



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 29, 2026 – 01:25 pm BST

PDB ID : 9SUQ / pdb_00009suq
Title : Crystal structure of the zymogen of cathepsin D from *Schistosoma mansoni*
Authors : Busa, M.; Brynda, J.; Houstecka, R.; Hanova, I.; Mares, M.
Deposited on : 2025-09-29
Resolution : 2.13 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 Phenix2.2
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

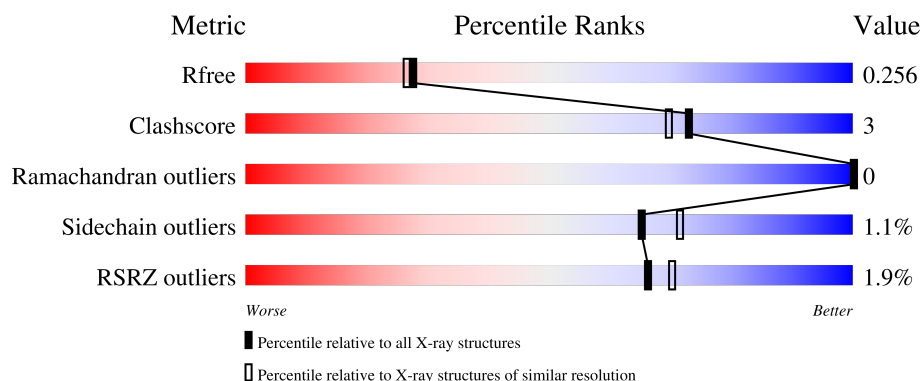
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.13 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.



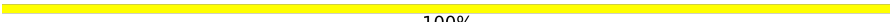
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	3689 (2.16-2.12)
Clashscore	190562	3812 (2.16-2.12)
Ramachandran outliers	187476	3773 (2.16-2.12)
Sidechain outliers	187428	3772 (2.16-2.12)
RSRZ outliers	180081	3691 (2.16-2.12)

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 $\leq 5\%$. 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	434	% 74% 10% • 16%
1	B	434	2% 74% 9% • 15%
2	C	5	100%
3	D	2	100%
3	E	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	2	 100%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6159 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.

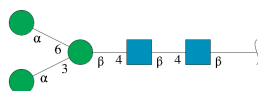
- Molecule 1 is a protein called Cathepsin D (A01 family).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	366	Total	C	N	O	S	0	2	0
			2832	1833	450	533	16			
1	B	367	Total	C	N	O	S	0	0	0
			2819	1823	448	532	16			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	416	HIS	-	expression tag	UNP G4VEV6
A	417	HIS	-	expression tag	UNP G4VEV6
A	418	HIS	-	expression tag	UNP G4VEV6
A	419	HIS	-	expression tag	UNP G4VEV6
A	420	HIS	-	expression tag	UNP G4VEV6
A	421	HIS	-	expression tag	UNP G4VEV6
B	416	HIS	-	expression tag	UNP G4VEV6
B	417	HIS	-	expression tag	UNP G4VEV6
B	418	HIS	-	expression tag	UNP G4VEV6
B	419	HIS	-	expression tag	UNP G4VEV6
B	420	HIS	-	expression tag	UNP G4VEV6
B	421	HIS	-	expression tag	UNP G4VEV6

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



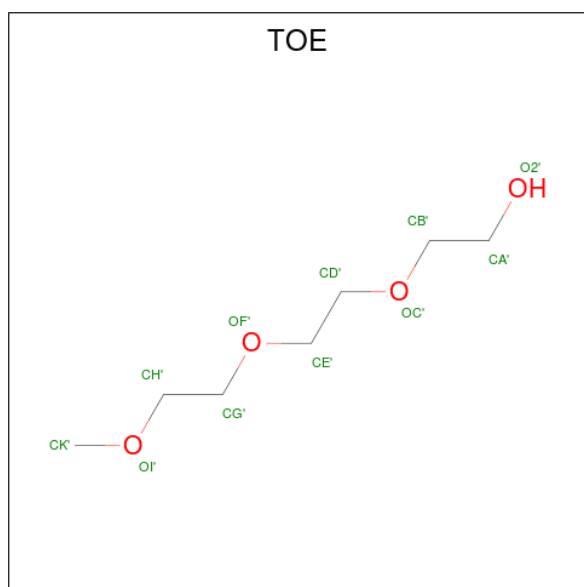
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	0	0	0
			61	34	2	25			

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



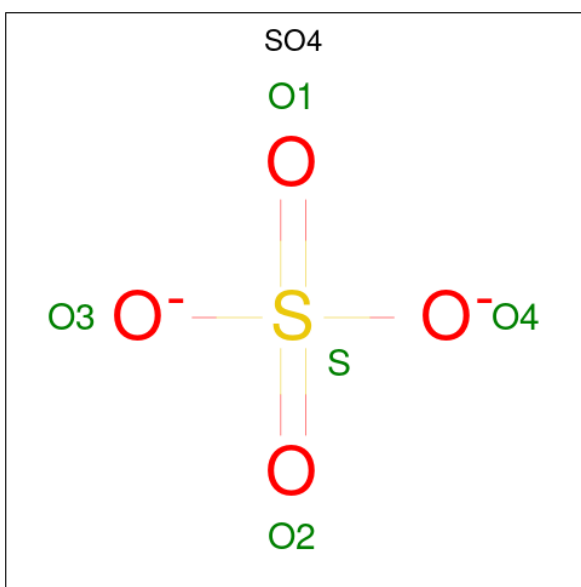
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	F	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is 2-[2-(2-METHOXY-ETHOXY)-ETHOXY]-ETHOXYL (CCD ID: TOE) (formula: C₇H₁₆O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	7	4		

