



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 11:01 PM GMT

PDB ID : 5F8R  
Title : Blood group antigen binding adhesin BabA of Helicobacter pylori strain S831  
in complex with blood group H Lewis b hexasaccharide  
Authors : Moonens, K.; Gideonsson, P.; Subedi, S.; Romao, E.; Oscarson, S.; Muylder-  
mans, S.; Boren, T.; Remaut, H.  
Deposited on : 2015-12-09  
Resolution : 2.44 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

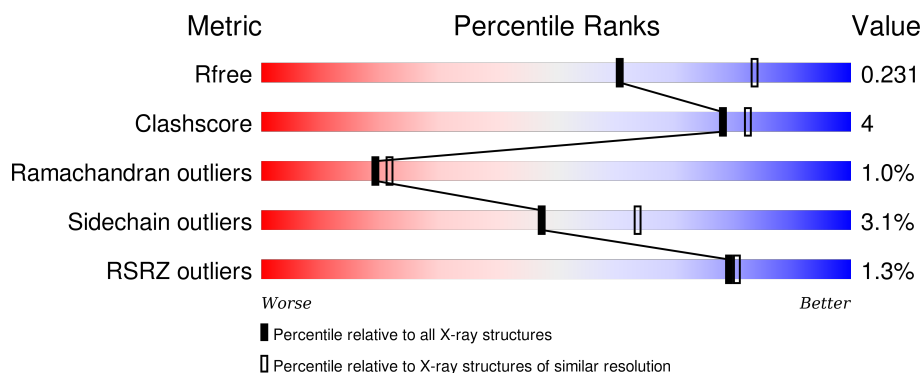
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.44 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	1003 (2.46-2.42)
Clashscore	102246	1071 (2.46-2.42)
Ramachandran outliers	100387	1065 (2.46-2.42)
Sidechain outliers	100360	1065 (2.46-2.42)
RSRZ outliers	91569	1005 (2.46-2.42)

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	466	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>8%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	466	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>8%</div> <div>•</div> <div>10%</div> </div> </div>
2	C	120	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>•</div> <div>6%</div> </div> </div>
2	D	120	<div> <div>%</div> <div> <div></div> <div>93%</div> <div>•</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8224 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 Adhesin binding fucosylated histo-blood group antigen,Adhesin,Adhesin binding fucosylated histo-blood group antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	417	Total	C	N	O	S	0	1	0
			3111	1920	536	642	13			
1	B	421	Total	C	N	O	S	0	0	0
			3139	1939	539	648	13			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	ALA	-	expression tag	UNP O52269
A	4	SER	-	expression tag	UNP O52269
A	5	TRP	-	expression tag	UNP O52269
A	6	SER	-	expression tag	UNP O52269
A	7	HIS	-	expression tag	UNP O52269
A	8	PRO	-	expression tag	UNP O52269
A	9	GLN	-	expression tag	UNP O52269
A	10	PHE	-	expression tag	UNP O52269
A	11	GLU	-	expression tag	UNP O52269
A	12	LYS	-	expression tag	UNP O52269
A	13	SER	-	expression tag	UNP O52269
A	14	GLY	-	expression tag	UNP O52269
A	15	GLY	-	expression tag	UNP O52269
A	16	GLY	-	expression tag	UNP O52269
A	17	GLY	-	expression tag	UNP O52269
A	18	GLY	-	expression tag	UNP O52269
A	19	LEU	-	expression tag	UNP O52269
A	20	VAL	-	expression tag	UNP O52269
A	21	PRO	-	expression tag	UNP O52269
A	22	ARG	-	expression tag	UNP O52269
A	23	GLY	-	expression tag	UNP O52269
A	24	SER	-	expression tag	UNP O52269
A	461	GLY	-	expression tag	UNP O52269
A	462	SER	-	expression tag	UNP O52269

