



# Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 09:09 am BST

PDB ID : 6VYE  
Title : 6-phosphogluconolactonase from Trypanosoma cruzi  
Authors : Merritt, E.A.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)  
Deposited on : 2020-02-26  
Resolution : 1.95 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](mailto: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

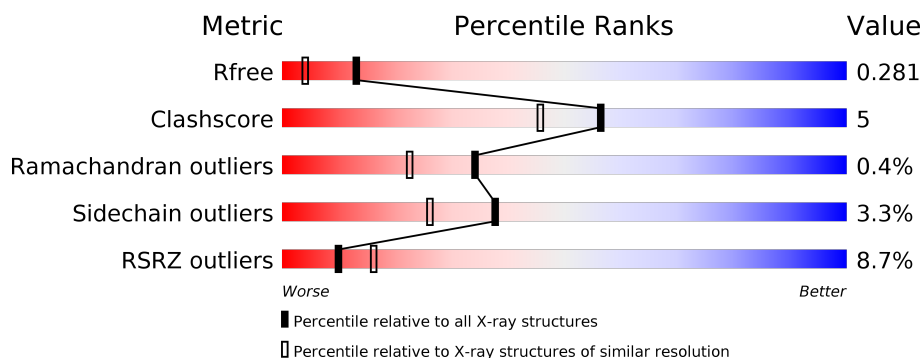
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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  $\geq 3$ , 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	271	<div> <div>8%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>••</div> </div> </div>
1	B	271	<div> <div>9%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>••</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4317 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 6-phosphogluconolactonase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	0	0	0
			2021	1286	351	372	12			
1	B	266	Total	C	N	O	S	0	1	0
			2047	1301	354	379	13			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP Q4CQA8
A	-6	ALA	-	expression tag	UNP Q4CQA8
A	-5	HIS	-	expression tag	UNP Q4CQA8
A	-4	HIS	-	expression tag	UNP Q4CQA8
A	-3	HIS	-	expression tag	UNP Q4CQA8
A	-2	HIS	-	expression tag	UNP Q4CQA8
A	-1	HIS	-	expression tag	UNP Q4CQA8
A	0	HIS	-	expression tag	UNP Q4CQA8
A	97	ARG	HIS	conflict	UNP Q4CQA8
A	259	GLU	GLY	conflict	UNP Q4CQA8
B	-7	MET	-	initiating methionine	UNP Q4CQA8
B	-6	ALA	-	expression tag	UNP Q4CQA8
B	-5	HIS	-	expression tag	UNP Q4CQA8
B	-4	HIS	-	expression tag	UNP Q4CQA8
B	-3	HIS	-	expression tag	UNP Q4CQA8
B	-2	HIS	-	expression tag	UNP Q4CQA8
B	-1	HIS	-	expression tag	UNP Q4CQA8
B	0	HIS	-	expression tag	UNP Q4CQA8
B	97	ARG	HIS	conflict	UNP Q4CQA8
B	259	GLU	GLY	conflict	UNP Q4CQA8

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

*Continued on next page...*

*Continued from previous page...*

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

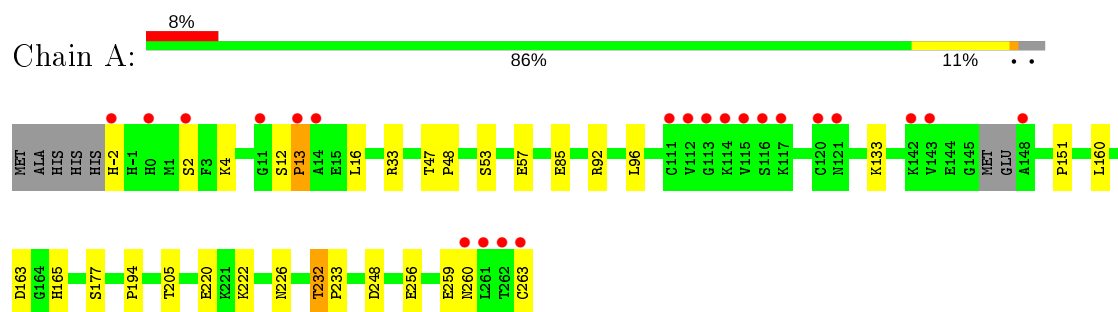
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	113	Total	O	0	0
			113	113		
4	B	103	Total	O	0	0
			103	103		

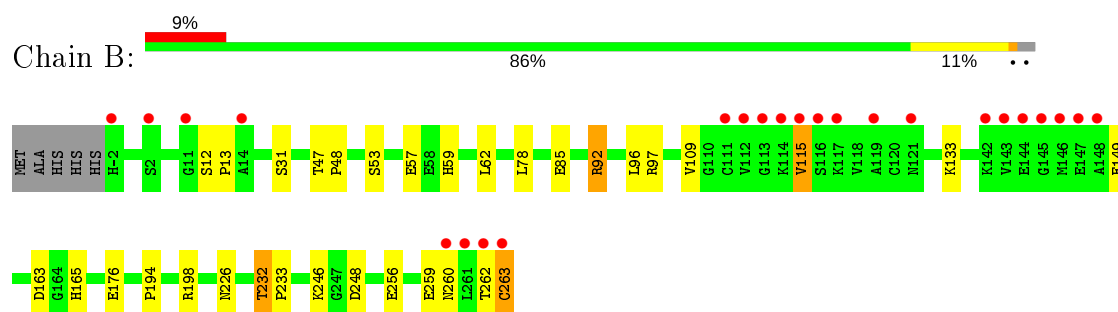
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: 6-phosphogluconolactonase



