



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 11:26 am GMT

PDB ID : 1QOT  
Title : LECTIN UEA-II COMPLEXED WITH FUCOSYLLACTOSE AND FUCOSYLGALACTOSE  
Authors : Loris, R.; De Greve, H.; Dao-Thi, M.-H.; Messens, J.; Imberty, A.; Wyns, L.  
Deposited on : 1999-11-16  
Resolution : 3.00 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 : 1.7.2 (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

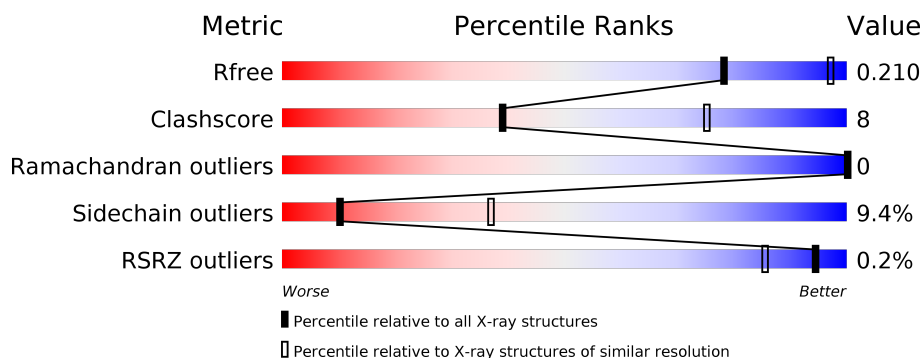
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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. 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	242	 77% 18% . .
1	B	242	 74% 21% . .
1	C	242	 75% 20% . .
1	D	242	 75% 20% . .

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
4	NAG	B	401	-	-	-	X
5	FUC	D	501	-	-	-	X
6	FUC	C	501	-	-	-	X
6	GLC	C	503	X	-	-	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7331 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 CHITIN BINDING LECTIN, UEA-II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	235	Total	C	N	O	S	0	0	1
			1783	1145	285	351	2			
1	B	237	Total	C	N	O	S	0	0	1
			1795	1153	287	353	2			
1	C	237	Total	C	N	O	S	0	0	1
			1795	1153	287	353	2			
1	D	237	Total	C	N	O	S	0	0	1
			1795	1153	287	353	2			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	ASP	ALA	CONFLICT	GB AF190633
A	62	ILE	THR	CONFLICT	GB AF190633
A	106	GLY	SER	CONFLICT	GB AF190633
A	112	SER	ASN	CONFLICT	GB AF190633
A	191	GLY	GLU	CONFLICT	GB AF190633
A	229	VAL	ILE	CONFLICT	GB AF190633
B	25	ASP	ALA	CONFLICT	GB AF190633
B	62	ILE	THR	CONFLICT	GB AF190633
B	106	GLY	SER	CONFLICT	GB AF190633
B	112	SER	ASN	CONFLICT	GB AF190633
B	191	GLY	GLU	CONFLICT	GB AF190633
B	229	VAL	ILE	CONFLICT	GB AF190633
C	25	ASP	ALA	CONFLICT	GB AF190633
C	62	ILE	THR	CONFLICT	GB AF190633
C	106	GLY	SER	CONFLICT	GB AF190633
C	112	SER	ASN	CONFLICT	GB AF190633
C	191	GLY	GLU	CONFLICT	GB AF190633
C	229	VAL	ILE	CONFLICT	GB AF190633
D	25	ASP	ALA	CONFLICT	GB AF190633
D	62	ILE	THR	CONFLICT	GB AF190633
D	106	GLY	SER	CONFLICT	GB AF190633

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	112	SER	ASN	CONFLICT	GB AF190633
D	191	GLY	GLU	CONFLICT	GB AF190633
D	229	VAL	ILE	CONFLICT	GB AF190633

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	1	Total	Mn	0	0
			1	1		
2	D	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca	0	0
			1	1		
3	A	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (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	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	2	Total	C	O	0	0
			22	12	10		
5	B	2	Total	C	O	0	0
			22	12	10		
5	D	2	Total	C	O	0	0
			22	12	10		

- Molecule 6 is a polymer of unknown type called SUGAR (3-MER).

