Full wwPDB X-ray Structure Validation Report (i)

Jan 22, 2023 – 12:18 am GMT

PDB ID	:	8AI7
Title	:	Structure of carbamoylated human butyrylcholinesterase upon reaction with 3-(((2-cycloheptylethyl)(methyl)amino)methyl)-1H-indol-7-yl N,N-dimethylcarbamate
Authors	:	Brazzolotto, X.; Meden, A.; Knez, D.; Gobec, S.; Nachon, F.
Deposited on	:	2022-07-25
Resolution	:	2.13 Å(reported)

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

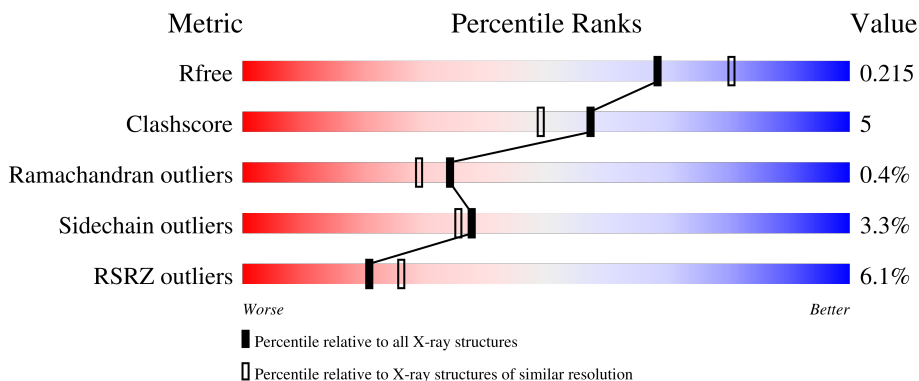|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance ⓘ

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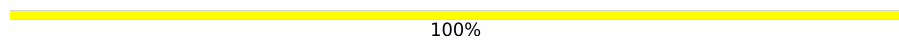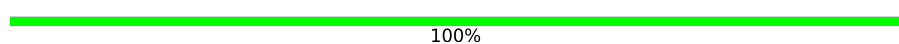*X-RAY DIFFRACTION*

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	Percentile Ranks	Value
Rfree		0.215
Clashscore		5
Ramachandran outliers		0.4%
Sidechain outliers		3.3%
RSRZ outliers		6.1%

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	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (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 <=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	F	529	6% / 85% / 13% / ••
2	B	2	100%
2	C	2	50% / 50%
2	E	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	3	67% 33%

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
2	FUC	B	2	-	-	-	X
2	FUC	C	2	-	-	-	X

2 Entry composition ⓘ

There are 10 unique types of molecules in this entry. The entry contains 4632 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 Cholinesterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	526	Total 4223	C 2726	N 711	O 771	S 15	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	17	GLN	ASN	engineered mutation	UNP P06276
F	455	GLN	ASN	engineered mutation	UNP P06276
F	481	GLN	ASN	engineered mutation	UNP P06276
F	486	GLN	ASN	engineered mutation	UNP P06276

- Molecule 2 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total 24	C 14	N 1	O 9	0	0	0
2	C	2	Total 24	C 14	N 1	O 9	0	0	0
2	E	2	Total 24	C 14	N 1	O 9	0	0	0

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.
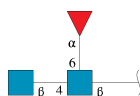
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	3	Total 38	C 22	N 2	O 14	0	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
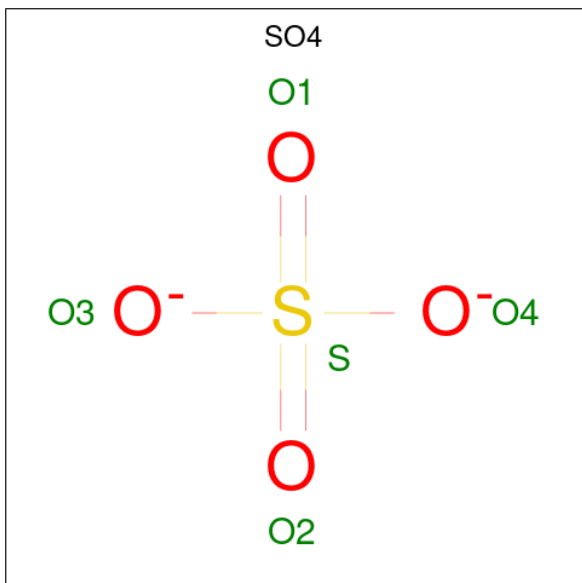
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	F	7	Total 7	Cl 7	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	1	Total 1	Na 1	0	0

- Molecule 7 is 3-[[2-cycloheptylethyl(methyl)amino]methyl]-1 {H}-indol-7-ol (three-letter code: M8X) (formula: $C_{19}H_{28}N_2O$) (labeled as "Ligand of Interest" by depositor).
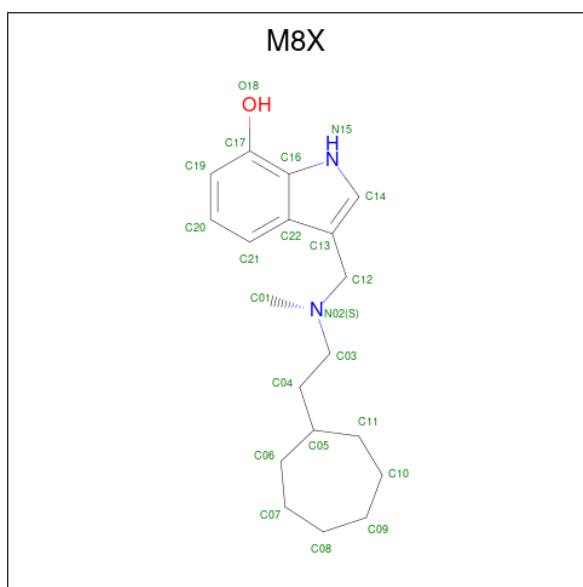
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	F	1	Total 22	C 19	N 2	O 1	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
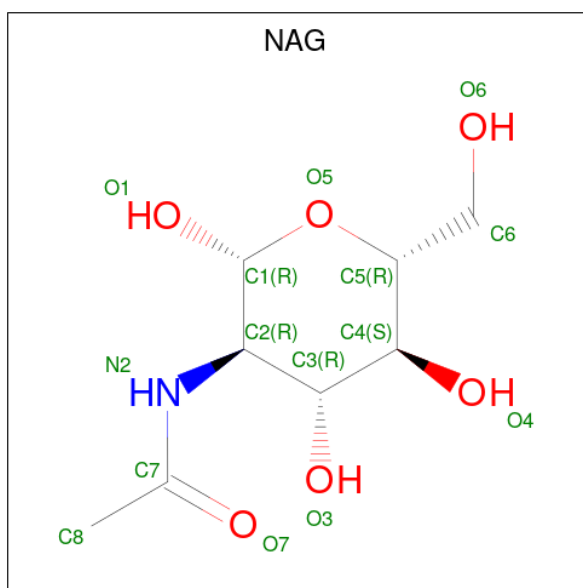
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	F	1	Total 14	C 8	N 1	O 5	0	0
8	F	1	Total 14	C 8	N 1	O 5	0	0

• Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	F	1	Total 6	C 3	O 3	0	0
9	F	1	Total 6	C 3	O 3	0	0
9	F	1	Total 6	C 3	O 3	0	0
9	F	1	Total 6	C 3	O 3	0	0

• Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	F	202	Total 202	O 202	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: Cholinesterase
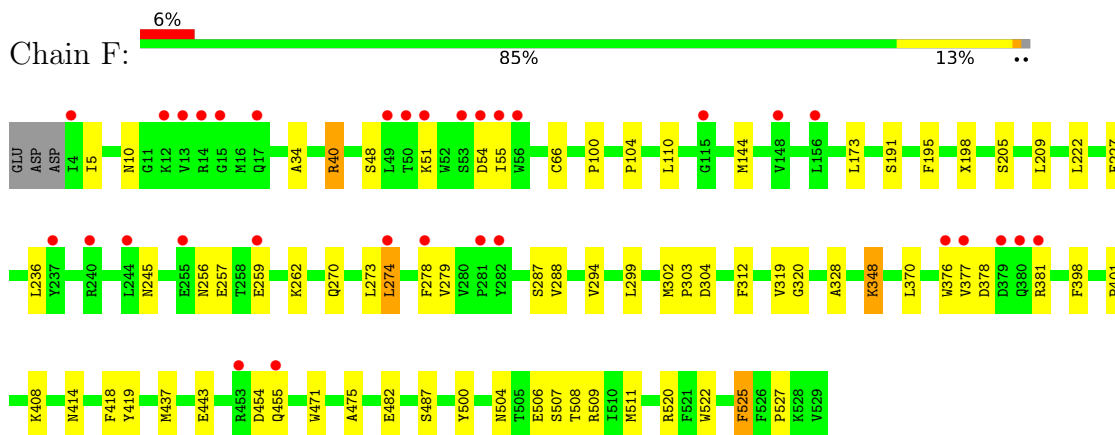
Chain F:

• Molecule 2: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

• Molecule 2: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:
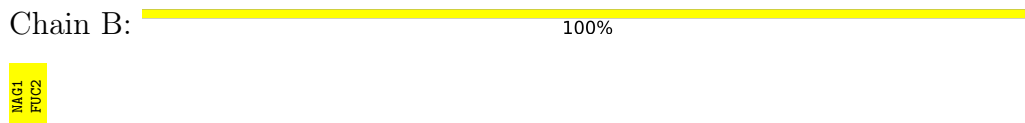
• Molecule 2: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

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

Chain D:

NAG1 NAG2 FUC3

4 Data and refinement statistics (i)

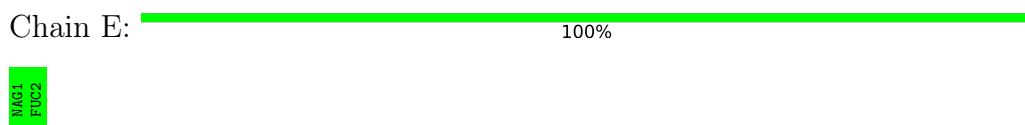| Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	154.46Å 154.46Å 128.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.34 – 2.13 49.34 – 2.13	Depositor EDS
% Data completeness (in resolution range)	98.2 (49.34-2.13) 98.2 (49.34-2.13)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.34 (at 2.14Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.177 , 0.218 0.177 , 0.215	Depositor DCC
R_{free} test set	2057 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	51.8	Xtriage
Anisotropy	0.334	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 60.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4632	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: FUC, GOL, NA, M8X, CL, SO4, MKF, NAG

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	F	0.42	0/4338	0.59	1/5887 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	274	LEU	CB-CG-CD2	5.54	120.41	111.00

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	F	4223	0	4110	43	0
2	B	24	0	22	0	0
2	C	24	0	22	1	0
2	E	24	0	22	0	0
3	D	38	0	34	0	0
4	F	15	0	0	0	0
5	F	7	0	0	0	0
6	F	1	0	0	0	0
7	F	22	0	0	0	0
8	F	28	0	26	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	F	24	0	32	3	0
10	F	202	0	0	4	0
All	All	4632	0	4268	43	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.

