wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 2, 2023 – 12:06 PM EDT

PDB ID	:	1ENM
Title	:	UDA TRISACCHARIDE COMPLEX. CRYSTAL STRUCTURE OF UR-TICA DIOICA AGGLUTININ, A SUPERANTIGEN PRESENTED BY MHC MOLECULES OF CLASS I AND CLASS II
Authors	:	Saul, F.A.; Rovira, P.; Boulot, G.; Van Damme, E.J.M.; Peumans, W.J.; Truffa-Bachi, P.; Bentley, G.A.
Deposited on	:	2000-03-21
Resolution	:	1.90 Å(reported)

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

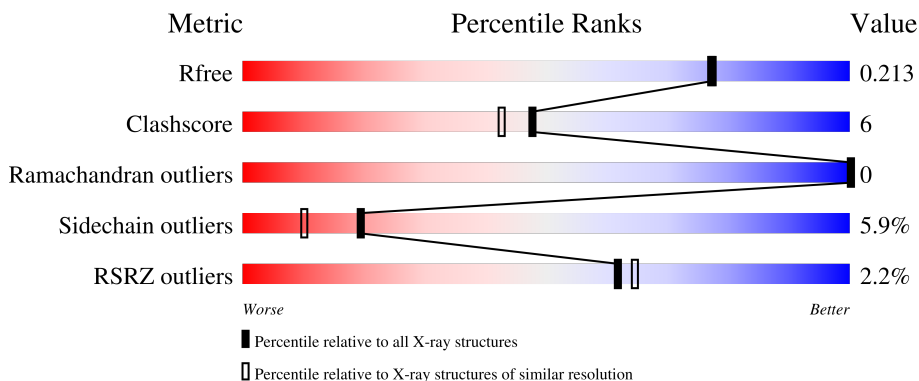|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.34
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

1 Overall quality at a glance ⓘ

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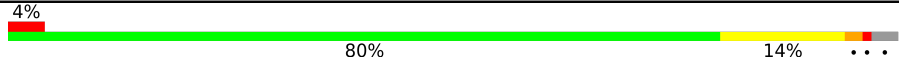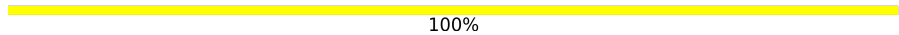*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	Percentile Ranks	Value
Rfree		0.213
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		5.9%
RSRZ outliers		2.2%

Worse   *Better*

▌ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 chain
1	A	94	4% 80% 14% · · ·
2	B	3	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	PCA	A	1	-	X	-	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 750 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 AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/ AGGLUTININ ISOLECTIN VI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	86	671	393	134	128	16	1	7	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	SER	GLY	microheterogeneity	GB 4138900
A	14	GLY	ALA	microheterogeneity	GB 4138900
A	16	ARG	TRP	microheterogeneity	GB 4138900
A	80	GLY	SER	microheterogeneity	GB 4138900
A	81	ASN	LYS	microheterogeneity	GB 4138900

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.
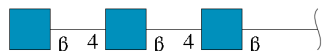
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	3	42	24	3	15	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	37	37	37	0	0

3 Residue-property plots ⓘ

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: AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/ AGGLUTININ ISOLECTIN VI

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

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	38.78Å 46.16Å 57.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 1.90 26.36 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.9 (25.00-1.90) 97.0 (26.36-1.80)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 1.80Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.201 , 0.240 0.186 , 0.213	Depositor DCC
R_{free} test set	423 reflections (4.37%)	wwPDB-VP
Wilson B-factor (Å^2)	21.0	Xtriage
Anisotropy	0.285	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 39.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	750	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.23% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical 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.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.95	4/681 (0.6%)	1.76	10/909 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	79[A]	GLY	C-N	-7.97	1.18	1.33
1	A	79[B]	GLY	C-N	-7.97	1.18	1.33
1	A	79[A]	GLY	N-CA	-5.22	1.38	1.46
1	A	79[B]	GLY	N-CA	-5.22	1.38	1.46

