W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

Full wwPDB X-ray Structure Validation Report (i)

Feb 13, 2017 – 11:37 am GMT

PDB ID	:	5JQB
Title	:	Crystal structure of Ebola glycoprotein in complex with ibuprofen
Authors	:	Zhao, Y.; Ren, J.; Stuart, D.I.
Deposited on	:	2016-05-04
Resolution	:	2.68 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.68 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.
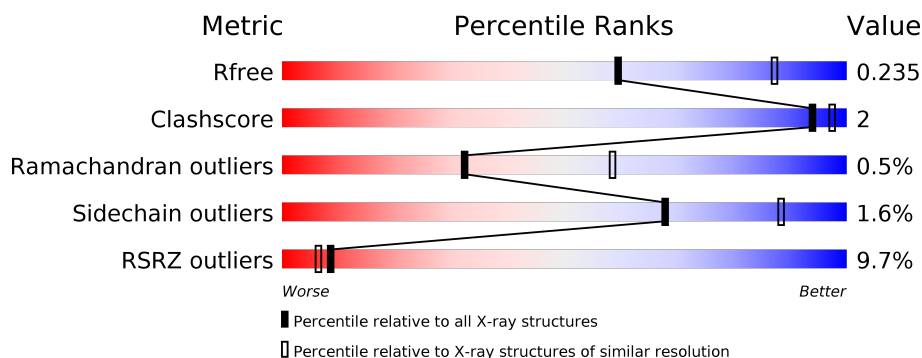
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3050 (2.70-2.66)
Clashscore	112137	3418 (2.70-2.66)
Ramachandran outliers	110173	3367 (2.70-2.66)
Sidechain outliers	110143	3367 (2.70-2.66)
RSRZ outliers	101464	3069 (2.70-2.66)

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. 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 <=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	330	5% / 74% / • / 22%
2	B	168	13% / 73% / 5% / 22%

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
3	NAG	A	604	-	-	-	X
4	GOL	A	605	-	-	-	X
4	GOL	A	607	-	-	-	X
4	GOL	B	701	-	-	-	X
4	GOL	B	702	-	-	-	X
6	IBP	B	706	-	-	-	X

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 3238 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 Envelope glycoprotein 1,Envelope glycoprotein 1,Envelope glycoprotein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	257	Total 1985	C 1266	N 334	O 380	S 5	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	GLU	-	expression tag	UNP Q05320
A	29	THR	-	expression tag	UNP Q05320
A	30	GLY	-	expression tag	UNP Q05320
A	31	ARG	-	expression tag	UNP Q05320
A	42	ALA	THR	engineered mutation	UNP Q05320

- Molecule 2 is a protein called Envelope glycoprotein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	131	Total 1040	C 663	N 179	O 192	S 6	0	0	0

There are 37 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	633	GLY	-	expression tag	UNP Q05320
B	634	SER	-	expression tag	UNP Q05320
B	635	GLY	-	expression tag	UNP Q05320
B	636	TYR	-	expression tag	UNP Q05320
B	637	ILE	-	expression tag	UNP Q05320
B	638	PRO	-	expression tag	UNP Q05320
B	639	GLU	-	expression tag	UNP Q05320
B	640	ALA	-	expression tag	UNP Q05320
B	641	PRO	-	expression tag	UNP Q05320
B	642	ARG	-	expression tag	UNP Q05320

Continued on next page...
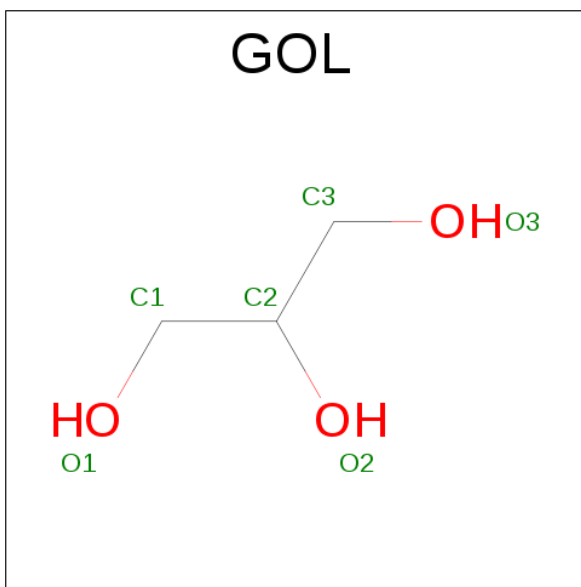
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	643	ASP	-	expression tag	UNP Q05320
B	644	GLY	-	expression tag	UNP Q05320
B	645	GLN	-	expression tag	UNP Q05320
B	646	ALA	-	expression tag	UNP Q05320
B	647	TYR	-	expression tag	UNP Q05320
B	648	VAL	-	expression tag	UNP Q05320
B	649	ARG	-	expression tag	UNP Q05320
B	650	LYS	-	expression tag	UNP Q05320
B	651	ASP	-	expression tag	UNP Q05320
B	652	GLY	-	expression tag	UNP Q05320
B	653	GLU	-	expression tag	UNP Q05320
B	654	TRP	-	expression tag	UNP Q05320
B	655	VAL	-	expression tag	UNP Q05320
B	656	LEU	-	expression tag	UNP Q05320
B	657	LEU	-	expression tag	UNP Q05320
B	658	SER	-	expression tag	UNP Q05320
B	659	THR	-	expression tag	UNP Q05320
B	660	PHE	-	expression tag	UNP Q05320
B	661	LEU	-	expression tag	UNP Q05320
B	662	GLY	-	expression tag	UNP Q05320
B	663	THR	-	expression tag	UNP Q05320
B	664	HIS	-	expression tag	UNP Q05320
B	665	HIS	-	expression tag	UNP Q05320
B	666	HIS	-	expression tag	UNP Q05320
B	667	HIS	-	expression tag	UNP Q05320
B	668	HIS	-	expression tag	UNP Q05320
B	669	HIS	-	expression tag	UNP Q05320