All (43) 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:F:348:LYS:H	1:F:348:LYS:HE2	1.48	0.78
1:F:508:THR:HG22	10:F:717:HOH:O	1.82	0.78
1:F:270:GLN:O	1:F:274:LEU:HD22	1.89	0.73
1:F:520:ARG:NH2	10:F:702:HOH:O	2.31	0.63
1:F:5:ILE:HD13	1:F:55:ILE:HD13	1.82	0.62
1:F:304:ASP:HB3	10:F:892:HOH:O	2.00	0.61
1:F:257:GLU:OE1	10:F:701:HOH:O	2.16	0.60
1:F:522:TRP:O	1:F:527:PRO:HD3	2.03	0.58
1:F:348:LYS:H	1:F:348:LYS:CE	2.17	0.58
1:F:227:PHE:CE1	1:F:303:PRO:HB2	2.39	0.57
1:F:236:LEU:HD21	1:F:294:VAL:O	2.05	0.57
1:F:191:SER:HB2	2:C:2:FUC:H3	1.87	0.56
1:F:40:ARG:HD2	1:F:40:ARG:H	1.69	0.56
1:F:304:ASP:OD2	1:F:304:ASP:N	2.40	0.54
1:F:378:ASP:HB3	1:F:381:ARG:HG3	1.90	0.52
1:F:328:ALA:HB2	1:F:437:MET:HE3	1.92	0.52
1:F:443:GLU:OE2	9:F:616:GOL:O1	2.29	0.51
1:F:500:TYR:CZ	1:F:511:MET:HB2	2.46	0.51
1:F:66:CYS:HB3	1:F:273:LEU:HD11	1.95	0.49
1:F:209:LEU:HD12	1:F:312:PHE:HB3	1.95	0.49
1:F:245:ASN:HB2	1:F:279:VAL:HG13	1.95	0.49
1:F:509:ARG:HH11	1:F:509:ARG:HG2	1.76	0.49
1:F:319:VAL:O	1:F:418:PHE:HA	2.14	0.48
1:F:54:ASP:OD1	1:F:54:ASP:N	2.35	0.48
1:F:525:PHE:C	1:F:525:PHE:CD1	2.87	0.48
1:F:525:PHE:C	1:F:525:PHE:HD1	2.17	0.47
1:F:509:ARG:HG2	1:F:509:ARG:NH1	2.28	0.47
1:F:100:PRO:HD2	1:F:104:PRO:HG3	1.97	0.46
1:F:482:GLU:OE1	1:F:487:SER:OG	2.29	0.45
1:F:302:MET:HB3	1:F:304:ASP:OD2	2.18	0.44
1:F:245:ASN:ND2	9:F:617:GOL:H2	2.33	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:205:SER:HB3	1:F:222:LEU:HD21	2.01	0.43
1:F:34:ALA:HB2	1:F:173:LEU:HD23	2.01	0.42
1:F:408:LYS:HD3	1:F:408:LYS:HA	1.82	0.42
1:F:10:ASN:HA	1:F:51:LYS:HG2	2.00	0.42
1:F:320:GLY:HA3	1:F:419:TYR:CE2	2.55	0.42
1:F:288:VAL:H	9:F:618:GOL:H2	1.85	0.42
1:F:299:LEU:HD12	1:F:299:LEU:HA	1.93	0.42
1:F:398:PHE:C	1:F:401:PRO:HD2	2.41	0.41
1:F:110:LEU:HD11	1:F:475:ALA:CB	2.51	0.41
1:F:256:ASN:HB3	1:F:259:GLU:HB2	2.03	0.40
1:F:504:ASN:OD1	1:F:506:GLU:HG2	2.20	0.40
1:F:227:PHE:CD1	1:F:227:PHE:C	2.95	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	F	524/529 (99%)	502 (96%)	20 (4%)	2 (0%)	34	29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	377	VAL
1	F	507	SER

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	F	453/453 (100%)	438 (97%)	15 (3%)	38	35

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	40	ARG
1	F	48	SER
1	F	144	MET
1	F	195	PHE
1	F	262	LYS
1	F	278	PHE
1	F	287	SER
1	F	348	LYS
1	F	370	LEU
1	F	376	TRP
1	F	414	ASN
1	F	454	ASP
1	F	455	GLN
1	F	471	TRP
1	F	525	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	35	GLN
1	F	414	ASN
1	F	517	GLN

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	MKF	F	198	1	9,10,11	1.28	2 (22%)	8,12,14	1.99	3 (37%)

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	MKF	F	198	1	-	2/9/11/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	198	MKF	OG-C	2.53	1.39	1.34
1	F	198	MKF	O-C	2.06	1.24	1.21

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	198	MKF	OG-C-N	4.55	115.60	111.27
1	F	198	MKF	CB-OG-C	2.45	120.24	115.72
1	F	198	MKF	O-C-N	-2.04	119.23	125.14

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	F	198	MKF	C3-CA-CB-OG
1	F	198	MKF	N1-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

5.5   Carbohydrates ⓘ

9 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	1,2	14,14,15	0.26	0	17,19,21	0.66	1 (5%)
2	FUC	B	2	2	10,10,11	1.01	0	14,14,16	1.20	1 (7%)
2	NAG	C	1	1,2	14,14,15	0.52	0	17,19,21	0.75	1 (5%)
2	FUC	C	2	2	10,10,11	1.98	2 (20%)	14,14,16	1.54	4 (28%)
3	NAG	D	1	1,3	14,14,15	0.44	0	17,19,21	0.50	0
3	NAG	D	2	3	14,14,15	0.28	0	17,19,21	0.70	1 (5%)
3	FUC	D	3	3	10,10,11	0.80	0	14,14,16	0.80	0
2	NAG	E	1	1,2	14,14,15	0.55	0	17,19,21	0.62	0
2	FUC	E	2	2	10,10,11	0.58	0	14,14,16	0.79	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
2	NAG	B	1	1,2	-	0/6/23/26	0/1/1/1
2	FUC	B	2	2	-	-	0/1/1/1
2	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1
2	FUC	C	2	2	-	-	0/1/1/1
3	NAG	D	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	4/6/23/26	0/1/1/1
3	FUC	D	3	3	-	-	0/1/1/1
2	NAG	E	1	1,2	-	2/6/23/26	0/1/1/1
2	FUC	E	2	2	-	-	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2	FUC	C1-C2	4.40	1.62	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2	FUC	C2-C3	3.36	1.57	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	FUC	C1-C2-C3	3.23	113.63	109.67
2	C	2	FUC	O2-C2-C1	2.65	114.57	109.15
3	D	2	NAG	C1-O5-C5	2.48	115.56	112.19
2	B	2	FUC	C1-O5-C5	2.30	118.00	112.78
2	B	1	NAG	C1-O5-C5	2.27	115.26	112.19
2	C	2	FUC	O5-C5-C4	2.25	113.57	109.52
2	C	2	FUC	C1-O5-C5	2.10	117.54	112.78
2	C	1	NAG	C1-O5-C5	2.06	114.99	112.19

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	1	NAG	O5-C5-C6-O6
3	D	2	NAG	C8-C7-N2-C2
3	D	2	NAG	O7-C7-N2-C2
2	E	1	NAG	C4-C5-C6-O6
3	D	2	NAG	O5-C5-C6-O6
3	D	2	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2	FUC	1	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.

