Full wwPDB X-ray Structure Validation Report ⓘ

Sep 8, 2025 – 11:39 PM JST

PDB ID : 7CFH / pdb_00007cfh
Title : Structure of the CBS domain of the bacterial CNNM/CorC family Mg2+ transporter
Authors : Huang, Y.; Jin, F.; Hattori, M.
Deposited on : 2020-06-25
Resolution : 2.00 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

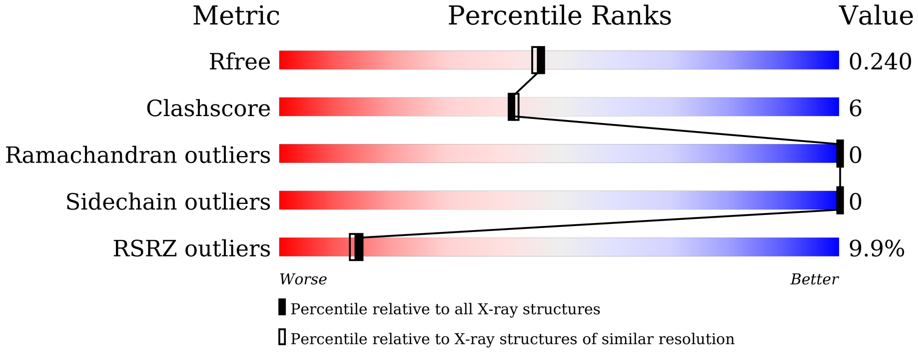
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

1 Overall quality at a glance ⓘ

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

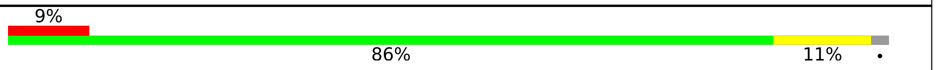
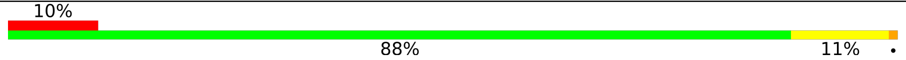
The reported resolution of this entry is 2.00 Å.

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	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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	169	
1	B	169	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2707 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.

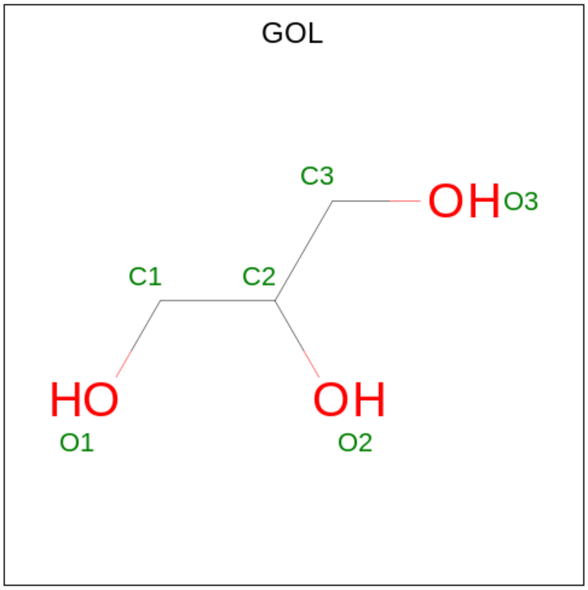
- Molecule 1 is a protein called Hemolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	165	1261	803	201	251	6	0	0	0
1	B	169	1293	822	213	252	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	255	ALA	TYR	conflict	UNP A0A109QFA5
B	255	ALA	TYR	conflict	UNP A0A109QFA5

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



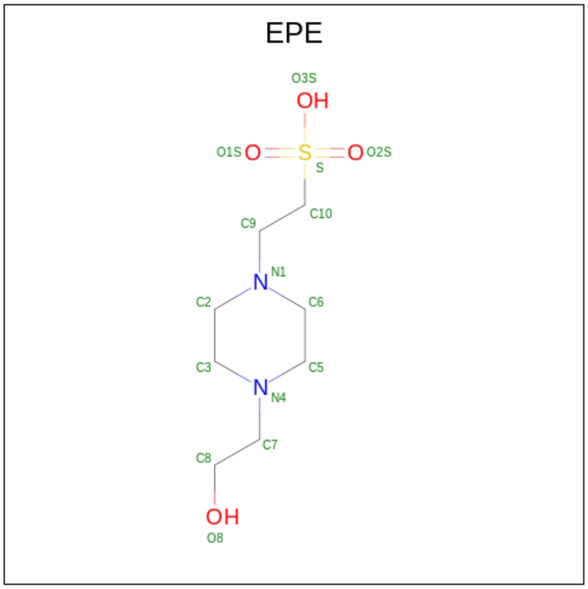
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	6	3	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 4 3 1	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (CCD ID: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C N O S 14 8 2 3 1	0	0

