



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 03:54 AM UTC

PDB ID : 4H34 / pdb_00004h34
Title : Crystal structure of the tyrosine phosphatase SHP-2 with Q506P mutation
Authors : Qiu, W.; Lin, A.; Hutchinson, A.; Romanov, V.; Ruzanov, M.; Thompson, C.;
Lam, K.; Kisselman, G.; Battaile, K.; Chirgadze, N.Y.
Deposited on : 2012-09-13
Resolution : 2.70 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

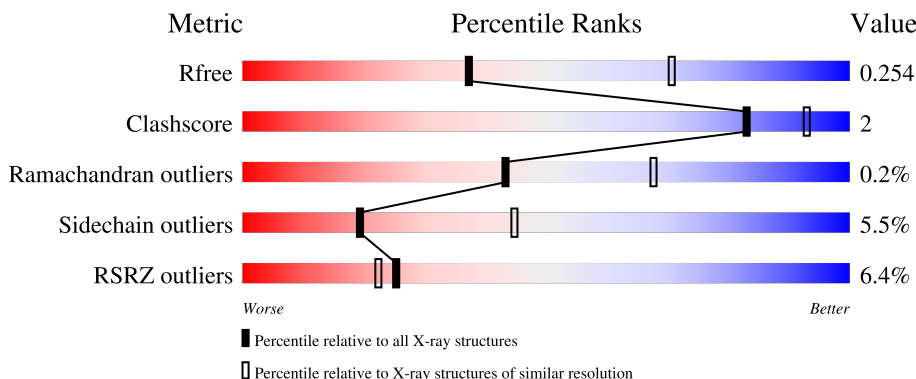
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	180053	3538 (2.70-2.70)
Clashscore	190562	3843 (2.70-2.70)
Ramachandran outliers	187476	3778 (2.70-2.70)
Sidechain outliers	187428	3778 (2.70-2.70)
RSRZ outliers	180081	3538 (2.70-2.70)

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 $\leq 5\%$. 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	560	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4041 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.

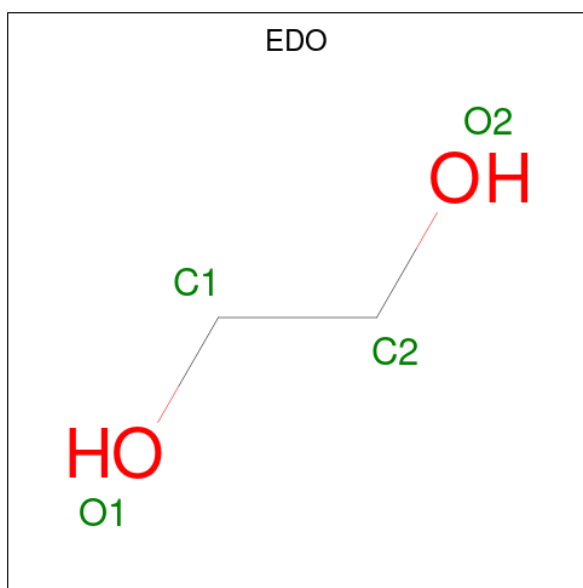
- Molecule 1 is a protein called Tyrosine-protein phosphatase non-receptor type 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	486	Total	C	N	O	S	0	0	0
			3951	2491	707	735	18			

There are 26 discrepancies between the modelled and reference sequences:

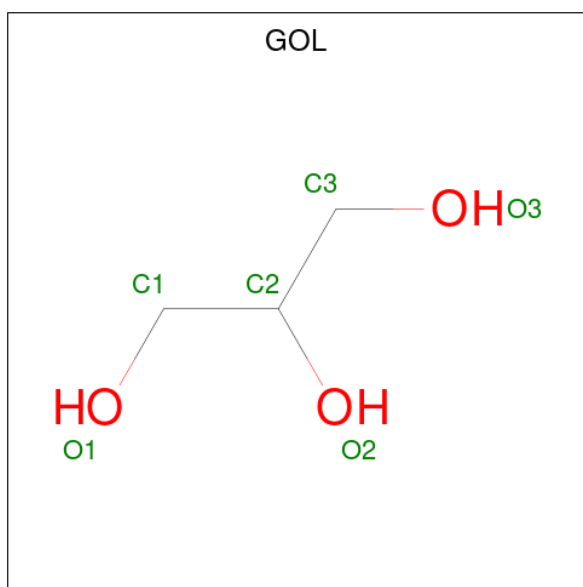
Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP Q06124
A	-19	GLY	-	expression tag	UNP Q06124
A	-18	SER	-	expression tag	UNP Q06124
A	-17	SER	-	expression tag	UNP Q06124
A	-16	HIS	-	expression tag	UNP Q06124
A	-15	HIS	-	expression tag	UNP Q06124
A	-14	HIS	-	expression tag	UNP Q06124
A	-13	HIS	-	expression tag	UNP Q06124
A	-12	HIS	-	expression tag	UNP Q06124
A	-11	HIS	-	expression tag	UNP Q06124
A	-10	SER	-	expression tag	UNP Q06124
A	-9	SER	-	expression tag	UNP Q06124
A	-8	GLY	-	expression tag	UNP Q06124
A	-7	ARG	-	expression tag	UNP Q06124
A	-6	GLU	-	expression tag	UNP Q06124
A	-5	ASN	-	expression tag	UNP Q06124
A	-4	LEU	-	expression tag	UNP Q06124
A	-3	TYR	-	expression tag	UNP Q06124
A	-2	PHE	-	expression tag	UNP Q06124
A	-1	GLN	-	expression tag	UNP Q06124
A	0	GLY	-	expression tag	UNP Q06124
A	?	-	GLN	deletion	UNP Q06124
A	?	-	ALA	deletion	UNP Q06124
A	?	-	LEU	deletion	UNP Q06124
A	?	-	LEU	deletion	UNP Q06124
A	506	PRO	GLN	engineered mutation	UNP Q06124

