



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 28, 2026 – 10:37 PM UTC

PDB ID : 2KR2 / pdb\_00002kr2  
Title : Xenopus laevis malectin complexed with maltose (Glcalpha1-4Glc)  
Authors : Schallus, T.; Feher, K.; Muhle-Goll, C.  
Deposited on : 2009-11-30

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
wwPDB-RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
wwPDB-ShiftChecker : v1.2  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

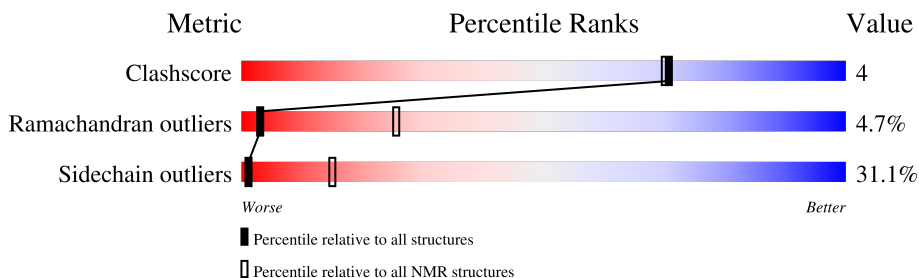
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	190	
2	B	2	

## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 14 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:6-A:33, A:49-A:64, A:69-A:180 (156)	0.49	14

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 4, 5, 7, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19
2	8, 20
Single-model clusters	6; 11

### 3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3015 atoms, of which 1498 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Malectin-A.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	190	2970	952	1476	252	285	5	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q6INX3
A	2	ALA	-	expression tag	UNP Q6INX3
A	3	MET	-	expression tag	UNP Q6INX3

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose.



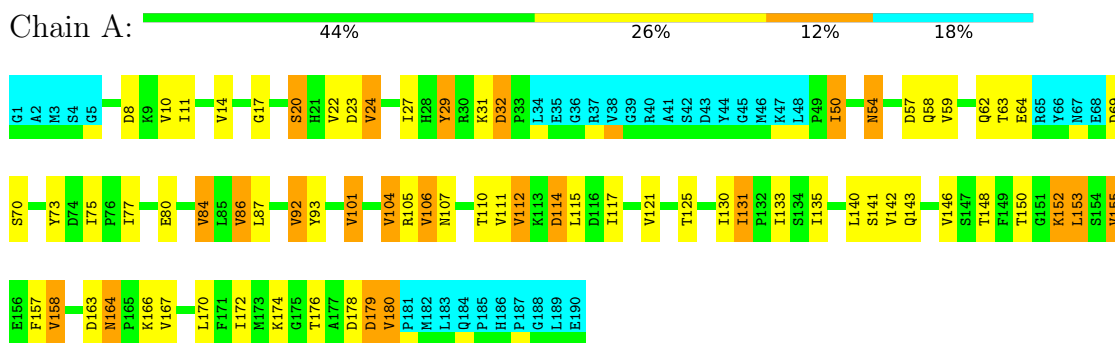
Mol	Chain	Residues	Atoms				Trace
			Total	C	H	O	
2	B	2	45	12	22	11	0

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

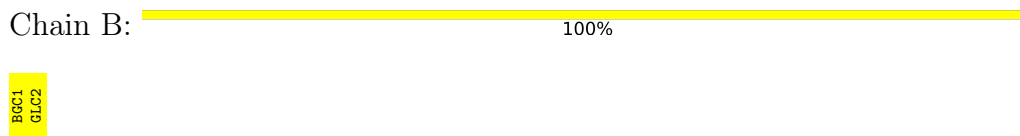
### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Malectin-A



- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

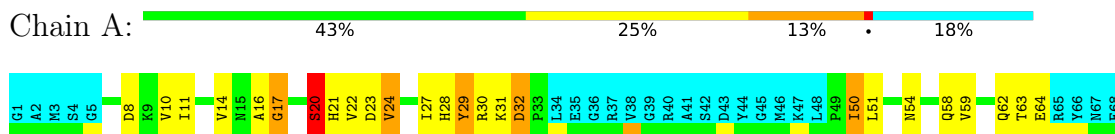


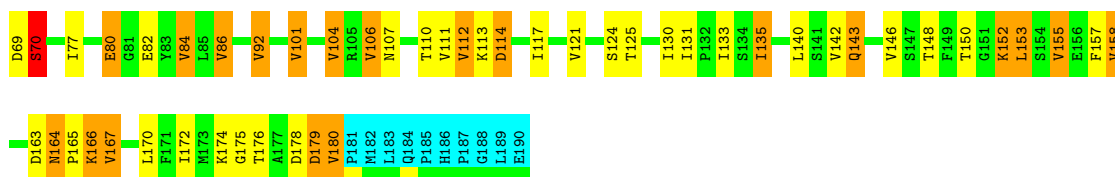
### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