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	463	HIS	-	expression tag	UNP O52269
A	464	HIS	-	expression tag	UNP O52269
A	465	HIS	-	expression tag	UNP O52269
A	466	HIS	-	expression tag	UNP O52269
A	467	HIS	-	expression tag	UNP O52269
A	468	HIS	-	expression tag	UNP O52269
B	3	ALA	-	expression tag	UNP O52269
B	4	SER	-	expression tag	UNP O52269
B	5	TRP	-	expression tag	UNP O52269
B	6	SER	-	expression tag	UNP O52269
B	7	HIS	-	expression tag	UNP O52269
B	8	PRO	-	expression tag	UNP O52269
B	9	GLN	-	expression tag	UNP O52269
B	10	PHE	-	expression tag	UNP O52269
B	11	GLU	-	expression tag	UNP O52269
B	12	LYS	-	expression tag	UNP O52269
B	13	SER	-	expression tag	UNP O52269
B	14	GLY	-	expression tag	UNP O52269
B	15	GLY	-	expression tag	UNP O52269
B	16	GLY	-	expression tag	UNP O52269
B	17	GLY	-	expression tag	UNP O52269
B	18	GLY	-	expression tag	UNP O52269
B	19	LEU	-	expression tag	UNP O52269
B	20	VAL	-	expression tag	UNP O52269
B	21	PRO	-	expression tag	UNP O52269
B	22	ARG	-	expression tag	UNP O52269
B	23	GLY	-	expression tag	UNP O52269
B	24	SER	-	expression tag	UNP O52269
B	461	GLY	-	expression tag	UNP O52269
B	462	SER	-	expression tag	UNP O52269
B	463	HIS	-	expression tag	UNP O52269
B	464	HIS	-	expression tag	UNP O52269
B	465	HIS	-	expression tag	UNP O52269
B	466	HIS	-	expression tag	UNP O52269
B	467	HIS	-	expression tag	UNP O52269
B	468	HIS	-	expression tag	UNP O52269

- Molecule 2 is a protein called Nanobody Nb-ER19.

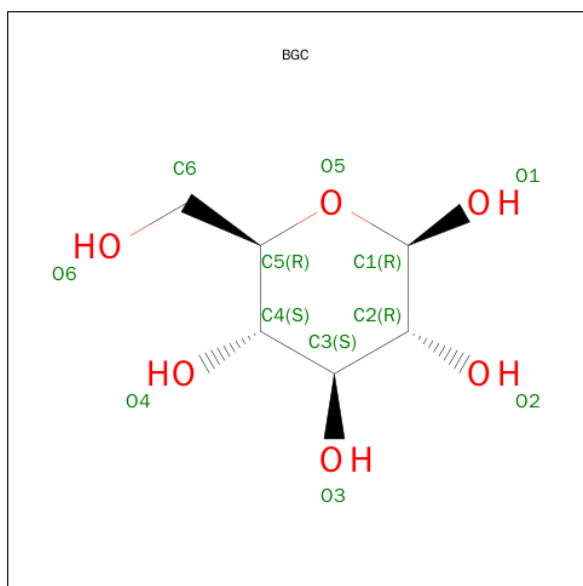
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	113	Total	C	N	O	S	0	0	0
			863	540	155	163	5			

*Continued on next page...*

*Continued from previous page...*

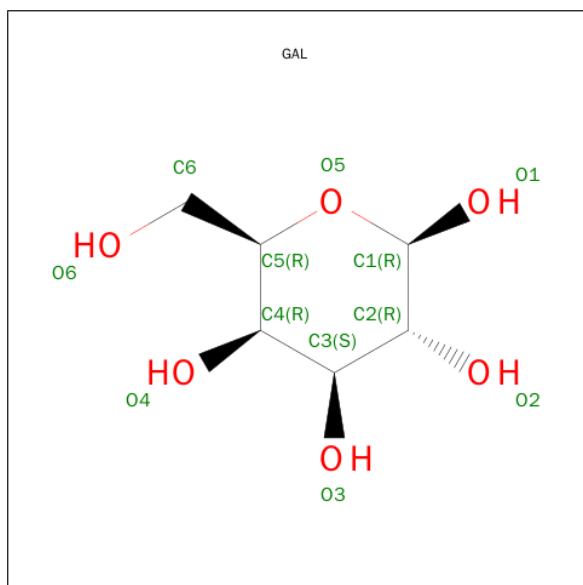
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	114	Total	C	N	O	S	0	0	0
			873	546	158	164	5			

- Molecule 3 is BETA-D-GLUCOSE (three-letter code: BGC) (formula:  $C_6H_{12}O_6$ ).



