



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 25, 2020 – 04:31 PM BST

PDB ID : 5VRB
Title : Crystal structure of a transketolase from *Neisseria gonorrhoeae*
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on : 2017-05-10
Resolution : 1.85 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.13
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) : 2.13

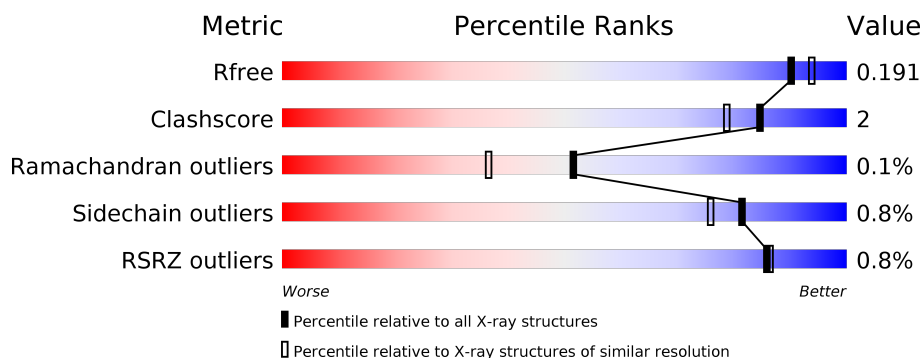
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	667	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 89%, grey 6%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 89% 5% 6% </div> </div>
1	B	667	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 91%, yellow 6%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 91% 6% • </div> </div>
1	C	667	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 87%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 87% 6% 8% </div> </div>
1	D	667	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 90%, grey 6%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 90% 6% • </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 21287 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 Transketolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	627	Total	C	N	O	S	0	14	0
			4786	3040	816	906	24			
1	B	646	Total	C	N	O	S	0	10	0
			4915	3118	840	933	24			
1	C	615	Total	C	N	O	S	0	6	0
			4651	2956	798	875	22			
1	D	640	Total	C	N	O	S	0	5	0
			4839	3075	832	909	23			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP B4RKU9
A	-6	ALA	-	expression tag	UNP B4RKU9
A	-5	HIS	-	expression tag	UNP B4RKU9
A	-4	HIS	-	expression tag	UNP B4RKU9
A	-3	HIS	-	expression tag	UNP B4RKU9
A	-2	HIS	-	expression tag	UNP B4RKU9
A	-1	HIS	-	expression tag	UNP B4RKU9
A	0	HIS	-	expression tag	UNP B4RKU9
B	-7	MET	-	initiating methionine	UNP B4RKU9
B	-6	ALA	-	expression tag	UNP B4RKU9
B	-5	HIS	-	expression tag	UNP B4RKU9
B	-4	HIS	-	expression tag	UNP B4RKU9
B	-3	HIS	-	expression tag	UNP B4RKU9
B	-2	HIS	-	expression tag	UNP B4RKU9
B	-1	HIS	-	expression tag	UNP B4RKU9
B	0	HIS	-	expression tag	UNP B4RKU9
C	-7	MET	-	initiating methionine	UNP B4RKU9
C	-6	ALA	-	expression tag	UNP B4RKU9
C	-5	HIS	-	expression tag	UNP B4RKU9
C	-4	HIS	-	expression tag	UNP B4RKU9
C	-3	HIS	-	expression tag	UNP B4RKU9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	HIS	-	expression tag	UNP B4RKU9
C	-1	HIS	-	expression tag	UNP B4RKU9
C	0	HIS	-	expression tag	UNP B4RKU9
D	-7	MET	-	initiating methionine	UNP B4RKU9
D	-6	ALA	-	expression tag	UNP B4RKU9
D	-5	HIS	-	expression tag	UNP B4RKU9
D	-4	HIS	-	expression tag	UNP B4RKU9
D	-3	HIS	-	expression tag	UNP B4RKU9
D	-2	HIS	-	expression tag	UNP B4RKU9
D	-1	HIS	-	expression tag	UNP B4RKU9
D	0	HIS	-	expression tag	UNP B4RKU9

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		

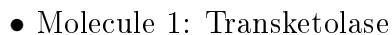
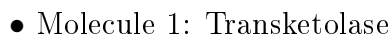
- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

