

wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2023 – 09:04 PM JST

PDB ID : 4ZFU

Title: Structural studies on a non-toxic homologue of type II RIPs from Momordica

charantia (bitter gourd) in complex with N-acetyl D galactosamine

Authors: Chandran, T.; Sharma, A.; Vijayan, M.

Deposited on : 2015-04-21

Resolution : 2.53 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*A user guide is available at

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

 $Mol Probity \quad : \quad 4.02b\text{--}467$

Mogul: 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13 EDS : 2.36

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

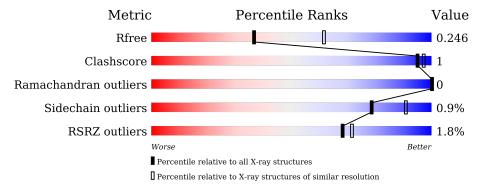
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

The reported resolution of this entry is 2.53 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},{\rm resolution\ range}(\mathring{\rm A})) \end{array}$
R_{free}	130704	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length		Qu	ality of chain	
1	A	247			98%	•
2	В	261	3%		97%	
3	С	2			100%	
4	D	5	20%	20%	60%	



2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 4345 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 rRNA N-glycosidase.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Λ	247	Total	С	N	О	S	0	0	0
1	A	241	1944	1248	308	377	11	0	0	U

• Molecule 2 is a protein called rRNA N-glycosidase.

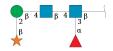
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	В	261	Total 2030	C 1253	N 361	O 402	S 14	0	0	0

• Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	A	Ator.	ns		ZeroOcc	AltConf	Trace
3	С	2	Total 28	C 16	N 2	O 10	0	0	0

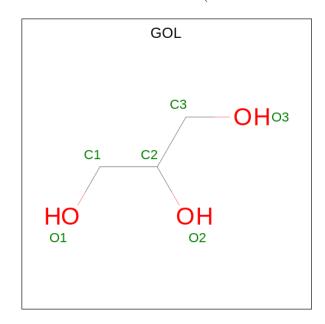
• Molecule 4 is an oligosaccharide called beta-D-xylopyranose-(1-2)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Ato	oms	ZeroOcc	AltConf	Trace
4	D	5	Total C		0	0	0

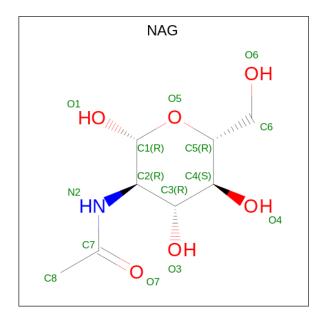


• Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	В	1	Total C O 6 3 3	0	0

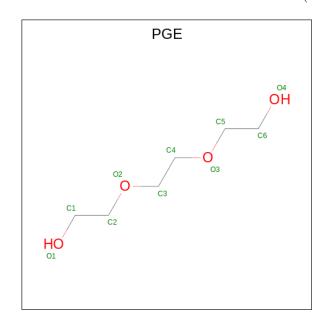
 \bullet Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $\rm C_8H_{15}NO_6).$





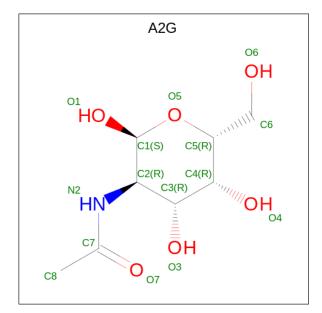
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total (N 1	O 5	0	0

 \bullet Molecule 7 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $\mathrm{C_6H_{14}O_4}).$



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

• Molecule 8 is 2-acetamido-2-deoxy-alpha-D-galactopyranose (three-letter code: A2G) (formula: $C_8H_{15}NO_6$).





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

• Molecule 9 is water.

