



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:02 PM GMT

PDB ID : 4UP4  
Title : Structure of the recombinant lectin PVL from *Psathyrella velutina* in complex with GlcNAcb-D-1,3Galactoside  
Authors : Audfray, A.; Beljoudi, M.; Hurbin, A.; Varrot, A.; Breiman, A.; Busser, B.; Lependu, J.; Coll, J.L.; Imberty, A.  
Deposited on : 2014-06-12  
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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

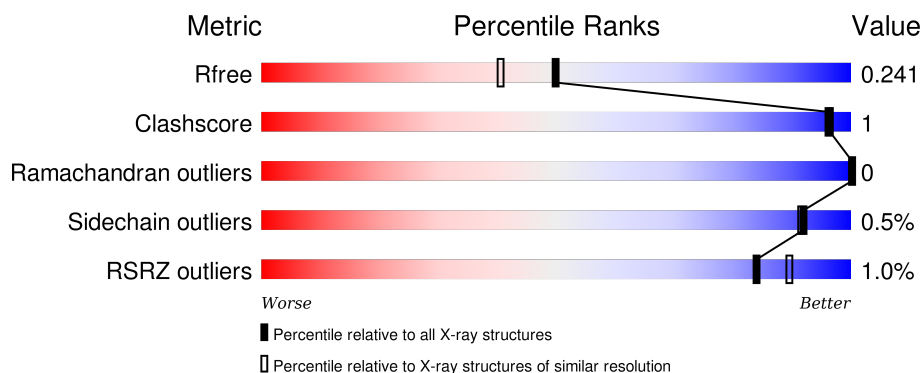
# 1 Overall quality at a glance

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))
$R_{free}$	91344	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (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. 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	401	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; left: 0; top: -10px;">%</span> </div> <div style="width: 97%; height: 10px; background-color: green; position: relative;"> <span style="position: absolute; left: 0; top: -10px;">%</span> </div> <div style="text-align: right; margin-top: 5px;">97%</div> </div>
1	B	401	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; left: 0; top: -10px;">%</span> </div> <div style="width: 96%; height: 10px; background-color: green; position: relative;"> <span style="position: absolute; left: 0; top: -10px;">%</span> </div> <div style="text-align: right; margin-top: 5px;">96%</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	CA	A	502	-	-	-	X
2	CA	B	502	-	-	-	X
3	NAG	A	582	-	-	-	X
3	GAL	B	551	-	-	-	X
3	GAL	B	591	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7251 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 LECTIN PVL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	400	Total	C	N	O	S	0	0	0
			3033	1924	545	555	9			
1	B	399	Total	C	N	O	S	0	4	0
			3057	1937	551	560	9			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	SEE REMARK 999	UNP Q309D1
A	2	VAL	-	SEE REMARK 999	UNP Q309D1
A	3	VAL	-	SEE REMARK 999	UNP Q309D1
A	4	VAL	-	SEE REMARK 999	UNP Q309D1
A	5	ILE	-	SEE REMARK 999	UNP Q309D1
A	6	SER	-	SEE REMARK 999	UNP Q309D1
A	72	PRO	SER	CONFLICT	UNP Q309D1
A	201	GLY	SER	CONFLICT	UNP Q309D1
A	376	GLN	LYS	CONFLICT	UNP Q309D1
B	1	SER	-	SEE REMARK 999	UNP Q309D1
B	2	VAL	-	SEE REMARK 999	UNP Q309D1
B	3	VAL	-	SEE REMARK 999	UNP Q309D1
B	4	VAL	-	SEE REMARK 999	UNP Q309D1
B	5	ILE	-	SEE REMARK 999	UNP Q309D1
B	6	SER	-	SEE REMARK 999	UNP Q309D1
B	72	PRO	SER	CONFLICT	UNP Q309D1
B	201	GLY	SER	CONFLICT	UNP Q309D1
B	376	GLN	LYS	CONFLICT	UNP Q309D1