Oligosaccharide Chain B

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain C

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings
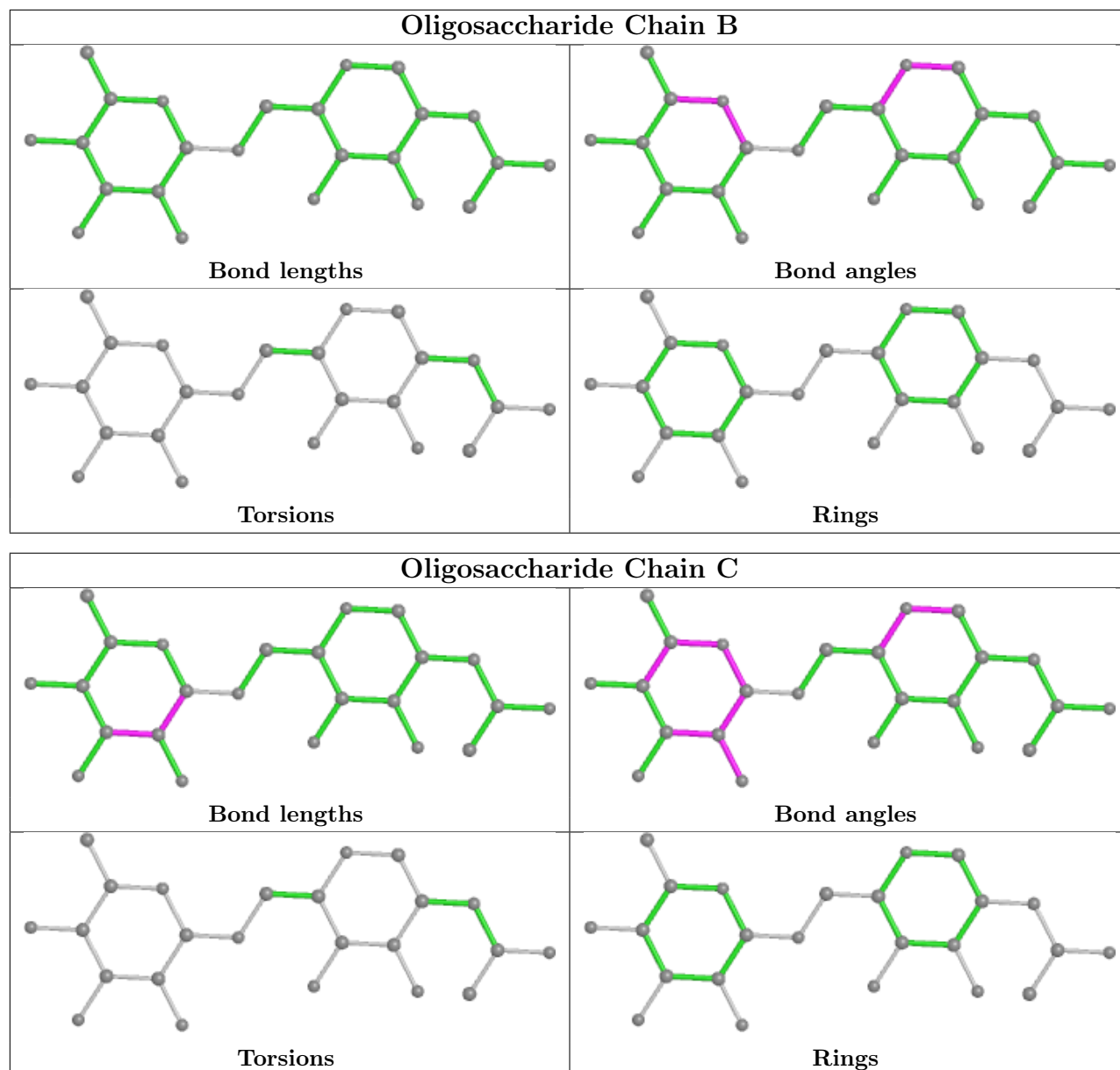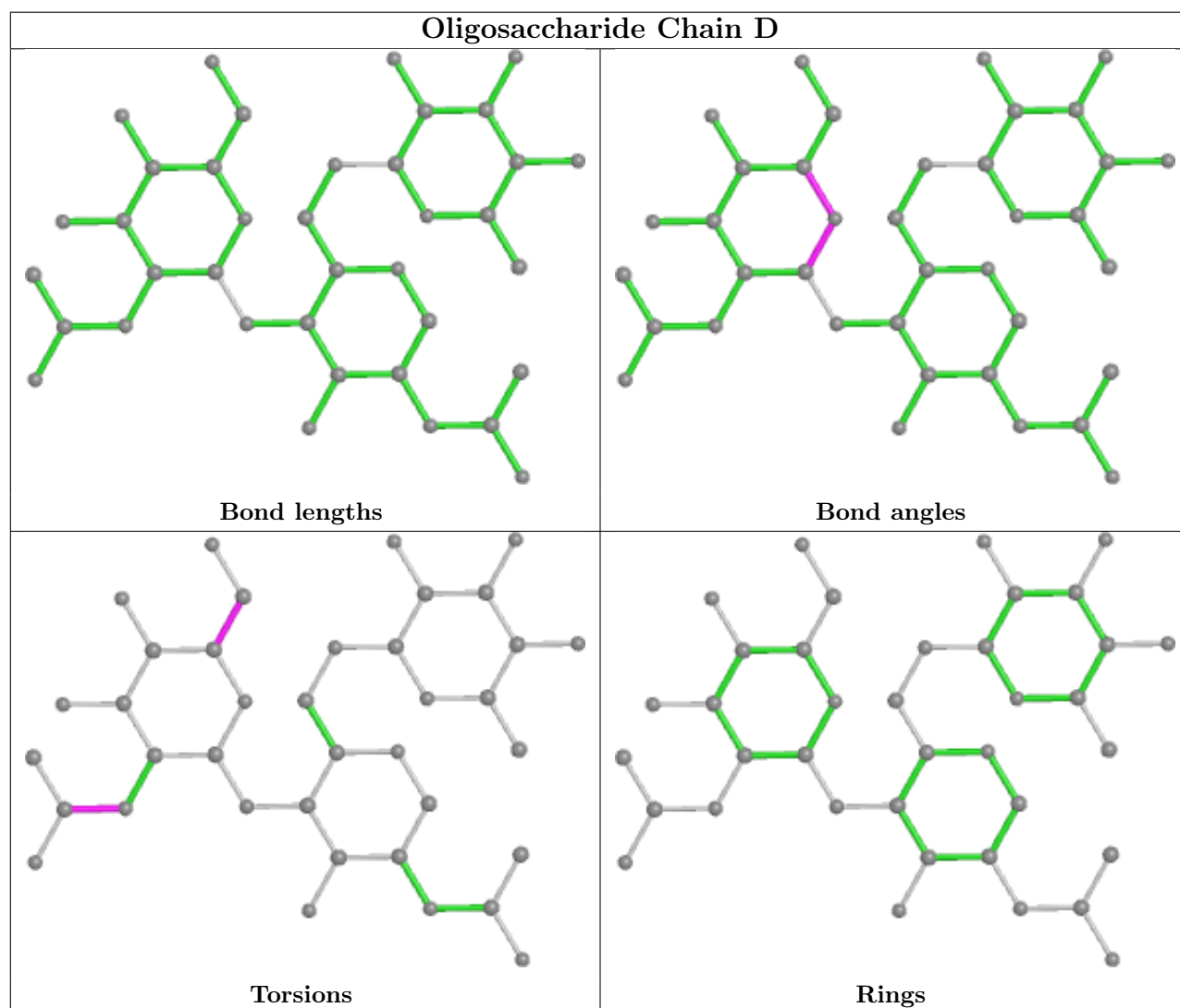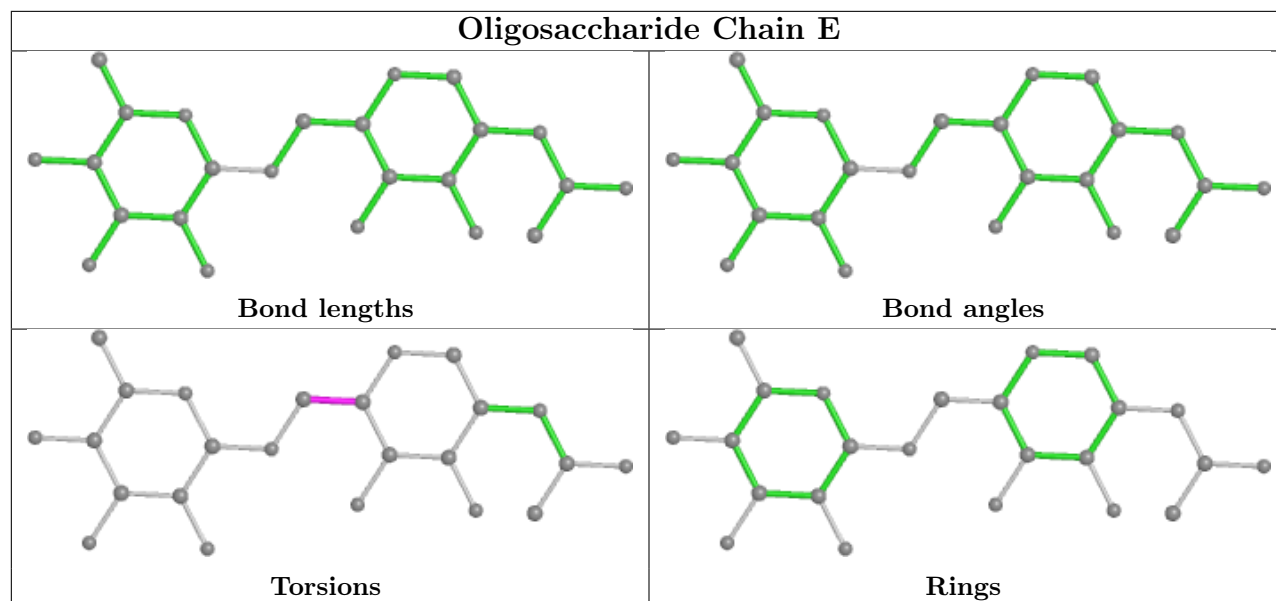
5.6 Ligand geometry ⓘ