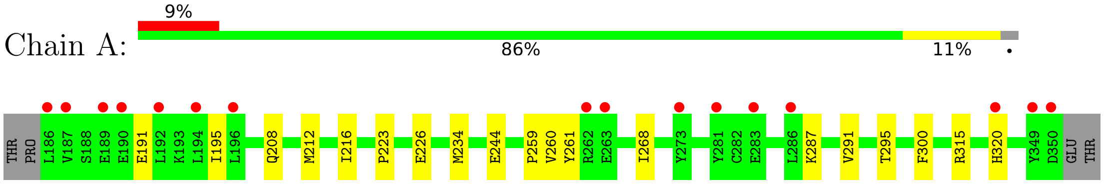
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	56	Total O 56 56	0	0
4	B	55	Total O 55 55	0	0

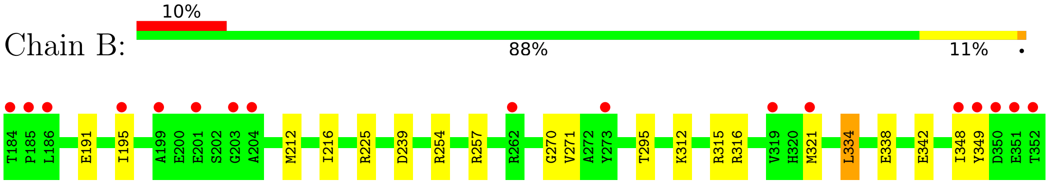
3 Residue-property plots ⓘ

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: Hemolysin



- Molecule 1: Hemolysin



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	89.52Å 61.31Å 81.83Å 90.00° 92.41° 90.00°	Depositor
Resolution (Å)	42.56 – 2.00 42.56 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.56-2.00) 95.5 (42.56-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.204 , 0.240 0.205 , 0.240	Depositor DCC
R_{free} test set	2005 reflections (6.67%)	wwPDB-VP
Wilson B-factor (Å^2)	37.2	Xtriage
Anisotropy	0.264	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 40.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.015 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2707	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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: GOL, EPE

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	0.40	0/1283	0.65	0/1750
1	B	0.40	0/1316	0.66	1/1794 (0.1%)
All	All	0.40	0/2599	0.66	1/3544 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	334	LEU	CB-CG-CD2	-6.94	89.88	110.70

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	1261	0	1210	16	0
1	B	1293	0	1246	21	0
2	A	6	0	8	1	0
2	B	22	0	25	1	0
3	B	14	0	16	0	0
4	A	56	0	0	1	0
4	B	55	0	0	6	0
All	All	2707	0	2505	32	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 (32) 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:239:ASP:OD1	4:B:501:HOH:O	2.03	0.76
2:B:403:GOL:O1	4:B:502:HOH:O	2.12	0.67
1:A:234:MET:HE3	1:A:259:PRO:HG2	1.80	0.63
1:A:320:HIS:CE1	1:B:349:TYR:HB3	2.36	0.60
1:B:257:ARG:HH22	1:B:338:GLU:CD	2.10	0.60
1:B:321:MET:HE2	1:B:334:LEU:HD22	1.84	0.58
1:A:212:MET:O	1:A:216:ILE:HG12	2.07	0.55
1:A:244:GLU:OE1	1:A:287:LYS:HE2	2.07	0.55
1:A:216:ILE:HD12	1:B:212:MET:CE	2.38	0.54
1:A:315:ARG:HA	1:B:348:ILE:HG23	1.89	0.54
1:A:216:ILE:HD12	1:B:212:MET:HE2	1.90	0.54
1:B:342:GLU:OE2	4:B:504:HOH:O	2.18	0.52
1:B:321:MET:HE2	1:B:334:LEU:CD2	2.40	0.51
1:B:271:VAL:HG11	1:B:321:MET:SD	2.53	0.48
1:B:191:GLU:O	1:B:195:ILE:HG13	2.14	0.48
1:A:320:HIS:ND1	1:B:349:TYR:HD2	2.11	0.48
1:B:315:ARG:NH1	4:B:507:HOH:O	2.43	0.47
1:A:320:HIS:HE1	1:B:349:TYR:HB3	1.77	0.47
1:A:208:GLN:NE2	4:A:502:HOH:O	2.48	0.46
1:B:254:ARG:HD3	1:B:254:ARG:HA	1.68	0.46
1:A:191:GLU:O	1:A:195:ILE:HG12	2.16	0.46
1:A:260:VAL:HG21	1:A:295:THR:HG22	1.98	0.46
1:A:291:VAL:O	1:A:295:THR:HG23	2.16	0.45
1:A:223:PRO:HG2	1:A:226:GLU:HG3	2.00	0.44
1:B:212:MET:O	1:B:216:ILE:HG12	2.18	0.44
1:B:225:ARG:NH1	4:B:503:HOH:O	2.15	0.43
1:B:334:LEU:HA	1:B:334:LEU:HD23	1.92	0.42
1:A:261:TYR:HB3	1:A:268:ILE:HD13	2.01	0.42
1:B:316:ARG:H	1:B:316:ARG:HG2	1.50	0.41
1:A:300:PHE:CE1	2:A:401:GOL:H2	2.55	0.41
1:B:270:GLY:HA3	1:B:295:THR:CG2	2.50	0.40
1:B:312:LYS:NZ	4:B:512:HOH:O	2.54	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	163/169 (96%)	159 (98%)	4 (2%)	0	100	100
1	B	167/169 (99%)	162 (97%)	5 (3%)	0	100	100
All	All	330/338 (98%)	321 (97%)	9 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	132/148 (89%)	132 (100%)	0	100	100
1	B	134/148 (90%)	134 (100%)	0	100	100
All	All	266/296 (90%)	266 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	206	GLN
1	A	267	HIS
1	B	253	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