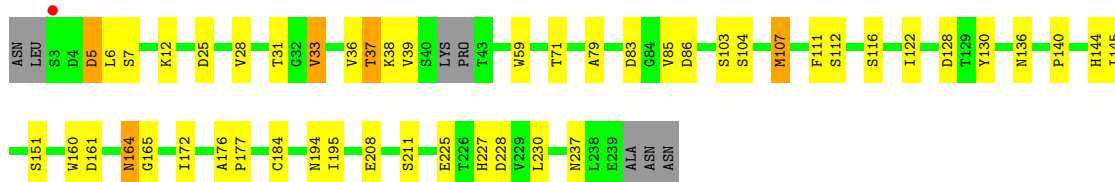
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	3	Total	C	O	0	0
			33	18	15		

### 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.

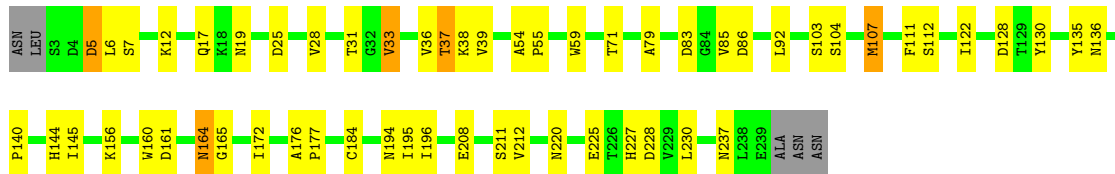
#### • Molecule 1: CHITIN BINDING LECTIN, UEA-II

Chain A: 



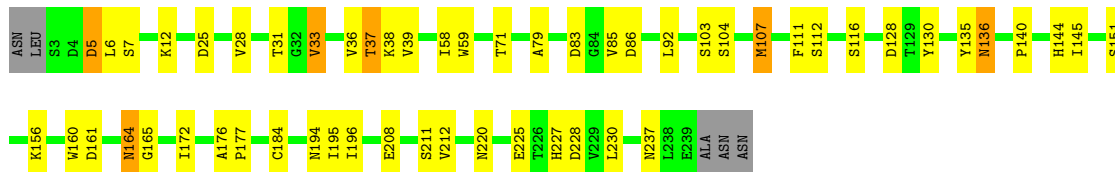
#### • Molecule 1: CHITIN BINDING LECTIN, UEA-II

Chain B: 



#### • Molecule 1: CHITIN BINDING LECTIN, UEA-II

Chain C: 



#### • Molecule 1: CHITIN BINDING LECTIN, UEA-II

Chain D: 







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.77Å 104.77Å 175.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 3.00 19.80 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.9 (30.00-3.00) 99.1 (19.80-2.99)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.04 (at 2.98Å)	Xtriage
Refinement program	X-PLOR 3.8	Depositor
R, $R_{free}$	0.192 , 0.215 0.188 , 0.210	Depositor DCC
$R_{free}$ test set	1786 reflections (8.59%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 30.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7331	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 5.78% 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: MN, NAG, CA, GLC, GAL, FUC

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	0.43	0/1827	0.69	0/2489
1	B	0.45	0/1841	0.70	0/2511
1	C	0.42	0/1841	0.68	0/2511
1	D	0.43	0/1841	0.69	0/2511
All	All	0.43	0/7350	0.69	0/10022

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
6	C	1	0

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	C	503	GLC	C1

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	1783	0	1721	24	0
1	B	1795	0	1731	31	0
1	C	1795	0	1731	32	0
1	D	1795	0	1731	33	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	14	0	13	0	0
4	B	14	0	13	0	0
4	C	14	0	13	0	0
4	D	14	0	13	0	0
5	A	22	0	21	0	0
5	B	22	0	21	4	0
5	D	22	0	21	5	0
6	C	33	0	30	8	0
All	All	7331	0	7059	122	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 8.