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total	Ca	0	0
			3	3		

*Continued on next page...*

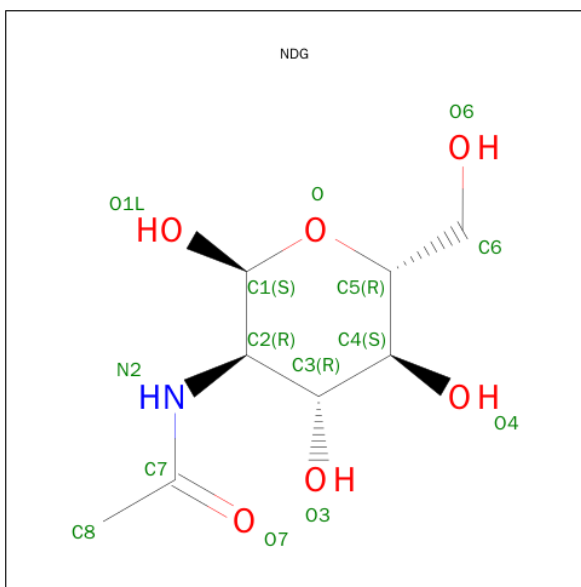
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	Ca	0	0
			3	3		

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

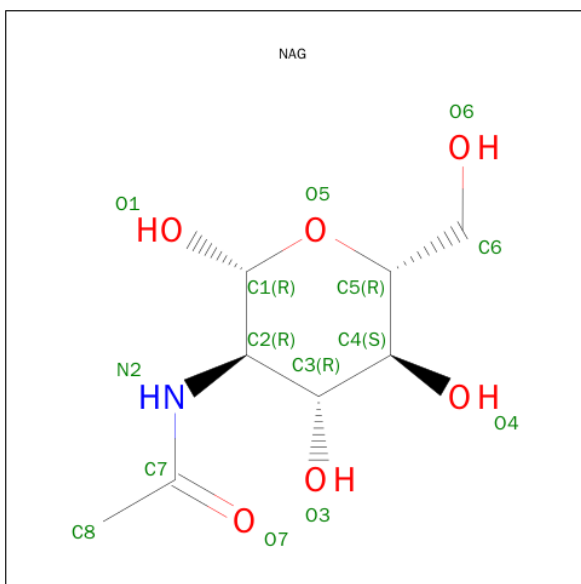
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			26	14	1	11		
3	A	2	Total	C	N	O	0	0
			26	14	1	11		
3	A	2	Total	C	N	O	0	0
			26	14	1	11		
3	A	2	Total	C	N	O	0	0
			26	14	1	11		
3	A	2	Total	C	N	O	0	0
			26	14	1	11		
3	B	2	Total	C	N	O	0	0
			26	14	1	11		
3	B	2	Total	C	N	O	0	0
			26	14	1	11		
3	B	2	Total	C	N	O	0	0
			26	14	1	11		
3	B	2	Total	C	N	O	0	0
			26	14	1	11		

- Molecule 4 is SUGAR (2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE) (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	1
			15	8	1	6		
4	B	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	1
			15	8	1	6		

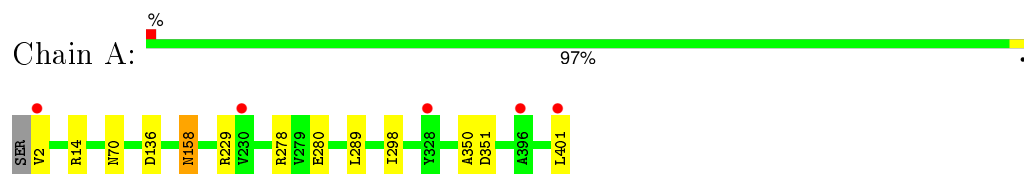
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	453	Total 453	O 453	0	0
6	B	397	Total 397	O 397	0	0

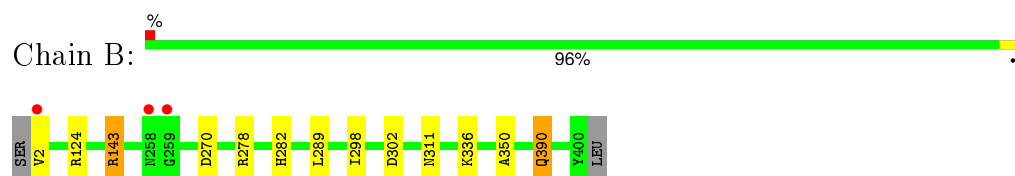
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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: LECTIN PVL



