

wwPDB X-ray Structure Validation Summary Report (i)

Jan 8, 2024 - 05:59 am GMT

PDB ID	:	5NNI
Title	:	Dimer structure of Sortilin ectodomain crystal form 2, 3.2 Angstrom
Authors	:	Leloup, N.O.L.; Janssen, B.J.C.
Deposited on		
Resolution	:	3.21 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:

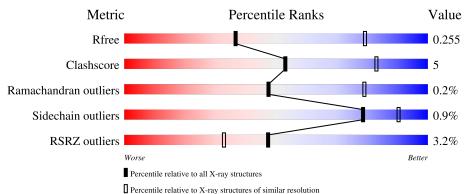
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 3.21 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	$\begin{array}{c} \textbf{Whole archive} \\ \textbf{(\#Entries)} \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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	Qualit	y of chain	
1	А	732	2% 76%	13	3% 11%
1	В	732	4%	12	% 11%
2	С	3	33%	67%	
2	D	3	67%	3	3%
2	F	3	1	00%	

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Е	2	100%



2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10409 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.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	651	Total 5120	C 3237	N 861	O 993	S 29	0	0	0
1	В	651	Total 5116	C 3236	N 861	O 990	S 29	0	0	0

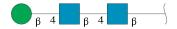
• Molecule 1 is a protein called Sortilin.

Chain	Residue	Modelled	Actual	Comment	Reference
А	723	ALA	-	expression tag	UNP Q6PHU5
A	724	ALA	-	expression tag	UNP Q6PHU5
А	725	ALA	-	expression tag	UNP Q6PHU5
А	726	HIS	-	expression tag	UNP Q6PHU5
A	727	HIS	-	expression tag	UNP Q6PHU5
А	728	HIS	-	expression tag	UNP Q6PHU5
А	729	HIS	-	expression tag	UNP Q6PHU5
A	730	HIS	-	expression tag	UNP Q6PHU5
А	731	HIS	-	expression tag	UNP Q6PHU5
В	723	ALA	-	expression tag	UNP Q6PHU5
В	724	ALA	-	expression tag	UNP Q6PHU5
В	725	ALA	-	expression tag	UNP Q6PHU5
В	726	HIS	-	expression tag	UNP Q6PHU5
В	727	HIS	-	expression tag	UNP Q6PHU5
В	728	HIS	-	expression tag	UNP Q6PHU5
В	729	HIS	-	expression tag	UNP Q6PHU5
В	730	HIS	-	expression tag	UNP Q6PHU5
В	731	HIS	-	expression tag	UNP Q6PHU5

There are 18 discrepancies between the modelled and reference sequences:

• Molecule 2 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-b eta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.





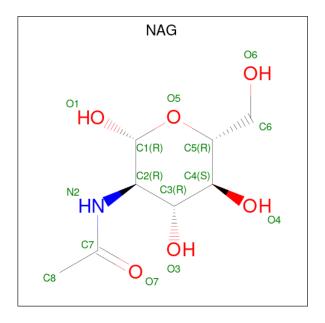
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	С	3	Total C N O 39 22 2 15	0	0	0
2	D	3	Total C N O 39 22 2 15	0	0	0
2	F	3	Total C N O 39 22 2 15	0	0	0

• 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	Ato	oms		ZeroOcc	AltConf	Trace
3	Е	2	Total C 28 1	C N 6 2	0 10	0	0	0

• Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



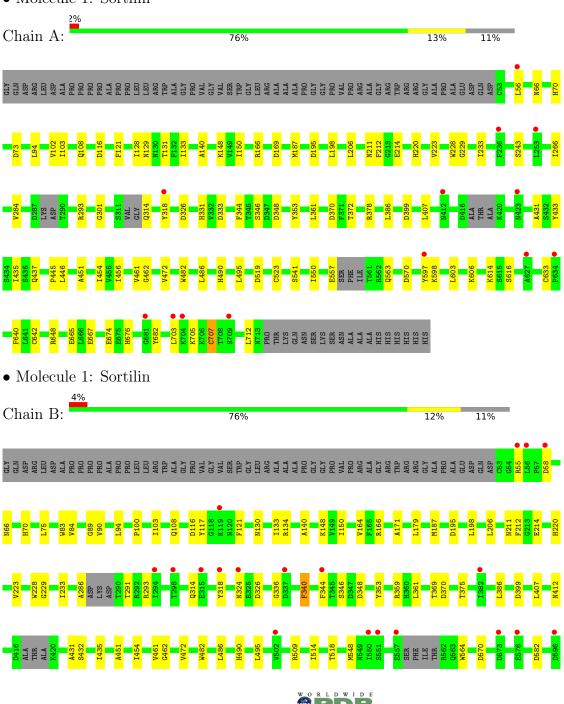


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	Total C N O 14 8 1 5	0	0
4	В	1	Total C N O 14 8 1 5	0	0



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.



