



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 02:30 PM BST

PDB ID : 3ZV0  
Title : Structure of the SHQ1P-CBF5P complex  
Authors : Walbott, H.; Machado-Pinilla, R.; Liger, D.; Blaud, M.; Rety, S.; Grozdanov, P.N.; Godin, K.; vanTilbeurgh, H.; Varani, G.; Meier, U.T.; Leulliot, N.  
Deposited on : 2011-07-22  
Resolution : 2.80 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 : 2.13.1  
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

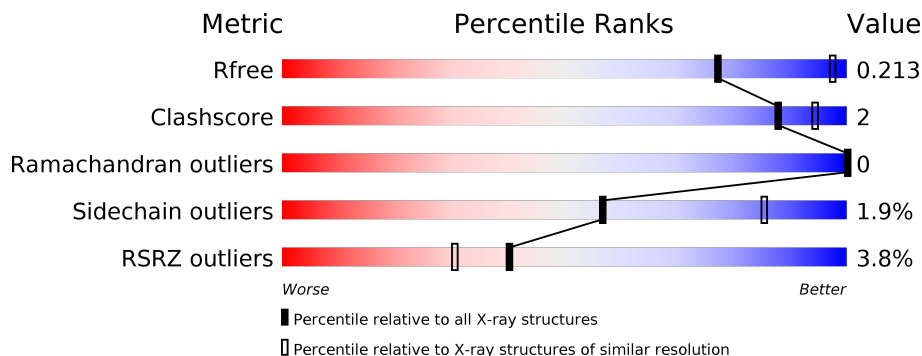
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	369	 3% 83% 5% 12%
1	B	369	 3% 84% 5% 11%
2	C	195	 4% 71% 7% 22%
2	D	195	 4% 73% 7% 19%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	B	1508	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7978 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 PROTEIN SHQ1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	325	Total	C	N	O	S	0	0	0
			2702	1726	439	528	9			
1	B	329	Total	C	N	O	S	0	0	0
			2734	1744	445	536	9			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	508	HIS	-	expression tag	UNP P40486
A	509	HIS	-	expression tag	UNP P40486
A	510	HIS	-	expression tag	UNP P40486
A	511	HIS	-	expression tag	UNP P40486
A	512	HIS	-	expression tag	UNP P40486
A	513	HIS	-	expression tag	UNP P40486
B	508	HIS	-	expression tag	UNP P40486
B	509	HIS	-	expression tag	UNP P40486
B	510	HIS	-	expression tag	UNP P40486
B	511	HIS	-	expression tag	UNP P40486
B	512	HIS	-	expression tag	UNP P40486
B	513	HIS	-	expression tag	UNP P40486

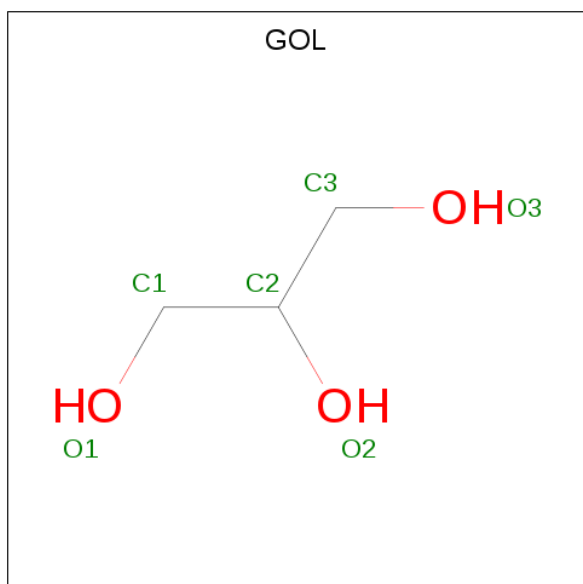
- Molecule 2 is a protein called H/ACA RIBONUCLEOPROTEIN COMPLEX SUBUNIT 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	152	Total	C	N	O	S	0	0	0
			1184	756	205	216	7			
2	D	157	Total	C	N	O	S	0	0	0
			1229	787	212	223	7			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-5	HIS	-	expression tag	UNP P33322
C	-4	HIS	-	expression tag	UNP P33322
C	-3	HIS	-	expression tag	UNP P33322
C	-2	HIS	-	expression tag	UNP P33322
C	-1	HIS	-	expression tag	UNP P33322
C	0	HIS	-	expression tag	UNP P33322
D	-5	HIS	-	expression tag	UNP P33322
D	-4	HIS	-	expression tag	UNP P33322
D	-3	HIS	-	expression tag	UNP P33322
D	-2	HIS	-	expression tag	UNP P33322
D	-1	HIS	-	expression tag	UNP P33322
D	0	HIS	-	expression tag	UNP P33322

