Full wwPDB X-ray Structure Validation Report ⓘ

Apr 3, 2025 – 04:04 PM EDT

PDB ID	:	9D2Q / pdb_00009d2q
Title	:	Crystal structure of E. coli Threonine dehydratase
Authors	:	Yun, M.; Rock, C.O.
Deposited on	:	2024-08-09
Resolution	:	1.87 Å(reported)

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

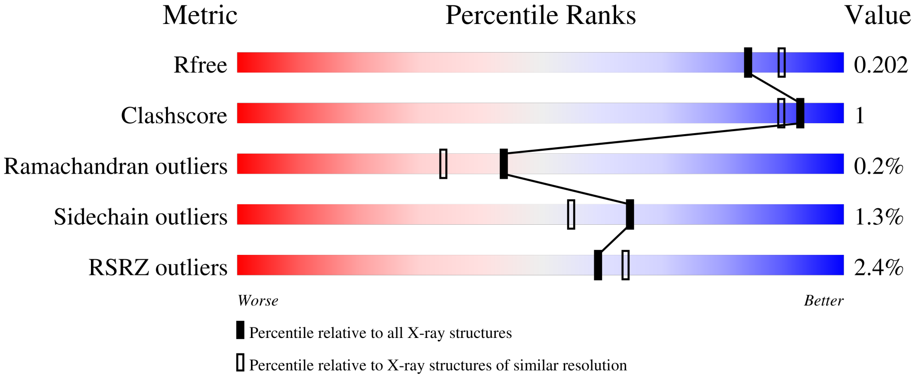
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

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 1.87 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	1090 (1.88-1.88)
Clashscore	180529	1144 (1.88-1.88)
Ramachandran outliers	177936	1135 (1.88-1.88)
Sidechain outliers	177891	1135 (1.88-1.88)
RSRZ outliers	164620	1090 (1.88-1.88)

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	533	

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 4082 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.

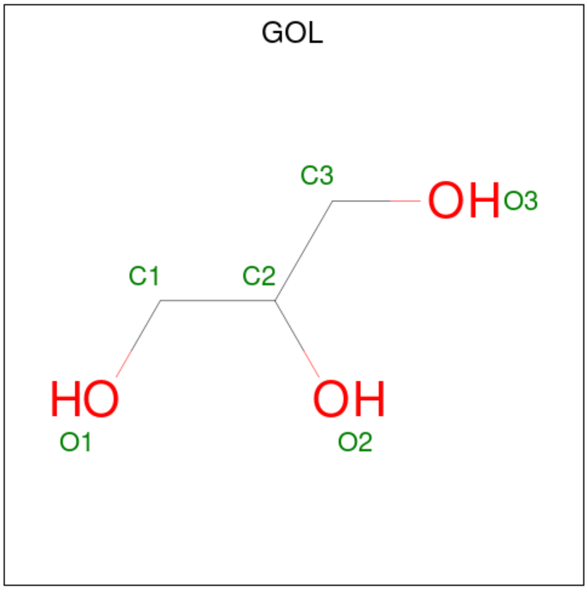
- Molecule 1 is a protein called L-threonine dehydratase biosynthetic IlvA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	494	3811	2418	669	705	1	18	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

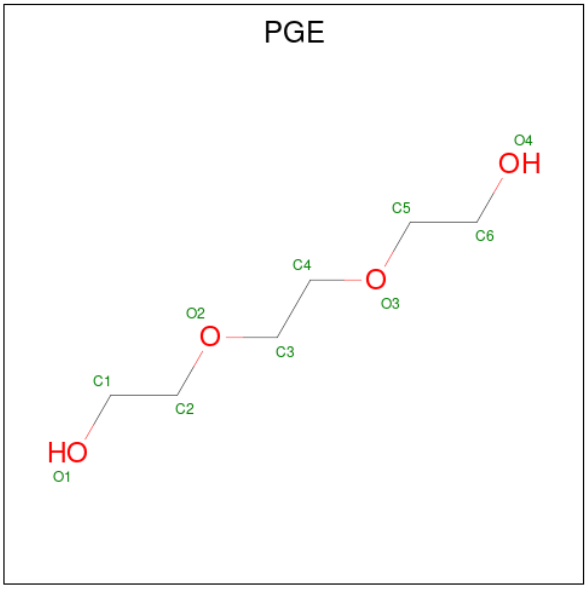
Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MET	-	initiating methionine	UNP P04968
A	-17	GLY	-	expression tag	UNP P04968
A	-16	SER	-	expression tag	UNP P04968
A	-15	SER	-	expression tag	UNP P04968
A	-14	HIS	-	expression tag	UNP P04968
A	-13	HIS	-	expression tag	UNP P04968
A	-12	HIS	-	expression tag	UNP P04968
A	-11	HIS	-	expression tag	UNP P04968
A	-10	HIS	-	expression tag	UNP P04968
A	-9	HIS	-	expression tag	UNP P04968
A	-8	SER	-	expression tag	UNP P04968
A	-7	SER	-	expression tag	UNP P04968
A	-6	GLY	-	expression tag	UNP P04968
A	-5	LEU	-	expression tag	UNP P04968
A	-4	VAL	-	expression tag	UNP P04968
A	-3	PRO	-	expression tag	UNP P04968
A	-2	ARG	-	expression tag	UNP P04968
A	-1	GLY	-	expression tag	UNP P04968
A	0	SER	-	expression tag	UNP P04968
A	1	HIS	-	expression tag	UNP P04968