- Molecule 1: LECTIN PVL





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.81Å 97.47Å 158.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.95 30.00 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-1.95) 99.8 (30.00-1.95)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.85 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, $R_{free}$	0.190 , 0.237 0.198 , 0.241	Depositor DCC
$R_{free}$ test set	3070 reflections (5.15%)	DCC
Wilson B-factor (Å <sup>2</sup> )	15.8	Xtriage
Anisotropy	0.792	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 52.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 62671 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7251	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CA, GAL, NAG, NDG

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	0.83	0/3101	0.87	4/4210 (0.1%)
1	B	0.80	0/3125	0.84	2/4243 (0.0%)
All	All	0.81	0/6226	0.85	6/8453 (0.1%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	351	ASP	CB-CG-OD2	5.53	123.28	118.30
1	B	270	ASP	CB-CG-OD1	5.37	123.13	118.30
1	A	278	ARG	NE-CZ-NH1	5.12	122.86	120.30
1	B	143	ARG	NE-CZ-NH1	5.07	122.84	120.30
1	A	136	ASP	CB-CG-OD1	5.05	122.85	118.30
1	A	229	ARG	NE-CZ-NH2	-5.02	117.79	120.30

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	3033	0	2960	8	0
1	B	3057	0	2980	10	0
2	A	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	3	0	0	0	0
3	A	130	0	120	0	0
3	B	130	0	120	0	0
4	A	15	0	15	0	0
4	B	15	0	15	0	0
5	A	15	0	15	0	0
6	A	453	0	0	2	0
6	B	397	0	0	2	0
All	All	7251	0	6225	18	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 1.

All (18) 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:280:GLU:OE2	6:A:2312:HOH:O	2.00	0.80
1:A:70:ASN:O	1:A:401:LEU:HB2	1.93	0.68
1:B:278:ARG:H	1:B:282:HIS:HD2	1.42	0.67
1:B:124:ARG:CZ	1:B:143:ARG:HG3	2.29	0.62
1:A:158:ASN:H	1:A:158:ASN:HD22	1.48	0.58
1:B:311:ASN:ND2	6:B:2317:HOH:O	2.38	0.56
1:B:2:VAL:O	1:B:2:VAL:HG13	2.06	0.55
1:B:282:HIS:HE1	1:B:302:ASP:OD1	1.94	0.50
1:A:2:VAL:HG21	6:A:2324:HOH:O	2.12	0.50
1:A:14:ARG:HG3	1:A:401:LEU:HD21	1.94	0.48
1:A:298:ILE:N	1:A:298:ILE:HD12	2.32	0.44
1:B:289:LEU:HD23	1:B:350:ALA:HB3	1.99	0.44
1:B:390[A]:GLN:CD	1:B:390[A]:GLN:N	2.71	0.44
1:A:158:ASN:N	1:A:158:ASN:HD22	2.13	0.43
1:B:298:ILE:HD12	1:B:298:ILE:N	2.35	0.41
1:B:124:ARG:CZ	1:B:143:ARG:CG	2.99	0.40
1:A:289:LEU:HD23	1:A:350:ALA:HB3	2.01	0.40
1:B:2:VAL:HA	6:B:2005:HOH:O	2.21	0.40

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	398/401 (99%)	392 (98%)	6 (2%)	0	100	100
1	B	401/401 (100%)	392 (98%)	9 (2%)	0	100	100
All	All	799/802 (100%)	784 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	310/311 (100%)	309 (100%)	1 (0%)	94	94
1	B	313/311 (101%)	310 (99%)	3 (1%)	82	80
All	All	623/622 (100%)	619 (99%)	4 (1%)	92	89

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	158	ASN
1	B	336	LYS
1	B	390[A]	GLN
1	B	390[B]	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	88	ASN
1	A	158	ASN
1	A	256	ASN
1	B	80	ASN
1	B	88	ASN
1	B	192	ASN
1	B	256	ASN
1	B	282	HIS
1	B	311	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 ⓘ