- 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		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

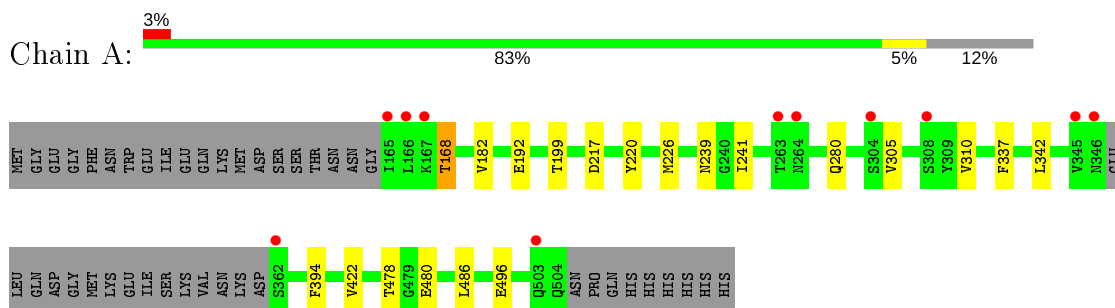
- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	43	Total 43	O 43	0	0
4	B	24	Total 24	O 24	0	0
4	C	19	Total 19	O 19	0	0
4	D	13	Total 13	O 13	0	0

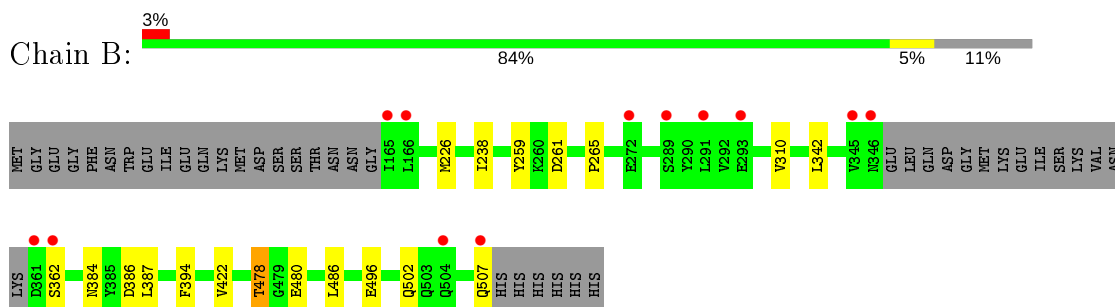
### 3 Residue-property plots [i](#)

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.