Of 18 ligands modelled in this entry, 8 are monoatomic - leaving 10 for Mogul analysis.

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	$\#|Z| > 2$
8	NAG	F	613	1	14,14,15	1.49	2 (14%)	17,19,21	1.70	2 (11%)
9	GOL	F	616	-	5,5,5	0.72	0	5,5,5	1.10	0
4	SO4	F	602	-	4,4,4	0.15	0	6,6,6	0.09	0
9	GOL	F	615	-	5,5,5	0.98	0	5,5,5	1.00	0
8	NAG	F	614	1	14,14,15	0.44	0	17,19,21	0.45	0
4	SO4	F	603	-	4,4,4	0.16	0	6,6,6	0.14	0
7	M8X	F	612	-	22,24,24	1.36	2 (9%)	21,32,32	0.85	1 (4%)
9	GOL	F	618	-	5,5,5	0.88	0	5,5,5	1.00	0
4	SO4	F	601	-	4,4,4	0.23	0	6,6,6	0.19	0
9	GOL	F	617	-	5,5,5	0.89	0	5,5,5	1.04	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
9	GOL	F	616	-	-	2/4/4/4	-
8	NAG	F	613	1	-	2/6/23/26	0/1/1/1
9	GOL	F	615	-	-	4/4/4/4	-
8	NAG	F	614	1	-	0/6/23/26	0/1/1/1
7	M8X	F	612	-	-	4/8/18/18	0/3/3/3
9	GOL	F	618	-	-	0/4/4/4	-
9	GOL	F	617	-	-	2/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	F	613	NAG	O5-C1	4.28	1.50	1.43
7	F	612	M8X	C12-N02	4.15	1.57	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	F	613	NAG	C1-C2	3.39	1.57	1.52
7	F	612	M8X	C03-N02	2.32	1.53	1.46