- Molecule 2 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



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

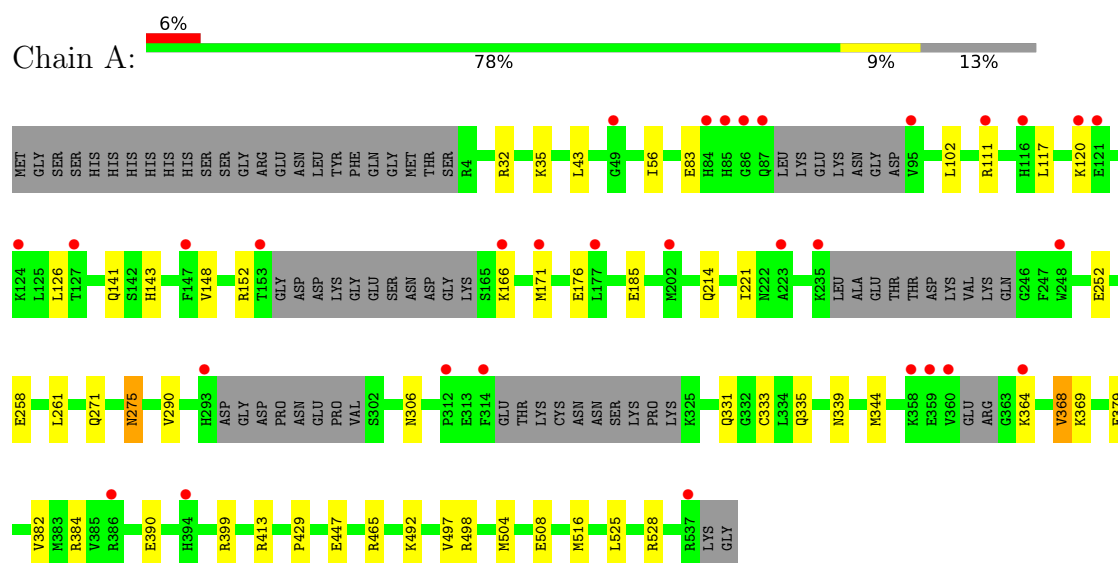
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	54	Total	O	0	0
			54	54		

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: Tyrosine-protein phosphatase non-receptor type 11



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	54.79Å 202.35Å 44.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.12 – 2.70 37.12 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (37.12-2.70) 99.9 (37.12-2.70)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.33 (at 2.68Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.226 , 0.242 0.240 , 0.254	Depositor DCC
R_{free} test set	1099 reflections (7.69%)	wwPDB-VP
Wilson B-factor (Å ²)	39.2	Xtriage
Anisotropy	0.402	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 37.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4041	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.75% 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: EDO, 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	0.67	0/4032	1.22	1/5433 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	275	ASN	CA-CB-CG	5.79	118.39	112.60

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	3951	0	3891	19	0
2	A	24	0	36	8	0
3	A	12	0	16	0	0
4	A	54	0	0	0	0
All	All	4041	0	3943	19	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 (19) 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:148:VAL:HG22	1:A:171:MET:HE2	1.79	0.65
1:A:339:ASN:HB3	2:A:607:EDO:H11	1.84	0.59
1:A:35:LYS:H	2:A:601:EDO:H21	1.68	0.57
1:A:335:GLN:HG2	1:A:368:VAL:HG21	1.89	0.55
1:A:382:VAL:H	2:A:607:EDO:H22	1.75	0.51
1:A:290:VAL:HG11	1:A:344:MET:HG3	1.92	0.51
1:A:126:LEU:O	1:A:152:ARG:NH1	2.44	0.50
1:A:32:ARG:HH22	2:A:601:EDO:H22	1.76	0.50
1:A:497:VAL:HG12	1:A:504:MET:HG3	1.94	0.49
1:A:306:ASN:HB3	1:A:331:GLN:HG2	1.97	0.45
1:A:379:GLU:HG2	1:A:384:ARG:HG3	1.99	0.45
1:A:43:LEU:HB2	1:A:56:ILE:HD11	1.99	0.44
1:A:413:ARG:HH22	2:A:606:EDO:H21	1.82	0.44
1:A:508:GLU:H	2:A:602:EDO:H12	1.82	0.44
1:A:258:GLU:HA	1:A:261:LEU:HD12	1.99	0.44
1:A:390:GLU:HG2	1:A:399:ARG:HG2	1.99	0.43
1:A:429:PRO:HG2	2:A:603:EDO:H21	2.01	0.42
1:A:32:ARG:NH2	2:A:601:EDO:H22	2.34	0.42
1:A:261:LEU:HD13	1:A:498:ARG:HB3	2.02	0.41