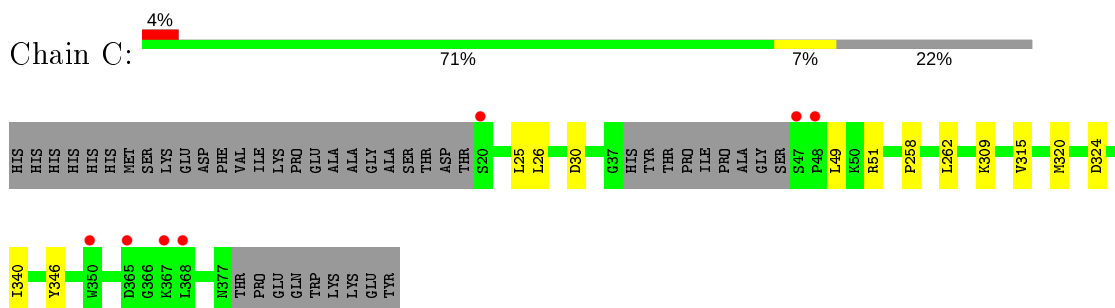
- Molecule 1: PROTEIN SHQ1



- Molecule 1: PROTEIN SHQ1

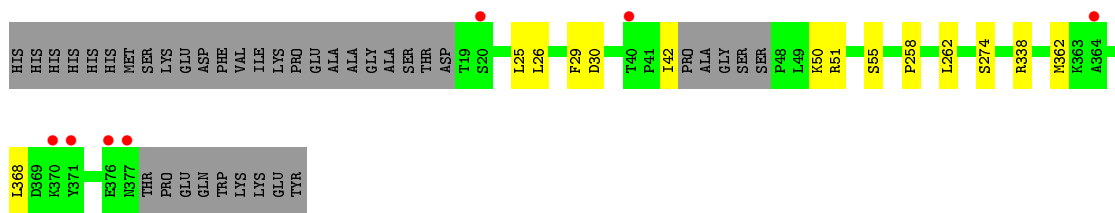


- Molecule 2: H/ACA RIBONUCLEOPROTEIN COMPLEX SUBUNIT 4



- Molecule 2: H/ACA RIBONUCLEOPROTEIN COMPLEX SUBUNIT 4







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.06Å 154.06Å 121.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.21 – 2.80 39.23 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (39.21-2.80) 99.9 (39.23-2.80)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 2.81Å)	Xtrriage
Refinement program	BUSTER 2.8.0	Depositor
R, $R_{free}$	0.178 , 0.208 0.186 , 0.213	Depositor DCC
$R_{free}$ test set	1830 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.1	Xtrriage
Anisotropy	0.473	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 52.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7978	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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

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	0.45	0/2758	0.60	0/3732
1	B	0.43	0/2791	0.62	0/3778
2	C	0.44	0/1201	0.69	0/1616
2	D	0.44	0/1249	0.70	0/1683
All	All	0.44	0/7999	0.64	0/10809

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	2702	0	2641	13	0
1	B	2734	0	2666	9	0
2	C	1184	0	1243	7	0
2	D	1229	0	1287	8	0
3	A	6	0	8	0	0
3	B	6	0	8	0	0
3	C	6	0	8	1	0
3	D	12	0	16	1	0
4	A	43	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	24	0	0	0	0
4	C	19	0	0	0	0
4	D	13	0	0	0	0
All	All	7978	0	7877	31	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 (31) 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:422:VAL:HG21	2:D:258:PRO:HB3	1.59	0.84
1:B:384:ASN:HD22	1:B:387:LEU:H	1.44	0.64
2:D:262:LEU:HB3	3:D:1379:GOL:H32	1.78	0.63
1:A:168:THR:HG21	1:A:199:THR:H	1.69	0.56
1:A:168:THR:HG21	1:A:199:THR:O	2.05	0.56
1:B:422:VAL:HG21	2:C:258:PRO:HB3	1.87	0.55
2:D:51:ARG:HH11	2:D:55:SER:HB3	1.72	0.55
1:A:239:ASN:HD22	1:A:241:ILE:H	1.56	0.53
1:A:182:VAL:HG22	2:C:49:LEU:HD11	1.91	0.52
2:D:29:PHE:HZ	2:D:42:ILE:HG21	1.75	0.52
1:A:310:VAL:HG11	1:A:394:PHE:HB2	1.93	0.51
1:B:259:TYR:CE2	1:B:265:PRO:HB2	2.46	0.51
1:B:310:VAL:HG11	1:B:394:PHE:HB2	1.94	0.50
1:B:502:GLN:HG2	1:B:507:GLN:HE22	1.78	0.48
1:A:226:MET:HE1	1:A:496:GLU:HG3	1.95	0.48
1:A:168:THR:O	1:A:168:THR:HG23	2.14	0.47
1:B:238:ILE:HD13	2:D:274:SER:HB2	1.97	0.47
1:A:192:GLU:HB3	1:A:337:PHE:CZ	2.50	0.46
1:B:226:MET:HE1	1:B:496:GLU:HA	1.98	0.45
1:B:384:ASN:HD21	1:B:386:ASP:HB2	1.80	0.45
1:A:226:MET:CE	1:A:496:GLU:HG3	2.47	0.45
1:B:478:THR:HG22	1:B:480:GLU:H	1.82	0.44
1:A:478:THR:HG22	1:A:480:GLU:H	1.82	0.44
1:A:217:ASP:OD2	1:A:220:TYR:HB2	2.17	0.44
2:C:309:LYS:HB3	2:D:50:LYS:HB3	2.01	0.43
1:A:168:THR:O	1:A:168:THR:CG2	2.67	0.42
2:C:320:MET:HB2	2:C:324:ASP:HB2	2.00	0.42
2:C:262:LEU:HG	2:D:338:ARG:HG2	2.03	0.41
2:C:315:VAL:HG23	2:C:340:ILE:HD11	2.03	0.41
2:D:362:MET:HB3	2:D:368:LEU:HD13	2.03	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:346:TYR:CE1	3:C:1378:GOL:H31	2.56	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	321/369 (87%)	313 (98%)	8 (2%)	0	100	100
1	B	325/369 (88%)	316 (97%)	9 (3%)	0	100	100
2	C	148/195 (76%)	148 (100%)	0	0	100	100
2	D	153/195 (78%)	153 (100%)	0	0	100	100
All	All	947/1128 (84%)	930 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	303/343 (88%)	298 (98%)	5 (2%)	60	87
1	B	307/343 (90%)	302 (98%)	5 (2%)	62	88
2	C	129/166 (78%)	125 (97%)	4 (3%)	40	74
2	D	134/166 (81%)	131 (98%)	3 (2%)	52	83

