306 (95%)	282 (97%)	10 (3%)	37	39
1	B	287/306 (94%)	281 (98%)	6 (2%)	53	59
1	C	291/306 (95%)	285 (98%)	6 (2%)	53	59
1	D	290/306 (95%)	278 (96%)	12 (4%)	30	31
All	All	1160/1224 (95%)	1126 (97%)	34 (3%)	46	46

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

Mol	Chain	Res	Type
1	A	461	GLN
1	A	469[A]	CYS
1	A	469[B]	CYS
1	A	497	SER
1	A	498	CYS
1	A	568[A]	ARG
1	A	568[B]	ARG
1	A	605	LEU
1	A	616	SER
1	A	617	SER
1	B	469	CYS
1	B	497	SER
1	B	498	CYS
1	B	605	LEU
1	B	650	GLN
1	B	769	GLU
1	C	469[A]	CYS
1	C	469[B]	CYS
1	C	497	SER
1	C	498	CYS
1	C	605	LEU
1	C	769	GLU
1	D	460	MET
1	D	469	CYS
1	D	497	SER
1	D	498	CYS
1	D	500	THR
1	D	568[A]	ARG
1	D	568[B]	ARG
1	D	576	GLN
1	D	605	LEU
1	D	616	SER
1	D	749[A]	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	749[B]	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
1	A	604	GLN
1	A	743	GLN
1	B	495	HIS
1	B	542	GLN
1	B	604	GLN
1	B	743	GLN
1	C	495	HIS
1	C	604	GLN
1	C	743	GLN
1	D	542	GLN
1	D	604	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 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	310/343 (90%)	-0.18	4 (1%)	77	80	25, 38, 60, 76	0
1	B	311/343 (90%)	-0.13	2 (0%)	89	91	23, 37, 56, 75	0
1	C	312/343 (90%)	-0.16	6 (1%)	66	71	23, 36, 57, 73	0
1	D	310/343 (90%)	-0.05	7 (2%)	60	65	29, 44, 64, 74	0
All	All	1243/1372 (90%)	-0.13	19 (1%)	73	77	23, 39, 60, 76	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	720	ASP	4.8
1	A	460	MET	2.9
1	D	460	MET	2.9
1	B	769	GLU	2.9
1	C	769	GLU	2.8
1	C	765	VAL	2.7
1	B	768	GLY	2.5
1	C	766	ILE	2.5
1	D	766	ILE	2.5
1	D	649	TYR	2.4
1	D	765	VAL	2.4
1	A	647[A]	GLU	2.3
1	A	506[A]	GLU	2.3
1	A	459	LEU	2.2
1	C	768	GLY	2.2
1	D	727	LEU	2.1
1	C	767	ARG	2.1
1	C	720	ASP	2.0
1	D	657	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^{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
3	MG	C	802	1/1	0.97	0.09	27,27,27,27	0
3	MG	B	802	1/1	0.99	0.09	30,30,30,30	0
3	MG	A	802	1/1	0.99	0.06	30,30,30,30	0
3	MG	D	802	1/1	0.99	0.07	37,37,37,37	0
2	ZN	A	801	1/1	1.00	0.11	33,33,33,33	0
2	ZN	B	801	1/1	1.00	0.10	31,31,31,31	0
2	ZN	C	801	1/1	1.00	0.10	31,31,31,31	0
2	ZN	D	801	1/1	1.00	0.08	38,38,38,38	0

6.5 Other polymers (i)

There are no such residues in this entry.