- Molecule 1: Malectin-A





- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

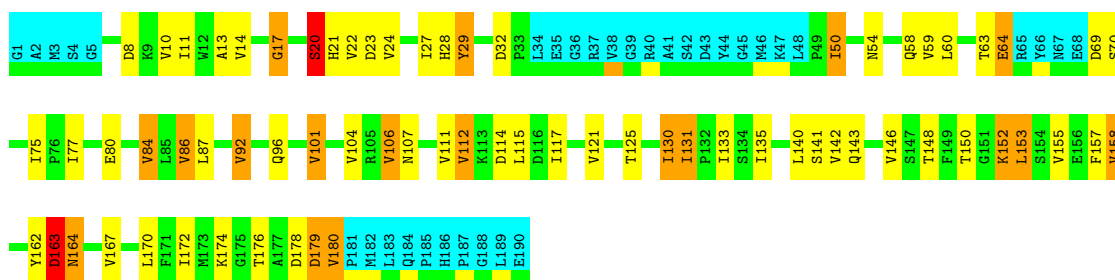
Chain B: 100%

BGC1  
GLC2

#### 4.2.2 Score per residue for model 2

- Molecule 1: Malectin-A

Chain A: 45% 26% 9% 18%



- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

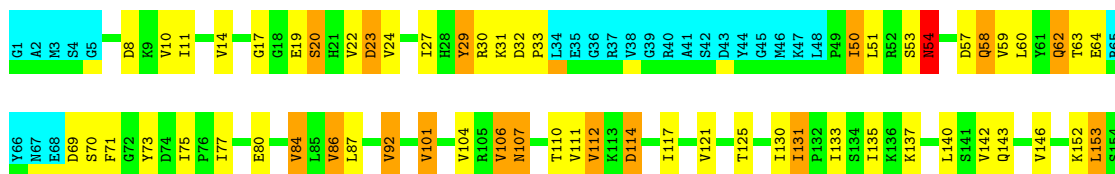
Chain B: 100%

BGC1  
GLC2

#### 4.2.3 Score per residue for model 3

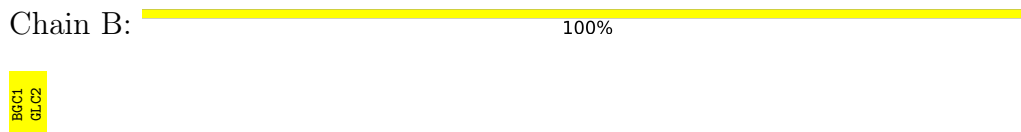
- Molecule 1: Malectin-A

Chain A: 44% 28% 10% 18%



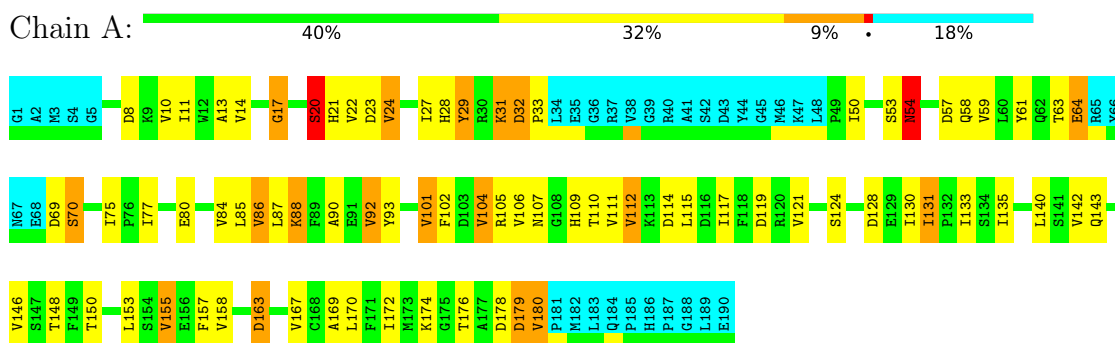


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

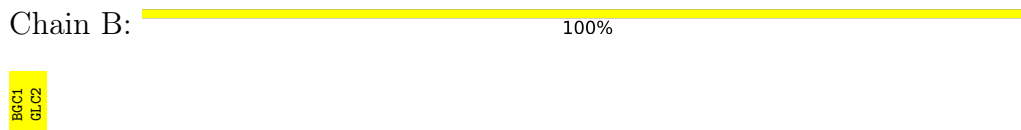


#### 4.2.4 Score per residue for model 4

- Molecule 1: Malectin-A

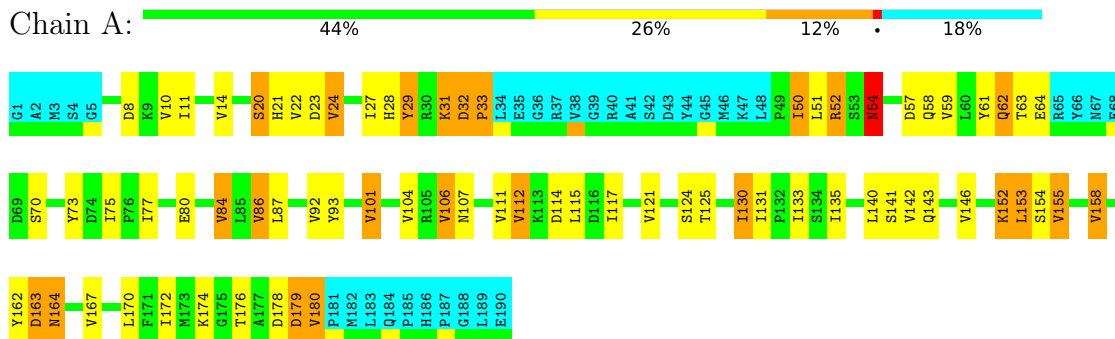


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose




#### 4.2.5 Score per residue for model 5

- Molecule 1: Malectin-A



- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B:  100%

BGC1  
GLC2

#### 4.2.6 Score per residue for model 6

- Molecule 1: Malectin-A


Chain A:  41% 31% 10% 18%

G1 A2 M3 S4 G5 D8 R9 V10 I11 V14 M15 S20 H21 V22 D23 V24 I27 H28 Y29 R30 K31 D32 F33 L34 E35 G36 R37 V38 G39 R40 A41 A44 S42 D43 Y44 G45 M46 K47 L48 P49 I50 S53 N54 D57 Q58 N59 L60 Y61 Q62 T63 E64 R65 Y66 N67 E68

D69 S70 I75 P76 I77 E80 V84 L85 V86 L87 K88 F89 A90 E91 V92 Y93 V101 V104 R105 V106 T110 V111 V112 K113 G114 L115 D116 I117 V121 D128 E129 I130 I131 P132 L133 S134 I135 L140 S141 V142 Q143 V146 F149 K152 L153 S154 V155 E156

F157 V158 K159 Y162 P163 R164 P165 K166 V167 A169 L170 F171 M172 I173 G174 G175 T176 A177 D178 D179 V180 M181 P182 L183 Q184 P185 H186 P187 G188 L189 E190

- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B:  100%

BGC1  
GLC2

#### 4.2.7 Score per residue for model 7

- Molecule 1: Malectin-A


Chain A:  44% 27% 9% 18%

G1 A2 M3 S4 G5 D8 R9 V10 I11 V14 M15 G17 S20 H21 V22 D23 V24 I27 H28 Y29 R30 K31 D32 F33 L34 E35 G36 R37 V38 G39 R40 A41 A44 S42 D43 Y44 G45 M46 K47 L48 P49 I50 L51 N54 D57 Q58 N59 L60 Y61 Q62 T63 E64 R65 Y66 N67 E68

D69 S70 Y73 D74 I75 P76 I77 E80 V84 L85 V86 L87 Y92 V101 V104 R105 L106 M107 T110 V111 V112 K113 D114 L115 D116 I117 F118 D119 R120 V121 T125 I130 I131 P132 I133 S134 I135 K139 L140 S141 V142 Q143 V146 F149 K152 L153 S154 V155 E156

F157 V158 R164 P165 K166 V167 L170 I171 I172 M173 K174 G175 T176 D178 D179 V180 P181 L183 Q184 P185 H186 P187 G188 L189 E190

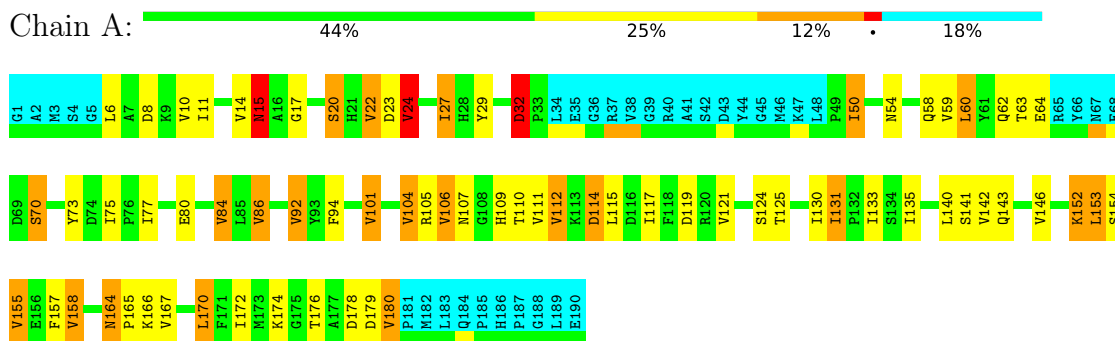
- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B:  100%