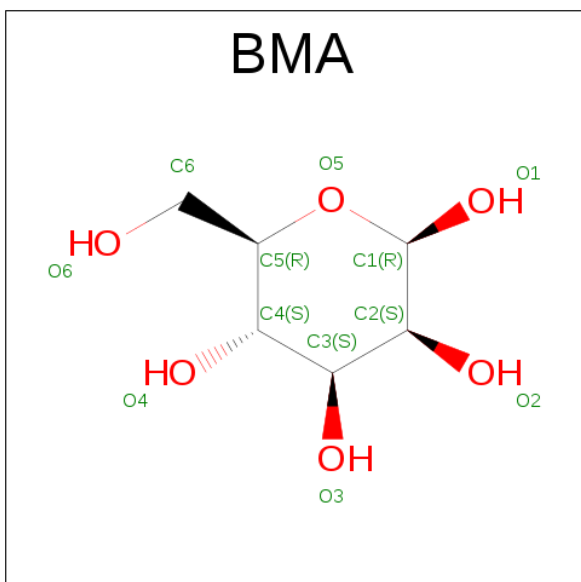- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

• Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

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

- Molecule 5 is BETA-D-MANNOSE (three-letter code: BMA) (formula: $C_6H_{12}O_6$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 11	C 6	O 5	0	0

- Molecule 6 is IBUPROFEN (three-letter code: IBP) (formula: $C_{13}H_{18}O_2$).
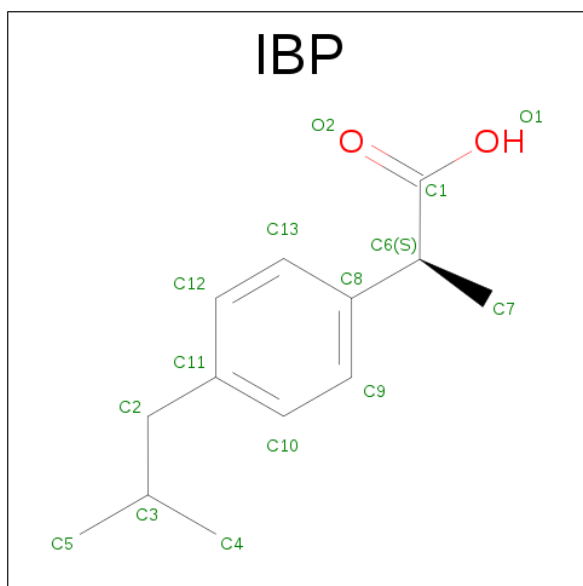
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total 15	C 13	O 2	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	40	Total 40	O 40	0	0
7	B	33	Total 33	O 33	0	0