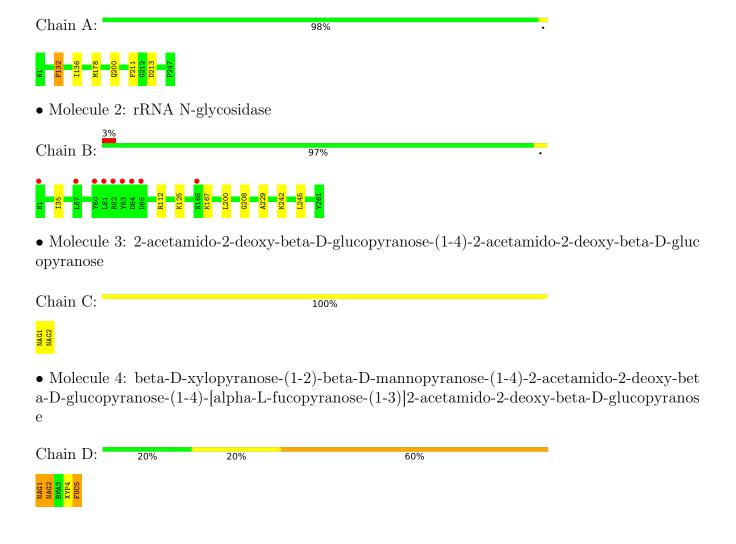
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	103	Total O 103 103	0	0
9	В	131	Total O 131 131	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants	144.38Å 136.33Å 44.87Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	60.01 - 2.53	Depositor
resolution (A)	68.17 - 2.53	EDS
% Data completeness	98.8 (60.01-2.53)	Depositor
(in resolution range)	98.8 (68.17-2.53)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.18	Depositor
$< I/\sigma(I) > 1$	1.92 (at 2.55Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
D D.	0.199 , 0.244	Depositor
R, R_{free}	0.203 , 0.246	DCC
R_{free} test set	1518 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	23.5	Xtriage
Anisotropy	0.566	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.36, 38.9	EDS
L-test for twinning ²	$ < L >=0.45, < L^2>=0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4345	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.90% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical values of <|L|>, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

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: GOL, A2G, BMA, XYP, FUC, NAG, PGE

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).

Mal	Mol Chain		lengths	Bond angles		
MIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.68	0/1988	0.75	0/2701	
2	В	0.68	0/2066	0.81	0/2810	
All	All	0.68	0/4054	0.78	0/5511	

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1944	0	1889	2	0
2	В	2030	0	1963	4	0
3	С	28	0	25	0	0
4	D	58	0	42	4	0
5	A	6	0	8	0	0
5	В	6	0	8	0	0
6	A	14	0	13	0	0
7	A	10	0	14	0	0
8	В	15	0	12	0	0
9	A	103	0	0	0	0
9	В	131	0	0	0	0
All	All	4345	0	3974	10	0



The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

The worst 5 of 10 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
4:D:1:NAG:O4	4:D:5:FUC:H5	2.02	0.59
2:B:35:ILE:HD12	2:B:112:ARG:HA	1.86	0.56
2:B:200:LEU:HD13	2:B:242:LYS:HD3	1.89	0.55
4:D:2:NAG:C1	4:D:5:FUC:H5	2.38	0.54
1:A:200:GLN:HB3	1:A:211:PHE:CG	2.48	0.49

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	Perce	Percentiles	
1	A	$245/247 \ (99\%)$	235 (96%)	10 (4%)	0	100	100
2	В	259/261 (99%)	253 (98%)	6 (2%)	0	100	100
All	All	504/508 (99%)	488 (97%)	16 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 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 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	217/217 (100%)	214 (99%)	3 (1%)	67 85		
2	В	229/229 (100%)	228 (100%)	1 (0%)	91 97		
All	All	446/446 (100%)	442 (99%)	4 (1%)	78 91		

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

Mol	Chain	Res	Type
1	A	132	PHE
1	A	178	MET
1	A	213	ASP
2	В	167	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

7 monosaccharides 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	Trme	Chain	Chain Res Link Bond lengths				В	ond ang	les	
IVIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NAG	С	1	2,3	14,14,15	0.64	0	17,19,21	0.91	1 (5%)
3	NAG	С	2	3	14,14,15	0.54	0	17,19,21	0.94	1 (5%)
4	NAG	D	1	2,4	14,14,15	1.08	1 (7%)	17,19,21	2.11	6 (35%)



Mol	Trino	Chain	Res	Link	Bo	Bond lengths		Bond angles		
MIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	NAG	D	2	4	14,14,15	0.89	0	17,19,21	2.33	8 (47%)
4	BMA	D	3	4	11,11,12	0.23	0	15,15,17	0.65	0
4	XYP	D	4	4	9,9,10	0.23	0	10,12,14	0.45	0
4	FUC	D	5	4	10,10,11	0.38	0	14,14,16	1.14	1 (7%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	С	1	2,3	-	2/6/23/26	0/1/1/1
3	NAG	С	2	3	-	2/6/23/26	0/1/1/1
4	NAG	D	1	2,4	-	1/6/23/26	0/1/1/1
4	NAG	D	2	4	-	2/6/23/26	0/1/1/1
4	BMA	D	3	4	-	1/2/19/22	0/1/1/1
4	XYP	D	4	4	-	-	0/1/1/1
4	FUC	D	5	4	-	-	0/1/1/1

