

wwPDB X-ray Structure Validation Summary Report (i)

Mar 6, 2023 – 06:06 PM EST

PDB ID : 8FRG

Title : Mouse acidic mammalian chitinase, catalytic domain in complex with N,N'-di

acetylchitobiose at pH 5.43

Authors : Diaz, R.E.; Fraser, J.S.

Deposited on : 2023-01-06

Resolution : 1.74 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 : 2.32.1

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)

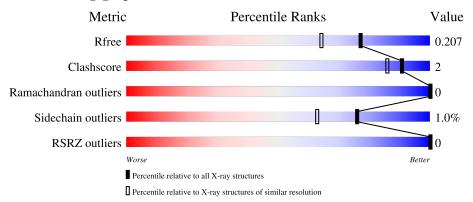
Validation Pipeline (wwPDB-VP) : 2.32.1

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 1.74 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},{\rm resolution\ range}({\rm \AA})) \end{array}$
R_{free}	130704	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)
RSRZ outliers	127900	3705 (1.76-1.72)

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	Quality of chai	n	
1	A	397	89%		• 7%
1	В	397	88%		• • 7%
2	С	2	50%	50%	
2	Е	2	50%	50%	
2	F	2	100%		

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality	of chain
3	D	2	50%	50%



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13545 atoms, of which 6214 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 Acidic mammalian chitinase.

Mol	Chain	Residues		Atoms						AltConf	Trace
1	A	369	Total 6211	C 2035	H 3040	N 518	O 603	S 15	0	60	0
1	В	369	Total 6093	C 1997	H 2978	N 506	O 598	S 14	0	57	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	392	HIS	-	expression tag	UNP Q91XA9
A	393	HIS	-	expression tag	UNP Q91XA9
A	394	HIS	-	expression tag	UNP Q91XA9
A	395	HIS	-	expression tag	UNP Q91XA9
A	396	HIS	-	expression tag	UNP Q91XA9
A	397	HIS	-	expression tag	UNP Q91XA9
В	392	HIS	-	expression tag	UNP Q91XA9
В	393	HIS	-	expression tag	UNP Q91XA9
В	394	HIS	-	expression tag	UNP Q91XA9
В	395	HIS	-	expression tag	UNP Q91XA9
В	396	HIS	_	expression tag	UNP Q91XA9
В	397	HIS	_	expression tag	UNP Q91XA9

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	С	9	Total	С	Н	N	О	0	9	0
		2	114	32	56	4	22	U	2	U

Continued on next page...



Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	9	Total	С	Н	N	О	0	2	0
	12	2	114	32	56	4	22			
9	2 F	E o	Total	С	Н	N	О	0	9	0
2		F	2	114	32	56	4	22	U	2

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	2	Total 57	C 16	H 28	N 2	O 11	0	2	0

• Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Mg 2 2	0	0

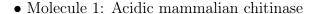
• Molecule 5 is water.

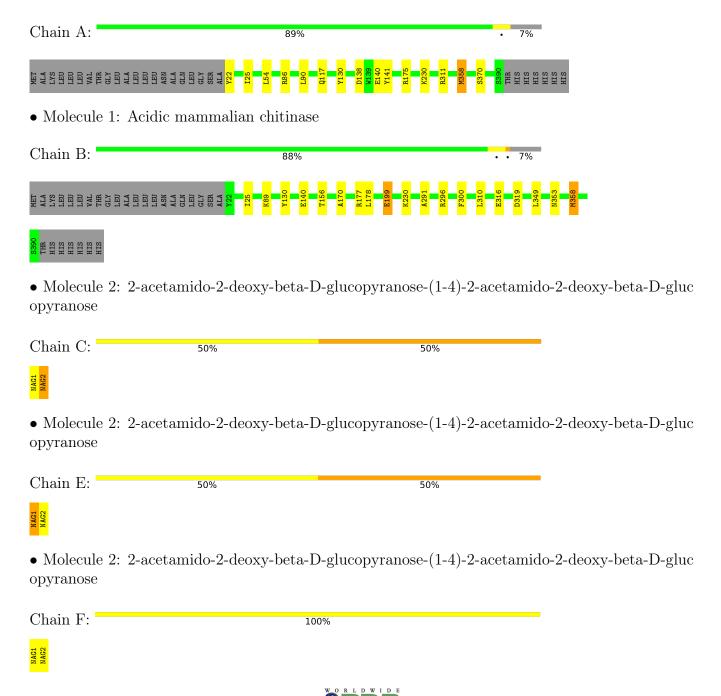
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	419	Total O 419 419	0	0
5	В	421	Total O 421 421	0	0



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.





