



# Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 02:50 pm BST

PDB ID : 5W56  
Title : Structure of Apo AztC  
Authors : Avalos, D.; Yukl, E.T.  
Deposited on : 2017-06-14  
Resolution : 2.03 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

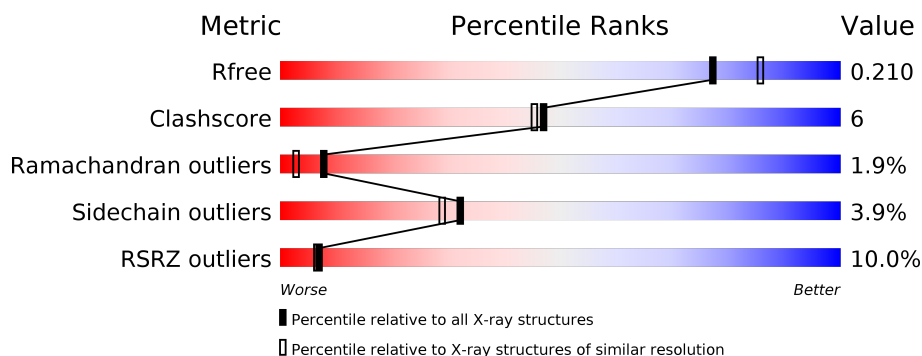
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.03 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	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

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	286	<div> <div>11%</div> <div> <div></div> <div>83%</div> <div>11%</div> <div>• •</div> </div> </div>
1	B	286	<div> <div>8%</div> <div> <div></div> <div>83%</div> <div>8%</div> <div>• 7%</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
2	GOL	A	401	-	-	X	-
2	GOL	B	402	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4174 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 Periplasmic solute binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	7	3	0
			2056	1293	366	392	5			
1	B	265	Total	C	N	O	S	0	1	0
			1956	1223	345	383	5			

- Molecule 2 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
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Na	0	0
			1	1		
3	A	1	Total	Na	0	0
			1	1		

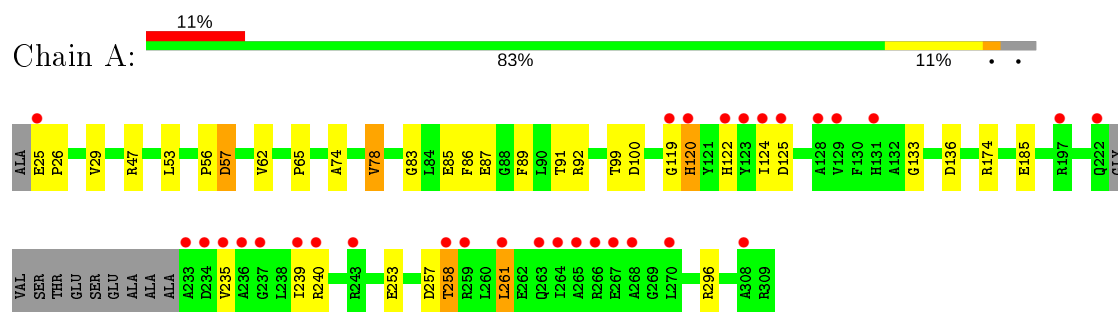
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	66	Total	O	0	0
			66	66		
4	B	70	Total	O	0	0
			70	70		