All (1) bond length outliers are listed below:

\mathbf{Mol}	Chain	Res	Type	Atoms	\mathbf{Z}	$\operatorname{Observed}(\operatorname{A})$	$ \operatorname{Ideal}(\check{\mathbf{A}}) $
4	D	1	NAG	C2-N2	-2.27	1.42	1.46

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^o)$
4	D	1	NAG	C1-O5-C5	5.11	119.12	112.19
4	D	2	NAG	C3-C4-C5	-4.64	101.95	110.24
4	D	2	NAG	C1-O5-C5	3.88	117.45	112.19
4	D	2	NAG	C4-C3-C2	3.85	116.65	111.02
4	D	1	NAG	O4-C4-C5	3.45	117.87	109.30

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	С	2	NAG	O5-C5-C6-O6
3	С	2	NAG	C4-C5-C6-O6
4	D	2	NAG	C8-C7-N2-C2

Continued on next page...



Continued from previous page...

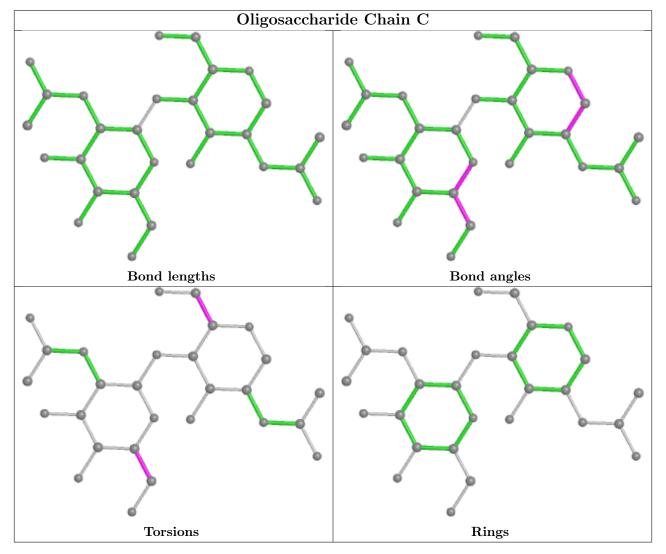
Mol	Chain	Res	Type	Atoms
4	D	3	BMA	O5-C5-C6-O6
4	D	2	NAG	O7-C7-N2-C2

There are no ring outliers.

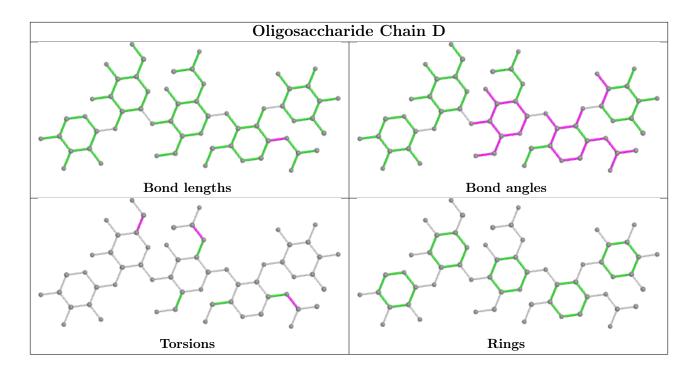
4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	4	XYP	1	0
4	D	2	NAG	3	0
4	D	1	NAG	1	0
4	D	5	FUC	3	0

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







5.6 Ligand geometry (i)

5 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	Bo	Bond lengths			Bond angles			
MIOI	туре		rtes	LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2		
8	A2G	В	301	-	15,15,15	0.70	0	21,21,21	0.99	0		
7	PGE	A	303	-	9,9,9	0.92	0	8,8,8	0.96	0		
5	GOL	В	302	-	5,5,5	0.59	0	5,5,5	0.50	0		
5	GOL	A	301	-	5,5,5	0.39	0	5,5,5	0.38	0		
6	NAG	A	302	1	14,14,15	0.90	0	17,19,21	2.19	6 (35%)		

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.