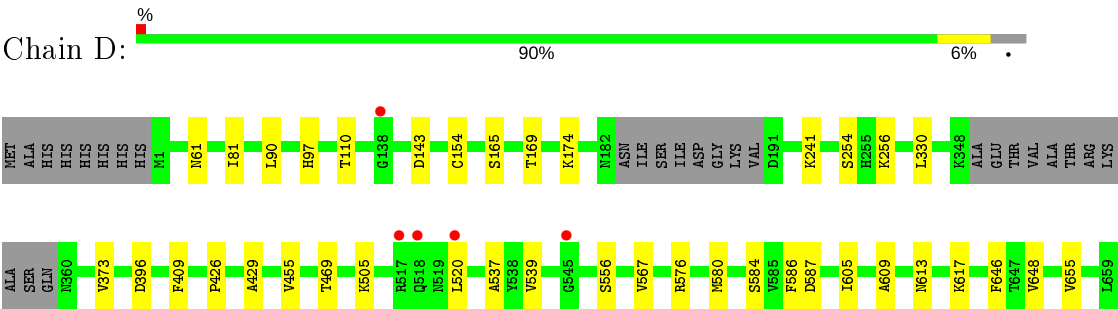
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	492	Total 498	O 498	0	6
4	B	583	Total 592	O 592	0	9
4	C	509	Total 513	O 513	0	4
4	D	455	Total 457	O 457	0	2

- Molecule 1: Transketolase



● Molecule 1: Transketolase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	82.45Å 114.21Å 125.69Å 90.00° 90.11° 90.00°	Depositor
Resolution (Å)	37.59 – 1.85 45.82 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.7 (37.59-1.85) 98.7 (45.82-1.85)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.67 (at 1.86Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.150 , 0.191 0.150 , 0.191	Depositor DCC
R_{free} test set	1970 reflections (1.01%)	wwPDB-VP
Wilson B-factor (Å ²)	18.4	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 39.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.087 for h,-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	21287	wwPDB-VP
Average B, all atoms (Å ²)	22.0	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.*

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: CA, OCS, EDO

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	0.36	0/4921	0.53	0/6696
1	B	0.36	0/5045	0.53	0/6855
1	C	0.34	0/4762	0.53	0/6476
1	D	0.34	0/4954	0.50	0/6731
All	All	0.35	0/19682	0.52	0/26758

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 [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	4786	0	4623	23	0
1	B	4915	0	4762	25	0
1	C	4651	0	4482	24	0
1	D	4839	0	4693	21	0
2	A	4	0	6	0	0
2	C	8	0	10	2	0
2	D	12	0	16	0	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	3	0	0	0	0
3	D	3	0	0	0	0
4	A	498	0	0	3	0
4	B	592	0	0	4	0
4	C	513	0	0	6	0
4	D	457	0	0	2	0
All	All	21287	0	18592	85	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.

The worst 5 of 85 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:639[B]:LEU:HD21	1:B:90:LEU:HD22	1.56	0.87
1:B:258:HIS:ND1	4:B:801:HOH:O	2.19	0.76
1:C:639:LEU:HD21	1:D:90:LEU:HD22	1.74	0.69
1:C:548:GLN:NE2	4:C:803:HOH:O	2.20	0.66
1:A:137:ASP:OD1	1:A:315:TYR:OH	2.14	0.65

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	634/667 (95%)	621 (98%)	12 (2%)	1 (0%)	47	33
1	B	649/667 (97%)	634 (98%)	14 (2%)	1 (0%)	47	33
1	C	608/667 (91%)	595 (98%)	12 (2%)	1 (0%)	47	33
1	D	638/667 (96%)	624 (98%)	14 (2%)	0	100	100
All	All	2529/2668 (95%)	2474 (98%)	52 (2%)	3 (0%)	51	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	64	ALA
1	C	111	GLY
1	A	111	GLY

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	486/528 (92%)	481 (99%)	5 (1%)	76	69
1	B	501/528 (95%)	498 (99%)	3 (1%)	86	83
1	C	468/528 (89%)	464 (99%)	4 (1%)	78	72
1	D	488/528 (92%)	484 (99%)	4 (1%)	81	76
All	All	1943/2112 (92%)	1927 (99%)	16 (1%)	81	76

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	584	SER
1	C	179	TYR
1	D	61	ASN
1	B	409	PHE
1	D	409	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	258	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

4 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	OCS	C	154	1	7,8,9	1.06	0	6,11,13	1.52	1 (16%)
1	OCS	D	154	1	7,8,9	1.31	1 (14%)	6,11,13	1.78	1 (16%)
1	OCS	A	154	1,3	7,8,9	1.17	0	6,11,13	1.39	1 (16%)
1	OCS	B	154	1	7,8,9	1.19	0	6,11,13	1.41	2 (33%)