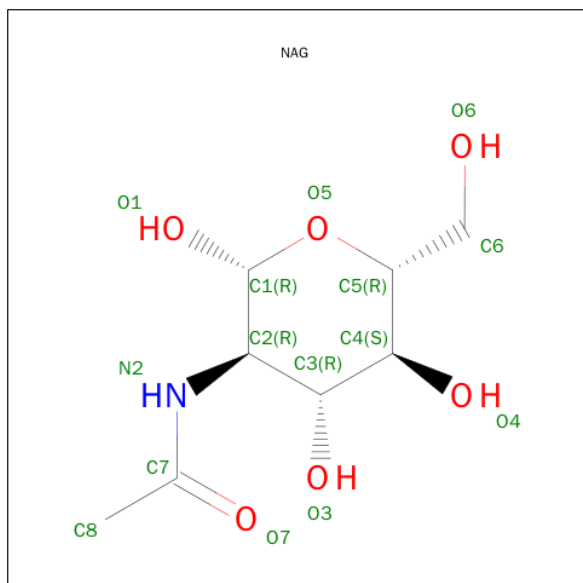
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			12	6	6		
3	B	1	Total	C	O	0	0
			12	6	6		

- Molecule 4 is BETA-D-GALACTOSE (three-letter code: GAL) (formula:  $C_6H_{12}O_6$ ).



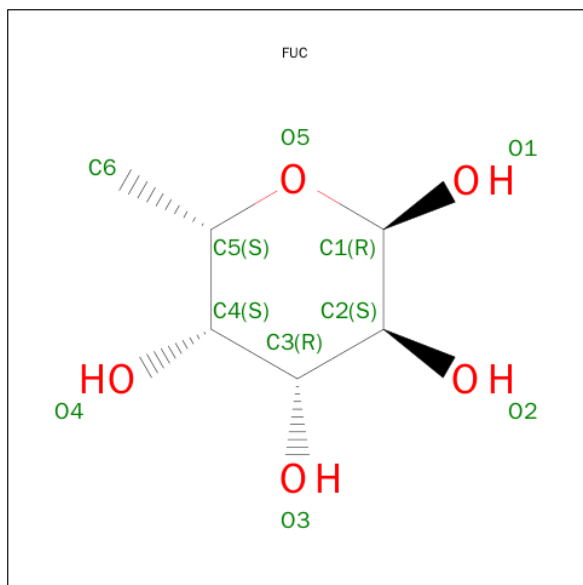
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	B	1	Total	C	O	0	0
			11	6	5		
4	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is 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	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			10	6	4		
6	A	1	Total	C	O	0	0
			10	6	4		
6	B	1	Total	C	O	0	0
			10	6	4		
6	B	1	Total	C	O	0	0
			10	6	4		

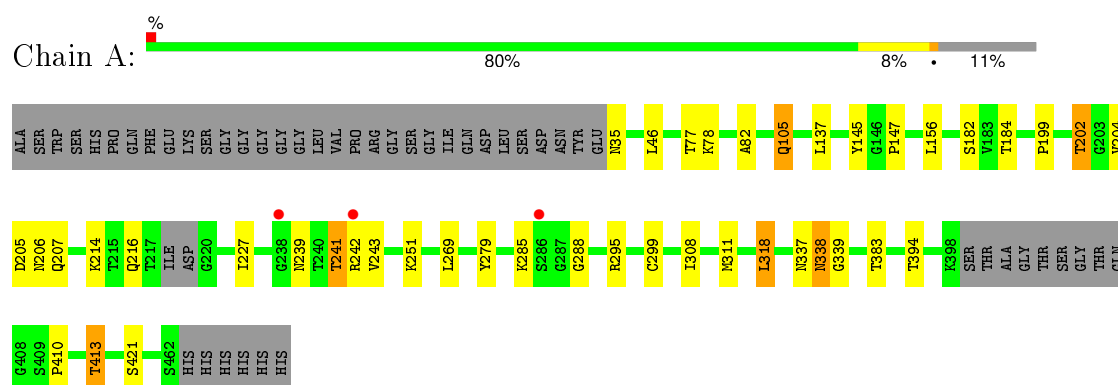
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	38	Total	O	0	0
			38	38		
7	B	28	Total	O	0	0
			28	28		
7	C	12	Total	O	0	0
			12	12		
7	D	24	Total	O	0	0
			24	24		