\mathbf{Mol}	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	A2G	В	301	_	-	1/6/26/26	0/1/1/1

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	PGE	A	303	-	-	5/7/7/7	-
5	GOL	В	302	-	-	2/4/4/4	-
5	GOL	A	301	-	-	2/4/4/4	-
6	NAG	A	302	1	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
6	A	302	NAG	O4-C4-C5	5.22	122.25	109.30
6	A	302	NAG	O4-C4-C3	3.48	118.39	110.35
6	A	302	NAG	C3-C4-C5	3.40	116.30	110.24
6	A	302	NAG	C2-N2-C7	2.96	127.12	122.90
6	A	302	NAG	C4-C3-C2	2.92	115.29	111.02

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	301	GOL	O1-C1-C2-C3
5	В	302	GOL	O1-C1-C2-C3
5	В	302	GOL	O1-C1-C2-O2
7	A	303	PGE	O1-C1-C2-O2
7	A	303	PGE	O3-C5-C6-O4

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.



6 Fit of model and data (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^{th} 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	$OWAB(A^2)$	Q < 0.9
1	A	247/247 (100%)	-0.21	0 100 100	13, 20, 33, 46	0
2	В	$261/261 \; (100\%)$	0.03	9 (3%) 45 49	14, 22, 47, 81	0
All	All	508/508 (100%)	-0.09	9 (1%) 68 71	13, 21, 38, 81	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	В	83	TYR	8.2
2	В	82	ARG	6.0
2	В	81	LEU	4.1
2	В	80	TYR	3.8
2	В	1	ASN	3.4

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. The B-factors column lists the minimum, median, 95^{th} 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
4	XYP	D	4	9/10	0.70	0.32	43,47,49,51	9
4	BMA	D	3	11/12	0.79	0.19	61,72,77,78	0
4	FUC	D	5	10/11	0.86	0.18	36,45,48,49	0
4	NAG	D	2	14/15	0.88	0.15	44,51,66,69	0
3	NAG	С	2	14/15	0.88	0.39	57,62,66,66	0

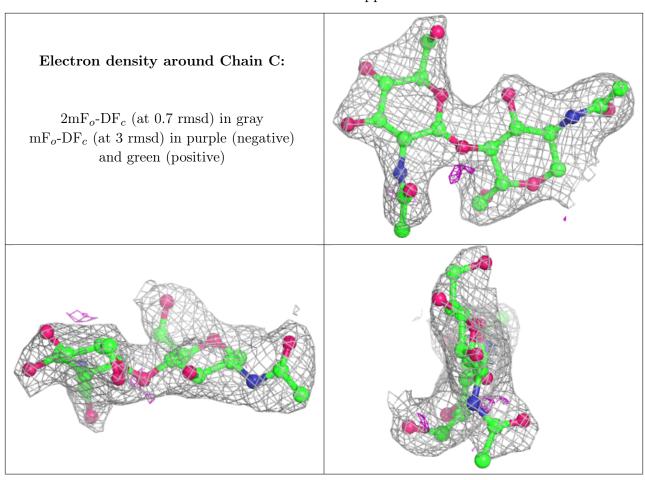
Continued on next page...



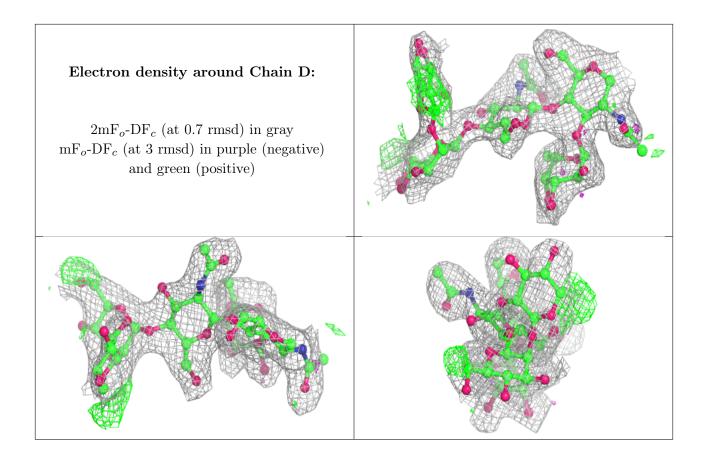
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
4	NAG	D	1	14/15	0.91	0.15	28,36,45,48	0
3	NAG	С	1	14/15	0.94	0.24	30,32,35,36	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.







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. The B-factors column lists the minimum, median, 95^{th} 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
6	NAG	A	302	14/15	0.78	0.20	47,58,63,66	0
7	PGE	A	303	10/10	0.78	0.28	37,46,57,59	0
5	GOL	В	302	6/6	0.82	0.19	33,35,36,37	0
5	GOL	A	301	6/6	0.83	0.20	41,43,46,48	0
8	A2G	В	301	15/15	0.94	0.13	19,23,25,26	0

6.5 Other polymers (i)

There are no such residues in this entry.