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	OCS	C	154	1	-	0/4/7/9	-
1	OCS	D	154	1	-	0/4/7/9	-
1	OCS	A	154	1,3	-	0/4/7/9	-
1	OCS	B	154	1	-	3/4/7/9	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	154	OCS	CB-CA	-2.58	1.51	1.53

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	154	OCS	OD3-SG-CB	3.95	111.63	106.94
1	C	154	OCS	OD3-SG-CB	2.88	110.36	106.94
1	A	154	OCS	OD3-SG-CB	2.82	110.30	106.94
1	B	154	OCS	OD3-SG-CB	2.26	109.62	106.94
1	B	154	OCS	OD1-SG-CB	2.09	109.42	106.94

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	154	OCS	CA-CB-SG-OD2
1	B	154	OCS	CA-CB-SG-OD1
1	B	154	OCS	CA-CB-SG-OD3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	154	OCS	1	0
1	B	154	OCS	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 12 are monoatomic - leaving 6 for Mogul analysis.

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	EDO	D	701	3	3,3,3	0.44	0	2,2,2	0.30	0
2	EDO	C	701	3	3,3,3	0.45	0	2,2,2	0.43	0
2	EDO	A	701	-	3,3,3	0.42	0	2,2,2	0.63	0
2	EDO	D	703	-	3,3,3	0.48	0	2,2,2	0.43	0
2	EDO	C	702	-	3,3,3	0.31	0	2,2,2	0.19	0
2	EDO	D	702	-	3,3,3	0.44	0	2,2,2	0.41	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	EDO	D	701	3	-	0/1/1/1	-
2	EDO	C	701	3	-	0/1/1/1	-
2	EDO	A	701	-	-	0/1/1/1	-
2	EDO	D	703	-	-	0/1/1/1	-
2	EDO	C	702	-	-	0/1/1/1	-
2	EDO	D	702	-	-	0/1/1/1	-

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	702	EDO	2	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	626/667 (93%)	-0.63	6 (0%) 82 82	9, 20, 37, 70	0
1	B	645/667 (96%)	-0.71	2 (0%) 94 93	9, 18, 35, 68	0
1	C	614/667 (92%)	-0.65	7 (1%) 80 81	9, 19, 40, 71	0
1	D	639/667 (95%)	-0.65	5 (0%) 86 86	11, 21, 39, 74	0
All	All	2524/2668 (94%)	-0.66	20 (0%) 86 86	9, 20, 38, 74	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	23	HIS	4.7
1	A	273	HIS	4.4
1	C	262	LEU	4.0
1	A	22	GLY	3.7
1	B	359	GLN	2.9

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	OCS	D	154	9/10	0.98	0.07	10,12,20,42	0
1	OCS	A	154	9/10	0.98	0.07	10,12,17,37	0
1	OCS	B	154	9/10	0.98	0.07	10,13,24,39	0
1	OCS	C	154	9/10	0.99	0.06	12,15,23,38	0

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, 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(\AA^2)	Q<0.9
3	CA	C	705	1/1	0.96	0.07	47,47,47,47	0
3	CA	B	703	1/1	0.96	0.05	30,30,30,30	0
2	EDO	C	701	4/4	0.97	0.06	21,23,24,26	0
2	EDO	D	701	4/4	0.97	0.11	27,27,28,33	0
3	CA	A	703	1/1	0.98	0.10	50,50,50,50	0
3	CA	D	706	1/1	0.98	0.04	25,25,25,25	0
3	CA	D	704	1/1	0.98	0.07	39,39,39,39	0
2	EDO	D	703	4/4	0.98	0.11	20,22,25,28	0
2	EDO	C	702	4/4	0.98	0.09	20,23,29,31	0
2	EDO	D	702	4/4	0.98	0.08	30,30,35,44	0
2	EDO	A	701	4/4	0.99	0.07	19,22,22,23	0
3	CA	A	704	1/1	0.99	0.03	31,31,31,31	0
3	CA	D	705	1/1	0.99	0.06	34,34,34,34	0
3	CA	B	702	1/1	0.99	0.04	22,22,22,22	0
3	CA	C	704	1/1	0.99	0.04	32,32,32,32	0
3	CA	B	701	1/1	0.99	0.03	25,25,25,25	0
3	CA	C	703	1/1	1.00	0.03	19,19,19,19	0
3	CA	A	702	1/1	1.00	0.03	22,22,22,22	0

6.5 Other polymers [i](#)

There are no such residues in this entry.