• Molecule 1: Sortilin

HIS K666 K666 K666

• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:	33%	67%

NAG1 NAG2 BMA3

• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:	67%	33%
MG2 MG2 BMA3		

• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

100%

NAG1 NAG2 BMA3

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-gluc opyranose

Chain E:

100%

NAG1 NAG2



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	79.84Å 137.17Å 147.56Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.20 - 3.21	Depositor
Resolution (A)	62.20 - 3.21	EDS
% Data completeness	99.5 (62.20-3.21)	Depositor
(in resolution range)	99.5 (62.20-3.21)	EDS
R _{merge}	0.08	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$3.62 (at 3.19 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.11_2567: ???)	Depositor
D D	0.227 , 0.255	Depositor
R, R_{free}	0.228 , 0.255	DCC
R_{free} test set	1365 reflections (5.04%)	wwPDB-VP
Wilson B-factor $(Å^2)$	94.3	Xtriage
Anisotropy	0.384	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.28,56.1	EDS
L-test for twinning ²	$ \langle L \rangle = 0.48, \langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10409	wwPDB-VP
Average B, all atoms $(Å^2)$	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.36% of the height of the origin peak. No significant pseudotranslation is detected.

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



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: NAG, BMA

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		
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.24	0/5240	0.43	0/7095	
1	В	0.25	0/5237	0.43	0/7092	
All	All	0.24	0/10477	0.43	0/14187	

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	А	5120	0	4903	55	0
1	В	5116	0	4905	56	0
2	С	39	0	34	0	0
2	D	39	0	34	0	0
2	F	39	0	34	0	0
3	Е	28	0	25	6	0
4	А	14	0	13	0	0
4	В	14	0	13	0	0
All	All	10409	0	9961	109	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 5.



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:324:ASN:HB3	3:E:1:NAG:H62	1.70	0.74
1:B:324:ASN:HB2	3:E:2:NAG:O7	1.92	0.70
1:B:94:LEU:HD11	1:B:133:ILE:HG13	1.73	0.69
1:B:140:ALA:HB3	1:B:150:ILE:HB	1.76	0.66
1:B:148:LYS:NZ	1:B:195:ASP:OD1	2.30	0.65

The worst 5 of 109 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

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	ntiles
1	А	641/732~(88%)	609~(95%)	32~(5%)	0	100	100
1	В	643/732~(88%)	599~(93%)	42~(6%)	2~(0%)	41	74
All	All	1284/1464~(88%)	1208 (94%)	74 (6%)	2~(0%)	47	79

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	340	PHE
1	В	626	VAL

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	Analysed Rotameric Outliers		Percentiles		
1	А	568/626~(91%)	565 (100%)	3(0%)	88 94		
1	В	567/626~(91%)	560~(99%)	7 (1%)	71 87		
All	All	1135/1252 (91%)	1125~(99%)	10 (1%)	78 90		

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

Mol	Chain	Res	Type
1	В	605	TYR
1	В	674	GLU
1	В	707	CYS
1	В	220	HIS
1	В	412	ASN

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

Mol	Chain	Res	Type
1	В	423	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

11 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	Turne	Chain	Res	Link	Bo	ond leng	ths	В	ond ang	les
10101	Type	Chain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
2	NAG	С	1	1,2	14,14,15	0.55	0	$17,\!19,\!21$	1.95	4 (23%)
2	NAG	С	2	2	14,14,15	0.24	0	17,19,21	2.09	<u>6 (35%)</u>
2	BMA	С	3	2	11,11,12	0.28	0	$15,\!15,\!17$	0.72	0
2	NAG	D	1	1,2	14,14,15	0.30	0	$17,\!19,\!21$	0.71	0
2	NAG	D	2	2	14,14,15	0.30	0	17,19,21	1.03	2 (11%)
2	BMA	D	3	2	11,11,12	0.30	0	$15,\!15,\!17$	0.93	0
3	NAG	Е	1	3,1	14,14,15	0.45	0	$17,\!19,\!21$	1.49	4 (23%)
3	NAG	Е	2	3	14,14,15	0.37	0	17,19,21	2.05	4 (23%)
2	NAG	F	1	1,2	14,14,15	0.36	0	17,19,21	1.63	3 (17%)
2	NAG	F	2	1,2	14,14,15	0.68	0	17,19,21	2.43	5 (29%)
2	BMA	F	3	2	11,11,12	0.34	0	$15,\!15,\!17$	1.56	3 (20%)