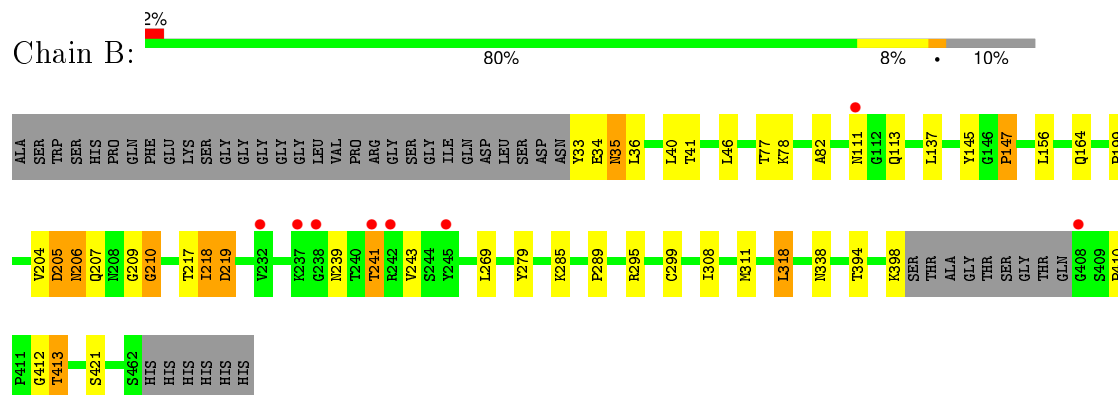
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ( $RSRZ > 2$ ). 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.

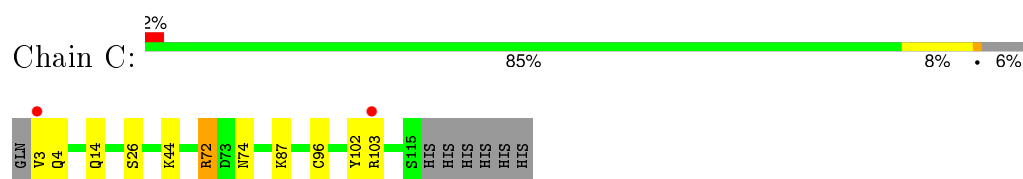
- Molecule 1: Adhesin binding fucosylated histo-blood group antigen,Adhesin,Adhesin binding fucosylated histo-blood group antigen



- Molecule 1: Adhesin binding fucosylated histo-blood group antigen,Adhesin,Adhesin binding fucosylated histo-blood group antigen