- Molecule 5 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		

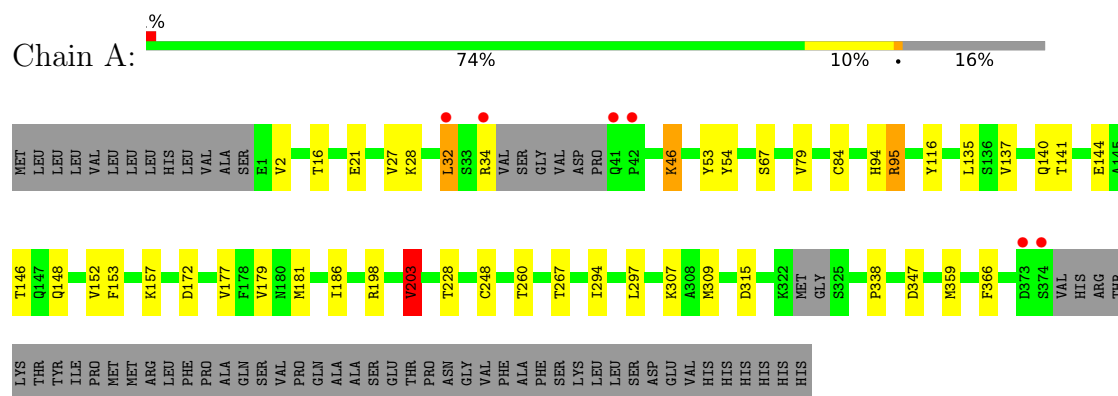
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	184	Total	O	0	0
			184	184		
6	B	158	Total	O	0	0
			158	158		

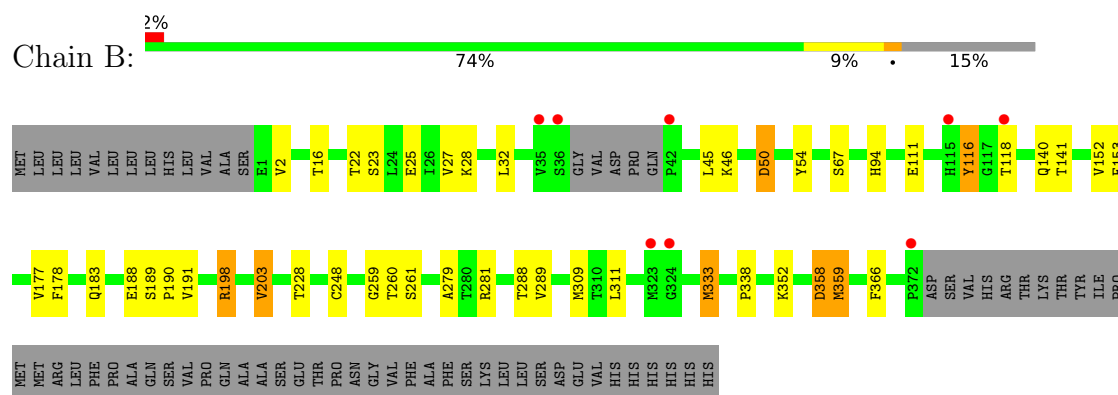
3 Residue-property plots

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.

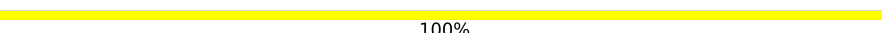
• Molecule 1: Cathepsin D (A01 family)



• Molecule 1: Cathepsin D (A01 family)



• Molecule 2: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C: 



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

Chain D:  100%

MAG1
MAG2

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

Chain E:  100%

MAG1
MAG2

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

Chain F:  100%

MAG1
MAG2

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	65.73Å 112.81Å 114.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.11 – 2.13 43.11 – 2.13	Depositor EDS
% Data completeness (in resolution range)	99.5 (43.11-2.13) 99.5 (43.11-2.13)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.73 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0431 (refmacat 0.4.105)	Depositor
R, R_{free}	0.200 , 0.256 0.200 , 0.256	Depositor DCC
R_{free} test set	2100 reflections (4.37%)	wwPDB-VP
Wilson B-factor (Å ²)	24.8	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 33.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.025 for -h,l,k	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6159	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, TOE, SO4, BMA, MAN

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	0.76	0/2903	1.38	25/3936 (0.6%)
1	B	0.73	0/2884	1.39	24/3911 (0.6%)
All	All	0.74	0/5787	1.38	49/7847 (0.6%)

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	1

There are no bond length outliers.

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	228	THR	CA-CB-OG1	-8.12	97.42	109.60
1	A	152	VAL	N-CA-CB	8.09	121.55	110.54
1	A	46	LYS	CB-CG-CD	7.90	129.46	111.30
1	A	228	THR	CA-CB-OG1	-7.78	97.94	109.60
1	A	203	VAL	N-CA-CB	-6.94	100.75	110.26
1	B	188	GLU	N-CA-CB	6.94	120.32	110.12
1	A	27	VAL	N-CA-CB	6.90	118.17	110.51
1	B	116	TYR	CA-C-N	6.66	127.49	120.03
1	B	116	TYR	C-N-CA	6.66	127.49	120.03
1	B	203	VAL	N-CA-CB	-6.63	101.17	110.26
1	B	50	ASP	CB-CA-C	-6.63	98.12	109.65
1	A	157	LYS	CB-CA-C	6.59	123.75	109.99
1	A	267	THR	CA-CB-OG1	-6.53	99.81	109.60
1	B	141	THR	CA-CB-OG1	-6.38	100.04	109.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	16	THR	CA-CB-OG1	-6.22	100.26	109.60
1	B	338	PRO	CB-CA-C	6.17	121.12	112.11
1	B	116	TYR	N-CA-CB	6.05	118.89	110.17
1	A	2	VAL	N-CA-CB	5.98	117.94	110.05
1	A	338	PRO	CB-CA-C	5.85	120.65	112.11
1	B	198	ARG	CA-CB-CG	5.85	125.81	114.10
1	A	116	TYR	N-CA-CB	5.84	118.58	110.17
1	B	359	MET	CA-C-N	5.78	126.35	119.94
1	B	359	MET	C-N-CA	5.78	126.35	119.94
1	A	307	LYS	CB-CG-CD	5.69	124.38	111.30
1	B	27	VAL	N-CA-CB	5.64	116.77	110.51
1	B	288	THR	CA-CB-OG1	-5.58	101.23	109.60
1	A	172	ASP	CA-CB-CG	5.56	118.16	112.60
1	A	347	ASP	CA-CB-CG	5.55	118.15	112.60
1	A	315	ASP	CA-CB-CG	5.54	118.14	112.60
1	B	190	PRO	CB-CA-C	5.53	119.95	112.04
1	B	152	VAL	N-CA-CB	5.46	118.76	110.58
1	B	111	GLU	CB-CG-CD	5.39	121.76	112.60
1	B	281	ARG	CD-NE-CZ	5.39	131.94	124.40
1	A	148	GLN	CB-CA-C	-5.37	103.37	111.27
1	B	22	THR	CA-CB-OG1	-5.32	101.62	109.60
1	A	34	ARG	CD-NE-CZ	5.28	131.79	124.40
1	B	333	MET	CG-SD-CE	5.27	112.50	100.90
1	B	359	MET	CB-CG-SD	5.25	128.46	112.70
1	A	141	THR	CA-CB-OG1	-5.24	101.73	109.60
1	A	203	VAL	CB-CA-C	5.22	118.75	111.34
1	A	79	VAL	N-CA-CB	5.19	116.08	111.67
1	A	307	LYS	CB-CA-C	5.19	119.21	109.71
1	A	21	GLU	CB-CG-CD	-5.17	103.81	112.60
1	A	16	THR	CA-CB-OG1	-5.05	102.02	109.60
1	A	144	GLU	CB-CA-C	5.03	118.83	110.78
1	A	116	TYR	CA-C-N	5.02	126.41	120.14
1	A	116	TYR	C-N-CA	5.02	126.41	120.14
1	B	358	ASP	CA-CB-CG	5.02	117.62	112.60
1	B	260	THR	OG1-CB-CG2	-5.01	99.29	109.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	95	ARG	Sidechain