All (122) 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:C:79:ALA:H	1:C:164:ASN:HD21	1.29	0.81
1:D:79:ALA:H	1:D:164:ASN:HD21	1.29	0.81
1:B:79:ALA:H	1:B:164:ASN:HD21	1.30	0.80
1:A:79:ALA:H	1:A:164:ASN:HD21	1.30	0.78
1:C:220:ASN:HB3	6:C:502:GAL:H61	1.68	0.75
1:B:161:ASP:H	1:B:194:ASN:HD21	1.35	0.75
1:C:161:ASP:H	1:C:194:ASN:HD21	1.33	0.74
1:A:161:ASP:H	1:A:194:ASN:HD21	1.36	0.74
1:D:161:ASP:H	1:D:194:ASN:HD21	1.34	0.74
1:C:220:ASN:CB	6:C:502:GAL:H61	2.20	0.71
1:B:135:TYR:HB3	5:B:501:FUC:H5	1.73	0.69
6:C:501:FUC:H61	6:C:503:GLC:C6	2.29	0.63
6:C:501:FUC:H61	6:C:503:GLC:H61	1.79	0.63
1:D:104:SER:HB3	5:D:501:FUC:H2	1.79	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:104:SER:HB3	6:C:501:FUC:H2	1.79	0.62
1:A:107:MET:HG2	1:A:111:PHE:HB2	1.82	0.62
1:C:79:ALA:H	1:C:164:ASN:ND2	1.98	0.61
1:D:41:LYS:CB	1:D:42:PRO:HD3	2.30	0.61
1:C:107:MET:HG2	1:C:111:PHE:HB2	1.82	0.61
1:D:107:MET:HG2	1:D:111:PHE:HB2	1.82	0.61
1:B:107:MET:HG2	1:B:111:PHE:HB2	1.81	0.61
1:B:85:VAL:HG12	1:B:130:TYR:HB2	1.84	0.60
1:B:79:ALA:H	1:B:164:ASN:ND2	1.98	0.60
1:C:59:TRP:O	1:C:208:GLU:HG2	2.02	0.59
1:D:85:VAL:HG12	1:D:130:TYR:HB2	1.84	0.59
1:D:79:ALA:H	1:D:164:ASN:ND2	1.98	0.59
1:A:85:VAL:HG12	1:A:130:TYR:HB2	1.84	0.58
1:B:37:THR:HG21	1:B:227:HIS:HD2	1.68	0.58
1:C:37:THR:HG21	1:C:227:HIS:HD2	1.68	0.58
1:B:59:TRP:O	1:B:208:GLU:HG2	2.03	0.58
1:A:37:THR:HG21	1:A:227:HIS:HD2	1.69	0.58
1:D:59:TRP:O	1:D:208:GLU:HG2	2.03	0.58
1:D:37:THR:HG21	1:D:227:HIS:HD2	1.68	0.57
1:A:59:TRP:O	1:A:208:GLU:HG2	2.04	0.57
1:A:79:ALA:H	1:A:164:ASN:ND2	1.99	0.57
1:C:85:VAL:HG12	1:C:130:TYR:HB2	1.85	0.56
1:B:220:ASN:HB3	5:B:502:GAL:H61	1.88	0.54
1:D:220:ASN:HB3	5:D:502:GAL:H61	1.90	0.53
1:B:79:ALA:N	1:B:164:ASN:HD21	2.03	0.53
1:B:6:LEU:HG	1:B:7:SER:N	2.25	0.52
1:C:6:LEU:HG	1:C:7:SER:N	2.24	0.52
1:A:6:LEU:HG	1:A:7:SER:N	2.25	0.52
1:A:164:ASN:HD22	1:A:165:GLY:N	2.08	0.51
1:D:79:ALA:N	1:D:164:ASN:HD21	2.03	0.51
1:D:6:LEU:HG	1:D:7:SER:N	2.25	0.51
1:B:164:ASN:HD22	1:B:165:GLY:N	2.08	0.51
1:B:104:SER:HB3	5:B:501:FUC:H2	1.93	0.51
1:C:104:SER:H	1:C:107:MET:HE2	1.76	0.50
1:C:79:ALA:N	1:C:164:ASN:HD21	2.02	0.50
1:D:161:ASP:N	1:D:194:ASN:HD21	2.06	0.50
1:D:164:ASN:HD22	1:D:165:GLY:N	2.09	0.50
1:C:161:ASP:N	1:C:194:ASN:HD21	2.06	0.50
1:B:5:ASP:H	1:B:237:ASN:HD22	1.59	0.50
1:A:79:ALA:N	1:A:164:ASN:HD21	2.03	0.49
1:A:71:THR:HG22	1:A:172:ILE:HB	1.94	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:25:ASP:HB3	1:D:38:LYS:HB2	1.94	0.49
1:B:161:ASP:N	1:B:194:ASN:HD21	2.06	0.49
1:C:164:ASN:HD22	1:C:165:GLY:N	2.10	0.49
1:C:25:ASP:HB3	1:C:38:LYS:HB2	1.95	0.49
1:D:5:ASP:H	1:D:237:ASN:HD22	1.61	0.49
1:A:104:SER:H	1:A:107:MET:HE2	1.78	0.49
1:A:25:ASP:HB3	1:A:38:LYS:HB2	1.94	0.49
1:B:71:THR:HG22	1:B:172:ILE:HB	1.95	0.49
1:C:71:THR:HG22	1:C:172:ILE:HB	1.94	0.49
1:A:5:ASP:H	1:A:237:ASN:HD22	1.60	0.48
1:D:135:TYR:HD2	5:D:501:FUC:H5	1.77	0.48
1:D:71:THR:HG22	1:D:172:ILE:HB	1.96	0.48