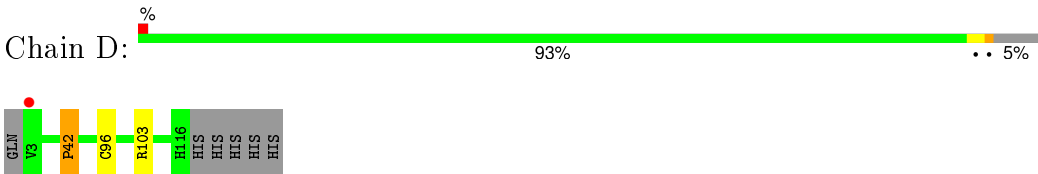


- Molecule 2: Nanobody Nb-ER19



- Molecule 2: Nanobody Nb-ER19





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.79Å 131.91Å 125.39Å 90.00° 94.88° 90.00°	Depositor
Resolution (Å)	48.40 – 2.44 48.36 – 2.44	Depositor EDS
% Data completeness (in resolution range)	99.4 (48.40-2.44) 99.5 (48.36-2.44)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.34 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.189 , 0.219 0.203 , 0.231	Depositor DCC
$R_{free}$ test set	3049 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	61.0	Xtriage
Anisotropy	0.285	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 37.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 60581 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8224	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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: BGC, GAL, NAG, 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	0.81	1/3157 (0.0%)	0.83	3/4292 (0.1%)
1	B	0.76	0/3187	0.81	2/4335 (0.0%)
2	C	0.81	0/882	0.89	2/1195 (0.2%)
2	D	0.94	0/893	0.96	1/1210 (0.1%)
All	All	0.81	1/8119 (0.0%)	0.85	8/11032 (0.1%)

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
1	A	0	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	105	GLN	CG-CD	5.12	1.62	1.51

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	299	CYS	CA-CB-SG	11.40	134.52	114.00
1	A	299	CYS	CB-CA-C	-8.42	93.56	110.40
2	C	72	ARG	NE-CZ-NH1	7.71	124.16	120.30
2	D	42	PRO	C-N-CA	-6.47	108.71	122.30
2	C	72	ARG	NE-CZ-NH2	-6.36	117.12	120.30
1	B	318	LEU	CA-CB-CG	5.99	129.08	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	318	LEU	CA-CB-CG	5.88	128.83	115.30
1	B	299	CYS	CA-CB-SG	5.08	123.15	114.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	242	ARG	Peptide
1	B	33	TYR	Peptide

## 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	3111	0	3038	30	0
1	B	3139	0	3064	22	0
2	C	863	0	845	6	0
2	D	873	0	852	0	0
3	A	12	0	11	0	0
3	B	12	0	11	0	0
4	A	22	0	18	0	0
4	B	22	0	18	0	0
5	A	14	0	11	0	0
5	B	14	0	11	0	0
6	A	20	0	20	0	0
6	B	20	0	20	0	0
7	A	38	0	0	0	0
7	B	28	0	0	0	0
7	C	12	0	0	0	0
7	D	24	0	0	0	0
All	All	8224	0	7919	58	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 4.