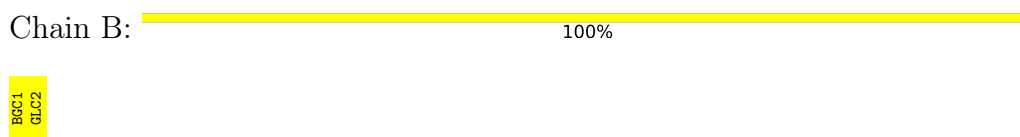
BGC1  
GLC2

### 4.2.8 Score per residue for model 8

- Molecule 1: Malectin-A

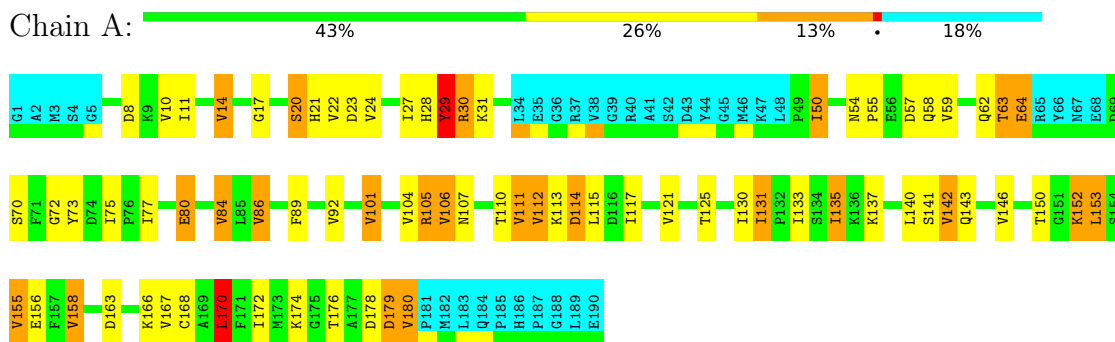


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

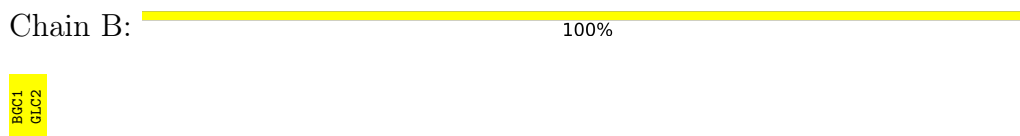


### 4.2.9 Score per residue for model 9

- Molecule 1: Malectin-A

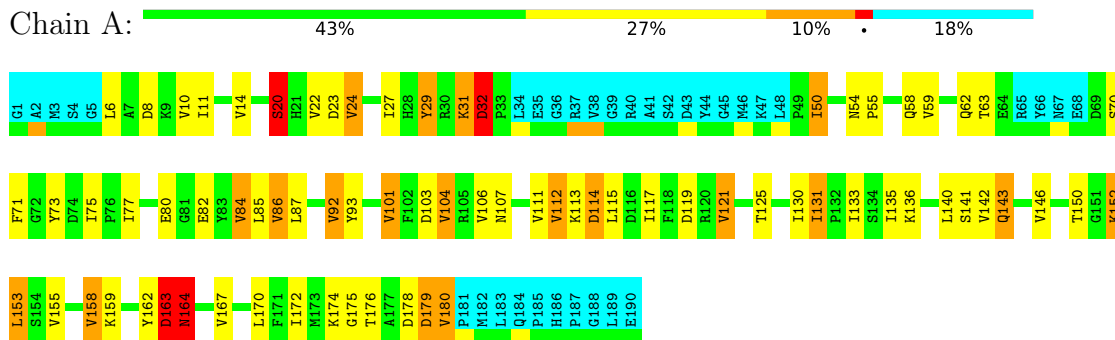


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

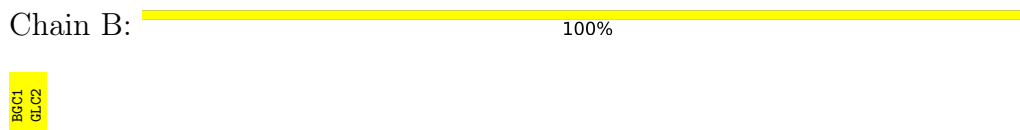


### 4.2.10 Score per residue for model 10

- Molecule 1: Malectin-A

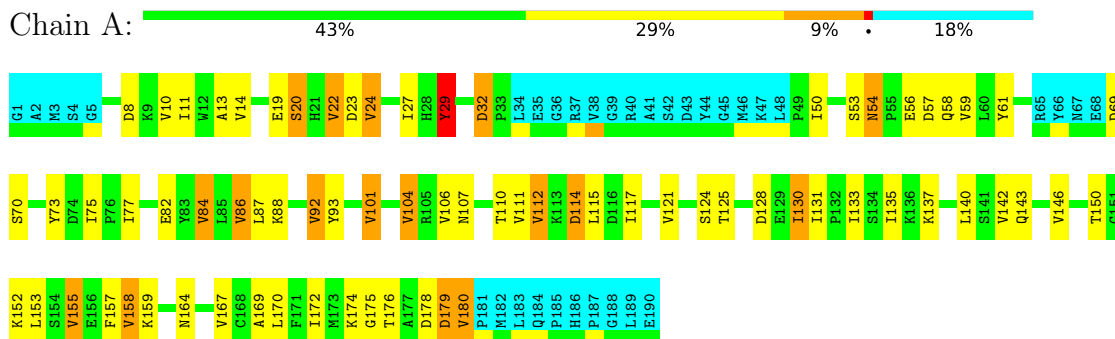


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

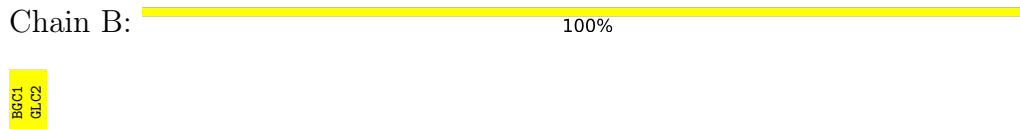


#### 4.2.11 Score per residue for model 11

- Molecule 1: Malectin-A

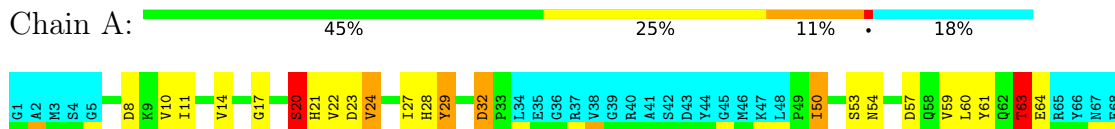


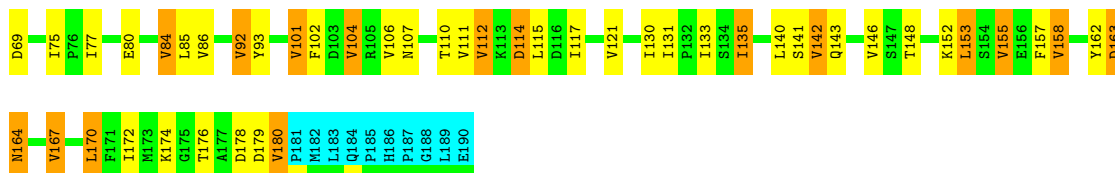
- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose



#### 4.2.12 Score per residue for model 12

- Molecule 1: Malectin-A





- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

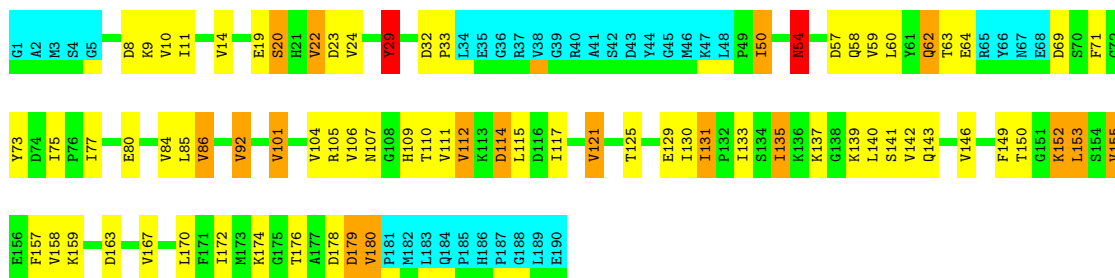
Chain B: 100%

BGC1  
GLC2

#### 4.2.13 Score per residue for model 13

- Molecule 1: Malectin-A

Chain A: 43% 29% 9% 18%



- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

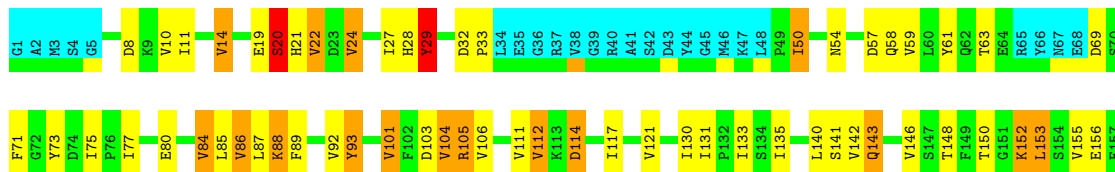
Chain B: 100%

BGC1  
GLC2

#### 4.2.14 Score per residue for model 14 (medoid)

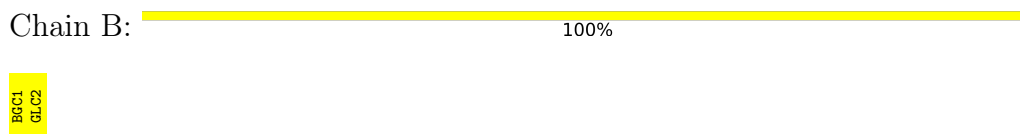
- Molecule 1: Malectin-A

Chain A: 45% 26% 11% 18%



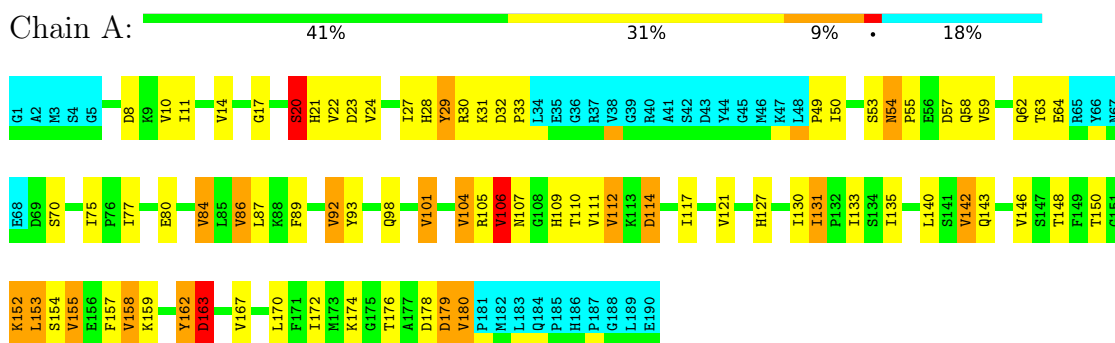


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

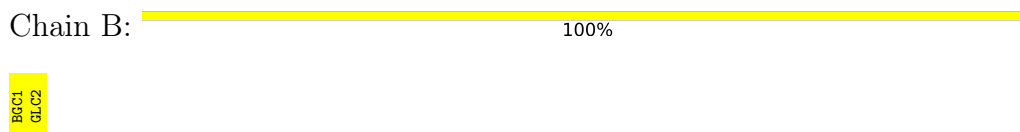


#### 4.2.15 Score per residue for model 15

- Molecule 1: Malectin-A

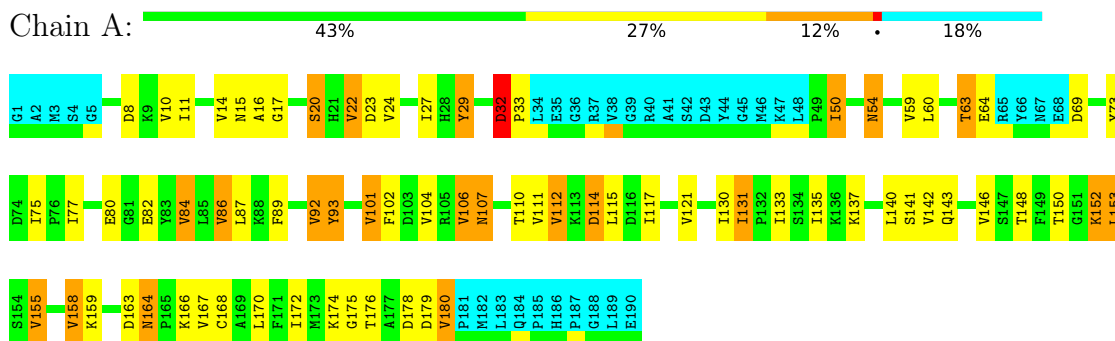


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose




#### 4.2.16 Score per residue for model 16

- Molecule 1: Malectin-A



- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B:  100%

BGC1  
GLC2

#### 4.2.17 Score per residue for model 17

- Molecule 1: Malectin-A


Chain A:  46% 25% 11% 18%

G1 A2 M3 S4 G5 D8 R9 V10 I11 V14 S20 H21 V22 D23 V24 T27 H28 Y29 R30 K31 D32 F33 L34 E35 G36 R37 V38 G39 R40 A41 S42 S44 D43 Y44 G45 M46 L47 L48 I50 N54 D58 V59 Q62 R65 Y66 M67 E68 D69 S70 Y73 D74

I75 P76 I77 E80 Y83 V84 L85 V86 L87 V82 Y83 V101 V104 R105 V106 M107 T110 V111 V112 K113 D114 L115 D116 I117 V121 T125 I130 I131 P132 I133 S134 I135 L140 S141 V142 Q143 V146 S147 T148 F149 T150 K151 L152 L153 S154 V155 E156 F157 V158