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	2832	0	2823	15	0
1	B	2819	0	2805	20	0
2	C	61	0	52	0	0
3	D	28	0	25	0	0
3	E	28	0	25	0	0
3	F	28	0	25	0	0
4	A	11	0	16	2	0
5	A	5	0	0	0	0
5	B	5	0	0	0	0
6	A	184	0	0	2	0
6	B	158	0	0	4	0
All	All	6159	0	5771	37	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 3.

All (37) 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:B:183:GLN:HB3	6:B:733:HOH:O	1.77	0.81
4:A:501:TOE:H3	4:A:501:TOE:H16	1.63	0.79
1:A:28:LYS:HA	1:A:32:LEU:HB2	1.88	0.55
1:A:46:LYS:HD3	1:A:260:THR:HB	1.89	0.55
1:B:309:MET:HE2	1:B:366:PHE:HB2	1.90	0.54
1:A:84:CYS:HA	1:A:146:THR:HA	1.92	0.52
1:A:94:HIS:HE1	1:A:153:PHE:O	1.92	0.51
4:A:501:TOE:H3	4:A:501:TOE:CK'	2.38	0.51
1:B:46:LYS:HE3	1:B:259:GLY:HA3	1.92	0.50
1:B:116:TYR:HB2	1:B:118:THR:HG22	1.94	0.50
1:A:54:TYR:CD1	1:A:67:SER:HB2	2.47	0.49
1:B:311:LEU:HD23	1:B:352:LYS:HG2	1.96	0.48
1:A:84:CYS:HB2	6:A:603:HOH:O	2.14	0.47
1:B:54:TYR:CD1	1:B:67:SER:HB2	2.50	0.46
1:A:294:ILE:HA	1:A:297:LEU:HD12	1.96	0.46
1:A:95:ARG:NH2	6:A:604:HOH:O	2.49	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:50:ASP:HB3	1:B:261:SER:H	1.81	0.45
1:B:309:MET:HE1	6:B:652:HOH:O	2.15	0.45
1:A:137:VAL:CG2	1:A:181:MET:HG2	2.46	0.45
1:A:248:CYS:SG	1:A:248:CYS:O	2.74	0.45
1:B:191:VAL:HG12	1:B:358:ASP:HA	1.98	0.45
1:B:333:MET:HB3	6:B:748:HOH:O	2.17	0.44
1:A:140:GLN:OE1	1:A:177:VAL:HA	2.18	0.44
1:B:248:CYS:SG	1:B:248:CYS:O	2.76	0.43
1:A:179:VAL:HG22	1:A:359:MET:HE1	2.00	0.43
1:B:28:LYS:HA	1:B:32:LEU:HB2	2.00	0.42
1:B:94:HIS:HE1	1:B:153:PHE:O	2.02	0.42
1:B:178:PHE:HD2	1:B:359:MET:HE3	1.85	0.42
1:B:140:GLN:OE1	1:B:177:VAL:HA	2.20	0.42
1:B:279:ALA:HB2	1:B:289:VAL:HG13	2.01	0.42
1:B:45:LEU:HD23	1:B:116:TYR:HA	2.01	0.41
1:B:309:MET:HE2	1:B:366:PHE:CB	2.51	0.41
1:B:359:MET:HE1	6:B:634:HOH:O	2.21	0.41
1:A:53:TYR:CZ	1:A:203:VAL:HG22	2.56	0.41
1:A:135:LEU:HD22	1:A:186:ILE:HG22	2.02	0.41
1:A:309:MET:HE2	1:A:366:PHE:HB2	2.03	0.40
1:B:23:SER:OG	1:B:25:GLU:HG2	2.22	0.40

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	Percentiles	
1	A	362/434 (83%)	357 (99%)	5 (1%)	0	100	100
1	B	363/434 (84%)	355 (98%)	8 (2%)	0	100	100
All	All	725/868 (84%)	712 (98%)	13 (2%)	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	314/376 (84%)	311 (99%)	3 (1%)	68	74
1	B	311/376 (83%)	307 (99%)	4 (1%)	61	67
All	All	625/752 (83%)	618 (99%)	7 (1%)	65	71