All (58) 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:279:TYR:OH	1:A:295:ARG:CZ	1.86	1.22
2:C:3:VAL:N	2:C:26:SER:HG	1.51	1.08
1:A:279:TYR:HH	1:A:295:ARG:NE	1.68	0.90
1:A:279:TYR:HH	1:A:295:ARG:CZ	1.85	0.88
1:A:279:TYR:OH	1:A:295:ARG:NH1	2.09	0.84
1:A:279:TYR:OH	1:A:295:ARG:NE	2.13	0.80
2:C:3:VAL:HG13	2:C:102:TYR:OH	1.86	0.75
1:A:337:ASN:O	1:A:339:GLY:N	2.21	0.74
1:B:410:PRO:O	1:B:413:THR:HG23	1.94	0.67
2:C:3:VAL:HG13	2:C:102:TYR:CZ	2.31	0.65
2:C:72:ARG:HD3	2:C:74:ASN:OD1	2.03	0.59
1:B:241:THR:HG22	1:B:243:VAL:H	1.69	0.57
1:B:205:ASP:O	1:B:206:ASN:HB2	2.05	0.56
1:A:410:PRO:O	1:A:413:THR:CG2	2.55	0.55
1:B:308:ILE:HA	1:B:311:MET:HE3	1.88	0.55
1:B:209:GLY:O	1:B:210:GLY:O	2.25	0.55
1:A:105:GLN:CD	1:A:279:TYR:HD2	2.11	0.55
1:B:218:ILE:HG22	1:B:219:ASP:N	2.22	0.54
1:A:239:ASN:OD1	1:A:241:THR:HG22	2.07	0.53
1:A:199:PRO:HD3	1:A:207:GLN:HG3	1.90	0.53
1:B:199:PRO:HD3	1:B:207:GLN:HG3	1.91	0.52
1:A:410:PRO:O	1:A:413:THR:HG23	2.09	0.52
1:A:308:ILE:HA	1:A:311:MET:HE3	1.91	0.52
1:A:394:THR:HG22	1:A:421:SER:HB2	1.94	0.50
1:A:35:ASN:OD1	1:A:35:ASN:C	2.49	0.50
2:C:3:VAL:N	2:C:26:SER:OG	2.32	0.50
1:B:218:ILE:CG2	1:B:219:ASP:N	2.76	0.49
1:B:35:ASN:C	1:B:35:ASN:OD1	2.51	0.48
1:A:202:THR:HG22	1:A:204:VAL:HB	1.95	0.48
1:A:279:TYR:OH	1:A:295:ARG:CG	2.62	0.48
1:B:156:LEU:C	1:B:156:LEU:HD23	2.34	0.48
1:A:202:THR:HG22	1:A:204:VAL:H	1.79	0.48
1:B:137:LEU:HD23	1:B:147:PRO:HG3	1.97	0.47
1:A:214:LYS:HE3	1:A:227:ILE:HD11	1.97	0.47
1:B:217:THR:O	1:B:217:THR:HG23	2.15	0.47
1:A:156:LEU:C	1:A:156:LEU:HD23	2.35	0.47
1:A:105:GLN:CD	1:A:279:TYR:CD2	2.89	0.46
1:A:156:LEU:HG	1:A:269:LEU:HD12	1.97	0.46
1:A:137:LEU:HD23	1:A:147:PRO:HG3	1.97	0.45
1:B:239:ASN:OD1	1:B:241:THR:HB	2.16	0.45
1:B:111:ASN:HB2	1:B:113:GLN:HG2	1.98	0.45
1:B:394:THR:HG22	1:B:421:SER:HB2	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:337:ASN:O	1:A:338:ASN:C	2.55	0.44
1:B:156:LEU:HG	1:B:269:LEU:HD12	1.99	0.44
1:B:164:GLN:OE1	1:B:412:GLY:HA2	2.18	0.43
1:A:202:THR:CG2	1:A:204:VAL:HB	2.49	0.43
2:C:3:VAL:CG1	2:C:102:TYR:OH	2.61	0.42
1:B:77:THR:O	1:B:78:LYS:C	2.57	0.42
1:B:36:LEU:O	1:B:40:LEU:HG	2.20	0.42
1:B:279:TYR:OH	1:B:295:ARG:HD2	2.20	0.41
1:A:205:ASP:OD1	1:A:205:ASP:N	2.54	0.41
1:A:241:THR:HG23	1:A:243:VAL:H	1.86	0.41
1:A:77:THR:O	1:A:78:LYS:C	2.58	0.41
1:B:204:VAL:HG12	1:B:206:ASN:H	1.84	0.41
1:A:82:ALA:HB1	1:A:318:LEU:HD21	2.03	0.41
1:A:184:THR:OG1	1:A:251:LYS:HD2	2.21	0.40
1:A:206:ASN:N	1:A:206:ASN:OD1	2.55	0.40
1:B:82:ALA:HB1	1:B:318:LEU:HD21	2.03	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	412/466 (88%)	396 (96%)	14 (3%)	2 (0%)	34	41
1	B	417/466 (90%)	397 (95%)	13 (3%)	7 (2%)	11	10
2	C	111/120 (92%)	108 (97%)	2 (2%)	1 (1%)	21	25
2	D	112/120 (93%)	109 (97%)	3 (3%)	0	100	100
All	All	1052/1172 (90%)	1010 (96%)	32 (3%)	10 (1%)	19	22

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	338	ASN
1	B	34	GLU
1	B	206	ASN
1	B	210	GLY
1	B	338	ASN
1	B	289	PRO
1	B	241	THR
2	C	4	GLN
1	B	218	ILE
1	A	288	GLY