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	F	613	NAG	C1-O5-C5	6.49	120.98	112.19
7	F	612	M8X	C12-C13-C14	-2.50	123.58	127.48
8	F	613	NAG	C4-C3-C2	2.06	114.04	111.02

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	F	612	M8X	C03-C04-C05-C06
7	F	612	M8X	C03-C04-C05-C11
7	F	612	M8X	C13-C12-N02-C01
9	F	615	GOL	C1-C2-C3-O3
9	F	615	GOL	O2-C2-C3-O3
8	F	613	NAG	O5-C5-C6-O6
8	F	613	NAG	C4-C5-C6-O6
9	F	615	GOL	O1-C1-C2-C3
9	F	616	GOL	O1-C1-C2-C3
9	F	615	GOL	O1-C1-C2-O2
9	F	617	GOL	O1-C1-C2-C3
7	F	612	M8X	N02-C03-C04-C05
9	F	616	GOL	O1-C1-C2-O2
9	F	617	GOL	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	F	616	GOL	1	0
9	F	618	GOL	1	0
9	F	617	GOL	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is

within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand M8X F 612

Bond lengths

Bond angles

Torsions

Rings
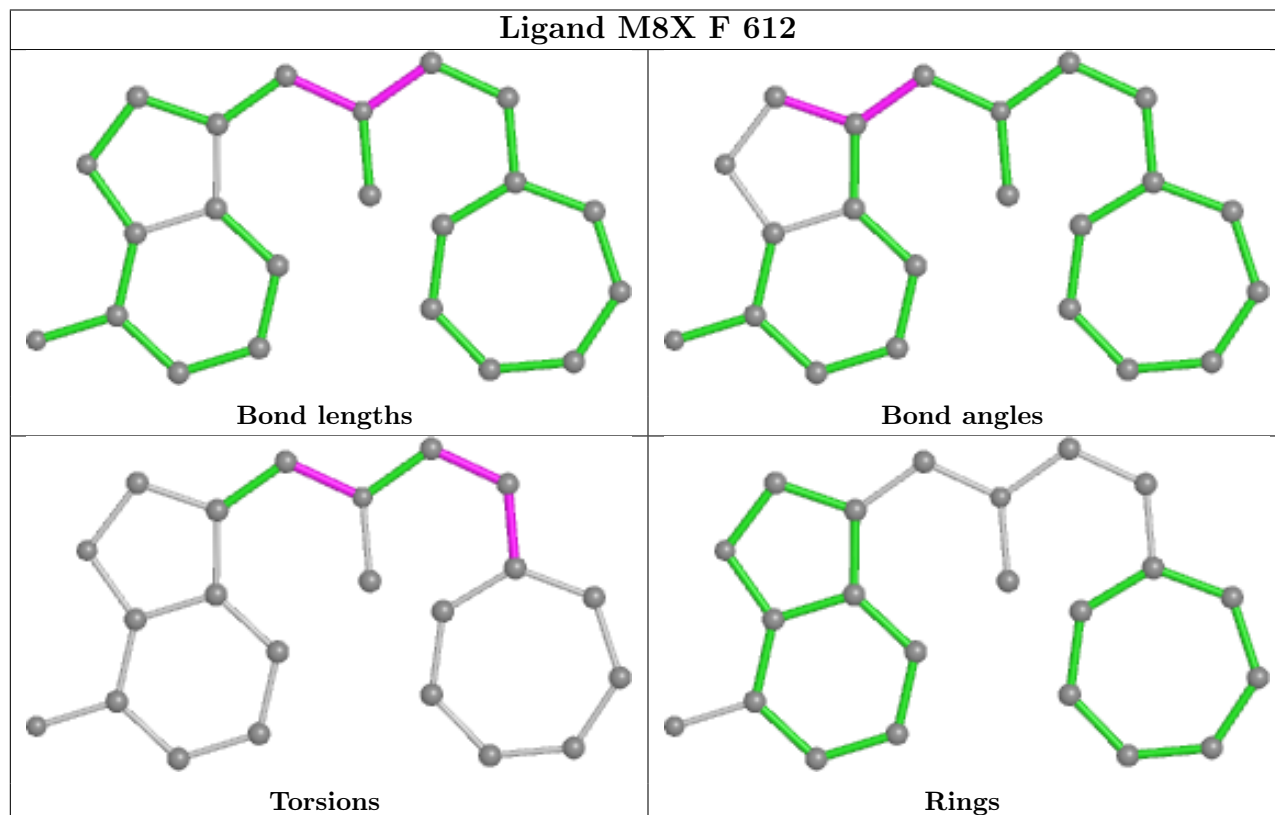
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

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(Å²)	Q<0.9
1	F	525/529 (99%)	0.30	32 (6%)	21	26	41, 59, 97, 144	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	282	TYR	8.5
1	F	255	GLU	5.2
1	F	4	ILE	4.9
1	F	53	SER	4.3
1	F	17	GLN	4.2
1	F	237	TYR	4.0
1	F	379	ASP	3.5
1	F	244	LEU	3.4
1	F	51	LYS	3.4
1	F	376	TRP	3.3
1	F	12	LYS	3.2
1	F	49	LEU	3.1
1	F	156	LEU	3.0
1	F	380	GLN	2.9
1	F	240	ARG	2.8
1	F	14	ARG	2.7
1	F	115	GLY	2.7
1	F	453	ARG	2.7
1	F	278	PHE	2.5
1	F	55	ILE	2.4
1	F	56	TRP	2.3
1	F	13	VAL	2.3
1	F	455	GLN	2.3
1	F	54	ASP	2.3
1	F	381	ARG	2.2
1	F	274	LEU	2.1
1	F	377	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	F	148	VAL	2.1
1	F	259	GLU	2.1
1	F	281	PRO	2.0
1	F	50	THR	2.0
1	F	15	GLY	2.0