 \bullet Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose

Chain D: 50% 50%





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants	92.89Å 105.04Å 70.81Å	Donositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	69.59 - 1.74	Depositor
Resolution (A)	69.59 - 1.74	EDS
% Data completeness	99.6 (69.59-1.74)	Depositor
(in resolution range)	99.6 (69.59-1.74)	EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.75 (at 1.74Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
D D.	0.166 , 0.208	Depositor
R, R_{free}	0.165 , 0.207	DCC
R_{free} test set	3562 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	15.7	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.40, 42.5	EDS
L-test for twinning ²	$ < L >=0.51, < L^2>=0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	13545	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The analyses of the Patterson function reveals a significant off-origin peak that is 58.16 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.1302e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

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: MG, NAG, NDG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with |Z| > 5 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond	lengths	Bond angles		
MIOI		RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.42	0/3552	0.61	0/4839	
1	В	0.45	0/3501	0.63	0/4771	
All	All	0.44	0/7053	0.62	0/9610	

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	3171	3040	2766	8	2
1	В	3115	2978	2687	12	1
2	С	58	56	54	1	0
2	Е	58	56	53	1	0
2	F	58	56	53	0	0
3	D	29	28	24	2	0
4	A	2	0	0	0	0
5	A	419	0	0	4	1
5	В	421	0	0	1	1
All	All	7331	6214	5637	21	4

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 2.

The worst 5 of 21 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$egin{aligned} ext{Clash} \ ext{overlap } (ext{Å}) \end{aligned}$
1:A:138[B]:ASP:OD2	3:D:1[B]:NDG:H8C2	1.86	0.76
1:B:140:GLU:OE2	2:E:1[C]:NAG:H2	2.12	0.49
1:B:156:THR:HG21	1:B:199[A]:GLU:HG2	1.94	0.49
1:A:54:LEU:HG	1:A:90:LEU:HD11	1.96	0.48
1:A:22:TYR:N	1:B:353:ASN:HD22	2.12	0.46

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	$\begin{array}{c} \text{Interatomic} \\ \text{distance (Å)} \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:A:130:TYR:OH	1:A:230[A]:LYS:NZ[4_456]	2.10	0.10
5:A:868:HOH:O	5:B:801:HOH:O[4_456]	2.13	0.07
1:A:311:ARG:NH2	1:A:370:SER:OG[2_565]	2.17	0.03
1:B:130:TYR:OH	1:B:230[A]:LYS:HZ1[4_555]	1.60	0.00

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	Analysed Favoured Allowed		Outliers	Perce	entiles
1	A	$425/397\ (107\%)$	416 (98%)	9 (2%)	0	100	100
1	В	$420/397\ (106\%)$	411 (98%)	9 (2%)	0	100	100
All	All	845/794 (106%)	827 (98%)	18 (2%)	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	366/331 (111%)	362 (99%)	4 (1%)	73	59	
1	В	360/331 (109%)	354 (98%)	6 (2%)	60	41	
All	All	726/662 (110%)	716 (99%)	10 (1%)	76	50	

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	349	LEU
1	В	358[A]	MET
1	В	358[B]	MET
1	A	358[B]	MET
1	В	199[A]	GLU

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)