1:B:25:ASP:HB3	1:B:38:LYS:HB2	1.94	0.48
1:C:5:ASP:H	1:C:237:ASN:HD22	1.60	0.48
1:B:104:SER:H	1:B:107:MET:HE2	1.79	0.48
1:A:161:ASP:N	1:A:194:ASN:HD21	2.08	0.47
1:C:85:VAL:O	1:C:225:GLU:HG2	2.14	0.47
1:B:195:ILE:HB	1:D:195:ILE:HB	1.95	0.47
1:C:161:ASP:H	1:C:194:ASN:ND2	2.08	0.46
1:A:85:VAL:O	1:A:225:GLU:HG2	2.17	0.45
1:D:85:VAL:O	1:D:225:GLU:HG2	2.16	0.45
1:D:104:SER:HB3	5:D:501:FUC:C2	2.47	0.45
1:A:161:ASP:H	1:A:194:ASN:ND2	2.10	0.44
1:B:85:VAL:O	1:B:225:GLU:HG2	2.17	0.44
1:C:128:ASP:HB3	1:C:144:HIS:CE1	2.53	0.44
1:A:128:ASP:HB3	1:A:144:HIS:CE1	2.53	0.44
1:C:220:ASN:HB2	6:C:502:GAL:H61	1.99	0.44
1:D:220:ASN:CB	5:D:502:GAL:H61	2.48	0.44
1:B:220:ASN:CB	5:B:502:GAL:H61	2.48	0.43
1:D:140:PRO:HD2	1:D:144:HIS:CE1	2.54	0.43
1:D:145:ILE:HG13	1:D:160:TRP:HB2	2.00	0.43
6:C:501:FUC:O5	6:C:503:GLC:C6	2.67	0.43
1:B:145:ILE:HG13	1:B:160:TRP:HB2	2.01	0.43
1:B:54:ALA:HA	1:B:55:PRO:HD3	1.91	0.43
1:A:195:ILE:HB	1:C:195:ILE:HB	2.01	0.42
1:B:156:LYS:HG2	1:B:196:ILE:HD11	2.00	0.42
1:C:145:ILE:HG13	1:C:160:TRP:HB2	2.01	0.42
1:D:156:LYS:HG2	1:D:196:ILE:HD11	2.01	0.42
1:B:128:ASP:HB3	1:B:144:HIS:CE1	2.54	0.42
1:D:176:ALA:N	1:D:177:PRO:HD2	2.35	0.42
1:B:161:ASP:H	1:B:194:ASN:ND2	2.09	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:33:VAL:CG2	1:B:228:ASP:HB3	2.49	0.42
1:D:128:ASP:HB3	1:D:144:HIS:CE1	2.54	0.42
1:B:92:LEU:HG	1:B:212:VAL:HG12	2.02	0.42
1:C:176:ALA:N	1:C:177:PRO:HD2	2.35	0.42
1:B:140:PRO:HD2	1:B:144:HIS:CE1	2.55	0.42
1:C:92:LEU:HG	1:C:212:VAL:HG12	2.01	0.42
6:C:501:FUC:O5	6:C:503:GLC:H62	2.20	0.42
1:D:33:VAL:CG2	1:D:228:ASP:HB3	2.50	0.42
1:A:140:PRO:HD2	1:A:144:HIS:CE1	2.54	0.41
1:D:54:ALA:HA	1:D:55:PRO:HD3	1.89	0.41
1:A:145:ILE:HG13	1:A:160:TRP:HB2	2.02	0.41
1:A:33:VAL:CG2	1:A:228:ASP:HB3	2.50	0.41
1:C:135:TYR:C	1:C:136:ASN:HD22	2.24	0.41
1:C:140:PRO:HD2	1:C:144:HIS:CE1	2.55	0.41
1:D:161:ASP:H	1:D:194:ASN:ND2	2.09	0.41
1:D:58:ILE:HA	1:D:58:ILE:HD12	1.91	0.41
1:D:85:VAL:HA	1:D:86:ASP:HA	1.96	0.41
1:A:176:ALA:N	1:A:177:PRO:HD2	2.36	0.41
1:C:156:LYS:HG2	1:C:196:ILE:HD11	2.02	0.41
1:C:58:ILE:HD12	1:C:58:ILE:HA	1.92	0.41
1:C:116:SER:HA	1:C:151:SER:HB2	2.02	0.40
1:A:116:SER:HA	1:A:151:SER:HB2	2.04	0.40
1:B:176:ALA:N	1:B:177:PRO:HD2	2.35	0.40
1:B:17:GLN:HG2	1:B:19:ASN:OD1	2.22	0.40
1:C:33:VAL:CG2	1:C:228:ASP:HB3	2.51	0.40
1:D:104:SER:H	1:D:107:MET:HE2	1.86	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	231/242 (96%)	223 (96%)	8 (4%)	0	100	100
1	B	235/242 (97%)	227 (97%)	8 (3%)	0	100	100
1	C	235/242 (97%)	227 (97%)	8 (3%)	0	100	100
1	D	235/242 (97%)	225 (96%)	10 (4%)	0	100	100
All	All	936/968 (97%)	902 (96%)	34 (4%)	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	197/205 (96%)	178 (90%)	19 (10%)	10	36
1	B	198/205 (97%)	179 (90%)	19 (10%)	10	36
1	C	198/205 (97%)	180 (91%)	18 (9%)	11	39
1	D	198/205 (97%)	180 (91%)	18 (9%)	11	39
All	All	791/820 (96%)	717 (91%)	74 (9%)	10	37