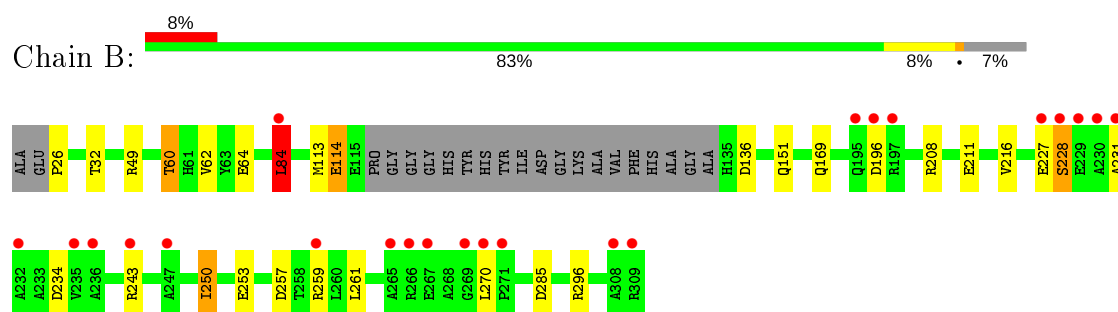
### 3 Residue-property plots [i](#)

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: Periplasmic solute binding protein



#### • Molecule 1: Periplasmic solute binding protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.47Å 106.26Å 61.67Å 90.00° 110.98° 90.00°	Depositor
Resolution (Å)	45.83 – 2.03 45.78 – 2.03	Depositor EDS
% Data completeness (in resolution range)	99.6 (45.83-2.03) 99.6 (45.78-2.03)	Depositor EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 2.03Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.165 , 0.201 0.174 , 0.210	Depositor DCC
$R_{free}$ test set	2008 reflections (4.22%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.3	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 59.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.147 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	4174	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	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.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NA

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	0.91	0/2097	0.94	2/2860 (0.1%)
1	B	0.94	0/1990	0.98	5/2714 (0.2%)
All	All	0.93	0/4087	0.96	7/5574 (0.1%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	136	ASP	CB-CG-OD1	7.06	124.65	118.30
1	B	136	ASP	CB-CG-OD2	-6.17	112.74	118.30
1	B	151	GLN	CB-CA-C	-5.52	99.36	110.40
1	B	84	LEU	CA-CB-CG	5.42	127.77	115.30
1	A	261	LEU	CA-CB-CG	5.16	127.17	115.30
1	A	136	ASP	CB-CG-OD2	-5.16	113.66	118.30
1	B	169	GLN	CB-CA-C	5.06	120.51	110.40

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	2056	0	1990	21	0
1	B	1956	0	1901	28	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	12	0	16	6	0
2	B	12	0	16	7	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	66	0	0	3	0
4	B	70	0	0	1	0
All	All	4174	0	3923	49	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 6.

All (49) 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:B:250:ILE:CG2	1:B:261:LEU:HD11	2.10	0.82
1:B:60:THR:HG21	1:B:64:GLU:HB3	1.64	0.80
1:B:60:THR:HG21	1:B:64:GLU:CB	2.13	0.78
1:B:114:GLU:O	1:B:208:ARG:NH2	2.27	0.68
1:A:99:THR:O	1:A:100:ASP:HB3	1.93	0.68
1:B:250:ILE:HD13	1:B:270:LEU:HB3	1.77	0.66
1:A:87:GLU:O	1:A:91:THR:HG23	1.96	0.65
1:A:92:ARG:NH1	4:A:501:HOH:O	2.20	0.65
1:B:250:ILE:HG21	1:B:261:LEU:HD11	1.78	0.65
1:A:296[A]:ARG:NE	2:A:401:GOL:O1	2.30	0.65
1:A:85:GLU:CD	1:A:85:GLU:H	2.01	0.64
1:B:32:THR:HG21	2:B:402:GOL:H32	1.82	0.62
1:A:99:THR:O	1:A:100:ASP:CB	2.45	0.61
1:B:60:THR:CG2	1:B:64:GLU:CB	2.80	0.59
1:A:47:ARG:HD2	4:A:550:HOH:O	2.02	0.59
1:A:296[A]:ARG:HD2	2:A:401:GOL:O1	2.01	0.59
1:A:185:GLU:OE1	1:A:296[A]:ARG:NH2	2.34	0.56
1:B:60:THR:CG2	1:B:64:GLU:HB3	2.35	0.56
1:B:60:THR:CG2	1:B:64:GLU:HB2	2.36	0.55
1:A:25:GLU:HB2	1:A:26:PRO:C	2.27	0.55
1:A:235:VAL:O	1:A:239:ILE:HD12	2.07	0.55
1:A:296[A]:ARG:CD	2:A:401:GOL:O1	2.55	0.55
1:B:60:THR:HG21	1:B:64:GLU:HB2	1.89	0.54
1:B:32:THR:HG21	2:B:402:GOL:C3	2.38	0.53
1:A:174:ARG:NH1	4:A:504:HOH:O	2.41	0.52
1:B:84:LEU:HG	2:B:402:GOL:O2	2.09	0.52
1:B:296:ARG:NH2	2:B:401:GOL:O1	2.31	0.52