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	79[A]	GLY	O-C-N	-18.18	92.28	123.20
1	A	79[B]	GLY	O-C-N	-18.18	92.28	123.20
1	A	79[A]	GLY	C-N-CA	17.47	158.99	122.30
1	A	79[B]	GLY	C-N-CA	17.47	158.99	122.30
1	A	79[A]	GLY	CA-C-N	14.29	144.78	116.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	79[A]	GLY	Mainchain,Peptide
1	A	81[B]	LYS	Mainchain

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	671	0	560	4	0
2	B	42	0	36	0	0
3	A	37	0	0	0	0
All	All	750	0	596	4	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 6.

All (4) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:78:SER:O	1:A:79[A]:GLY:C	2.54	0.47
1:A:55:ASN:N	1:A:56:PRO:CD	2.80	0.45
1:A:15:LEU:HD22	1:A:15:LEU:N	2.33	0.43
1:A:2:ARG:HD3	1:A:23:TRP:CH2	2.56	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	81/94 (86%)	81 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

5.3.2   Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	69/71 (97%)	65 (94%)	4 (6%)	20	10

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

Mol	Chain	Res	Type
1	A	10[A]	SER
1	A	16[A]	ARG
1	A	16[B]	TRP
1	A	80[B]	SER

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

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
1	PCA	A	1	1	7,8,9	5.18	4 (57%)	9,10,12	6.84	8 (88%)

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
1	PCA	A	1	1	-	0/0/11/13	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	PCA	CA-N	-11.93	1.32	1.46
1	A	1	PCA	OE-CD	-4.55	1.14	1.23
1	A	1	PCA	CG-CD	3.72	1.60	1.50
1	A	1	PCA	CD-N	-2.01	1.28	1.34

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1	PCA	OE-CD-CG	-14.40	101.64	126.76
1	A	1	PCA	OE-CD-N	11.50	151.67	124.86
1	A	1	PCA	CA-N-CD	5.04	130.84	113.58
1	A	1	PCA	O-C-CA	-4.02	114.26	124.78
1	A	1	PCA	CB-CA-C	-3.90	107.34	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

3 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	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	NAG	B	1	2	14,14,15	1.35	1 (7%)	17,19,21	0.86	0
2	NAG	B	2	2	14,14,15	1.30	2 (14%)	17,19,21	1.12	1 (5%)
2	NAG	B	3	2	14,14,15	1.89	5 (35%)	17,19,21	2.20	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.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	1	2	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	NAG	B	3	2	-	0/6/23/26	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	NAG	O7-C7	-3.62	1.15	1.23
2	B	3	NAG	O7-C7	-3.49	1.15	1.23
2	B	2	NAG	O7-C7	-3.19	1.16	1.23
2	B	3	NAG	O5-C1	2.92	1.48	1.43
2	B	3	NAG	C1-C2	2.65	1.56	1.52

The worst 5 of 7 bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	3	NAG	C4-C3-C2	-4.55	104.36	111.02
2	B	3	NAG	O5-C1-C2	-4.37	104.39	111.29
2	B	3	NAG	C1-O5-C5	-3.20	107.86	112.19
2	B	3	NAG	C2-N2-C7	-2.54	119.29	122.90
2	B	2	NAG	O5-C1-C2	-2.37	107.54	111.29