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	1,2	-	1/6/23/26	0/1/1/1
2	NAG	С	2	2	-	3/6/23/26	0/1/1/1
2	BMA	С	3	2	-	1/2/19/22	0/1/1/1
2	NAG	D	1	$1,\!2$	-	2/6/23/26	0/1/1/1
2	NAG	D	2	2	-	2/6/23/26	0/1/1/1
2	BMA	D	3	2	-	1/2/19/22	0/1/1/1
3	NAG	Ε	1	3,1	-	3/6/23/26	0/1/1/1
3	NAG	Е	2	3	-	3/6/23/26	0/1/1/1
2	NAG	F	1	1,2	-	4/6/23/26	0/1/1/1
2	NAG	F	2	1,2	-	0/6/23/26	0/1/1/1
2	BMA	F	3	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
3	Ε	2	NAG	C2-N2-C7	5.79	131.14	122.90
2	С	2	NAG	C1-O5-C5	-5.49	104.75	112.19
2	F	2	NAG	O4-C4-C5	-5.21	96.36	109.30
2	F	2	NAG	O4-C4-C3	-4.85	99.15	110.35
2	F	1	NAG	O4-C4-C3	-4.80	99.25	110.35



There are no chirality outliers.

Mol	Chain	Res	Type	Atoms
2	С	2	NAG	C3-C2-N2-C7
2	С	2	NAG	C8-C7-N2-C2
2	С	2	NAG	O7-C7-N2-C2
2	F	1	NAG	C8-C7-N2-C2
2	F	1	NAG	O7-C7-N2-C2

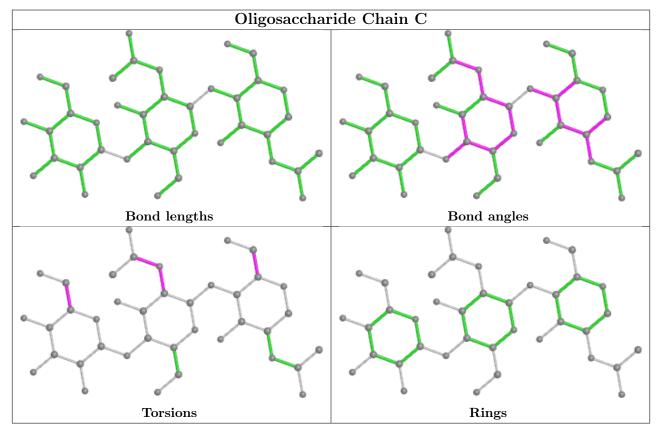
5 of 20 torsion outliers are listed below:

There are no ring outliers.

2 monomers are involved in 6 short contacts:

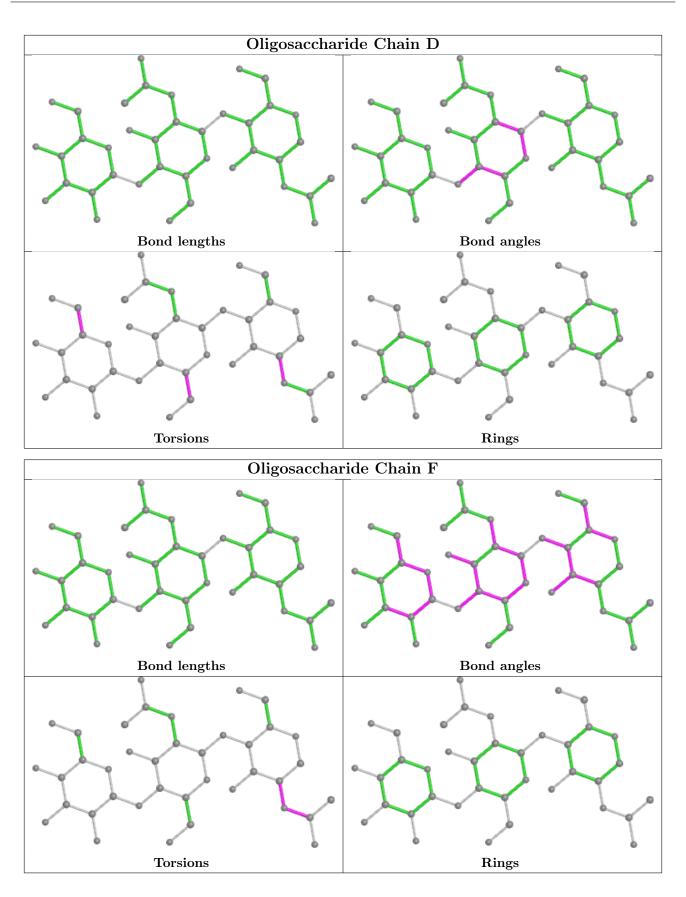
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	Е	2	NAG	2	0
3	Е	1	NAG	4	0