20 carbohydrates 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	GAL	A	551	3	12,12,12	0.93	1 (8%)	17,17,17	1.54	3 (17%)
3	NAG	A	552	3	14,14,15	0.90	1 (7%)	15,19,21	0.94	0
3	GAL	A	571	3	12,12,12	0.80	0	17,17,17	1.77	5 (29%)
3	NAG	A	572	3	14,14,15	1.21	2 (14%)	15,19,21	1.63	2 (13%)
3	GAL	A	581	3	12,12,12	0.87	0	17,17,17	1.49	2 (11%)
3	NAG	A	582	3	14,14,15	0.75	0	15,19,21	1.19	2 (13%)
3	GAL	A	591	3	12,12,12	0.57	0	17,17,17	1.46	4 (23%)
3	NAG	A	592	3	14,14,15	0.82	0	15,19,21	1.67	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GAL	A	601	3	12,12,12	0.93	0	17,17,17	1.50	3 (17%)
3	NAG	A	602	3	14,14,15	0.87	1 (7%)	15,19,21	1.69	5 (33%)
3	GAL	B	551	3	12,12,12	0.64	0	17,17,17	1.39	3 (17%)
3	NAG	B	552	3	14,14,15	1.22	1 (7%)	15,19,21	1.01	1 (6%)
3	GAL	B	571	3	12,12,12	0.67	0	17,17,17	0.99	0
3	NAG	B	572	3	14,14,15	0.86	0	15,19,21	1.56	4 (26%)
3	GAL	B	581	3	12,12,12	0.57	0	17,17,17	1.03	1 (5%)
3	NAG	B	582	3	14,14,15	0.74	0	15,19,21	1.18	1 (6%)
3	GAL	B	591	3	12,12,12	0.57	0	17,17,17	1.24	2 (11%)
3	NAG	B	592	3	14,14,15	0.70	0	15,19,21	1.46	3 (20%)
3	GAL	B	601	3	12,12,12	0.74	0	17,17,17	1.76	5 (29%)
3	NAG	B	602	3	14,14,15	0.78	0	15,19,21	1.44	3 (20%)

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	GAL	A	551	3	-	0/2/22/22	0/1/1/1
3	NAG	A	552	3	-	0/6/23/26	0/1/1/1
3	GAL	A	571	3	-	0/2/22/22	0/1/1/1
3	NAG	A	572	3	-	0/6/23/26	0/1/1/1
3	GAL	A	581	3	-	0/2/22/22	0/1/1/1
3	NAG	A	582	3	-	0/6/23/26	0/1/1/1
3	GAL	A	591	3	-	0/2/22/22	0/1/1/1
3	NAG	A	592	3	-	0/6/23/26	0/1/1/1
3	GAL	A	601	3	-	0/2/22/22	0/1/1/1
3	NAG	A	602	3	-	0/6/23/26	0/1/1/1
3	GAL	B	551	3	-	0/2/22/22	0/1/1/1
3	NAG	B	552	3	-	0/6/23/26	0/1/1/1
3	GAL	B	571	3	-	0/2/22/22	0/1/1/1
3	NAG	B	572	3	-	0/6/23/26	0/1/1/1
3	GAL	B	581	3	-	0/2/22/22	0/1/1/1
3	NAG	B	582	3	-	0/6/23/26	0/1/1/1
3	GAL	B	591	3	-	0/2/22/22	0/1/1/1
3	NAG	B	592	3	-	0/6/23/26	0/1/1/1
3	GAL	B	601	3	-	0/2/22/22	0/1/1/1
3	NAG	B	602	3	-	0/6/23/26	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	572	NAG	O5-C1	-3.20	1.38	1.43
3	B	552	NAG	O5-C1	-2.62	1.39	1.43
3	A	551	GAL	O5-C1	-2.14	1.39	1.43
3	A	602	NAG	C1-C2	2.25	1.55	1.52
3	A	552	NAG	C1-C2	2.47	1.55	1.52
3	A	572	NAG	O7-C7	2.91	1.30	1.23

