



wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 07:05 pm BST

PDB ID : 6JLQ
Title : Crystal structure of human USP46-WDR48-WDR20 complex
Authors : Zhu, H.; Zhang, T.; Ding, J.
Deposited on : 2019-03-06
Resolution : 3.10 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 : 2.11
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) : 2.11

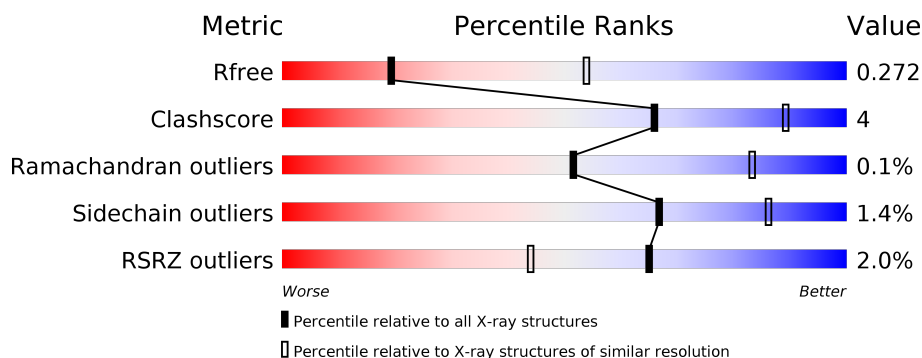
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 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	352	<div> <div>3%</div> <div> <div></div> <div>73%</div> <div>9%</div> <div>17%</div> </div> </div>
2	B	620	<div> <div>%</div> <div> <div></div> <div>77%</div> <div>6%</div> <div>17%</div> </div> </div>
3	C	444	<div> <div>%</div> <div> <div></div> <div>79%</div> <div>11%</div> <div>10%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	B	606	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9327 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 Ubiquitin carboxyl-terminal hydrolase 46.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	0	0
			2217	1403	372	426	16			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	initiating methionine	UNP P62068
A	367	LEU	-	expression tag	UNP P62068
A	368	GLU	-	expression tag	UNP P62068
A	369	HIS	-	expression tag	UNP P62068
A	370	HIS	-	expression tag	UNP P62068
A	371	HIS	-	expression tag	UNP P62068
A	372	HIS	-	expression tag	UNP P62068
A	373	HIS	-	expression tag	UNP P62068
A	374	HIS	-	expression tag	UNP P62068

- Molecule 2 is a protein called WD repeat-containing protein 48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	512	Total	C	N	O	S	0	0	0
			3954	2493	691	749	21			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-39	MET	-	initiating methionine	UNP Q8TAF3
B	-38	SER	-	expression tag	UNP Q8TAF3
B	-37	TYR	-	expression tag	UNP Q8TAF3
B	-36	TYR	-	expression tag	UNP Q8TAF3
B	-35	ASP	-	expression tag	UNP Q8TAF3
B	-34	TYR	-	expression tag	UNP Q8TAF3
B	-33	LYS	-	expression tag	UNP Q8TAF3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-32	ASP	-	expression tag	UNP Q8TAF3
B	-31	ASP	-	expression tag	UNP Q8TAF3
B	-30	ASP	-	expression tag	UNP Q8TAF3
B	-29	ASP	-	expression tag	UNP Q8TAF3
B	-28	LYS	-	expression tag	UNP Q8TAF3
B	-27	SER	-	expression tag	UNP Q8TAF3
B	-26	GLY	-	expression tag	UNP Q8TAF3
B	-25	SER	-	expression tag	UNP Q8TAF3
B	-24	GLY	-	expression tag	UNP Q8TAF3
B	-23	HIS	-	expression tag	UNP Q8TAF3
B	-22	HIS	-	expression tag	UNP Q8TAF3
B	-21	HIS	-	expression tag	UNP Q8TAF3
B	-20	HIS	-	expression tag	UNP Q8TAF3
B	-19	HIS	-	expression tag	UNP Q8TAF3
B	-18	HIS	-	expression tag	UNP Q8TAF3
B	-17	ASP	-	expression tag	UNP Q8TAF3
B	-16	TYR	-	expression tag	UNP Q8TAF3
B	-15	ASP	-	expression tag	UNP Q8TAF3
B	-14	ILE	-	expression tag	UNP Q8TAF3
B	-13	PRO	-	expression tag	UNP Q8TAF3
B	-12	THR	-	expression tag	UNP Q8TAF3
B	-11	THR	-	expression tag	UNP Q8TAF3
B	-10	GLU	-	expression tag	UNP Q8TAF3
B	-9	ASN	-	expression tag	UNP Q8TAF3
B	-8	LEU	-	expression tag	UNP Q8TAF3
B	-7	TYR	-	expression tag	UNP Q8TAF3
B	-6	PHE	-	expression tag	UNP Q8TAF3
B	-5	GLN	-	expression tag	UNP Q8TAF3
B	-4	GLY	-	expression tag	UNP Q8TAF3
B	-3	ALA	-	expression tag	UNP Q8TAF3
B	-2	MET	-	expression tag	UNP Q8TAF3
B	-1	GLY	-	expression tag	UNP Q8TAF3
B	0	SER	-	expression tag	UNP Q8TAF3