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	873/1018 (86%)	856 (98%)	17 (2%)	57 85

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	168	THR
1	A	280	GLN
1	A	305	VAL
1	A	342	LEU
1	A	486	LEU
1	B	261	ASP
1	B	342	LEU
1	B	362	SER
1	B	478	THR
1	B	486	LEU
2	C	25	LEU
2	C	26	LEU
2	C	30	ASP
2	C	51	ARG
2	D	25	LEU
2	D	26	LEU
2	D	30	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	232	ASN
1	A	239	ASN
1	B	191	ASN
1	B	384	ASN
1	B	507	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

5 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	D	1379	-	5,5,5	0.56	0	5,5,5	0.98	0
3	GOL	A	1505	-	5,5,5	0.75	0	5,5,5	0.46	0
3	GOL	B	1508	-	5,5,5	1.10	0	5,5,5	0.77	0
3	GOL	D	1378	-	5,5,5	0.54	0	5,5,5	0.50	0
3	GOL	C	1378	-	5,5,5	0.73	0	5,5,5	0.74	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	D	1379	-	-	1/4/4/4	-
3	GOL	A	1505	-	-	2/4/4/4	-
3	GOL	B	1508	-	-	0/4/4/4	-
3	GOL	D	1378	-	-	0/4/4/4	-
3	GOL	C	1378	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1505	GOL	O1-C1-C2-O2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	C	1378	GOL	C1-C2-C3-O3
3	C	1378	GOL	O2-C2-C3-O3
3	D	1379	GOL	C1-C2-C3-O3
3	A	1505	GOL	O1-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	1379	GOL	1	0
3	C	1378	GOL	1	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	325/369 (88%)	0.10	11 (3%) 45 35	30, 53, 98, 138	0
1	B	329/369 (89%)	0.01	12 (3%) 42 32	31, 55, 93, 131	0
2	C	152/195 (77%)	0.03	7 (4%) 32 22	29, 55, 114, 137	0
2	D	157/195 (80%)	-0.08	7 (4%) 33 23	30, 53, 104, 119	0
All	All	963/1128 (85%)	0.03	37 (3%) 40 30	29, 54, 102, 138	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	165	ILE	6.0
1	B	166	LEU	5.6
2	D	20	SER	4.9
1	B	362	SER	4.6
1	A	362	SER	4.5
1	A	166	LEU	3.9
1	B	165	ILE	3.8
2	D	377	ASN	3.6
1	B	345	VAL	3.5
2	C	47	SER	3.4
1	A	263	THR	3.4
1	B	361	ASP	3.4
1	A	264	ASN	3.3
2	C	20	SER	3.1
1	A	503	GLN	3.1
2	C	350	TRP	3.1
1	A	167	LYS	3.0
1	B	346	ASN	2.9
1	B	507	GLN	2.7
2	D	371	TYR	2.7
2	D	40	THR	2.7

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	368	LEU	2.6
1	B	272	GLU	2.5
1	A	308	SER	2.5
1	B	293	GLU	2.5
1	B	504	GLN	2.5
2	C	365	ASP	2.5
1	B	291	LEU	2.4
1	A	304	SER	2.4
2	D	370	LYS	2.3
1	A	345	VAL	2.2
2	C	367	LYS	2.2
2	C	48	PRO	2.1
2	D	364	ALA	2.0
2	D	376	GLU	2.0
1	A	346	ASN	2.0
1	B	289	SER	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 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, 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(Å <sup>2</sup> )	Q<0.9
3	GOL	B	1508	6/6	0.65	0.43	74,76,76,76	0
3	GOL	A	1505	6/6	0.75	0.23	82,83,83,84	0
3	GOL	C	1378	6/6	0.78	0.40	73,75,76,76	0
3	GOL	D	1378	6/6	0.83	0.37	94,95,95,96	0
3	GOL	D	1379	6/6	0.88	0.26	71,73,73,74	0

## 6.5 Other polymers

There are no such residues in this entry.