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	32	LEU
1	A	198	ARG
1	A	203	VAL
1	B	2	VAL
1	B	189	SER
1	B	198	ARG
1	B	203	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	131	GLN
1	A	296	ASN
1	A	362	ASN
1	B	148	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	NAG	C	1	1,2	14,14,15	0.43	0	17,19,21	1.62	3 (17%)
2	NAG	C	2	2	14,14,15	0.56	0	17,19,21	1.31	1 (5%)
2	BMA	C	3	2	11,11,12	1.04	1 (9%)	15,15,17	0.99	1 (6%)
2	MAN	C	4	2	11,11,12	0.95	1 (9%)	15,15,17	1.44	2 (13%)
2	MAN	C	5	2	11,11,12	1.18	1 (9%)	15,15,17	1.65	3 (20%)
3	NAG	D	1	3,1	14,14,15	0.60	0	17,19,21	1.44	2 (11%)
3	NAG	D	2	3	14,14,15	0.61	0	17,19,21	1.35	5 (29%)
3	NAG	E	1	3,1	14,14,15	0.56	0	17,19,21	1.44	3 (17%)
3	NAG	E	2	3	14,14,15	0.52	0	17,19,21	1.46	2 (11%)
3	NAG	F	1	3,1	14,14,15	0.68	0	17,19,21	1.51	2 (11%)
3	NAG	F	2	3	14,14,15	0.56	0	17,19,21	1.95	4 (23%)

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 Torsions and Rings columns. '-' means no outliers of that kind were identified.

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

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	D	2	3	-	0/6/23/26	0/1/1/1
3	NAG	E	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1
3	NAG	F	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	2/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	3	BMA	O5-C5	2.70	1.48	1.43
2	C	5	MAN	O5-C5	2.58	1.48	1.43
2	C	4	MAN	C4-C5	2.06	1.57	1.53

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	2	NAG	C2-N2-C7	5.65	130.95	122.90
2	C	2	NAG	C2-N2-C7	4.18	128.86	122.90
3	E	2	NAG	C1-O5-C5	4.02	117.64	112.19
3	E	1	NAG	C2-N2-C7	3.99	128.58	122.90
2	C	4	MAN	C1-O5-C5	3.85	117.40	112.19
3	F	2	NAG	C1-O5-C5	3.84	117.40	112.19
3	F	1	NAG	O4-C4-C5	3.76	118.64	109.30
3	D	1	NAG	C1-C2-N2	-3.73	104.11	110.49
2	C	1	NAG	O3-C3-C2	3.65	117.02	109.47
3	E	2	NAG	C2-N2-C7	3.60	128.03	122.90
2	C	5	MAN	C1-O5-C5	3.59	117.06	112.19
2	C	5	MAN	O2-C2-C3	-2.93	104.26	110.14
2	C	4	MAN	O2-C2-C3	-2.79	104.55	110.14
3	D	1	NAG	C4-C3-C2	2.67	114.93	111.02
3	F	1	NAG	O5-C5-C4	-2.64	104.41	110.83
2	C	1	NAG	C1-O5-C5	2.55	115.64	112.19
3	F	2	NAG	C1-C2-N2	-2.48	106.25	110.49
3	D	2	NAG	O4-C4-C5	2.45	115.39	109.30
3	E	1	NAG	C1-O5-C5	2.39	115.44	112.19
3	D	2	NAG	C2-N2-C7	2.34	126.24	122.90
3	D	2	NAG	O5-C5-C4	-2.33	105.17	110.83
2	C	5	MAN	O4-C4-C5	-2.31	103.56	109.30
2	C	1	NAG	C1-C2-N2	-2.30	106.56	110.49
3	F	2	NAG	C4-C3-C2	2.27	114.34	111.02
2	C	3	BMA	O4-C4-C5	-2.26	103.69	109.30
3	D	2	NAG	C6-C5-C4	2.24	118.24	113.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1	NAG	C4-C3-C2	2.19	114.22	111.02
3	D	2	NAG	C1-C2-N2	2.09	114.07	110.49

There are no chirality outliers.

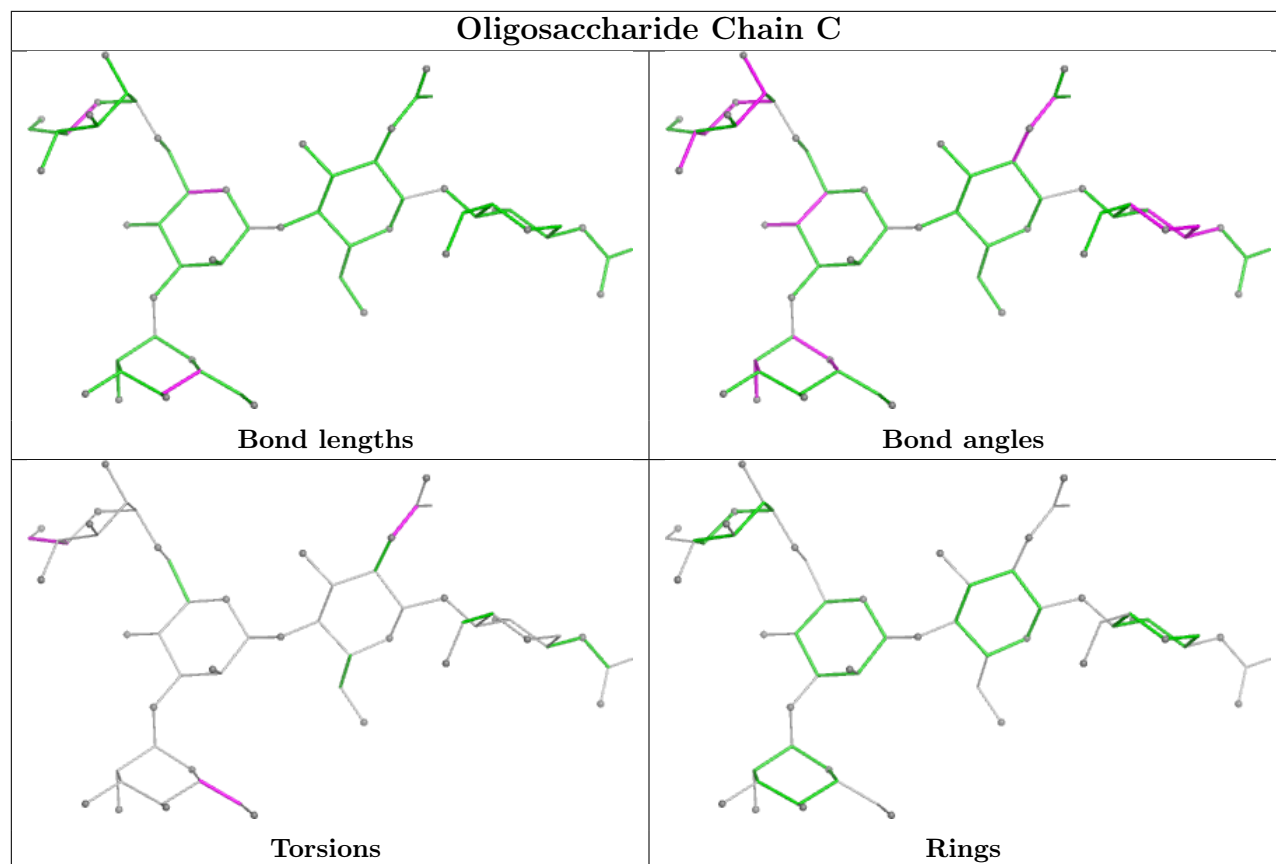
All (8) torsion outliers are listed below:

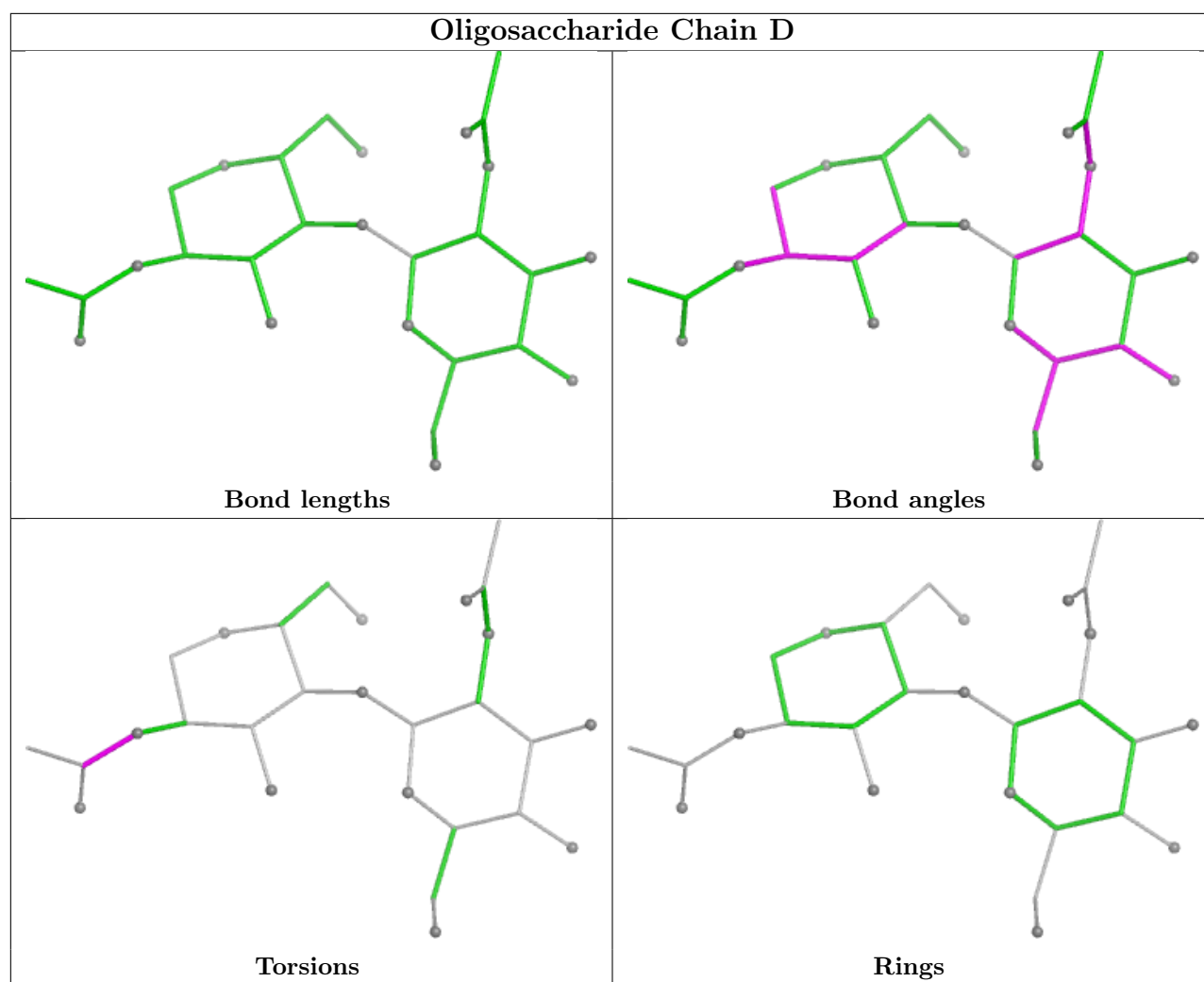
Mol	Chain	Res	Type	Atoms
3	F	2	NAG	C8-C7-N2-C2
2	C	2	NAG	C8-C7-N2-C2
2	C	2	NAG	O7-C7-N2-C2
3	F	2	NAG	O7-C7-N2-C2
3	D	1	NAG	C8-C7-N2-C2
3	D	1	NAG	O7-C7-N2-C2
2	C	5	MAN	O5-C5-C6-O6
2	C	4	MAN	O5-C5-C6-O6