14 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	Type	Chain	Res	Link	Вс	Bond lengths			Bond angles			
WIOI	Type	Chain	rtes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2		
2	NAG	С	1[A]	2	15,15,15	1.89	4 (26%)	21,21,21	1.33	2 (9%)		
2	NAG	С	1[B]	2	15,15,15	1.83	5 (33%)	21,21,21	0.93	0		
2	NAG	С	2[A]	2	14,14,15	2.01	5 (35%)	17,19,21	1.04	1 (5%)		
2	NAG	С	2[B]	2	14,14,15	2.13	4 (28%)	17,19,21	1.25	3 (17%)		
3	NDG	D	1[B]	3	15,15,15	1.98	6 (40%)	21,21,21	1.57	3 (14%)		
3	NAG	D	2[B]	3	14,14,15	2.19	6 (42%)	17,19,21	1.51	3 (17%)		
2	NAG	Е	1[A]	2	15,15,15	2.13	7 (46%)	21,21,21	1.19	1 (4%)		
2	NAG	Е	1[C]	2	15,15,15	2.81	5 (33%)	21,21,21	2.17	4 (19%)		
2	NAG	Е	2[A]	2	14,14,15	2.26	7 (50%)	17,19,21	1.03	1 (5%)		
2	NAG	Е	2[C]	2	14,14,15	2.10	5 (35%)	17,19,21	0.95	1 (5%)		
2	NAG	F	1[A]	2	15,15,15	2.24	5 (33%)	21,21,21	0.97	0		
2	NAG	F	1[B]	2	15,15,15	2.00	4 (26%)	21,21,21	0.95	0		
2	NAG	F	2[A]	2	14,14,15	2.08	5 (35%)	17,19,21	1.16	1 (5%)		
2	NAG	F	2[B]	2	14,14,15	2.05	4 (28%)	17,19,21	1.28	2 (11%)		

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	NAG	С	1[A]	2	-	2/6/26/26	0/1/1/1
2	NAG	С	1[B]	2	-	0/6/26/26	0/1/1/1
2	NAG	С	2[A]	2	-	4/6/23/26	0/1/1/1
2	NAG	С	2[B]	2	-	3/6/23/26	0/1/1/1
3	NDG	D	1[B]	3	-	0/6/26/26	0/1/1/1
3	NAG	D	2[B]	3	-	2/6/23/26	0/1/1/1
2	NAG	Ε	1[A]	2	-	0/6/26/26	0/1/1/1
2	NAG	E	1[C]	2	-	2/6/26/26	0/1/1/1
2	NAG	Е	2[A]	2	-	0/6/23/26	0/1/1/1
2	NAG	Е	2[C]	2	-	0/6/23/26	0/1/1/1
2	NAG	F	1[A]	2	-	2/6/26/26	0/1/1/1
2	NAG	F	1[B]	2	-	0/6/26/26	0/1/1/1
2	NAG	F	2[A]	2	-	4/6/23/26	0/1/1/1
2	NAG	F	2[B]	2	-	2/6/23/26	0/1/1/1



The worst	5	of 72	bond	length	outliers	are	listed	below:
THE WOLDS	\circ	01 12	DOM	10115 011	Outilities	$\alpha_{\rm L}$	mouca	DCIOW.

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}(ext{\AA})$
2	Ε	1[C]	NAG	C1-C2	8.39	1.63	1.52
2	С	2[B]	NAG	C1-C2	5.01	1.59	1.52
2	F	2[B]	NAG	C1-C2	4.90	1.59	1.52
2	F	1[A]	NAG	C1-C2	4.88	1.58	1.52
2	С	2[A]	NAG	C1-C2	4.39	1.58	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	Е	1[C]	NAG	O5-C1-C2	8.15	117.70	109.52
3	D	1[B]	NDG	C1-C2-N2	-4.21	105.85	110.73
3	D	2[B]	NAG	C4-C3-C2	-3.53	105.84	111.02
3	D	1[B]	NDG	C3-C2-N2	3.30	116.85	110.62
3	D	2[B]	NAG	C1-O5-C5	3.24	116.59	112.19

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	Е	1[C]	NAG	C1-C2-N2-C7
3	D	2[B]	NAG	C4-C5-C6-O6
2	F	1[A]	NAG	O5-C5-C6-O6
3	D	2[B]	NAG	O5-C5-C6-O6
2	F	1[A]	NAG	C4-C5-C6-O6