All (52) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	592	NAG	C6-C5-C4	-3.59	104.15	113.02
3	B	601	GAL	O3-C3-C2	-3.24	103.03	110.34
3	A	571	GAL	O1-C1-C2	-3.18	100.68	109.21
3	B	551	GAL	C1-C2-C3	-3.04	105.90	110.43
3	B	601	GAL	O4-C4-C3	-2.93	103.74	110.34
3	B	601	GAL	C4-C3-C2	-2.90	105.38	110.79
3	B	582	NAG	C4-C3-C2	-2.85	106.80	111.23
3	B	552	NAG	C4-C3-C2	-2.84	106.82	111.23
3	A	601	GAL	O3-C3-C2	-2.64	104.39	110.34
3	A	571	GAL	C6-C5-C4	-2.63	106.52	113.02
3	A	602	NAG	O7-C7-C8	-2.62	117.25	122.06
3	B	551	GAL	O5-C1-C2	-2.47	105.86	109.80
3	B	601	GAL	O5-C5-C4	-2.45	105.09	109.68
3	A	591	GAL	C1-O5-C5	-2.42	108.99	113.47
3	B	592	NAG	O3-C3-C4	-2.30	105.15	110.34
3	A	601	GAL	O2-C2-C1	-2.27	104.82	109.82
3	B	572	NAG	C2-N2-C7	-2.26	120.13	123.04
3	A	601	GAL	O2-C2-C3	-2.25	105.27	110.34
3	B	602	NAG	C4-C3-C2	-2.16	107.86	111.23
3	B	601	GAL	O1-C1-O5	-2.14	104.41	110.25
3	A	591	GAL	O4-C4-C3	-2.10	105.61	110.34
3	B	581	GAL	O3-C3-C2	-2.07	105.67	110.34
3	B	592	NAG	O7-C7-C8	-2.04	118.33	122.06
3	A	582	NAG	C6-C5-C4	-2.04	107.99	113.02
3	B	572	NAG	C3-C4-C5	-2.02	106.67	110.20
3	A	571	GAL	C3-C4-C5	-2.01	106.69	110.20
3	A	571	GAL	C1-C2-C3	2.10	113.55	110.43
3	A	582	NAG	O3-C3-C2	2.14	113.34	109.11
3	A	602	NAG	O5-C5-C6	2.14	111.98	107.35
3	B	591	GAL	O5-C5-C4	2.15	113.72	109.68
3	A	592	NAG	C1-O5-C5	2.19	115.02	112.25
3	A	591	GAL	C4-C3-C2	2.21	114.92	110.79

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	591	GAL	C1-C2-C3	2.32	113.88	110.43
3	A	551	GAL	O4-C4-C5	2.34	115.43	109.24
3	B	592	NAG	O7-C7-N2	2.34	126.63	121.86
3	B	602	NAG	O3-C3-C2	2.34	113.76	109.11
3	B	551	GAL	O5-C5-C4	2.39	114.17	109.68
3	A	602	NAG	O7-C7-N2	2.42	126.80	121.86
3	A	551	GAL	C1-O5-C5	2.46	118.02	113.47
3	A	572	NAG	C3-C2-N2	2.50	116.55	110.56
3	B	572	NAG	C6-C5-C4	2.60	119.43	113.02
3	B	572	NAG	O4-C4-C5	2.63	116.22	109.24
3	A	571	GAL	O5-C5-C4	2.72	114.79	109.68
3	A	591	GAL	O5-C5-C6	2.80	113.43	106.36
3	A	602	NAG	O3-C3-C2	2.94	114.93	109.11
3	A	581	GAL	C1-C2-C3	3.09	115.03	110.43
3	A	581	GAL	O2-C2-C1	3.17	116.80	109.82
3	A	602	NAG	C2-N2-C7	3.17	127.12	123.04
3	B	602	NAG	O5-C5-C6	3.20	114.28	107.35
3	A	551	GAL	O5-C5-C4	3.28	115.84	109.68
3	A	592	NAG	O5-C5-C6	3.30	114.50	107.35
3	A	572	NAG	C1-O5-C5	3.36	116.51	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 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	$\# Z  > 2$
4	NDG	A	562[A]	-	15,15,15	0.43	0	17,21,21	2.60	3 (17%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	A	563[B]	-	15,15,15	0.39	0	17,21,21	1.19	1 (5%)
4	NDG	B	562	-	15,15,15	0.89	0	17,21,21	1.53	3 (17%)