- Molecule 3 is a protein called WD repeat-containing protein 20, highly similar to WD repeat protein 20, WD repeat-containing protein 20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	401	Total	C	N	O	S	0	0	0
			3102	1981	526	577	18			

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	MET	-	initiating methionine	UNP Q8TBZ3
C	-6	GLY	-	expression tag	UNP Q8TBZ3
C	-5	HIS	-	expression tag	UNP Q8TBZ3
C	-4	HIS	-	expression tag	UNP Q8TBZ3
C	-3	HIS	-	expression tag	UNP Q8TBZ3
C	-2	HIS	-	expression tag	UNP Q8TBZ3
C	-1	HIS	-	expression tag	UNP Q8TBZ3
C	0	HIS	-	expression tag	UNP Q8TBZ3
C	319	GLY	-	linker	UNP Q8TBZ3
C	320	GLY	-	linker	UNP Q8TBZ3
C	321	GLY	-	linker	UNP Q8TBZ3
C	361	GLY	-	linker	UNP Q8TBZ3
C	362	SER	-	linker	UNP Q8TBZ3
C	489	GLY	-	linker	UNP B3KQX8
C	490	GLY	-	linker	UNP B3KQX8
C	491	GLY	-	linker	UNP B3KQX8
C	492	GLY	-	linker	UNP B3KQX8
C	493	SER	-	linker	UNP B3KQX8
C	494	GLY	-	linker	UNP B3KQX8
C	495	GLY	-	linker	UNP B3KQX8
C	496	GLY	-	linker	UNP B3KQX8
C	497	GLY	-	linker	UNP B3KQX8
C	498	SER	-	linker	UNP B3KQX8
C	499	GLY	-	linker	UNP B3KQX8
C	500	GLY	-	linker	UNP B3KQX8
C	501	GLY	-	linker	UNP B3KQX8
C	502	GLY	-	linker	UNP B3KQX8
C	503	SER	-	linker	UNP B3KQX8
C	504	GLY	-	linker	UNP B3KQX8
C	505	GLY	-	linker	UNP B3KQX8
C	506	GLY	-	linker	UNP B3KQX8
C	507	GLY	-	linker	UNP B3KQX8
C	508	SER	-	linker	UNP B3KQX8

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

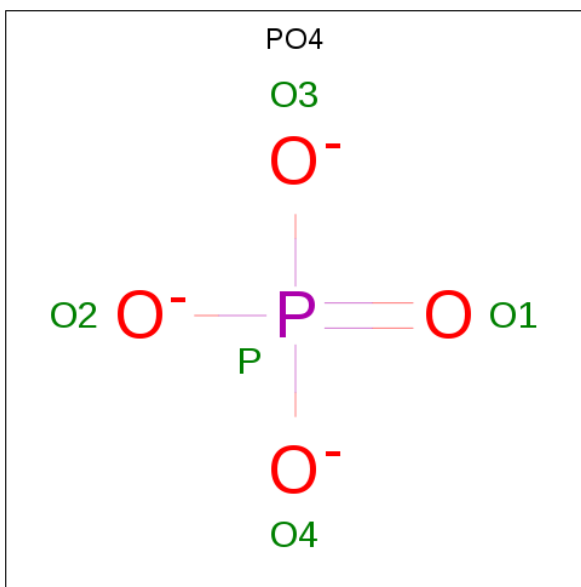
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).

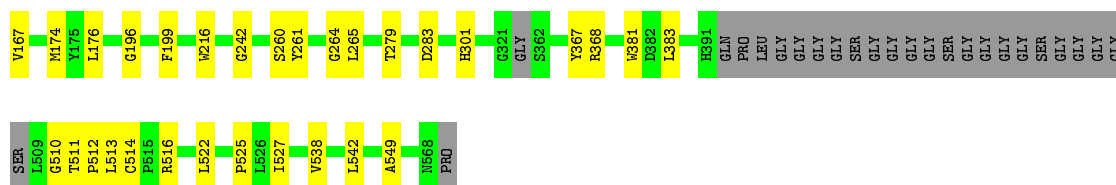


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	O	P	0	0
			5	4	1		



4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	217.23Å 217.23Å 223.93Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.99 – 3.10 49.76 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.0 (43.99-3.10) 99.0 (49.76-3.10)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.50 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.247 , 0.272 0.247 , 0.272	Depositor DCC
R_{free} test set	2826 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	88.1	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 53.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	9327	wwPDB-VP
Average B, all atoms (Å ²)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% 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: GOL, ZN, PO4

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	0.27	0/2256	0.44	0/3069
2	B	0.25	0/4035	0.44	0/5496
3	C	0.28	0/3181	0.47	0/4323
All	All	0.27	0/9472	0.45	0/12888

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	2217	0	2036	20	0
2	B	3954	0	3848	24	0
3	C	3102	0	2986	32	0
4	A	1	0	0	0	0
5	A	6	0	8	0	0
5	B	36	0	48	4	0
5	C	6	0	8	1	0
6	B	5	0	0	0	0
All	All	9327	0	8934	73	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 4.