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	472/560 (84%)	453 (96%)	18 (4%)	1 (0%)	43 68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	369	LYS

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	433/498 (87%)	409 (94%)	24 (6%)	19	45

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	GLU
1	A	102	LEU
1	A	111	ARG
1	A	117	LEU
1	A	120	LYS
1	A	141	GLN
1	A	143	HIS
1	A	166	LYS
1	A	176	GLU
1	A	185	GLU
1	A	214	GLN
1	A	221	ILE
1	A	252	GLU
1	A	271	GLN
1	A	275	ASN
1	A	333	CYS
1	A	364	LYS
1	A	368	VAL
1	A	447	GLU
1	A	465	ARG
1	A	492	LYS
1	A	516	MET
1	A	525	LEU
1	A	528	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	85	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	87	GLN
1	A	169	HIS
1	A	217	ASN
1	A	256	GLN
1	A	281	ASN
1	A	495	GLN
1	A	520	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 ⓘ

8 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	$\# Z > 2$
3	GOL	A	604	-	5,5,5	0.05	0	5,5,5	0.13	0
2	EDO	A	607	-	3,3,3	0.51	0	2,2,2	0.29	0
2	EDO	A	601	-	3,3,3	0.47	0	2,2,2	0.36	0
2	EDO	A	602	-	3,3,3	0.48	0	2,2,2	0.35	0
2	EDO	A	603	-	3,3,3	0.46	0	2,2,2	0.38	0
3	GOL	A	608	-	5,5,5	0.06	0	5,5,5	0.16	0
2	EDO	A	606	-	3,3,3	0.47	0	2,2,2	0.42	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	EDO	A	605	-	3,3,3	0.46	0	2,2,2	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
3	GOL	A	604	-	-	0/4/4/4	-
2	EDO	A	607	-	-	1/1/1/1	-
2	EDO	A	601	-	-	0/1/1/1	-
2	EDO	A	602	-	-	1/1/1/1	-
2	EDO	A	603	-	-	0/1/1/1	-
3	GOL	A	608	-	-	0/4/4/4	-
2	EDO	A	606	-	-	0/1/1/1	-
2	EDO	A	605	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	607	EDO	O1-C1-C2-O2
2	A	602	EDO	O1-C1-C2-O2

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	607	EDO	2	0
2	A	601	EDO	3	0
2	A	602	EDO	1	0
2	A	603	EDO	1	0
2	A	606	EDO	1	0

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	486/560 (86%)	0.65	31 (6%) 25 22	23, 40, 63, 90	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	87	GLN	4.9
1	A	86	GLY	4.4
1	A	235	LYS	4.4
1	A	85	HIS	4.0
1	A	248	TRP	3.8
1	A	116	HIS	3.3
1	A	153	THR	3.2
1	A	359	GLU	3.1
1	A	360	VAL	2.8
1	A	120	LYS	2.8
1	A	49	GLY	2.7
1	A	177	LEU	2.7
1	A	111	ARG	2.7
1	A	364	LYS	2.6
1	A	394	HIS	2.6
1	A	127	THR	2.4
1	A	95	VAL	2.4
1	A	121	GLU	2.3
1	A	124	LYS	2.3
1	A	171	MET	2.3
1	A	314	PHE	2.3
1	A	223	ALA	2.2
1	A	84	HIS	2.2
1	A	386	ARG	2.1
1	A	293	HIS	2.1
1	A	147	PHE	2.1
1	A	312	PRO	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	358	LYS	2.1
1	A	166	LYS	2.0
1	A	202	MET	2.0
1	A	537	ARG	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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
3	GOL	A	608	6/6	0.60	0.13	46,46,46,46	0
2	EDO	A	605	4/4	0.73	0.19	47,47,47,47	0
2	EDO	A	606	4/4	0.75	0.15	47,47,48,48	0
3	GOL	A	604	6/6	0.76	0.13	49,50,50,50	0
2	EDO	A	607	4/4	0.79	0.19	46,46,47,47	0
2	EDO	A	603	4/4	0.80	0.15	51,51,51,52	0
2	EDO	A	601	4/4	0.80	0.15	42,42,42,42	0
2	EDO	A	602	4/4	0.83	0.15	48,49,49,49	0

6.5 Other polymers [i](#)

There are no such residues in this entry.