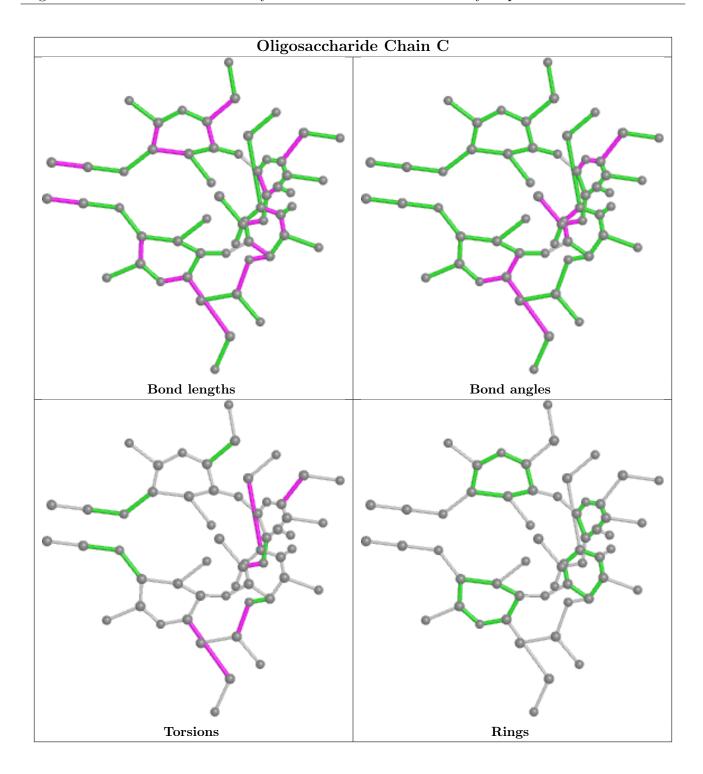
There are no ring outliers.

3 monomers are involved in 4 short contacts:

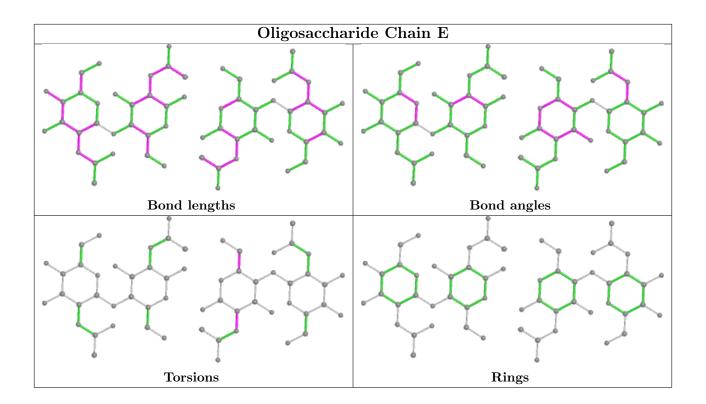
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	Е	1[C]	NAG	1	0
2	С	2[A]	NAG	1	0
3	D	1[B]	NDG	2	0