6 ligands 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	$\#\|Z\| > 2$
2	GOL	B	403	-	5,5,5	0.85	0	5,5,5	0.99	0
3	EPE	B	405	-	12,14,15	0.39	0	15,17,20	2.32	5 (33%)
2	GOL	A	401	-	5,5,5	1.10	0	5,5,5	1.08	0
2	GOL	B	404	-	5,5,5	0.95	0	5,5,5	1.02	0
2	GOL	B	402	-	5,5,5	0.55	0	5,5,5	1.13	0
2	GOL	B	401	-	3,3,5	1.10	0	3,3,5	0.58	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
2	GOL	B	403	-	-	4/4/4/4	-
3	EPE	B	405	-	-	4/6/18/19	0/1/1/1
2	GOL	A	401	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	B	404	-	-	0/4/4/4	-
2	GOL	B	402	-	-	2/4/4/4	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	405	EPE	C5-N4-C3	6.26	122.93	108.83
3	B	405	EPE	C7-N4-C5	2.84	118.50	111.23
3	B	405	EPE	C7-N4-C3	2.67	118.06	111.23
3	B	405	EPE	C2-C3-N4	2.49	115.75	110.64
3	B	405	EPE	O1S-S-C10	2.27	111.12	105.69

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	402	GOL	O1-C1-C2-C3
2	B	403	GOL	O1-C1-C2-C3
2	B	403	GOL	C1-C2-C3-O3
2	A	401	GOL	C1-C2-C3-O3
2	B	402	GOL	O1-C1-C2-O2
2	B	403	GOL	O1-C1-C2-O2
2	B	403	GOL	O2-C2-C3-O3
3	B	405	EPE	C8-C7-N4-C5
3	B	405	EPE	C10-C9-N1-C2
3	B	405	EPE	C10-C9-N1-C6
2	A	401	GOL	O2-C2-C3-O3
3	B	405	EPE	N4-C7-C8-O8

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	403	GOL	1	0
2	A	401	GOL	1	0

5.7 Other polymers ⓘ

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, 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	165/169 (97%)	0.55	16 (9%)	15	13	29, 42, 68, 80	0
1	B	169/169 (100%)	0.61	17 (10%)	14	13	29, 41, 73, 86	0
All	All	334/338 (98%)	0.58	33 (9%)	14	13	29, 42, 73, 86	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	184	THR	5.1
1	B	352	THR	4.8
1	B	349	TYR	4.7
1	B	185	PRO	4.2
1	A	273	TYR	4.1
1	B	348	ILE	3.8
1	A	349	TYR	3.5
1	B	203	GLY	3.4
1	B	199	ALA	3.4
1	A	189	GLU	3.4
1	A	194	LEU	3.0
1	B	321	MET	3.0
1	A	196	LEU	2.9
1	B	186	LEU	2.9
1	A	190	GLU	2.9
1	B	350	ASP	2.8
1	A	187	VAL	2.7
1	A	262	ARG	2.6
1	B	204	ALA	2.5
1	A	283	GLU	2.5
1	B	319	VAL	2.5
1	B	262	ARG	2.5
1	B	201	GLU	2.4
1	A	281	TYR	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	286	LEU	2.4
1	A	263	GLU	2.4
1	B	195	ILE	2.4
1	A	186	LEU	2.3
1	A	192	LEU	2.2
1	A	350	ASP	2.2
1	A	320	HIS	2.2
1	B	351	GLU	2.1
1	B	273	TYR	2.0

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 oligosaccharides 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, 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
3	EPE	B	405	14/15	0.73	0.18	65,72,77,84	0
2	GOL	B	404	6/6	0.78	0.13	72,75,79,80	0
2	GOL	B	403	6/6	0.84	0.13	62,65,67,72	0
2	GOL	B	401	4/6	0.85	0.19	40,49,52,54	0
2	GOL	B	402	6/6	0.86	0.14	44,50,52,58	0
2	GOL	A	401	6/6	0.91	0.11	41,43,47,51	0

6.5 Other polymers ⓘ

There are no such residues in this entry.