All (74) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	ASP
1	A	12	LYS
1	A	28	VAL
1	A	31	THR
1	A	33	VAL
1	A	36	VAL
1	A	37	THR
1	A	39	VAL
1	A	83	ASP
1	A	86	ASP
1	A	103	SER
1	A	107	MET
1	A	112	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	122	ILE
1	A	136	ASN
1	A	164	ASN
1	A	184	CYS
1	A	211	SER
1	A	230	LEU
1	B	5	ASP
1	B	12	LYS
1	B	28	VAL
1	B	31	THR
1	B	33	VAL
1	B	36	VAL
1	B	37	THR
1	B	39	VAL
1	B	83	ASP
1	B	86	ASP
1	B	103	SER
1	B	107	MET
1	B	112	SER
1	B	122	ILE
1	B	136	ASN
1	B	164	ASN
1	B	184	CYS
1	B	211	SER
1	B	230	LEU
1	C	5	ASP
1	C	12	LYS
1	C	28	VAL
1	C	31	THR
1	C	33	VAL
1	C	36	VAL
1	C	37	THR
1	C	39	VAL
1	C	83	ASP
1	C	86	ASP
1	C	103	SER
1	C	107	MET
1	C	112	SER
1	C	136	ASN
1	C	164	ASN
1	C	184	CYS
1	C	211	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	C	230	LEU
1	D	5	ASP
1	D	12	LYS
1	D	28	VAL
1	D	31	THR
1	D	33	VAL
1	D	36	VAL
1	D	37	THR
1	D	39	VAL
1	D	83	ASP
1	D	86	ASP
1	D	103	SER
1	D	107	MET
1	D	112	SER
1	D	136	ASN
1	D	164	ASN
1	D	184	CYS
1	D	211	SER
1	D	230	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (27) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	35	GLN
1	A	57	GLN
1	A	150	ASN
1	A	164	ASN
1	A	194	ASN
1	A	220	ASN
1	A	227	HIS
1	A	237	ASN
1	B	35	GLN
1	B	150	ASN
1	B	164	ASN
1	B	194	ASN
1	B	220	ASN
1	B	227	HIS
1	B	237	ASN
1	C	35	GLN
1	C	150	ASN
1	C	164	ASN
1	C	194	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	227	HIS
1	C	237	ASN
1	D	35	GLN
1	D	150	ASN
1	D	164	ASN
1	D	194	ASN
1	D	227	HIS
1	D	237	ASN