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



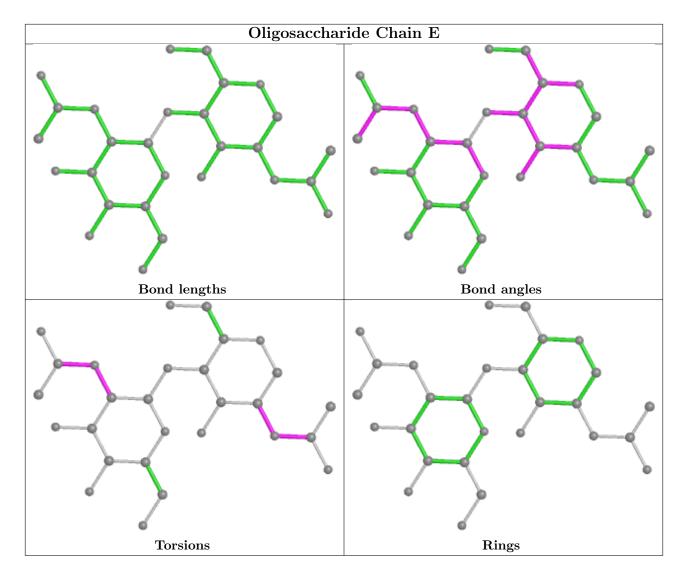












5.6 Ligand geometry (i)

2 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	Turne	Chain	Dec	Link	Bond lengths			Bond angles		
	Type	Chain	Res		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2
4	NAG	В	801	1	14,14,15	0.28	0	17,19,21	0.69	0
4	NAG	А	801	1	14,14,15	0.38	0	17,19,21	1.14	1 (5%)



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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	В	801	1	-	2/6/23/26	0/1/1/1
4	NAG	А	801	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Ι	Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
	4	А	801	NAG	C1-O5-C5	3.51	116.95	112.19

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	А	801	NAG	C3-C2-N2-C7
4	А	801	NAG	C8-C7-N2-C2
4	А	801	NAG	O7-C7-N2-C2
4	В	801	NAG	C4-C5-C6-O6
4	В	801	NAG	O5-C5-C6-O6

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	А	651/732~(88%)	0.16	13 (1%) 65 52	45, 96, 167, 240	0
1	В	651/732~(88%)	0.36	29 (4%) 33 21	71, 119, 193, 278	0
All	All	1302/1464~(88%)	0.26	42 (3%) 47 33	45, 108, 184, 278	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	668	PHE	6.2
1	А	423	ASN	4.8
1	В	56	LEU	4.6
1	В	700	VAL	4.5
1	В	337	ASP	4.4

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	B-factors(Å ²)	Q < 0.9
2	BMA	D	3	11/12	0.56	0.23	127,132,137,142	0
3	NAG	Е	2	14/15	0.66	0.36	$145,\!155,\!158,\!159$	0
2	NAG	F	2	14/15	0.74	0.28	149,159,164,165	0
2	BMA	F	3	11/12	0.77	0.24	162, 165, 167, 169	0
3	NAG	Е	1	14/15	0.79	0.39	140,154,163,164	0

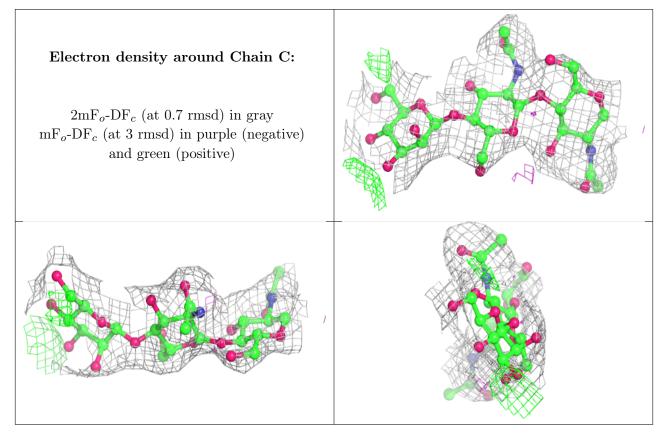
Continued on next page...