3 Residue-property plots ⓘ

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: Envelope glycoprotein 1,Envelope glycoprotein 1,Envelope glycoprotein 1
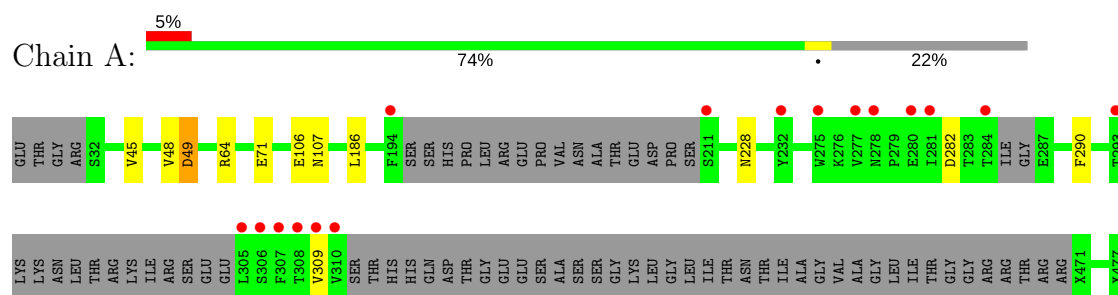
• Molecule 2: Envelope glycoprotein 2
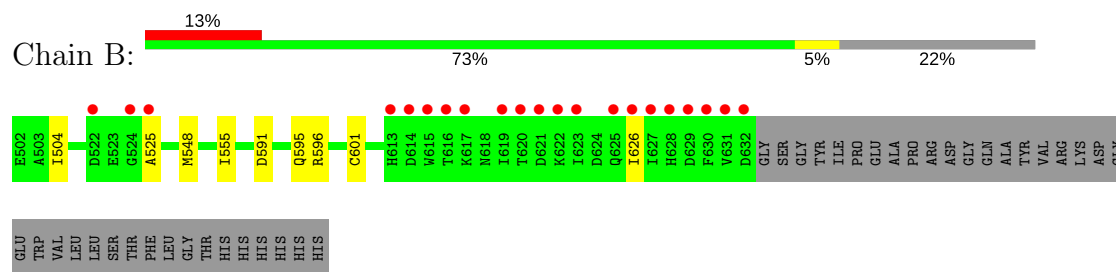
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	113.76Å 113.76Å 306.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	82.84 – 2.68 82.84 – 2.68	Depositor EDS
% Data completeness (in resolution range)	99.9 (82.84-2.68) 99.9 (82.84-2.68)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.199 , 0.234 0.200 , 0.235	Depositor DCC
R_{free} test set	1107 reflections (5.34%)	DCC
Wilson B-factor (Å^2)	59.3	Xtriage
Anisotropy	0.398	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 50.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3238	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: IBP, GOL, BMA, NAG

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	0.39	0/1996	0.63	0/2714
2	B	0.40	0/1066	0.56	0/1453
All	All	0.39	0/3062	0.61	0/4167

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	1985	0	1881	8	0
2	B	1040	0	1003	6	0
3	A	56	0	52	0	0
3	B	28	0	24	0	0
4	A	18	0	24	0	0
4	B	12	0	16	0	0
5	B	11	0	10	0	0
6	B	15	0	17	0	0
7	A	40	0	0	0	0
7	B	33	0	0	1	0
All	All	3238	0	3027	10	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 (10) 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:45:VAL:HG23	2:B:504:ILE:HD13	1.85	0.58
1:A:106:GLU:HG3	1:A:290:PHE:CD1	2.39	0.57
1:A:106:GLU:HG3	1:A:290:PHE:CG	2.40	0.56
1:A:49:ASP:HA	2:B:595:GLN:HG3	1.92	0.50
1:A:64:ARG:HG3	1:A:186:LEU:HD11	1.96	0.47
1:A:48:VAL:HG21	2:B:591:ASP:HB3	2.00	0.44
2:B:548:MET:HG2	2:B:555:ILE:HD11	2.00	0.43
2:B:596:ARG:NE	7:B:801:HOH:O	2.52	0.42
1:A:71:GLU:OE2	1:A:107:ASN:N	2.45	0.42
1:A:45:VAL:HG23	2:B:504:ILE:CD1	2.49	0.42

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	242/330 (73%)	235 (97%)	7 (3%)	0	100	100
2	B	129/168 (77%)	119 (92%)	8 (6%)	2 (2%)	11	26
All	All	371/498 (74%)	354 (95%)	15 (4%)	2 (0%)	32	58

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	525	ALA
2	B	626	ILE

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	209/273 (77%)	205 (98%)	4 (2%)	62	86
2	B	110/140 (79%)	109 (99%)	1 (1%)	82	94
All	All	319/413 (77%)	314 (98%)	5 (2%)	68	88

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	ASP
1	A	228	ASN
1	A	282	ASP
1	A	309	VAL
2	B	601	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	ASN
1	A	139	HIS
2	B	508	GLN