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



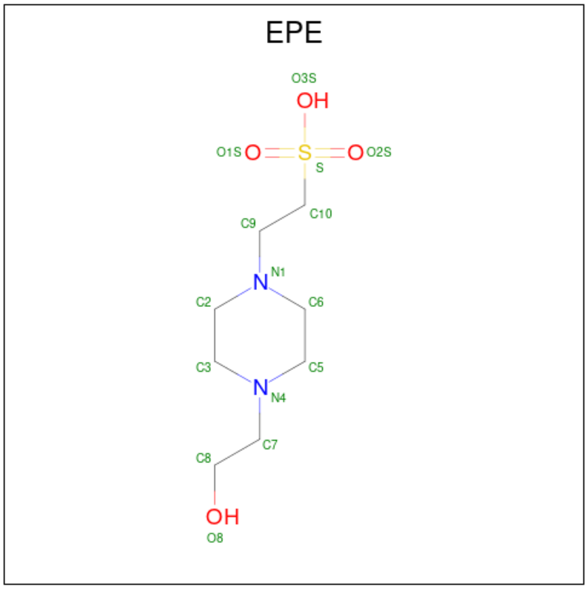
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0

- Molecule 3 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 10 6 4	0	0

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	A	1	15	8	2	4	1	0	0

- Molecule 5 is water.

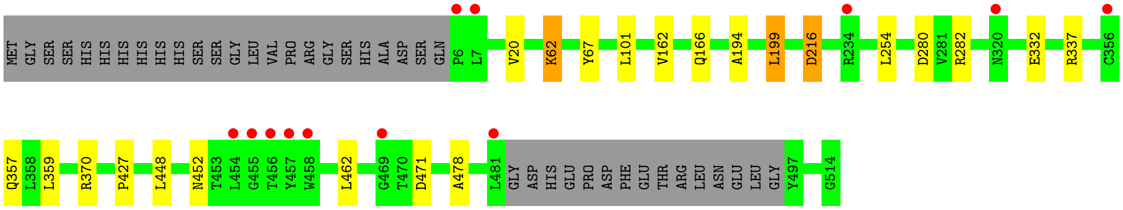
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	228	228	228	0	0

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: L-threonine dehydratase biosynthetic IlvA

Chain A:



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	80.02Å 93.10Å 161.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	60.68 – 1.87 60.68 – 1.87	Depositor EDS
% Data completeness (in resolution range)	99.9 (60.68-1.87) 95.7 (60.68-1.87)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 1.87Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.176 , 0.202 0.177 , 0.202	Depositor DCC
R_{free} test set	48182 reflections (3.97%)	wwPDB-VP
Wilson B-factor (Å^2)	34.0	Xtriage
Anisotropy	0.536	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 44.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4082	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% 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: EPE, PGE, LLP, GOL

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	0.35	0/3861	0.52	0/5218

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	3811	0	3789	11	0
2	A	18	0	24	0	0
3	A	10	0	14	0	0
4	A	15	0	17	0	0
5	A	228	0	0	1	0
All	All	4082	0	3844	11	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 1.