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:113:MET:HE1	1:B:208:ARG:HB3	1.91	0.52
1:B:285:ASP:HB2	4:B:518:HOH:O	2.09	0.51
1:B:296:ARG:HD2	2:B:401:GOL:C1	2.40	0.50
1:B:257:ASP:OD2	1:B:259:ARG:NH2	2.46	0.49
1:A:86:PHE:N	2:A:402:GOL:O3	2.43	0.49
1:A:83:GLY:H	2:A:402:GOL:H2	1.78	0.48
1:A:29:VAL:HG22	1:A:78[B]:VAL:CG1	2.43	0.48
1:A:296[A]:ARG:HD2	2:A:401:GOL:C1	2.44	0.47
1:B:211:GLU:HG3	1:B:216:VAL:O	2.14	0.47
1:B:84:LEU:HB2	2:B:402:GOL:H2	1.98	0.45
1:B:227:GLU:HB3	1:B:259:ARG:HD2	1.98	0.44
1:B:257:ASP:C	1:B:257:ASP:OD1	2.56	0.44
1:B:270:LEU:HD23	1:B:270:LEU:HA	1.95	0.43
1:B:113:MET:HB3	1:B:113:MET:HE3	1.87	0.42
1:A:257:ASP:O	1:A:258:THR:CB	2.68	0.42
1:B:26:PRO:HG2	1:B:49:ARG:HH21	1.85	0.42
1:B:231:ALA:N	1:B:234:ASP:OD2	2.51	0.41
1:B:296:ARG:HD2	2:B:401:GOL:O1	2.21	0.41
1:A:56:PRO:O	1:A:57:ASP:C	2.59	0.41
1:A:65:PRO:HB3	1:A:89:PHE:CE2	2.57	0.40
1:A:74:ALA:O	1:A:99:THR:OG1	2.38	0.40
1:B:250:ILE:HG23	1:B:261:LEU:HD11	1.98	0.40

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	274/286 (96%)	258 (94%)	9 (3%)	7 (3%)	5	1
1	B	262/286 (92%)	250 (95%)	9 (3%)	3 (1%)	14	7
All	All	536/572 (94%)	508 (95%)	18 (3%)	10 (2%)	8	3

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	124	ILE
1	B	84	LEU
1	A	57	ASP
1	A	119	GLY
1	A	133	GLY
1	A	258	THR
1	B	228	SER
1	A	62	VAL
1	A	120	HIS
1	B	62	VAL

### 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	196/200 (98%)	187 (95%)	9 (5%)	27	22
1	B	189/200 (94%)	182 (96%)	7 (4%)	34	31
All	All	385/400 (96%)	369 (96%)	16 (4%)	32	26

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	LEU
1	A	78[A]	VAL
1	A	78[B]	VAL
1	A	120	HIS
1	A	122	HIS
1	A	125	ASP
1	A	240	ARG
1	A	253	GLU
1	A	261	LEU
1	B	60	THR
1	B	114	GLU
1	B	196	ASP
1	B	228	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	243	ARG
1	B	250	ILE
1	B	253	GLU