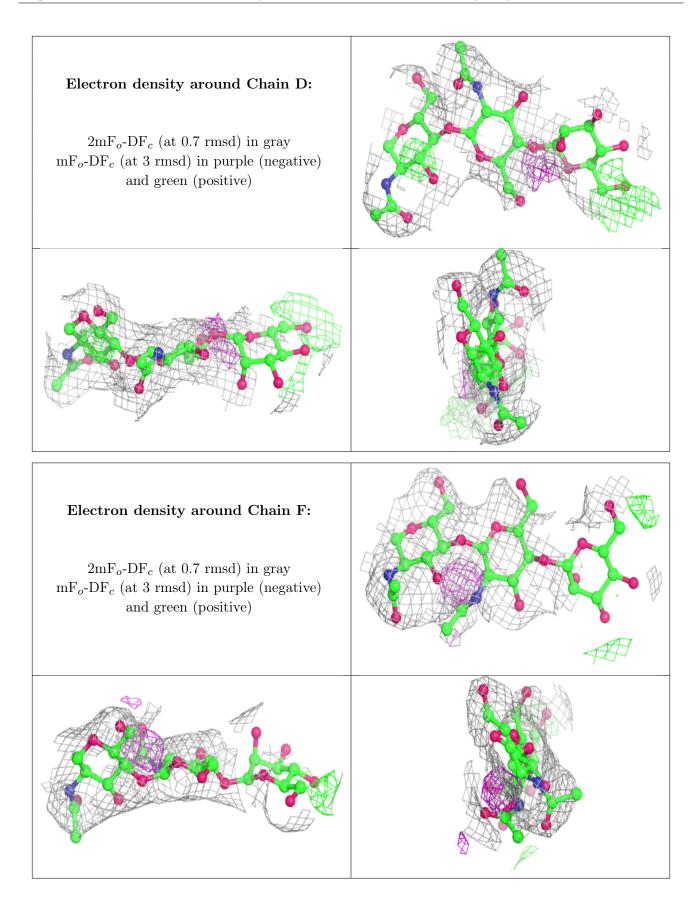
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\operatorname{B-factors}(\operatorname{\AA}^2)$	Q < 0.9
2	BMA	С	3	11/12	0.80	0.30	$142,\!149,\!153,\!153$	0
2	NAG	С	1	14/15	0.89	0.27	87,102,111,117	0
2	NAG	D	1	14/15	0.90	0.19	92,101,109,113	0
2	NAG	F	1	14/15	0.90	0.20	$107,\!124,\!135,\!143$	0
2	NAG	С	2	14/15	0.91	0.18	92,116,126,134	0
2	NAG	D	2	14/15	0.92	0.18	84,109,116,121	0

Continued from previous page...

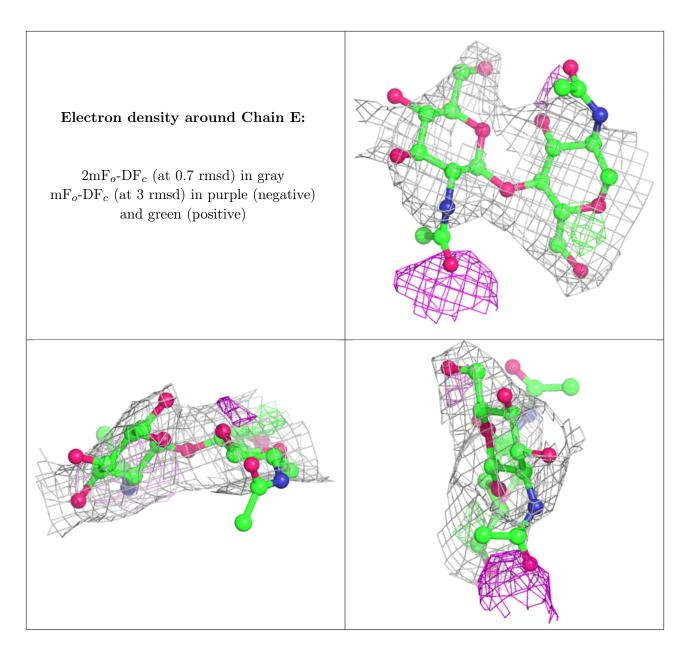
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.











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	B-factors(Å ²)	Q < 0.9
4	NAG	А	801	14/15	0.84	0.24	$149,\!157,\!170,\!172$	0
4	NAG	В	801	14/15	0.87	0.19	110,118,125,125	0



6.5 Other polymers (i)

There are no such residues in this entry.