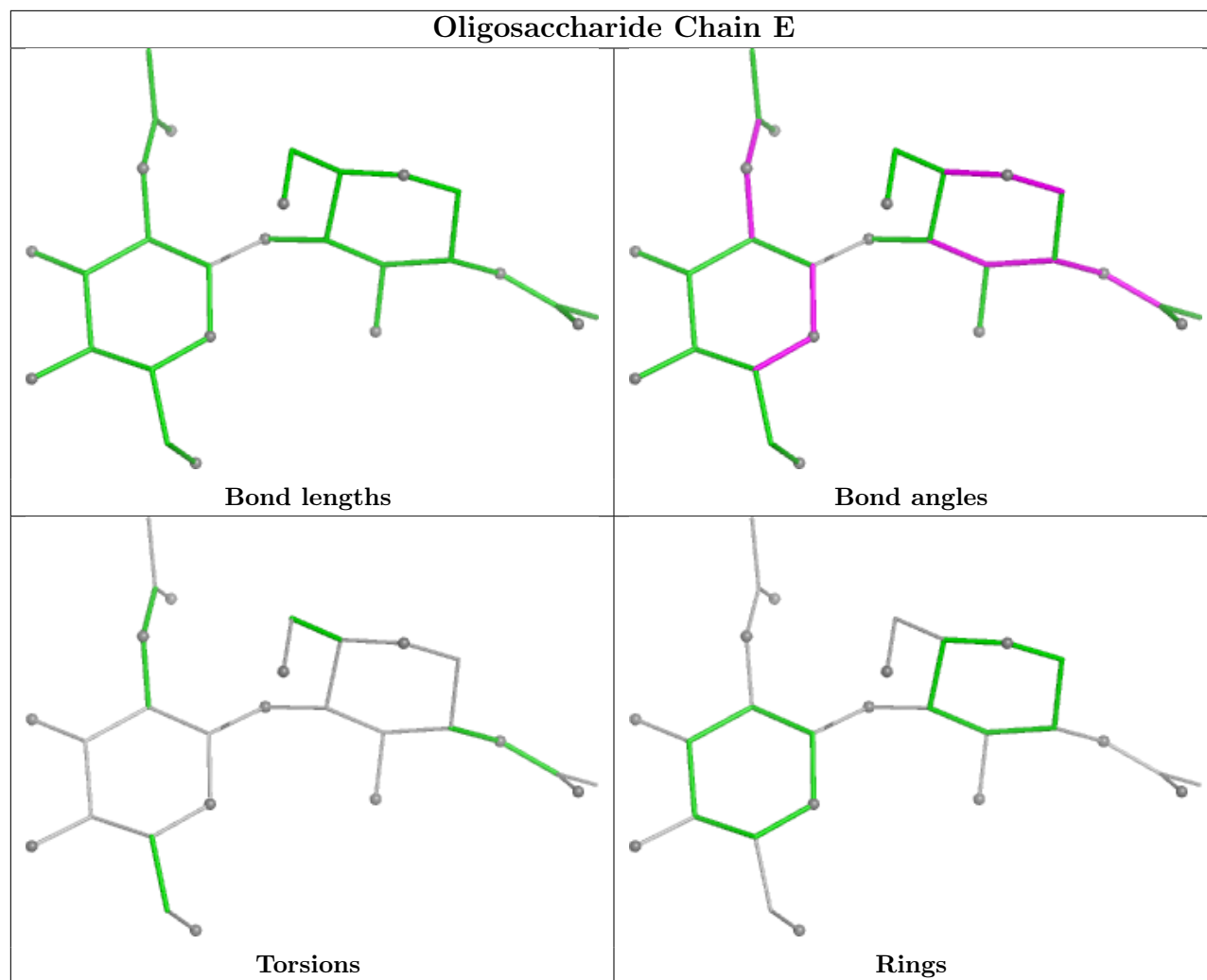
There are no ring outliers.

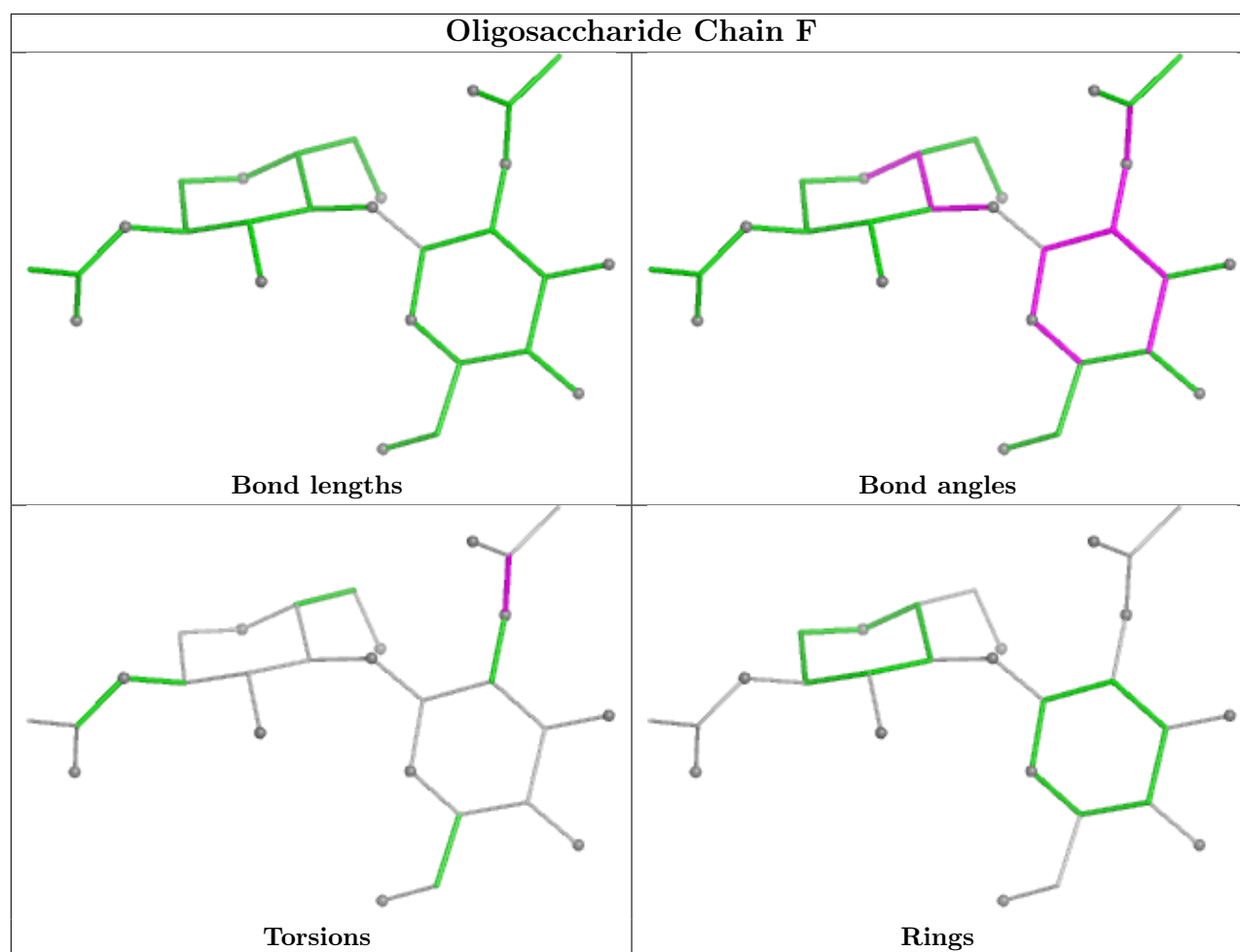
No monomer is involved in short contacts.