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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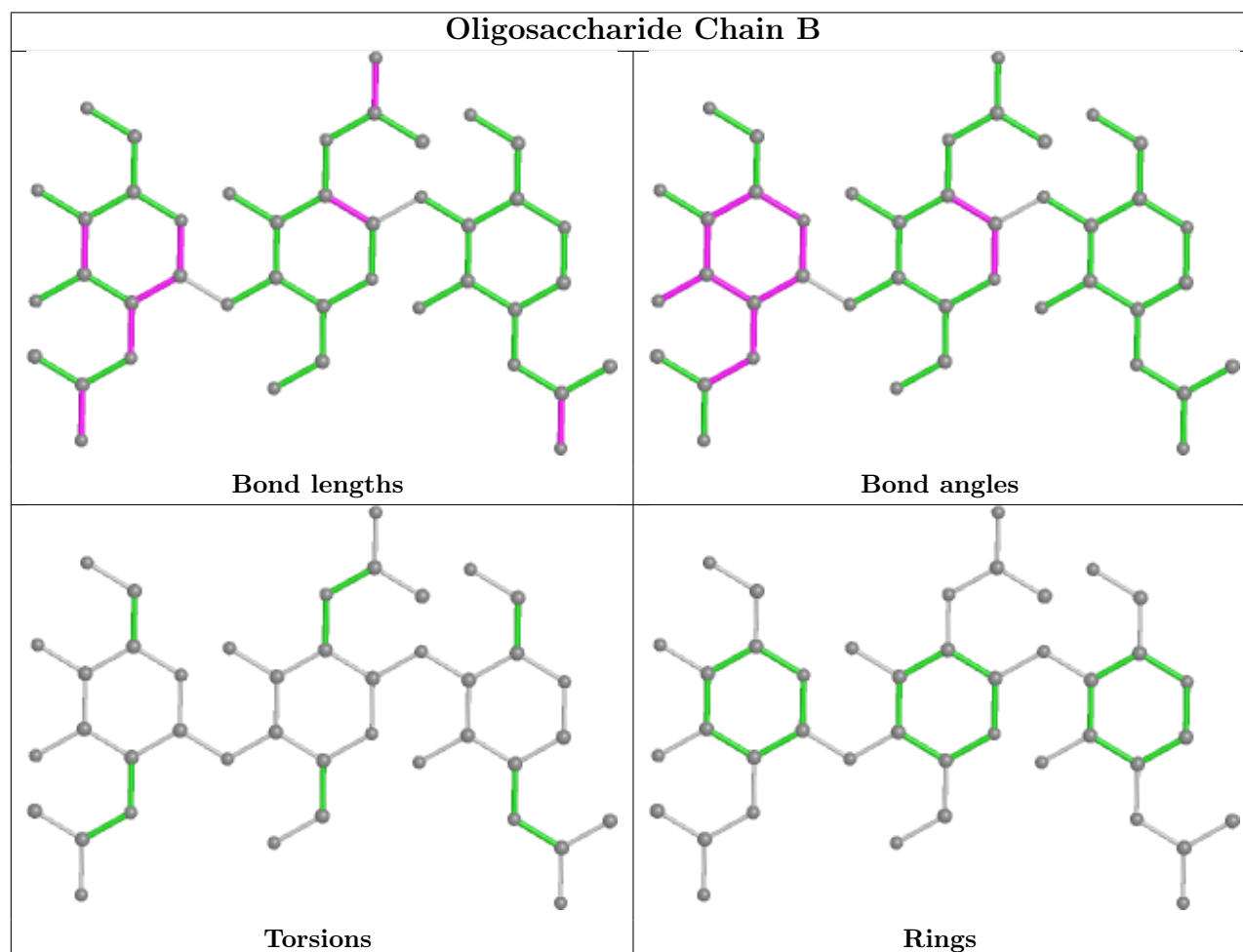
5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	79[A]:GLY	C	80[A]:GLY	N	1.18
1	A	79[B]:GLY	C	80[B]:SER	N	0.74

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	90/94 (95%)	-0.21	2 (2%)	62	64	16, 28, 45, 50	10 (11%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	80[A]	GLY	1.6
1	A	80[A]	GLY	1.6