- Molecule 1: 6-phosphogluconolactonase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.77Å 63.39Å 79.32Å 90.00° 94.91° 90.00°	Depositor
Resolution (Å)	49.50 – 1.95 49.45 – 1.95	Depositor EDS
% Data completeness (in resolution range)	95.3 (49.50-1.95) 95.3 (49.45-1.95)	Depositor EDS
$R_{merge}$	0.24	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.52 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, $R_{free}$	0.231 , 0.270 0.244 , 0.281	Depositor DCC
$R_{free}$ test set	2543 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.9	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 51.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	4317	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.21% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 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	0.76	0/2061	0.86	0/2791
1	B	0.74	0/2088	0.87	0/2828
All	All	0.75	0/4149	0.86	0/5619

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	2021	0	2070	21	0
1	B	2047	0	2091	21	0
2	A	10	0	0	1	0
2	B	5	0	0	1	0
3	A	12	0	16	0	0
3	B	6	0	8	2	0
4	A	113	0	0	8	0
4	B	103	0	0	6	0
All	All	4317	0	4185	41	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 5.



All (41) 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:151:PRO:HG2	4:A:468:HOH:O	1.65	0.96
1:A:226:ASN:HB3	4:A:498:HOH:O	1.67	0.93
1:B:62:LEU:HB2	3:B:302:GOL:H2	1.56	0.87
1:B:226:ASN:HB3	4:B:492:HOH:O	1.74	0.84
1:B:246:LYS:HE3	4:B:474:HOH:O	1.78	0.83
1:B:176:GLU:OE1	4:B:402:HOH:O	2.13	0.67
1:B:47:THR:HB	1:B:48:PRO:HD3	1.78	0.65
1:B:59:HIS:ND1	3:B:302:GOL:H12	2.13	0.64
1:B:246:LYS:CE	4:B:474:HOH:O	2.41	0.62
1:A:53:SER:O	1:A:57:GLU:HG2	2.01	0.60
1:B:53:SER:O	1:B:57:GLU:HG2	2.00	0.60
1:A:205:THR:HG23	4:A:468:HOH:O	2.02	0.59
1:A:256:GLU:O	1:A:259:GLU:HG3	2.04	0.57
1:B:226:ASN:HD21	1:B:260:ASN:HD22	1.54	0.56
1:B:256:GLU:O	1:B:259:GLU:HG3	2.08	0.53
1:A:160:LEU:HD13	4:A:492:HOH:O	2.09	0.52
1:B:163:ASP:OD1	1:B:165:HIS:ND1	2.27	0.52
1:B:133:LYS:HE3	4:B:498:HOH:O	2.08	0.51
1:A:232:THR:HG23	1:A:232:THR:O	2.10	0.51
1:B:232:THR:HG23	1:B:232:THR:O	2.11	0.51
1:B:78:LEU:HD23	1:B:109:VAL:HG12	1.93	0.49
1:A:92:ARG:HA	1:A:96:LEU:HB2	1.94	0.49
1:A:256:GLU:OE1	1:B:97:ARG:NH1	2.44	0.48
1:B:92:ARG:HA	1:B:96:LEU:HB2	1.93	0.48
1:B:198:ARG:NH2	2:B:301:PO4:O2	2.41	0.48
1:A:163:ASP:OD1	1:A:165:HIS:ND1	2.28	0.47
1:A:226:ASN:HD21	1:A:260:ASN:HD22	1.61	0.47
1:A:220:GLU:HG3	4:A:404:HOH:O	2.14	0.47
1:B:85:GLU:HG2	1:B:194:PRO:HG3	1.97	0.47
1:A:33:ARG:HD2	2:A:301:PO4:O4	2.15	0.47
1:A:85:GLU:HG2	1:A:194:PRO:HG3	1.97	0.46
1:B:115:VAL:HG22	4:B:473:HOH:O	2.15	0.45
1:A:220:GLU:CG	4:A:404:HOH:O	2.63	0.45
1:A:47:THR:HB	1:A:48:PRO:HD3	1.98	0.45
1:B:262:THR:O	1:B:263:CYS:OXT	2.35	0.44
1:A:205:THR:CG2	4:A:468:HOH:O	2.62	0.44
1:A:2:SER:HB3	1:A:4:LYS:HE3	1.98	0.44
1:A:232:THR:N	1:A:233:PRO:CD	2.81	0.44
1:B:232:THR:N	1:B:233:PRO:CD	2.82	0.43
1:A:13:PRO:HA	1:A:16:LEU:HD13	2.02	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:133:LYS:HE3	4:A:499:HOH:O	2.19	0.42