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 [i](#)

3 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	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	SO4	A	502	-	4,4,4	0.83	0	6,6,6	0.33	0
4	TOE	A	501	-	10,10,10	0.43	0	9,9,9	0.59	0
5	SO4	B	501	-	4,4,4	0.30	0	6,6,6	0.08	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
4	TOE	A	501	-	-	5/8/8/8	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	501	TOE	O2'-CA'-CB'-OC'
4	A	501	TOE	OF'-CG'-CH'-OI'
4	A	501	TOE	CH'-CG'-OF'-CE'
4	A	501	TOE	CE'-CD'-OC'-CB'
4	A	501	TOE	CD'-CE'-OF'-CG'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	501	TOE	2	0

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, 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	366/434 (84%)	-0.49	6 (1%) 70 74	14, 22, 45, 85	2 (0%)
1	B	367/434 (84%)	-0.29	8 (2%) 62 66	14, 27, 47, 82	0
All	All	733/868 (84%)	-0.39	14 (1%) 66 70	14, 25, 47, 85	2 (0%)

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	374	SER	4.4
1	B	42	PRO	4.4
1	A	42	PRO	3.7
1	B	35	VAL	3.6
1	B	324	GLY	3.3
1	B	36	SER	2.8
1	A	41	GLN	2.8
1	A	373	ASP	2.6
1	A	32	LEU	2.4
1	B	323	MET	2.2
1	B	118	THR	2.2
1	A	34	ARG	2.2
1	B	115	HIS	2.1
1	B	372	PRO	2.1

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

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

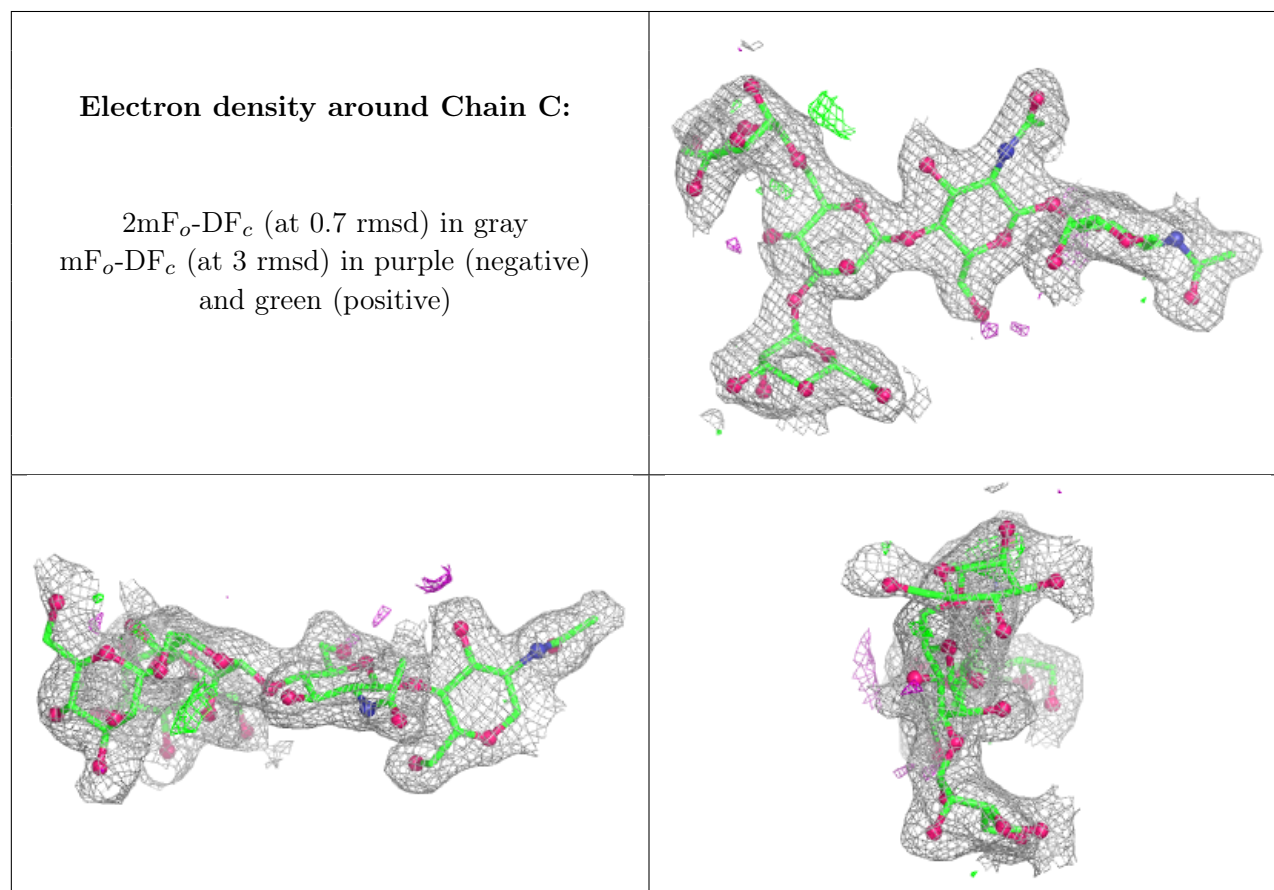
6.3 Carbohydrates [i](#)

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, 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.