### 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	345/382 (90%)	336 (97%)	9 (3%)	54	70
1	B	348/382 (91%)	338 (97%)	10 (3%)	50	66
2	C	92/99 (93%)	87 (95%)	5 (5%)	27	39
2	D	93/99 (94%)	90 (97%)	3 (3%)	46	62
All	All	878/962 (91%)	851 (97%)	27 (3%)	47	64

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	46	LEU
1	A	145	TYR
1	A	182	SER
1	A	202	THR
1	A	216	GLN
1	A	241	THR
1	A	285	LYS
1	A	383	THR
1	A	413	THR
1	B	35	ASN
1	B	41	THR
1	B	46	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	145	TYR
1	B	147	PRO
1	B	205	ASP
1	B	219	ASP
1	B	285	LYS
1	B	398	LYS
1	B	413	THR
2	C	14	GLN
2	C	44	LYS
2	C	87	LYS
2	C	96	CYS
2	C	103	ARG
2	D	42	PRO
2	D	96	CYS
2	D	103	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	105	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

12 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	BGC	A	501	4	12,12,12	0.56	0	17,17,17	1.18	1 (5%)
4	GAL	A	502	3,5	11,11,12	0.55	0	15,15,17	1.66	6 (40%)
5	NAG	A	503	4,6	14,14,15	1.17	2 (14%)	15,19,21	1.67	4 (26%)
4	GAL	A	504	5,6	11,11,12	0.63	0	15,15,17	1.00	1 (6%)
6	FUC	A	505	4	10,10,11	1.03	0	13,14,16	1.38	1 (7%)
6	FUC	A	506	5	10,10,11	0.89	1 (10%)	13,14,16	1.25	1 (7%)
3	BGC	B	501	4	12,12,12	0.62	0	17,17,17	1.76	3 (17%)
4	GAL	B	502	3,5	11,11,12	0.61	0	15,15,17	1.82	4 (26%)
5	NAG	B	503	4,6	14,14,15	0.82	1 (7%)	15,19,21	1.26	4 (26%)
4	GAL	B	504	5,6	11,11,12	0.87	0	15,15,17	1.11	1 (6%)
6	FUC	B	505	4	10,10,11	0.69	0	13,14,16	1.53	1 (7%)
6	FUC	B	506	5	10,10,11	0.58	0	13,14,16	1.82	4 (30%)

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	BGC	A	501	4	-	0/2/22/22	0/1/1/1
4	GAL	A	502	3,5	-	0/2/19/22	0/1/1/1
5	NAG	A	503	4,6	-	0/6/23/26	0/1/1/1
4	GAL	A	504	5,6	-	0/2/19/22	0/1/1/1
6	FUC	A	505	4	-	0/0/17/20	0/1/1/1
6	FUC	A	506	5	-	0/0/17/20	0/1/1/1
3	BGC	B	501	4	-	0/2/22/22	0/1/1/1
4	GAL	B	502	3,5	-	0/2/19/22	0/1/1/1
5	NAG	B	503	4,6	-	0/6/23/26	0/1/1/1
4	GAL	B	504	5,6	-	0/2/19/22	0/1/1/1
6	FUC	B	505	4	-	0/0/17/20	0/1/1/1
6	FUC	B	506	5	-	0/0/17/20	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	503	NAG	O5-C1	-2.91	1.39	1.43
6	A	506	FUC	O5-C1	-2.39	1.39	1.43
5	A	503	NAG	O5-C5	-2.38	1.38	1.43
5	B	503	NAG	O5-C1	-2.15	1.40	1.43