### 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](#)

9 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	$\# Z  > 2$
5	FUC	A	501	5	9,10,11	0.49	0	13,14,16	0.44	0
5	GAL	A	502	5	12,12,12	0.21	0	17,17,17	0.62	0
5	FUC	B	501	5	9,10,11	0.47	0	13,14,16	0.33	0
5	GAL	B	502	5	12,12,12	0.38	0	17,17,17	0.56	0
6	FUC	C	501	6	9,10,11	0.53	0	13,14,16	0.40	0
6	GAL	C	502	6	11,11,12	0.37	0	13,15,17	0.67	0
6	GLC	C	503	6	12,12,12	0.37	0	17,17,17	0.48	0
5	FUC	D	501	5	9,10,11	0.49	0	13,14,16	0.36	0
5	GAL	D	502	5	12,12,12	0.41	0	17,17,17	0.57	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
5	FUC	A	501	5	-	0/0/17/20	0/1/1/1
5	GAL	A	502	5	-	0/2/22/22	0/1/1/1
5	FUC	B	501	5	-	0/0/17/20	0/1/1/1
5	GAL	B	502	5	-	0/2/22/22	0/1/1/1
6	FUC	C	501	6	-	0/0/17/20	0/1/1/1
6	GAL	C	502	6	-	0/2/19/22	0/1/1/1
6	GLC	C	503	6	1/1/5/5	0/2/22/22	0/1/1/1
5	FUC	D	501	5	-	0/0/17/20	0/1/1/1
5	GAL	D	502	5	-	0/2/22/22	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	C	503	GLC	C1

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	501	FUC	2	0
5	B	502	GAL	2	0
6	C	501	FUC	5	0
6	C	502	GAL	3	0
6	C	503	GLC	4	0
5	D	501	FUC	3	0
5	D	502	GAL	2	0