K159 M164 P165 K166 V167 I172 M173 K174 G175 A176 A177 D178 D179 V180 M181 L182 L183 Q184 P185 H186 P187 G188 L189 E190

- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B:  100%

BGC1  
GLC2

#### 4.2.18 Score per residue for model 18

- Molecule 1: Malectin-A


Chain A:  40% 33% 8% 18%

G1 A2 M3 S4 G5 D8 R9 V10 I11 V14 M15 A16 G17 G18 S20 H21 V22 D23 V24 I27 H28 Y29 D32 P33 L34 E35 G36 R37 V38 G39 R40 A41 S42 D43 Y44 M46 G45 K47 L48 P49 I50 L51 R52 S53 N54 D57 Q58 V59 L60 Y61 Q62 T63 E64 R65

Y66 M67 E68 S70 I75 P76 I77 E80 Y84 L85 V86 L87 K88 Y92 Y93 Q96 Q99 V101 V104 R105 V106 M107 L108 H109 T110 V111 V112 D114 L115 I117 V121 T125 I130 I131 P132 I133 S134 I135 L140 V141 S141 V142 Q143 V146 S147

T148 K152 L153 S154 V155 E156 F157 V158 K159 D163 M164 P165 V167 L170 F171 L172 M173 K174 G175 A176 A177 D178 D179 V180 P181 M182 L183 Q184 P185 H186 P187 G188 L189 E190

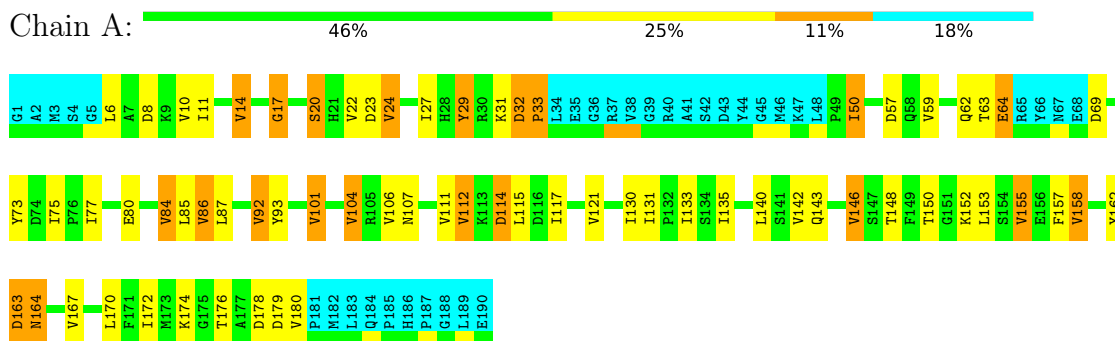
- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B:  100%

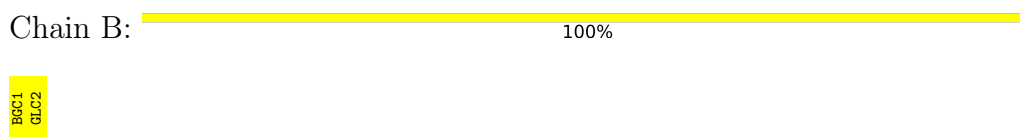
BGC1  
GLC2

### 4.2.19 Score per residue for model 19

- Molecule 1: Malectin-A

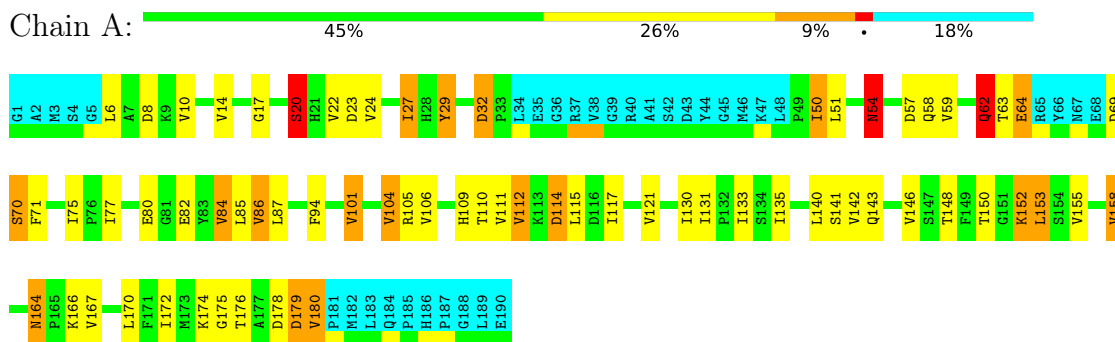


- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

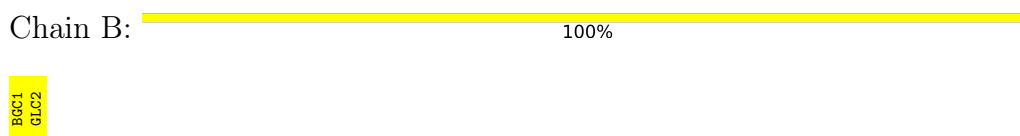


### 4.2.20 Score per residue for model 20

- Molecule 1: Malectin-A



- Molecule 2: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics*.

Of the 200 calculated structures, 20 were deposited, based on the following criterion: ?.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
ARIA	structure solution	1.2
ARIA	refinement	1.2

No chemical shift data was provided.

## 6 Model quality i

### 6.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, BGC

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 (average) 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	1.95±0.02	36±3/1264 ( 2.9± 0.2%)	1.29±0.02	5±1/1714 ( 0.3± 0.1%)
All	All	1.95	724/25280 ( 2.9%)	1.29	107/34280 ( 0.3%)

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	Planarity
1	A	0.0±0.0	0.6±0.7
All	All	0	12