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	505	FUC	O5-C1-C2	-3.47	105.35	110.89
3	B	501	BGC	C1-C2-C3	-3.38	105.14	110.68
5	A	503	NAG	O5-C5-C6	-3.01	100.90	107.34
5	A	503	NAG	C4-C3-C2	-2.99	106.69	111.34
4	B	502	GAL	C3-C4-C5	-2.99	104.90	110.23
4	B	502	GAL	O2-C2-C3	-2.76	104.62	110.19
4	A	502	GAL	O3-C3-C2	-2.38	105.64	110.01
5	A	503	NAG	O7-C7-C8	-2.37	117.70	122.07
6	B	506	FUC	C6-C5-C4	-2.36	108.45	113.02
4	B	504	GAL	O5-C5-C4	-2.24	106.42	110.13
4	A	502	GAL	O5-C5-C4	-2.21	106.48	110.13
4	A	502	GAL	O2-C2-C3	-2.07	106.01	110.19
5	A	503	NAG	O3-C3-C4	-2.07	105.69	110.36
5	B	503	NAG	O7-C7-C8	-2.01	118.36	122.07
5	B	503	NAG	C1-O5-C5	2.03	115.13	112.14
5	B	503	NAG	O7-C7-N2	2.07	126.06	121.84
4	A	504	GAL	C1-O5-C5	2.11	115.25	112.14
4	B	502	GAL	O3-C3-C4	2.16	115.22	110.36
3	A	501	BGC	O5-C5-C6	2.16	111.98	106.38
6	A	506	FUC	O5-C5-C6	2.33	110.46	106.28
4	A	502	GAL	O5-C1-C2	2.38	114.70	110.89
4	A	502	GAL	C1-C2-C3	2.48	112.56	109.55
6	B	506	FUC	O5-C1-C2	2.49	114.88	110.89
5	B	503	NAG	C2-N2-C7	2.63	126.52	123.11
4	A	502	GAL	C1-O5-C5	2.63	116.01	112.14
6	B	506	FUC	C1-C2-C3	2.95	113.12	109.55
6	A	505	FUC	C1-C2-C3	3.06	113.25	109.55
3	B	501	BGC	O5-C5-C6	3.06	114.31	106.38
6	B	506	FUC	O5-C5-C6	3.44	112.45	106.28
4	B	502	GAL	O5-C5-C6	3.95	115.79	107.34
3	B	501	BGC	C3-C4-C5	3.95	117.27	110.23

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 ⓘ

### 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	417/466 (89%)	-0.08	3 (0%) 89 90	44, 61, 105, 142	0
1	B	421/466 (90%)	-0.01	8 (1%) 70 69	45, 66, 112, 135	0
2	C	113/120 (94%)	-0.13	2 (1%) 71 70	47, 62, 80, 103	0
2	D	114/120 (95%)	-0.02	1 (0%) 85 87	43, 54, 73, 107	0
All	All	1065/1172 (90%)	-0.05	14 (1%) 79 80	43, 62, 107, 142	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	3	VAL	5.1
2	D	3	VAL	3.2
1	B	242	ARG	3.2
1	B	238	GLY	3.0
1	A	238	GLY	3.0
1	B	241	THR	2.9
1	B	111	ASN	2.9
1	B	408	GLY	2.8
1	B	237	LYS	2.8
1	A	242	ARG	2.6
1	B	245	TYR	2.5
1	B	232	VAL	2.2
2	C	103	ARG	2.1
1	A	286	SER	2.0

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
6	FUC	B	505	10/11	0.94	0.26	0.98	29,31,31,32	0
6	FUC	A	505	10/11	0.95	0.18	0.42	24,27,30,30	0
6	FUC	A	506	10/11	0.92	0.28	-	34,36,36,37	0
4	GAL	A	502	11/12	0.92	0.27	-	42,46,50,51	0
5	NAG	A	503	14/15	0.94	0.19	-	32,35,37,39	0
4	GAL	B	504	11/12	0.86	0.30	-	26,29,31,31	0
5	NAG	B	503	14/15	0.91	0.27	-	34,38,40,43	0
3	BGC	B	501	12/12	0.91	0.27	-	54,61,65,66	0
3	BGC	A	501	12/12	0.91	0.28	-	57,68,76,80	0
4	GAL	A	504	11/12	0.94	0.22	-	29,31,33,33	0
4	GAL	B	502	11/12	0.85	0.31	-	43,48,52,56	0
6	FUC	B	506	10/11	0.91	0.37	-	32,35,37,37	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.