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
4	NDG	A	562[A]	-	-	0/6/26/26	0/1/1/1
5	NAG	A	563[B]	-	-	0/6/26/26	0/1/1/1
4	NDG	B	562	-	-	0/6/26/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	562[A]	NDG	O1L-C1-O	-8.23	87.74	110.25
4	B	562	NDG	O1L-C1-O	2.28	116.49	110.25
4	B	562	NDG	O3-C3-C2	2.30	114.34	109.66
4	A	562[A]	NDG	O3-C3-C2	3.02	115.82	109.66
5	A	563[B]	NAG	O3-C3-C2	3.40	116.59	109.66
4	B	562	NDG	O-C5-C6	3.54	115.30	106.36
4	A	562[A]	NDG	C1-O-C5	4.80	122.36	113.47

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 ⓘ

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 [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, 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	400/401 (99%)	0.10	5 (1%) 79 86	8, 15, 24, 37	2 (0%)
1	B	399/401 (99%)	0.17	3 (0%) 87 92	9, 17, 25, 34	3 (0%)
All	All	799/802 (99%)	0.14	8 (1%) 84 89	8, 16, 25, 37	5 (0%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	401	LEU	3.6
1	B	2	VAL	3.5
1	B	259	GLY	2.3
1	A	396	ALA	2.3
1	A	328	TYR	2.2
1	A	230	VAL	2.2
1	B	258	ASN	2.1
1	A	2	VAL	2.1

### 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](#)

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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GAL	B	591	12/12	0.89	0.19	6.10	24,29,32,36	0
3	NAG	A	582	14/15	0.91	0.14	4.28	13,19,25,26	0
3	GAL	B	551	12/12	0.92	0.14	3.06	20,24,27,29	0
3	GAL	B	601	12/12	0.91	0.17	1.55	22,29,31,35	0
3	GAL	B	571	12/12	0.87	0.17	1.20	21,25,27,31	0
3	NAG	B	582	14/15	0.82	0.15	1.07	21,25,29,31	0
3	GAL	A	551	12/12	0.90	0.13	1.01	19,22,26,33	0
3	GAL	A	571	12/12	0.92	0.12	0.41	15,21,25,31	0
3	NAG	A	552	14/15	0.95	0.09	0.00	13,15,17,19	0
3	NAG	B	552	14/15	0.94	0.10	-0.14	14,17,18,19	0
3	NAG	A	592	14/15	0.93	0.11	-0.22	16,20,26,29	0
3	NAG	B	602	14/15	0.96	0.11	-0.26	17,19,22,27	0
3	NAG	A	572	14/15	0.96	0.09	-0.45	10,12,17,24	0
3	NAG	B	572	14/15	0.93	0.09	-0.57	15,19,23,27	0
3	NAG	B	592	14/15	0.95	0.09	-0.99	15,21,24,24	0
3	NAG	A	602	14/15	0.95	0.08	-1.68	13,15,18,24	0
3	GAL	A	591	12/12	0.82	0.17	-	22,27,30,32	12
3	GAL	B	581	12/12	0.88	0.16	-	29,37,42,46	0
3	GAL	A	581	12/12	0.73	0.26	-	28,35,42,43	0
3	GAL	A	601	12/12	0.90	0.14	-	18,22,24,27	0

## 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	A	502	1/1	0.98	0.19	5.96	29,29,29,29	0
2	CA	B	502	1/1	0.98	0.20	2.42	30,30,30,30	0
5	NAG	A	563[B]	15/15	0.94	0.11	1.91	16,18,22,24	15
4	NDG	A	562[A]	15/15	0.94	0.11	1.59	16,17,21,23	15
2	CA	A	501	1/1	0.98	0.15	1.17	24,24,24,24	0
2	CA	B	501	1/1	0.97	0.17	0.95	33,33,33,33	0
4	NDG	B	562	15/15	0.95	0.11	0.43	17,18,20,20	0
2	CA	B	500	1/1	1.00	0.03	-2.41	15,15,15,15	0
2	CA	A	500	1/1	1.00	0.02	-4.70	13,13,13,13	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.