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

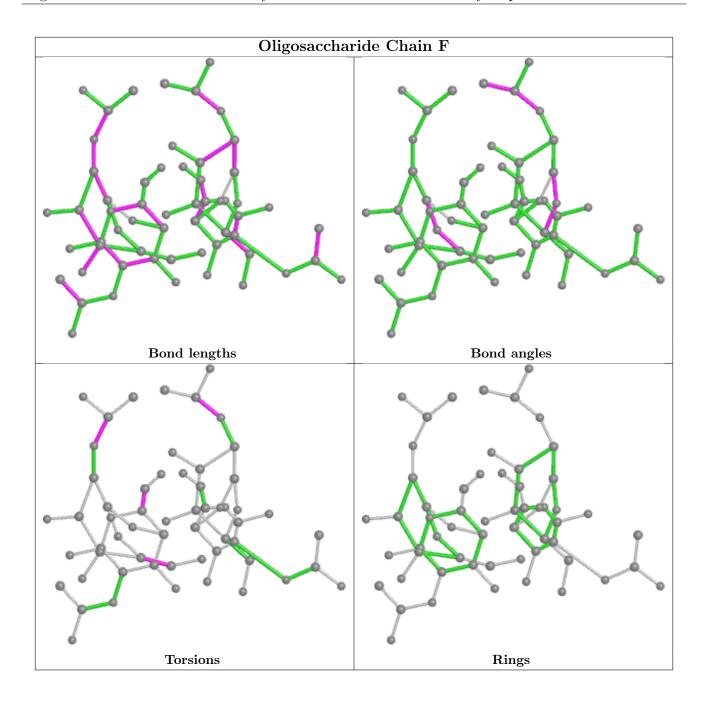




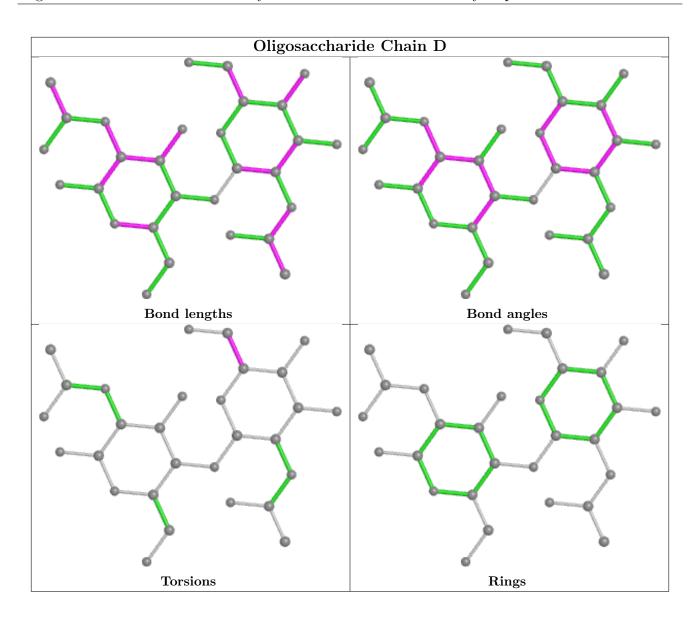












5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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 (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		Z>2	$OWAB(Å^2)$	Q < 0.9
1	A	369/397~(92%)	-0.40	0	100	100	9, 16, 30, 47	0
1	В	369/397~(92%)	-0.46	0	100	100	9, 14, 24, 34	0
All	All	738/794 (92%)	-0.43	0	100	100	9, 15, 27, 47	0

There are no RSRZ outliers to report.

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	$\operatorname{B-factors}(\mathring{\mathbf{A}}^2)$	Q < 0.9
2	NAG	Е	1[A]	15/15	0.80	0.20	16,25,30,32	29
2	NAG	Е	1[C]	15/15	0.80	0.20	13,18,23,25	29
2	NAG	Е	2[A]	14/15	0.85	0.17	20,25,31,36	28
2	NAG	Е	2[C]	14/15	0.85	0.17	16,25,29,31	28
2	NAG	F	2[A]	14/15	0.87	0.16	20,25,29,30	28
2	NAG	F	2[B]	14/15	0.87	0.16	15,22,28,28	28
3	NDG	D	1[B]	15/15	0.89	0.16	20,26,35,41	0
2	NAG	С	2[B]	14/15	0.90	0.15	17,24,29,31	28
2	NAG	С	2[A]	14/15	0.90	0.15	20,24,29,33	28
3	NAG	D	2[B]	14/15	0.91	0.12	15,26,39,47	0
2	NAG	С	1[A]	15/15	0.94	0.13	19,25,29,38	29

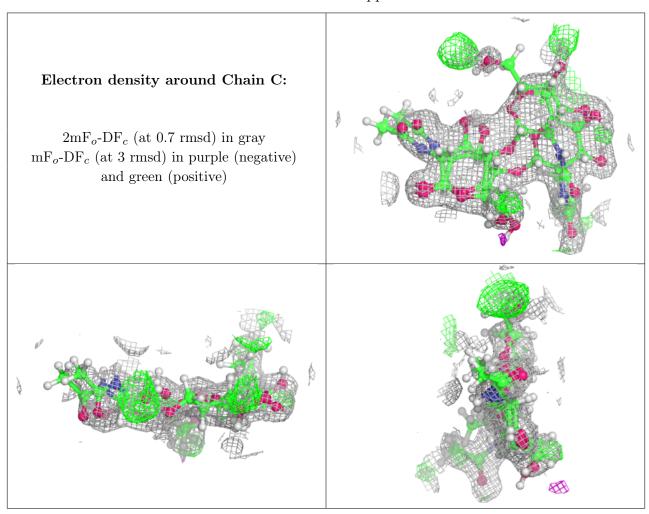
Continued on next page...