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

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, 95th 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.
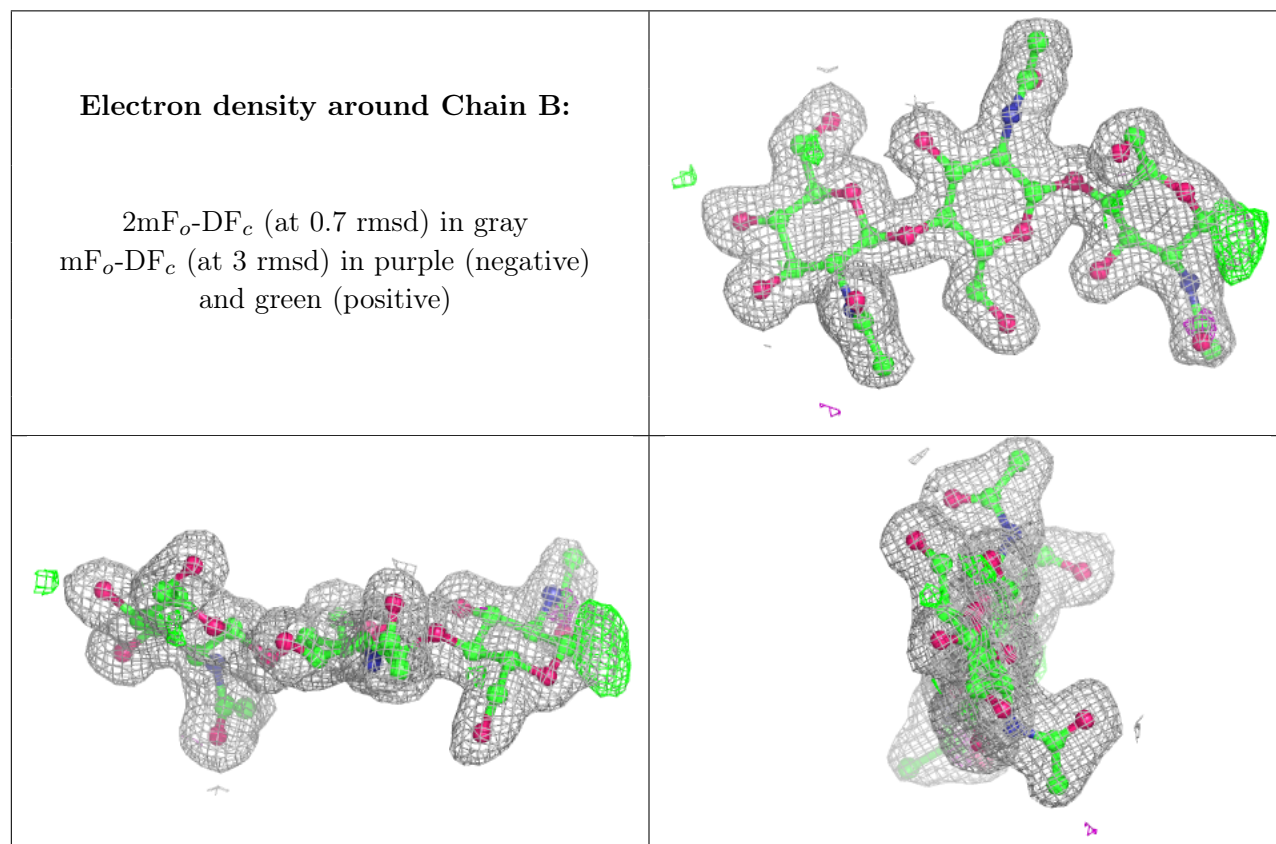
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	PCA	A	1	8/9	0.97	0.07	18,23,27,31	0

6.3 Carbohydrates ⓘ

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, 95th 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	NAG	B	1	14/15	0.89	0.10	25,33,42,42	0
2	NAG	B	2	14/15	0.97	0.08	17,21,25,29	0
2	NAG	B	3	14/15	0.97	0.09	19,20,22,23	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.

Electron density around Chain B:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.