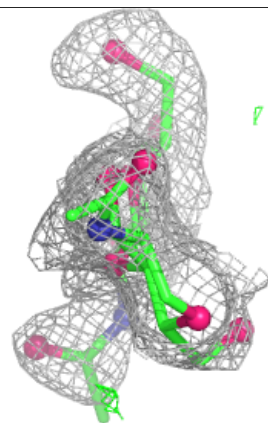
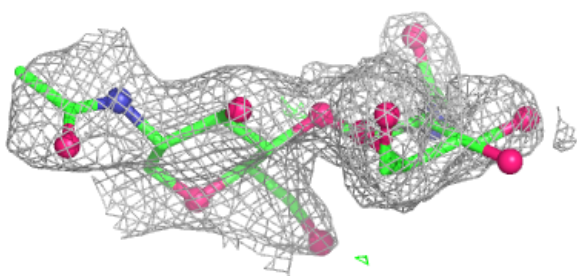
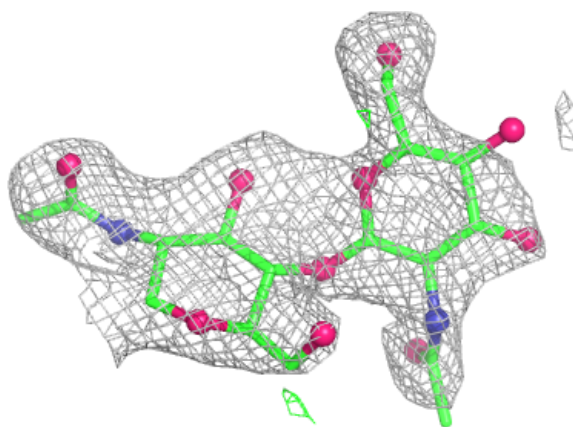
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	NAG	E	2	14/15	0.71	0.13	51,67,77,80	0
3	NAG	D	2	14/15	0.72	0.12	43,63,70,72	0
2	MAN	C	5	11/12	0.73	0.14	52,59,63,64	0
3	NAG	F	2	14/15	0.79	0.12	44,49,54,56	0
3	NAG	D	1	14/15	0.87	0.11	39,48,63,64	0
2	BMA	C	3	11/12	0.87	0.10	31,36,49,53	0
3	NAG	E	1	14/15	0.90	0.08	34,38,41,52	0
3	NAG	F	1	14/15	0.91	0.09	30,40,49,51	0
2	MAN	C	4	11/12	0.91	0.08	31,38,43,51	0
2	NAG	C	2	14/15	0.94	0.08	20,24,30,43	0
2	NAG	C	1	14/15	0.96	0.05	17,21,22,23	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.



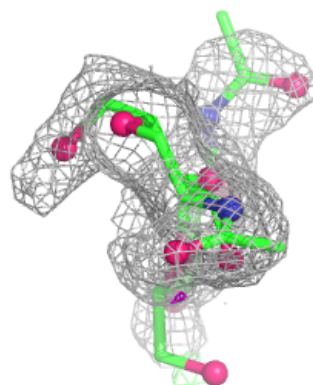
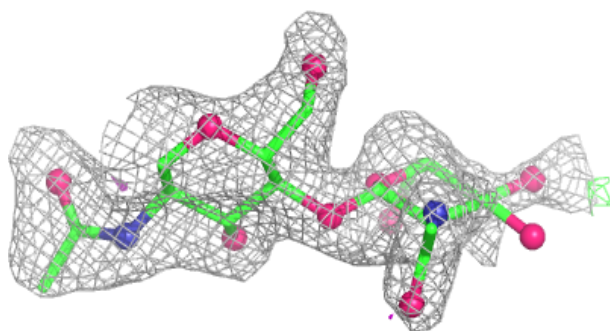
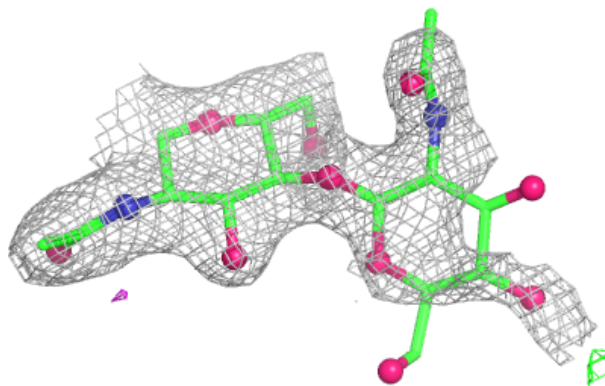
Electron density around Chain D:

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

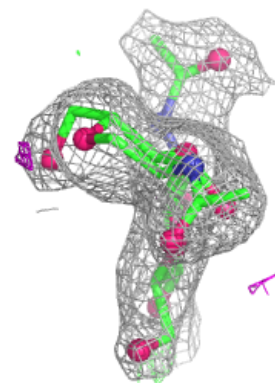
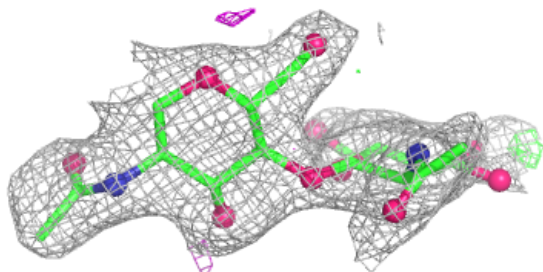
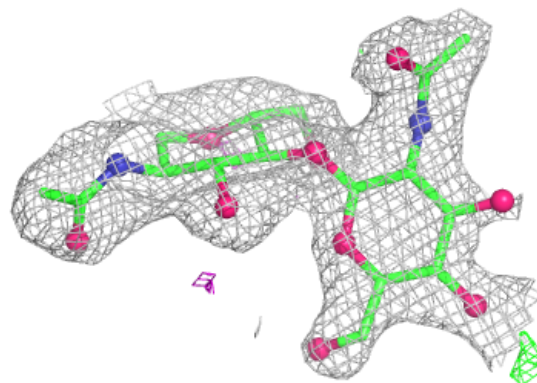


Electron density around Chain E:

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

**Electron density around Chain F:**

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



6.4 Ligands [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, 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
5	SO4	B	501	5/5	0.73	0.14	82,83,90,93	0
4	TOE	A	501	11/11	0.83	0.15	48,55,71,74	0
5	SO4	A	502	5/5	0.98	0.06	26,32,34,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.