## 5.6 Ligand geometry

Of 12 ligands modelled in this entry, 8 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$
4	NAG	A	401	1	14,14,15	0.72	0	15,19,21	1.46	3 (20%)
4	NAG	B	401	1	14,14,15	0.65	0	15,19,21	1.38	2 (13%)
4	NAG	C	401	1	14,14,15	0.82	0	15,19,21	1.31	1 (6%)
4	NAG	D	401	1	14,14,15	0.74	0	15,19,21	1.36	2 (13%)

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	NAG	A	401	1	-	0/6/23/26	0/1/1/1
4	NAG	B	401	1	-	0/6/23/26	0/1/1/1
4	NAG	C	401	1	-	0/6/23/26	0/1/1/1
4	NAG	D	401	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	401	NAG	O5-C1-C2	-2.47	108.03	111.47
4	D	401	NAG	O5-C1-C2	-2.45	108.07	111.47
4	D	401	NAG	C3-C4-C5	2.54	114.69	110.22
4	A	401	NAG	C3-C4-C5	2.54	114.69	110.22
4	B	401	NAG	C3-C4-C5	2.56	114.72	110.22
4	C	401	NAG	C3-C4-C5	2.60	114.79	110.22
4	A	401	NAG	C1-O5-C5	2.95	116.24	112.17
4	B	401	NAG	C1-O5-C5	3.26	116.66	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

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	235/242 (97%)	-0.65	1 (0%) 92 77	10, 22, 47, 66	0
1	B	237/242 (97%)	-0.67	0 100 100	10, 22, 49, 66	0
1	C	237/242 (97%)	-0.67	0 100 100	10, 22, 49, 66	0
1	D	237/242 (97%)	-0.66	1 (0%) 92 77	10, 22, 50, 66	0
All	All	946/968 (97%)	-0.66	2 (0%) 94 85	10, 22, 49, 66	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	3	SER	3.0
1	A	3	SER	2.2

### 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(Å <sup>2</sup> )	Q<0.9
5	FUC	D	501	10/11	0.90	0.31	3.22	53,53,54,54	0
6	FUC	C	501	10/11	0.83	0.29	2.41	53,54,54,54	0
5	FUC	B	501	10/11	0.88	0.25	1.80	53,53,54,54	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	GAL	C	502	11/12	0.90	0.24	1.66	52,53,53,54	0
5	GAL	B	502	12/12	0.93	0.23	1.05	52,53,54,54	0
5	GAL	A	502	12/12	0.95	0.23	1.04	52,53,53,54	0
5	GAL	D	502	12/12	0.90	0.23	0.59	52,53,53,54	0
5	FUC	A	501	10/11	0.92	0.19	0.51	53,54,54,54	0
6	GLC	C	503	12/12	0.86	0.37	-	54,55,56,56	0

## 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. 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
4	NAG	B	401	14/15	0.74	0.31	2.70	60,65,70,71	0
4	NAG	A	401	14/15	0.75	0.32	1.48	59,63,68,69	0
3	CA	B	304	1/1	0.97	0.14	-0.44	22,22,22,22	0
2	MN	A	301	1/1	0.98	0.11	-0.81	10,10,10,10	0
3	CA	A	302	1/1	0.97	0.11	-1.44	23,23,23,23	0
2	MN	D	307	1/1	0.99	0.10	-1.50	9,9,9,9	0
3	CA	C	306	1/1	0.95	0.12	-1.60	22,22,22,22	0
2	MN	C	305	1/1	0.98	0.11	-1.87	10,10,10,10	0
2	MN	B	303	1/1	0.97	0.09	-2.35	7,7,7,7	0
3	CA	D	308	1/1	0.83	0.10	-2.69	22,22,22,22	0
4	NAG	C	401	14/15	0.77	0.35	-	59,63,68,69	0
4	NAG	D	401	14/15	0.77	0.36	-	59,63,69,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.