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	32	ASP	N-CA	11.48	1.54	1.46	18	14
1	A	180	VAL	CA-CB	11.35	1.61	1.53	19	20
1	A	10	VAL	CA-CB	9.93	1.64	1.54	13	20
1	A	142	VAL	CA-CB	9.68	1.64	1.53	10	20
1	A	50	ILE	CA-CB	9.58	1.65	1.54	3	20
1	A	112	VAL	CA-CB	9.03	1.63	1.53	1	20
1	A	29	TYR	CE1-CZ	8.96	1.59	1.38	13	4
1	A	29	TYR	CE2-CZ	-8.45	1.18	1.38	13	4
1	A	22	VAL	CA-CB	8.33	1.63	1.53	15	18
1	A	63	THR	N-CA	7.97	1.56	1.46	16	17
1	A	14	VAL	CA-CB	7.90	1.65	1.54	4	20
1	A	146	VAL	CA-CB	7.80	1.63	1.54	1	19
1	A	32	ASP	CA-C	7.73	1.58	1.53	13	5
1	A	167	VAL	CA-CB	7.54	1.64	1.54	16	20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	111	VAL	CA-CB	7.51	1.63	1.55	3	3
1	A	104	VAL	CA-CB	7.47	1.64	1.54	13	20
1	A	121	VAL	CA-C	7.41	1.60	1.52	7	2
1	A	180	VAL	CA-C	7.29	1.58	1.52	19	19
1	A	172	ILE	CA-CB	7.29	1.62	1.54	20	20
1	A	59	VAL	CA-CB	7.28	1.63	1.54	12	20
1	A	86	VAL	CA-CB	7.19	1.62	1.54	12	20
1	A	75	ILE	CA-CB	7.14	1.60	1.54	13	19
1	A	84	VAL	CA-CB	7.14	1.62	1.54	13	20
1	A	158	VAL	CA-CB	7.09	1.63	1.54	18	20
1	A	24	VAL	N-CA	7.06	1.55	1.46	11	19
1	A	92	VAL	CA-CB	6.99	1.64	1.54	5	19
1	A	155	VAL	CA-CB	6.90	1.62	1.53	10	20
1	A	101	VAL	CA-CB	6.88	1.62	1.54	5	20
1	A	121	VAL	CA-CB	6.66	1.62	1.54	7	5
1	A	106	VAL	CA-CB	6.63	1.63	1.54	5	20
1	A	142	VAL	CA-C	6.63	1.59	1.52	10	18
1	A	157	PHE	CE1-CZ	6.57	1.58	1.38	12	1
1	A	157	PHE	CE2-CZ	-6.53	1.19	1.38	12	1
1	A	10	VAL	CA-C	6.37	1.60	1.52	9	15
1	A	27	ILE	CA-CB	6.30	1.61	1.53	2	17
1	A	135	ILE	CA-CB	6.08	1.61	1.54	7	20
1	A	59	VAL	CA-C	6.06	1.60	1.52	12	1
1	A	92	VAL	N-CA	6.06	1.53	1.46	14	1
1	A	180	VAL	C-N	6.03	1.40	1.33	4	2
1	A	125	THR	N-CA	6.01	1.53	1.46	1	10
1	A	60	LEU	CA-C	5.97	1.60	1.52	8	1
1	A	92	VAL	CA-C	5.82	1.60	1.52	6	3
1	A	59	VAL	N-CA	5.80	1.53	1.46	6	13
1	A	61	TYR	CE2-CZ	-5.79	1.24	1.38	5	1
1	A	164	ASN	C-N	5.77	1.40	1.33	2	8
1	A	31	LYS	N-CA	5.76	1.53	1.46	10	3
1	A	170	LEU	N-CA	5.66	1.53	1.46	6	3
1	A	24	VAL	CA-CB	5.60	1.62	1.54	13	19
1	A	63	THR	CA-C	5.59	1.60	1.52	16	1
1	A	131	ILE	CA-CB	5.58	1.61	1.54	5	3
1	A	14	VAL	CA-C	5.54	1.59	1.52	20	1
1	A	146	VAL	CA-C	5.51	1.59	1.52	8	13
1	A	131	ILE	C-N	5.50	1.40	1.33	18	13
1	A	104	VAL	CA-C	5.45	1.59	1.52	1	12
1	A	131	ILE	CA-C	5.44	1.58	1.52	5	2

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	10	VAL	N-CA	5.38	1.52	1.46	17	10
1	A	20	SER	N-CA	5.35	1.53	1.46	20	4
1	A	15	ASN	CA-CB	5.34	1.60	1.52	8	1
1	A	54	ASN	C-N	5.33	1.40	1.34	13	9
1	A	64	GLU	N-CA	5.32	1.53	1.46	2	1
1	A	50	ILE	CA-C	5.32	1.59	1.52	10	7
1	A	61	TYR	CE1-CZ	5.28	1.50	1.38	5	1
1	A	57	ASP	CA-C	5.28	1.59	1.52	11	1
1	A	62	GLN	N-CA	5.24	1.52	1.46	20	1
1	A	55	PRO	N-CA	5.22	1.53	1.47	15	1
1	A	54	ASN	CA-C	5.21	1.59	1.52	12	1
1	A	130	ILE	CA-CB	5.18	1.60	1.54	7	4
1	A	164	ASN	CA-C	5.15	1.59	1.52	18	1
1	A	148	THR	CA-CB	5.12	1.61	1.53	14	10
1	A	75	ILE	C-N	5.08	1.38	1.33	19	1
1	A	117	ILE	CA-C	5.05	1.59	1.52	17	1
1	A	54	ASN	N-CA	5.04	1.51	1.46	4	1

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	180	VAL	N-CA-CB	8.29	115.72	110.50	19	1
1	A	23	ASP	CA-CB-CG	7.36	119.96	112.60	7	16
1	A	54	ASN	CA-CB-CG	7.06	119.66	112.60	16	7
1	A	50	ILE	CB-CA-C	6.87	119.27	110.96	19	15
1	A	179	ASP	CA-C-N	6.58	124.39	120.24	19	1
1	A	179	ASP	C-N-CA	6.58	124.39	120.24	19	1
1	A	29	TYR	CE1-CZ-OH	-6.50	100.40	119.90	13	3
1	A	31	LYS	N-CA-C	6.06	119.00	109.96	4	3
1	A	22	VAL	N-CA-CB	6.04	118.28	111.21	8	5
1	A	15	ASN	CA-CB-CG	6.00	118.60	112.60	8	1
1	A	114	ASP	CA-CB-CG	5.98	118.58	112.60	7	10
1	A	164	ASN	CA-CB-CG	5.89	118.49	112.60	17	4
1	A	24	VAL	N-CA-C	5.76	119.12	111.17	1	4
1	A	174	LYS	N-CA-C	5.74	118.91	109.85	3	19
1	A	14	VAL	N-CA-C	5.64	115.80	108.35	20	1
1	A	57	ASP	CA-CB-CG	5.36	117.96	112.60	15	1
1	A	170	LEU	N-CA-C	5.32	117.80	110.35	8	1
1	A	69	ASP	CA-CB-CG	5.30	117.90	112.60	12	3
1	A	103	ASP	CA-CB-CG	5.27	117.87	112.60	10	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	157	PHE	CA-CB-CG	5.26	119.06	113.80	12	1
1	A	164	ASN	CB-CA-C	5.26	120.53	110.17	18	1
1	A	104	VAL	CB-CA-C	5.16	117.38	110.42	18	1
1	A	101	VAL	CA-CB-CG1	5.11	119.08	110.40	5	1
1	A	101	VAL	CB-CA-C	5.09	118.42	110.83	5	1
1	A	180	VAL	CB-CA-C	-5.09	108.95	114.35	19	1
1	A	112	VAL	N-CA-CB	5.08	117.73	112.65	2	1
1	A	111	VAL	CB-CA-C	5.02	116.21	111.44	2	1
1	A	155	VAL	N-CA-CB	5.01	118.00	111.83	9	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	29	TYR	Sidechain	4
1	A	149	PHE	Sidechain	2
1	A	89	PHE	Sidechain	2
1	A	102	PHE	Sidechain	2
1	A	73	TYR	Sidechain	1
1	A	71	PHE	Sidechain	1

## 6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1235	1221	1215	9±2
All	All	25160	24860	24720	175

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 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:54:ASN:O	1:A:58:GLN:HB2	0.79	1.76	8	15