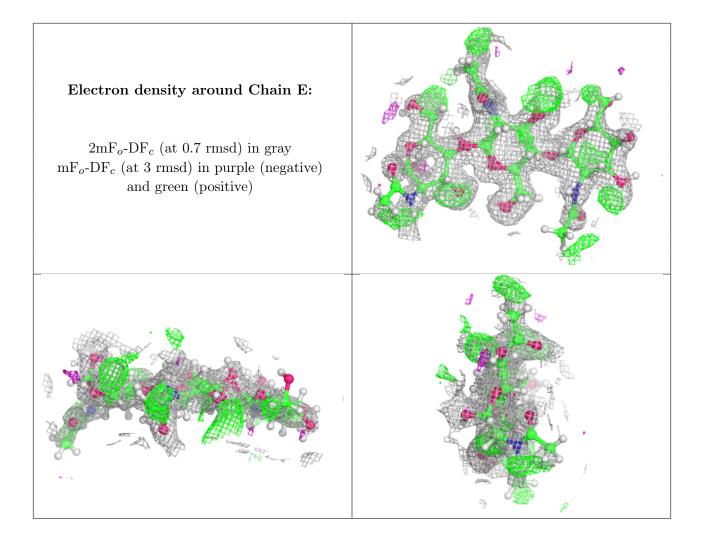
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	NAG	С	1[B]	15/15	0.94	0.13	22,26,29,31	29
2	NAG	F	1[A]	15/15	0.94	0.12	22,26,31,32	29
2	NAG	F	1[B]	15/15	0.94	0.12	21,26,32,33	29

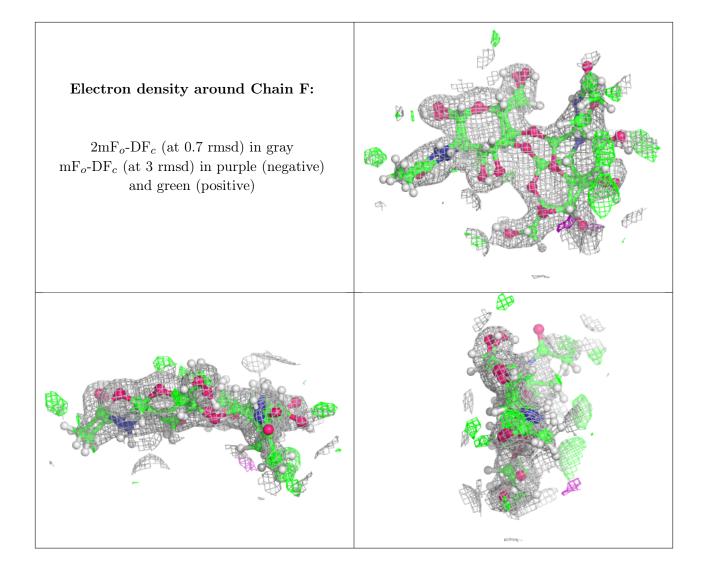
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



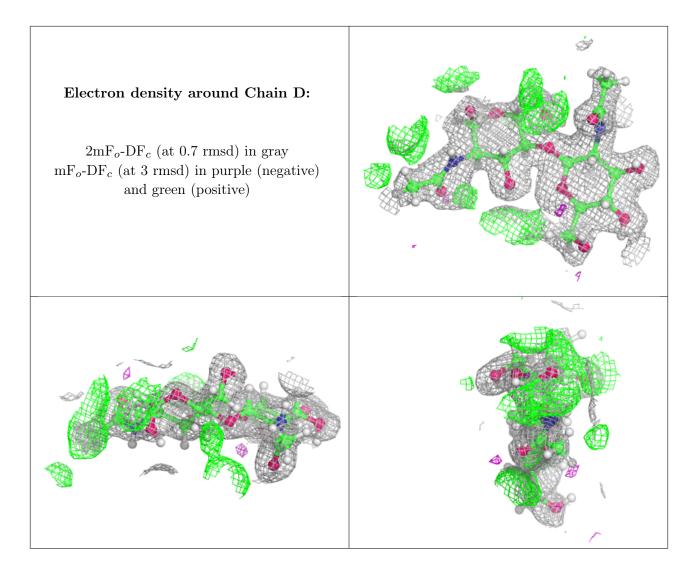












6.4 Ligands (i)

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	$\operatorname{B-factors}(\check{\mathbf{A}}^2)$	Q<0.9
4	MG	A	401	1/1	0.94	0.18	29,29,29,29	0
4	MG	A	402	1/1	0.95	0.17	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.