All (11) 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:20:VAL:HG11	1:A:199:LEU:HD13	1.73	0.70
1:A:357:GLN:HG3	1:A:448:LEU:HD11	1.75	0.68
1:A:216:ASP:OD1	1:A:216:ASP:N	2.33	0.62
1:A:194:ALA:HB1	1:A:254:LEU:HD12	1.94	0.50
1:A:332:GLU:OE2	1:A:337:ARG:NH1	2.50	0.43
1:A:162:VAL:O	1:A:166:GLN:HG2	2.18	0.43
1:A:370:ARG:NH2	1:A:427:PRO:HD3	2.35	0.42
1:A:67:TYR:CE1	1:A:101:LEU:HG	2.55	0.42
1:A:462:LEU:HB3	1:A:478:ALA:HB3	2.02	0.41
1:A:359:LEU:HD23	1:A:359:LEU:HA	1.93	0.41
1:A:62:LLP:HD3	5:A:762:HOH:O	2.21	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	489/533 (92%)	477 (98%)	11 (2%)	1 (0%)	44 34

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	471	ASP

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	389/423 (92%)	384 (99%)	5 (1%)	65 55

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	199	LEU
1	A	216	ASP
1	A	280	ASP
1	A	282	ARG
1	A	452	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	LLP	A	62	1	23,24,25	1.24	2 (8%)	25,32,34	1.35	2 (8%)

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	LLP	A	62	1	-	4/16/17/19	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	62	LLP	C4'-NZ	3.71	1.39	1.27
1	A	62	LLP	C4-C4'	2.21	1.51	1.46

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	62	LLP	CE-NZ-C4'	4.42	132.89	118.72
1	A	62	LLP	C4-C4'-NZ	-3.12	109.62	124.04

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	62	LLP	C4-C4'-NZ-CE
1	A	62	LLP	CE-CD-CG-CB
1	A	62	LLP	C5'-OP4-P-OP1
1	A	62	LLP	C3-C4-C4'-NZ

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	62	LLP	1	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

5 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	A	605	-	5,5,5	0.97	0	5,5,5	1.01	0
2	GOL	A	602	-	5,5,5	0.93	0	5,5,5	1.12	0
3	PGE	A	603	-	9,9,9	0.30	0	8,8,8	0.30	0
4	EPE	A	604	-	15,15,15	0.89	1 (6%)	19,20,20	1.70	5 (26%)
2	GOL	A	601	-	5,5,5	0.92	0	5,5,5	1.11	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	A	605	-	-	0/4/4/4	-
2	GOL	A	602	-	-	2/4/4/4	-
3	PGE	A	603	-	-	2/7/7/7	-
4	EPE	A	604	-	-	7/9/19/19	0/1/1/1
2	GOL	A	601	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	604	EPE	C10-S	3.08	1.81	1.77

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	604	EPE	C5-N4-C3	4.35	118.21	108.84
4	A	604	EPE	C7-N4-C5	2.98	119.19	111.24
4	A	604	EPE	C7-N4-C3	2.95	119.11	111.24
4	A	604	EPE	O1S-S-C10	2.43	110.41	106.73
4	A	604	EPE	O3S-S-C10	2.05	110.01	106.00

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	602	GOL	O1-C1-C2-C3
4	A	604	EPE	C10-C9-N1-C6
4	A	604	EPE	C8-C7-N4-C5
4	A	604	EPE	C9-C10-S-O2S

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	A	604	EPE	C9-C10-S-O3S
2	A	602	GOL	O1-C1-C2-O2
3	A	603	PGE	O2-C3-C4-O3
4	A	604	EPE	C10-C9-N1-C2
4	A	604	EPE	C9-C10-S-O1S
4	A	604	EPE	N4-C7-C8-O8
3	A	603	PGE	C3-C4-O3-C5

There are no ring outliers.

No monomer is involved in short contacts.

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	493/533 (92%)	0.16	12 (2%)	59	64	30, 43, 68, 96	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	481	LEU	5.3
1	A	6	PRO	3.3
1	A	320	ASN	2.7
1	A	7	LEU	2.6
1	A	458	TRP	2.5
1	A	455	GLY	2.5
1	A	469	GLY	2.4
1	A	457	TYR	2.4
1	A	456	THR	2.3
1	A	356	CYS	2.0
1	A	454	LEU	2.0
1	A	234	ARG	2.0

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

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	LLP	A	62	24/25	0.97	0.07	29,34,46,58	0

6.3 Carbohydrates ⓘ

There are no monosaccharides 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**
4	EPE	A	604	15/15	0.85	0.17	72,78,90,97	0
2	GOL	A	605	6/6	0.86	0.15	47,74,82,86	0
2	GOL	A	601	6/6	0.86	0.14	49,57,59,64	0
3	PGE	A	603	10/10	0.91	0.13	57,59,75,77	0
2	GOL	A	602	6/6	0.94	0.10	47,54,58,64	0

6.5 Other polymers ⓘ

There are no such residues in this entry.