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:21:HIS:O	1:A:28:HIS:HA	0.66	1.90	15	9
1:A:57:ASP:O	1:A:61:TYR:HB2	0.60	1.95	12	4
1:A:58:GLN:O	1:A:62:GLN:HB3	0.59	1.97	15	3
1:A:105:ARG:HB3	1:A:109:HIS:O	0.58	1.97	13	6
1:A:20:SER:HA	1:A:29:TYR:O	0.57	2.00	11	18
1:A:83:TYR:O	1:A:133:ILE:HG13	0.56	2.00	17	1
1:A:14:VAL:O	1:A:170:LEU:HB2	0.56	1.99	9	2
1:A:70:SER:H	1:A:164:ASN:ND2	0.55	2.00	17	5
1:A:19:GLU:N	1:A:29:TYR:OH	0.54	2.40	13	3
1:A:17:GLY:HA2	1:A:64:GLU:C	0.54	2.28	3	12
1:A:15:ASN:HD22	1:A:15:ASN:H	0.52	1.47	8	1
1:A:162:TYR:O	1:A:163:ASP:HB2	0.52	2.04	2	3
1:A:70:SER:H	1:A:164:ASN:HD21	0.52	1.48	17	4
1:A:60:LEU:O	1:A:60:LEU:HG	0.51	2.05	2	4
1:A:15:ASN:OD1	1:A:60:LEU:HA	0.51	2.06	16	1
1:A:107:ASN:OD1	1:A:153:LEU:HA	0.49	2.07	18	4
1:A:88:LYS:HB3	1:A:169:ALA:HB3	0.48	1.84	4	3
1:A:64:GLU:HB3	1:A:167:VAL:HG22	0.47	1.86	12	2
1:A:88:LYS:HD3	1:A:88:LYS:C	0.47	2.34	6	1
1:A:152:LYS:HG2	1:A:153:LEU:N	0.47	2.25	10	12
1:A:58:GLN:O	1:A:62:GLN:CB	0.46	2.62	15	9
1:A:82:GLU:HB2	1:A:175:GLY:O	0.46	2.11	20	5
1:A:55:PRO:O	1:A:58:GLN:HB3	0.46	2.11	9	2
1:A:51:LEU:O	1:A:52:ARG:HB2	0.45	2.12	18	2
1:A:163:ASP:O	1:A:164:ASN:HB2	0.45	2.10	10	1
1:A:88:LYS:HD2	1:A:89:PHE:N	0.45	2.27	14	1
1:A:29:TYR:CD1	1:A:29:TYR:C	0.44	2.94	11	2
1:A:13:ALA:O	1:A:29:TYR:HA	0.44	2.13	11	3
1:A:106:VAL:HA	1:A:154:SER:O	0.44	2.12	6	4
1:A:135:ILE:HA	1:A:139:LYS:O	0.44	2.13	7	2
1:A:15:ASN:ND2	1:A:168:CYS:O	0.44	2.50	16	1
1:A:107:ASN:OD1	1:A:153:LEU:HG	0.43	2.14	16	2
1:A:6:LEU:HG	1:A:27:ILE:HD11	0.43	1.91	10	4
1:A:49:PRO:HD2	1:A:62:GLN:NE2	0.43	2.29	15	1
1:A:61:TYR:O	1:A:88:LYS:HE3	0.42	2.14	11	1
1:A:80:GLU:HB3	1:A:135:ILE:O	0.42	2.15	9	3
1:A:88:LYS:HE2	1:A:90:ALA:HB2	0.42	1.90	4	2
1:A:16:ALA:O	1:A:64:GLU:HA	0.42	2.15	1	1
1:A:105:ARG:HG3	1:A:156:GLU:O	0.42	2.14	14	2
1:A:159:LYS:HG3	1:A:163:ASP:O	0.42	2.15	18	1
1:A:102:PHE:HE1	1:A:163:ASP:O	0.41	1.98	4	1

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:89:PHE:O	1:A:127:HIS:HB3	0.41	2.15	15	1
1:A:19:GLU:HA	1:A:29:TYR:OH	0.41	2.15	3	1
1:A:17:GLY:HA3	1:A:63:THR:O	0.41	2.15	12	1
1:A:58:GLN:O	1:A:62:GLN:HB2	0.41	2.16	20	1
1:A:73:TYR:CD1	1:A:73:TYR:N	0.41	2.89	7	1
1:A:17:GLY:HA3	1:A:63:THR:C	0.41	2.41	9	1
1:A:71:PHE:CE1	1:A:164:ASN:ND2	0.41	2.89	14	1
1:A:15:ASN:HB2	1:A:60:LEU:HD12	0.40	1.92	18	1
1:A:88:LYS:C	1:A:88:LYS:HD3	0.40	2.42	4	1
1:A:159:LYS:HB3	1:A:164:ASN:OD1	0.40	2.16	6	1
1:A:9:LYS:O	1:A:174:LYS:HG2	0.40	2.17	13	1
1:A:96:GLN:NE2	1:A:99:GLN:NE2	0.40	2.70	18	1

## 6.3 Torsion angles [i](#)

### 6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	156/190 (82%)	123±3 (79±2%)	26±3 (17±2%)	7±2 (5±1%)	3	25
All	All	3120/3800 (82%)	2455 (79%)	519 (17%)	146 (5%)	3	25

All 21 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	20	SER	20
1	A	114	ASP	20
1	A	143	GLN	20
1	A	93	TYR	12
1	A	164	ASN	11
1	A	166	LYS	8
1	A	33	PRO	8
1	A	31	LYS	7
1	A	165	PRO	6
1	A	163	ASP	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	30	ARG	5
1	A	17	GLY	4
1	A	124	SER	4
1	A	63	THR	3
1	A	162	TYR	3
1	A	70	SER	2
1	A	52	ARG	2
1	A	94	PHE	2
1	A	64	GLU	1
1	A	125	THR	1
1	A	24	VAL	1

### 6.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. 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	136/162 (84%)	94±3 (69±2%)	42±3 (31±2%)	1	15
All	All	2720/3240 (84%)	1874 (69%)	846 (31%)	1	15

All 89 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	8	ASP	20
1	A	77	ILE	20
1	A	101	VAL	20
1	A	112	VAL	20
1	A	117	ILE	20
1	A	130	ILE	20
1	A	133	ILE	20
1	A	140	LEU	20
1	A	153	LEU	20
1	A	176	THR	20
1	A	178	ASP	20
1	A	11	ILE	19
1	A	29	TYR	19
1	A	152	LYS	19