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	260/271 (96%)	252 (97%)	7 (3%)	1 (0%)	34	22
1	B	265/271 (98%)	255 (96%)	9 (3%)	1 (0%)	34	22
All	All	525/542 (97%)	507 (97%)	16 (3%)	2 (0%)	34	22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	PRO
1	B	13	PRO

### 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	224/230 (97%)	217 (97%)	7 (3%)	40	28
1	B	227/230 (99%)	219 (96%)	8 (4%)	36	24
All	All	451/460 (98%)	436 (97%)	15 (3%)	38	26

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-2	HIS
1	A	12	SER
1	A	177	SER
1	A	222	LYS
1	A	232	THR
1	A	248	ASP
1	A	263	CYS
1	B	12	SER
1	B	31	SER
1	B	92	ARG
1	B	115	VAL
1	B	149	GLU
1	B	232	THR
1	B	248	ASP
1	B	263	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	93	GLN
1	A	244	ASN
1	A	260	ASN
1	B	93	GLN
1	B	244	ASN
1	B	260	ASN

### 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 carbohydrates 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	PO4	B	301	-	4,4,4	0.65	0	6,6,6	0.73	0
2	PO4	A	301	-	4,4,4	1.67	1 (25%)	6,6,6	1.06	0
2	PO4	A	302	-	4,4,4	1.48	1 (25%)	6,6,6	0.65	0
3	GOL	A	304	-	5,5,5	0.09	0	5,5,5	0.29	0
3	GOL	B	302	-	5,5,5	0.16	0	5,5,5	0.35	0
3	GOL	A	303	-	5,5,5	0.21	0	5,5,5	0.45	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	303	-	-	2/4/4/4	-
3	GOL	A	304	-	-	2/4/4/4	-
3	GOL	B	302	-	-	4/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	302	PO4	P-O1	2.24	1.56	1.50
2	A	301	PO4	P-O4	-2.18	1.48	1.54

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	304	GOL	C1-C2-C3-O3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	B	302	GOL	C1-C2-C3-O3
3	B	302	GOL	O2-C2-C3-O3
3	A	303	GOL	O1-C1-C2-O2
3	A	303	GOL	O1-C1-C2-C3
3	A	304	GOL	O2-C2-C3-O3
3	B	302	GOL	O1-C1-C2-O2
3	B	302	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	301	PO4	1	0
2	A	301	PO4	1	0
3	B	302	GOL	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

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	264/271 (97%)	0.63	22 (8%) 11 17	12, 22, 56, 88	0
1	B	266/271 (98%)	0.68	24 (9%) 9 15	13, 23, 60, 87	0
All	All	530/542 (97%)	0.66	46 (8%) 10 16	12, 22, 60, 88	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	113	GLY	9.1
1	B	262	THR	8.8
1	A	263	CYS	8.7
1	B	263	CYS	8.5
1	B	116	SER	8.5
1	B	148	ALA	8.2
1	A	262	THR	7.9
1	A	113	GLY	7.9
1	B	115	VAL	7.7
1	A	148	ALA	7.4
1	A	116	SER	5.8
1	A	260	ASN	5.5
1	B	146	MET	5.5
1	B	11	GLY	5.0
1	A	112	VAL	4.8
1	B	147	GLU	4.6
1	A	261	LEU	4.5
1	B	145	GLY	4.5
1	A	11	GLY	4.4
1	A	111	CYS	4.2
1	A	115	VAL	4.1
1	A	114	LYS	4.0
1	B	143	VAL	3.9
1	A	14	ALA	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	111	CYS	3.6
1	A	-2	HIS	3.5
1	B	2	SER	3.5
1	B	121	ASN	3.5
1	B	261	LEU	3.4
1	A	2	SER	3.4
1	B	-2	HIS	3.3
1	B	260	ASN	3.2
1	B	112	VAL	3.2
1	A	142	LYS	3.1
1	A	143	VAL	3.1
1	A	121	ASN	3.0
1	B	14	ALA	2.8
1	B	144	GLU	2.8
1	A	120	CYS	2.7
1	B	114	LYS	2.6
1	A	117	LYS	2.6
1	B	117	LYS	2.5
1	B	142	LYS	2.2
1	A	13	PRO	2.1
1	B	119	ALA	2.1
1	A	0	HIS	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 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, 95<sup>th</sup> 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	B	302	6/6	0.83	0.19	23,31,38,39	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GOL	A	304	6/6	0.85	0.14	34,36,37,37	6
3	GOL	A	303	6/6	0.90	0.17	28,35,38,38	0
2	PO4	A	302	5/5	0.95	0.13	33,35,36,36	0
2	PO4	B	301	5/5	0.95	0.13	34,35,39,40	0
2	PO4	A	301	5/5	0.99	0.16	19,19,22,23	5

## 6.5 Other polymers [i](#)

There are no such residues in this entry.