The worst 5 of 73 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)
2:B:377:THR:HG1	2:B:387:TRP:HE1	1.29	0.80
3:C:511:THR:N	3:C:512:PRO:HD2	1.96	0.80
1:A:353:ASN:N	1:A:353:ASN:OD1	2.19	0.72
1:A:326:LEU:C	1:A:326:LEU:HD13	2.13	0.69
1:A:317:ILE:HD13	1:A:326:LEU:HD23	1.75	0.68

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	284/352 (81%)	267 (94%)	16 (6%)	1 (0%)	34	69
2	B	504/620 (81%)	475 (94%)	29 (6%)	0	100	100
3	C	395/444 (89%)	377 (95%)	18 (5%)	0	100	100
All	All	1183/1416 (84%)	1119 (95%)	63 (5%)	1 (0%)	51	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	311	ARG

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	230/325 (71%)	221 (96%)	9 (4%)	32	65
2	B	428/541 (79%)	426 (100%)	2 (0%)	88	94
3	C	339/373 (91%)	336 (99%)	3 (1%)	78	91
All	All	997/1239 (80%)	983 (99%)	14 (1%)	67	86

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	326	LEU
1	A	330	ASP
3	C	69	VAL
1	A	320	SER
2	B	169	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	167	ASN
3	C	43	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 1 is monoatomic - leaving 9 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	$\# Z > 2$
5	GOL	A	402	-	5,5,5	0.37	0	5,5,5	0.22	0
5	GOL	B	604	-	5,5,5	0.36	0	5,5,5	0.24	0
5	GOL	C	601	-	5,5,5	0.37	0	5,5,5	0.25	0
5	GOL	B	602	-	5,5,5	0.37	0	5,5,5	0.24	0
5	GOL	B	605	-	5,5,5	0.38	0	5,5,5	0.25	0
5	GOL	B	606	-	5,5,5	0.37	0	5,5,5	0.29	0
5	GOL	B	603	-	5,5,5	0.37	0	5,5,5	0.30	0
6	PO4	B	601	-	4,4,4	0.93	0	6,6,6	0.43	0
5	GOL	B	607	-	5,5,5	0.37	0	5,5,5	0.34	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
5	GOL	A	402	-	-	2/4/4/4	-
5	GOL	B	604	-	-	0/4/4/4	-
5	GOL	C	601	-	-	2/4/4/4	-
5	GOL	B	602	-	-	0/4/4/4	-
5	GOL	B	605	-	-	2/4/4/4	-
5	GOL	B	606	-	-	2/4/4/4	-
5	GOL	B	603	-	-	2/4/4/4	-
5	GOL	B	607	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	402	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	C	601	GOL	O1-C1-C2-C3
5	B	605	GOL	O1-C1-C2-C3
5	B	606	GOL	C1-C2-C3-O3
5	B	603	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	601	GOL	1	0
5	B	602	GOL	1	0
5	B	606	GOL	3	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	292/352 (82%)	0.24	12 (4%) 37 18	59, 126, 165, 176	0
2	B	512/620 (82%)	0.05	9 (1%) 68 47	51, 86, 142, 165	0
3	C	401/444 (90%)	-0.01	3 (0%) 87 75	50, 81, 122, 143	0
All	All	1205/1416 (85%)	0.08	24 (1%) 65 44	50, 90, 150, 176	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	344	ASN	4.7
1	A	98	PRO	3.6
2	B	517	PHE	3.2
2	B	345	PRO	3.2
3	C	59	SER	3.0

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 carbohydrates 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
5	GOL	B	606	6/6	0.53	0.41	82,86,90,90	0
5	GOL	B	603	6/6	0.70	0.22	78,91,95,97	0
5	GOL	B	604	6/6	0.77	0.21	87,102,112,112	0
5	GOL	C	601	6/6	0.80	0.34	76,82,92,93	0
5	GOL	B	607	6/6	0.81	0.21	105,108,116,119	0
5	GOL	B	605	6/6	0.83	0.62	96,99,101,102	0
5	GOL	A	402	6/6	0.85	0.13	86,107,107,111	0
5	GOL	B	602	6/6	0.94	0.10	75,81,84,85	0
6	PO4	B	601	5/5	0.95	0.35	66,69,88,100	0
4	ZN	A	401	1/1	0.99	0.17	66,66,66,66	0

6.5 Other polymers ⓘ

There are no such residues in this entry.