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	180	VAL	19
1	A	80	GLU	18
1	A	84	VAL	18
1	A	111	VAL	18
1	A	121	VAL	18
1	A	170	LEU	18
1	A	86	VAL	17
1	A	131	ILE	17
1	A	158	VAL	17
1	A	179	ASP	17
1	A	115	LEU	16
1	A	50	ILE	15
1	A	70	SER	15
1	A	92	VAL	15
1	A	110	THR	15
1	A	87	LEU	15
1	A	104	VAL	13
1	A	107	ASN	13
1	A	150	THR	13
1	A	155	VAL	13
1	A	157	PHE	13
1	A	141	SER	13
1	A	32	ASP	12
1	A	163	ASP	12
1	A	73	TYR	12
1	A	24	VAL	10
1	A	62	GLN	10
1	A	69	ASP	9
1	A	54	ASN	9
1	A	57	ASP	9
1	A	85	LEU	9
1	A	159	LYS	8
1	A	20	SER	7
1	A	22	VAL	7
1	A	106	VAL	6
1	A	53	SER	6
1	A	114	ASP	6
1	A	64	GLU	6
1	A	164	ASN	5
1	A	137	LYS	5
1	A	51	LEU	4
1	A	113	LYS	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	143	GLN	4
1	A	23	ASP	4
1	A	119	ASP	4
1	A	166	LYS	3
1	A	71	PHE	3
1	A	31	LYS	3
1	A	88	LYS	3
1	A	128	ASP	3
1	A	60	LEU	3
1	A	105	ARG	3
1	A	142	VAL	3
1	A	14	VAL	3
1	A	125	THR	2
1	A	146	VAL	2
1	A	33	PRO	2
1	A	15	ASN	2
1	A	162	TYR	2
1	A	156	GLU	2
1	A	93	TYR	2
1	A	96	GLN	1
1	A	58	GLN	1
1	A	116	ASP	1
1	A	30	ARG	1
1	A	168	CYS	1
1	A	136	LYS	1
1	A	56	GLU	1
1	A	124	SER	1
1	A	63	THR	1
1	A	135	ILE	1
1	A	129	GLU	1
1	A	19	GLU	1
1	A	98	GLN	1
1	A	148	THR	1

### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.5 Carbohydrates i

2 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Counts	Bond lengths	
						RMSZ	#Z>2
2	BGC	B	1	2	12,12,12	2.05±0.01	6±1 (45±7%)
2	GLC	B	2	2	11,11,12	2.57±0.02	6±0 (54±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Counts	Bond angles	
						RMSZ	#Z>2
2	BGC	B	1	2	17,17,17	0.94±0.02	1±0 (5±1%)
2	GLC	B	2	2	15,15,17	1.34±0.01	2±0 (13±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
2	BGC	B	1	2	-	0±0,2,22,22	0±0,1,1,1
2	GLC	B	2	2	-	0±0,2,19,22	0±0,1,1,1

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	2	GLC	C2-C3	5.09	1.60	1.52	14	20

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	1	BGC	C4-C5	3.72	1.61	1.53	7	20
2	B	1	BGC	C4-C3	3.58	1.61	1.52	14	20
2	B	2	GLC	O5-C5	3.32	1.49	1.43	15	20
2	B	2	GLC	O5-C1	3.32	1.49	1.43	12	20
2	B	2	GLC	C4-C5	3.30	1.60	1.53	2	20
2	B	2	GLC	C4-C3	3.07	1.60	1.52	2	20
2	B	2	GLC	C1-C2	2.82	1.58	1.52	6	20
2	B	1	BGC	C3-C2	2.76	1.59	1.52	11	20
2	B	1	BGC	O4-C4	2.52	1.49	1.43	6	20
2	B	1	BGC	O5-C1	2.28	1.48	1.42	3	16
2	B	1	BGC	O5-C5	2.10	1.49	1.44	8	13
2	B	1	BGC	C1-C2	2.00	1.56	1.52	18	1

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

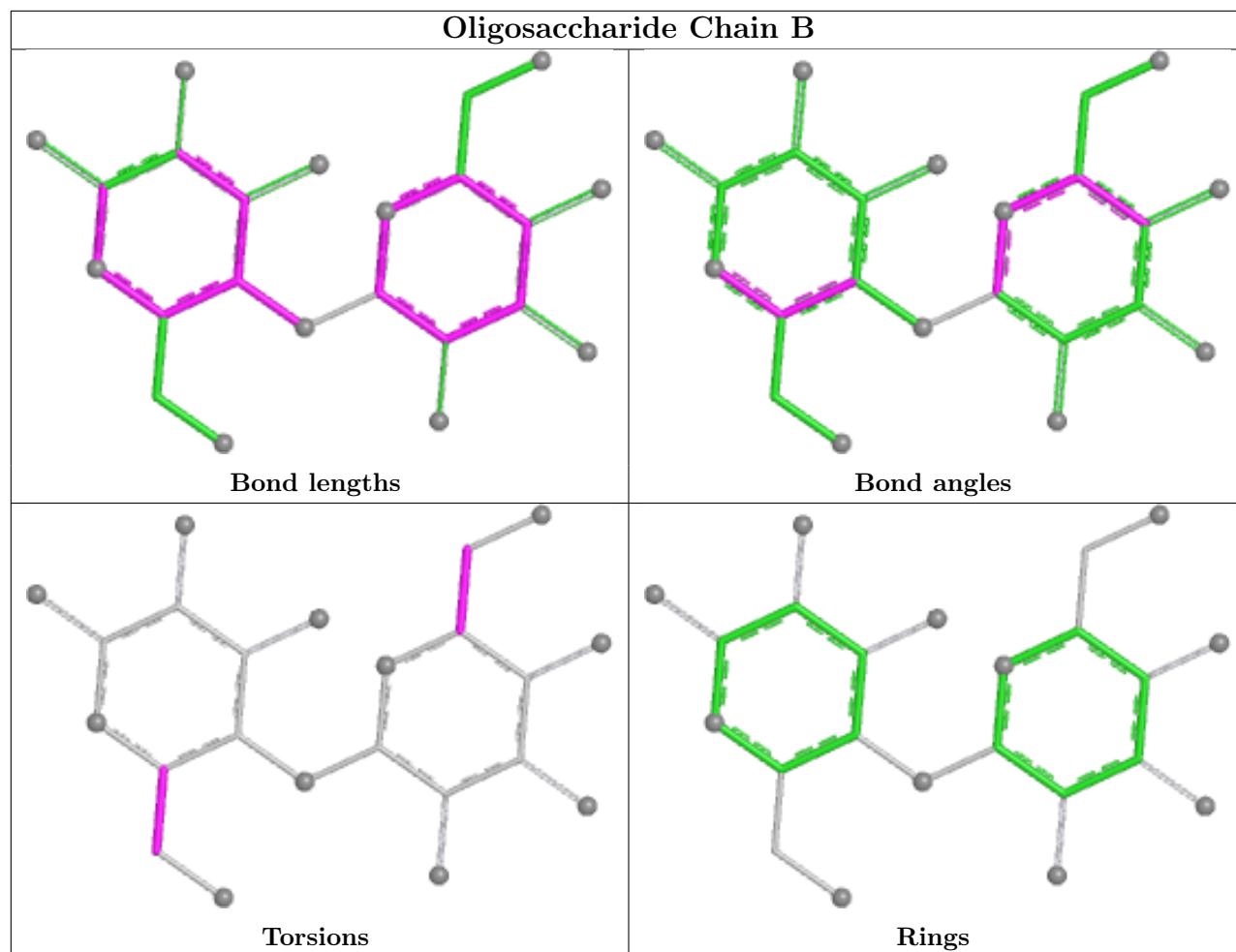
Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
2	B	2	GLC	C1-O5-C5	3.47	116.83	112.19	11	20
2	B	1	BGC	O5-C5-C4	2.76	104.72	109.70	18	19
2	B	2	GLC	O5-C5-C4	2.46	104.85	110.83	5	20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



## 6.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 6.7 Other polymers [i](#)

There are no such molecules in this entry.

## 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation

No chemical shift data were provided