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 ⓘ

13 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	NAG	A	601	1	14,14,15	0.23	0	15,19,21	0.86	1 (6%)
3	NAG	A	602	1	14,14,15	0.39	0	15,19,21	0.99	0
3	NAG	A	603	1	14,14,15	0.29	0	15,19,21	0.70	0
3	NAG	A	604	1	14,14,15	0.57	0	15,19,21	2.51	3 (20%)
4	GOL	A	605	-	5,5,5	0.32	0	5,5,5	0.33	0
4	GOL	A	606	-	5,5,5	0.36	0	5,5,5	0.24	0
4	GOL	A	607	-	5,5,5	0.34	0	5,5,5	0.26	0
4	GOL	B	701	-	5,5,5	0.21	0	5,5,5	0.31	0
4	GOL	B	702	-	5,5,5	0.26	0	5,5,5	0.20	0
3	NAG	B	703	3,2	14,14,15	0.30	0	15,19,21	1.05	1 (6%)
3	NAG	B	704	3,5	14,14,15	0.38	0	15,19,21	0.95	1 (6%)
5	BMA	B	705	3	11,11,12	0.36	0	13,15,17	0.58	0
6	IBP	B	706	-	12,15,15	0.81	0	16,20,20	0.90	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	NAG	A	601	1	-	0/6/23/26	0/1/1/1
3	NAG	A	602	1	-	0/6/23/26	0/1/1/1
3	NAG	A	603	1	-	0/6/23/26	0/1/1/1
3	NAG	A	604	1	-	0/6/23/26	0/1/1/1
4	GOL	A	605	-	-	0/4/4/4	0/0/0/0
4	GOL	A	606	-	-	0/4/4/4	0/0/0/0
4	GOL	A	607	-	-	0/4/4/4	0/0/0/0
4	GOL	B	701	-	-	0/4/4/4	0/0/0/0
4	GOL	B	702	-	-	0/4/4/4	0/0/0/0
3	NAG	B	703	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	704	3,5	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	BMA	B	705	3	-	0/2/19/22	0/1/1/1
6	IBP	B	706	-	-	0/8/12/12	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	703	NAG	O5-C1-C2	-2.47	108.04	111.47
3	A	604	NAG	C4-C3-C2	-2.23	107.74	111.02
3	A	601	NAG	O5-C1-C2	-2.18	108.44	111.47
3	A	604	NAG	O5-C1-C2	2.65	115.16	111.47
3	B	704	NAG	C1-O5-C5	2.75	115.96	112.17
3	A	604	NAG	C1-O5-C5	8.66	124.10	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	250/330 (75%)	0.46	16 (6%)	20	18	42, 63, 130, 145	0
2	B	131/168 (77%)	0.83	21 (16%)	2	1	46, 59, 159, 166	0
All	All	381/498 (76%)	0.58	37 (9%)	8	6	42, 62, 142, 166	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	626	ILE	6.4
2	B	627	ILE	6.3
2	B	615	TRP	6.2
2	B	619	ILE	5.8
1	A	307	PHE	5.7
2	B	630	PHE	5.5
2	B	613	HIS	5.3
1	A	309	VAL	5.2
2	B	632	ASP	4.9
2	B	620	THR	4.8
1	A	305	LEU	4.6
2	B	631	VAL	4.4
2	B	628	HIS	4.3
2	B	623	ILE	4.1
1	A	306	SER	4.1
1	A	293	THR	4.0
2	B	525	ALA	3.9
2	B	621	ASP	3.9
2	B	522	ASP	3.5
2	B	616	THR	3.3
1	A	308	THR	3.2
1	A	277	VAL	3.0
1	A	284	THR	3.0
2	B	625	GLN	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	629	ASP	2.8
1	A	211	SER	2.7
1	A	281	ILE	2.7
1	A	310	VAL	2.6
2	B	622	LYS	2.6
1	A	232	TYR	2.5
1	A	278	ASN	2.5
1	A	280	GLU	2.5
2	B	524	GLY	2.2
1	A	194	PHE	2.2
2	B	617	LYS	2.1
2	B	614	ASP	2.1
1	A	275	TRP	2.1

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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
6	IBP	B	706	15/15	0.87	0.46	22.13	85,93,106,107	0
4	GOL	A	605	6/6	0.96	0.38	8.33	56,58,59,60	0
4	GOL	A	607	6/6	0.79	0.31	6.52	84,85,87,87	0
4	GOL	B	701	6/6	0.93	0.24	3.00	70,71,72,72	0
4	GOL	B	702	6/6	0.89	0.27	2.57	87,89,90,91	0
3	NAG	A	604	14/15	0.82	0.46	1.79	100,106,109,110	0
3	NAG	B	703	14/15	0.96	0.19	0.23	54,57,63,79	0
4	GOL	A	606	6/6	0.81	0.17	-0.57	83,86,87,87	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	NAG	A	601	14/15	0.96	0.17	-1.10	69,74,77,78	0
5	BMA	B	705	11/12	0.64	0.66	-	130,134,137,138	0
3	NAG	B	704	14/15	0.80	0.53	-	97,104,113,124	0
3	NAG	A	602	14/15	0.70	0.43	-	124,130,133,133	0
3	NAG	A	603	14/15	0.89	0.17	-	93,101,107,108	0

6.5 Other polymers ⓘ

There are no such residues in this entry.