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

### 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 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	GOL	B	402	-	5,5,5	0.58	0	5,5,5	1.07	0
2	GOL	A	401	-	5,5,5	0.41	0	5,5,5	0.85	0
2	GOL	A	402	-	5,5,5	0.68	0	5,5,5	1.04	0
2	GOL	B	401	-	5,5,5	0.38	0	5,5,5	0.82	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	B	402	-	-	4/4/4/4	-
2	GOL	A	401	-	-	4/4/4/4	-
2	GOL	A	402	-	-	2/4/4/4	-
2	GOL	B	401	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	402	GOL	O1-C1-C2-O2
2	B	402	GOL	O1-C1-C2-C3
2	B	402	GOL	C1-C2-C3-O3
2	A	401	GOL	O1-C1-C2-C3
2	A	401	GOL	O1-C1-C2-O2
2	A	402	GOL	C1-C2-C3-O3
2	B	402	GOL	O2-C2-C3-O3
2	A	401	GOL	O2-C2-C3-O3
2	A	402	GOL	O2-C2-C3-O3
2	A	401	GOL	C1-C2-C3-O3

There are no ring outliers.

4 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	402	GOL	4	0
2	A	401	GOL	4	0
2	A	402	GOL	2	0
2	B	401	GOL	3	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, 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	275/286 (96%)	0.72	31 (11%)	5 4	35, 52, 96, 121	0
1	B	265/286 (92%)	0.56	23 (8%)	10 9	36, 51, 94, 127	0
All	All	540/572 (94%)	0.64	54 (10%)	7 6	35, 52, 94, 127	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	233	ALA	6.3
1	A	259	ARG	6.1
1	A	235	VAL	6.1
1	A	119	GLY	5.7
1	B	231	ALA	5.4
1	A	120	HIS	4.8
1	B	243	ARG	4.7
1	B	197	ARG	4.7
1	B	228	SER	4.6
1	A	122	HIS	4.4
1	A	125	ASP	4.4
1	A	261	LEU	4.4
1	B	84	LEU	4.1
1	A	124	ILE	4.1
1	A	234	ASP	3.9
1	B	308	ALA	3.9
1	A	222	GLN	3.9
1	B	230	ALA	3.8
1	A	236	ALA	3.6
1	B	271	PRO	3.5
1	B	266	ARG	3.4
1	A	243	ARG	3.3
1	B	269	GLY	3.3
1	A	197	ARG	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	240	ARG	3.1
1	A	268	ALA	3.0
1	B	270	LEU	3.0
1	B	235	VAL	2.9
1	A	266	ARG	2.8
1	A	258	THR	2.8
1	A	263	GLN	2.7
1	B	227	GLU	2.7
1	B	265	ALA	2.7
1	A	270	LEU	2.7
1	B	259	ARG	2.6
1	B	267	GLU	2.6
1	B	232	ALA	2.5
1	A	129	VAL	2.5
1	A	265	ALA	2.5
1	B	309	ARG	2.4
1	B	236	ALA	2.4
1	A	308	ALA	2.4
1	A	239	ILE	2.3
1	B	229	GLU	2.3
1	A	131	HIS	2.2
1	B	247	ALA	2.2
1	B	196	ASP	2.2
1	A	237	GLY	2.2
1	A	25	GLU	2.1
1	A	123	TYR	2.1
1	B	195	GLN	2.1
1	A	267	GLU	2.1
1	A	264	ILE	2.1
1	A	128	ALA	2.0

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

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

## 6.3 Carbohydrates ⓘ

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
2	GOL	B	401	6/6	0.81	0.27	76,82,87,106	0
2	GOL	A	401	6/6	0.84	0.26	63,76,82,86	0
2	GOL	A	402	6/6	0.90	0.33	48,69,89,92	0
3	NA	A	403	1/1	0.92	0.09	60,60,60,60	0
3	NA	B	403	1/1	0.94	0.07	58,58,58,58	0
2	GOL	B	402	6/6	0.94	0.47	48,82,100,102	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.