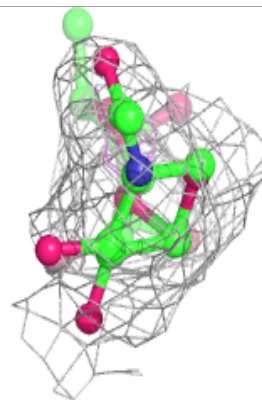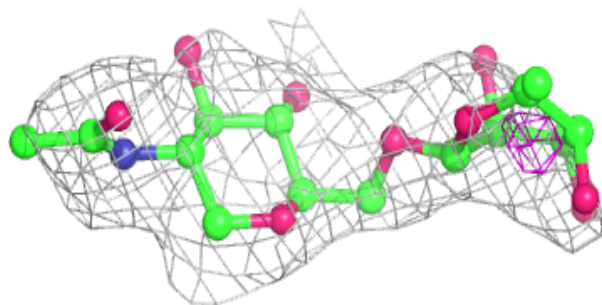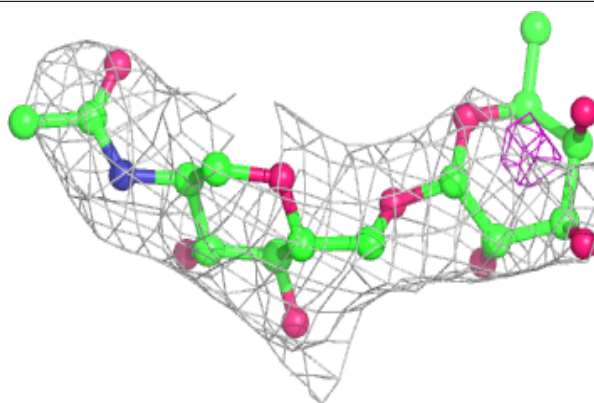## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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	MKF	F	198	11/12	0.98	0.14	42,45,50,51	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.

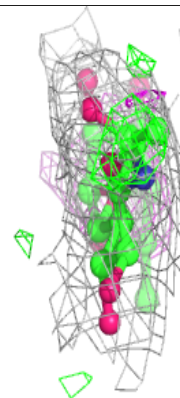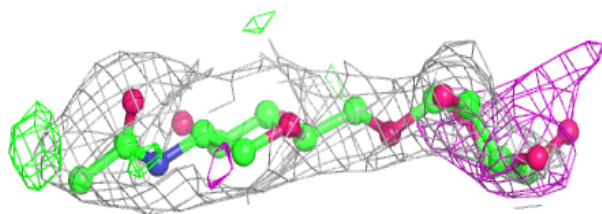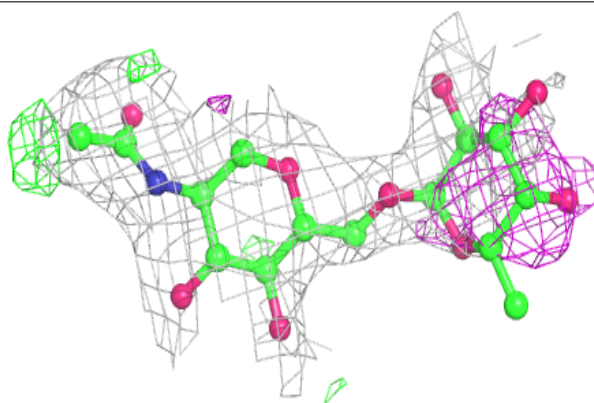| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	FUC	C	2	10/11	0.74	0.51	89,102,111,114	0
2	FUC	E	2	10/11	0.76	0.32	147,165,168,171	0
2	NAG	E	1	14/15	0.78	0.19	68,95,127,138	0
2	FUC	B	2	10/11	0.78	0.42	142,149,153,156	0
2	NAG	C	1	14/15	0.81	0.12	89,93,100,104	0
3	FUC	D	3	10/11	0.89	0.27	78,90,93,100	0
3	NAG	D	2	14/15	0.93	0.20	91,105,112,118	0
2	NAG	B	1	14/15	0.93	0.25	100,110,120,130	0
3	NAG	D	1	14/15	0.96	0.17	75,79,91,95	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.
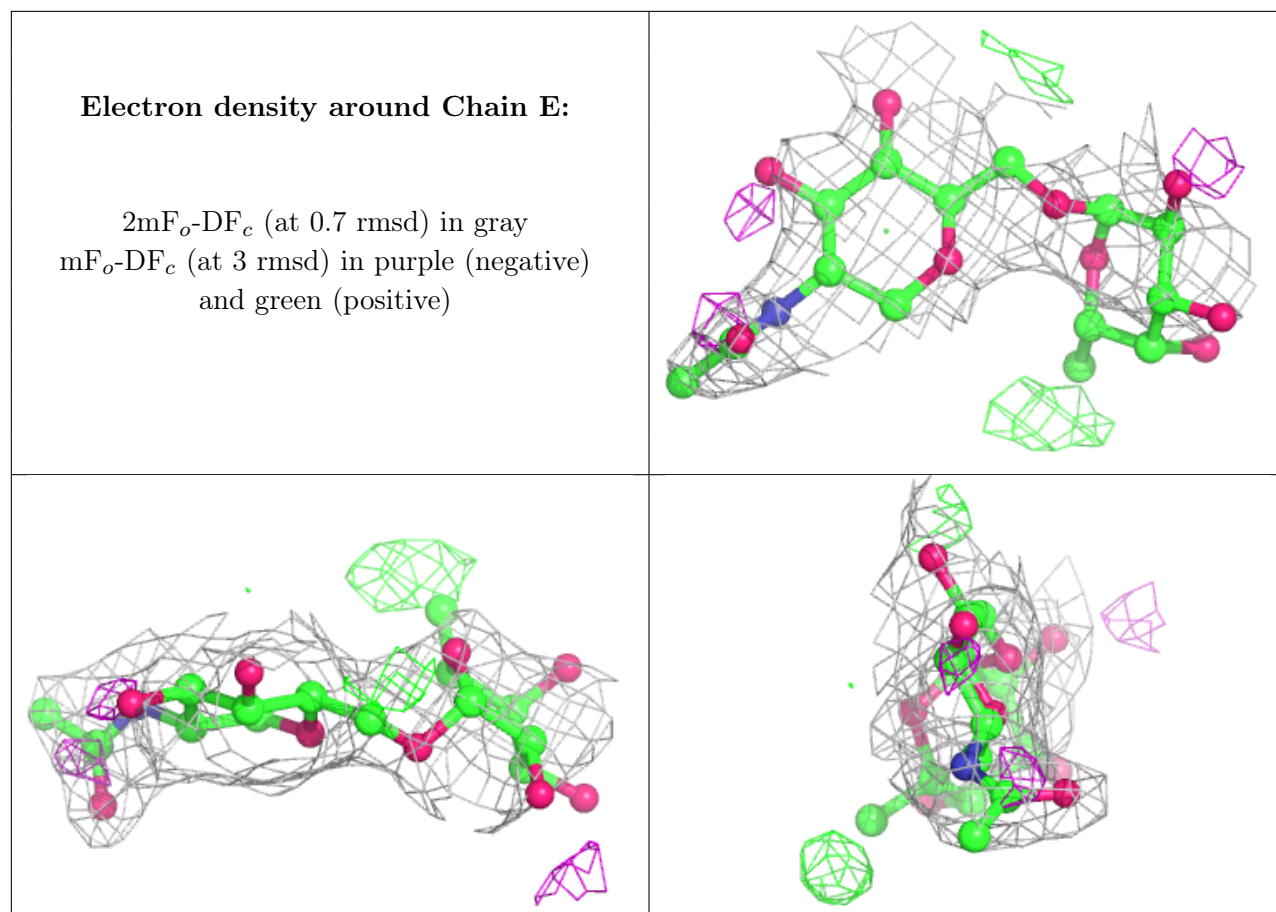
Electron density around Chain B:

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

Electron density around Chain C:

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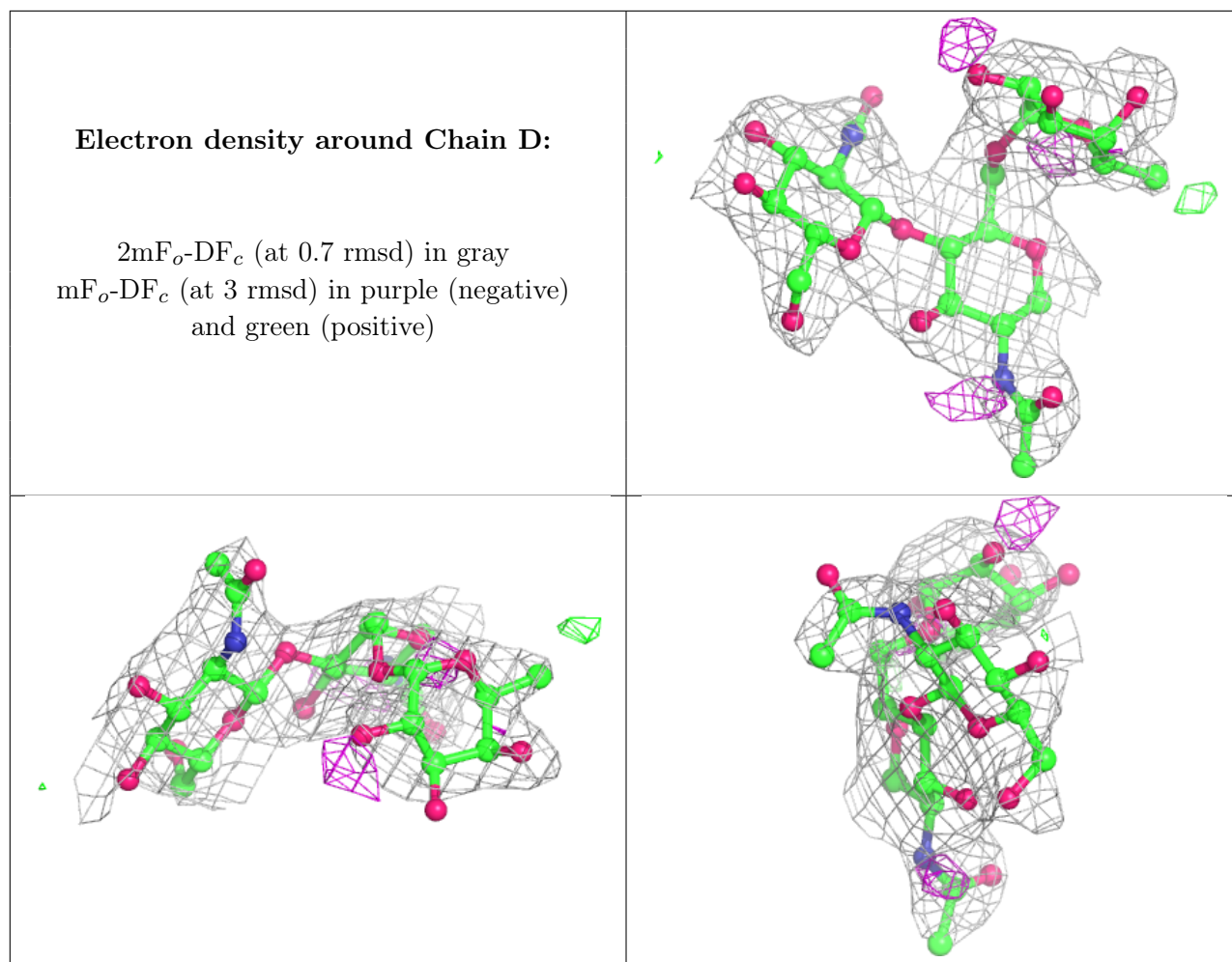and green (positive)

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 D:

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

6.4 Ligands ⓘ

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.
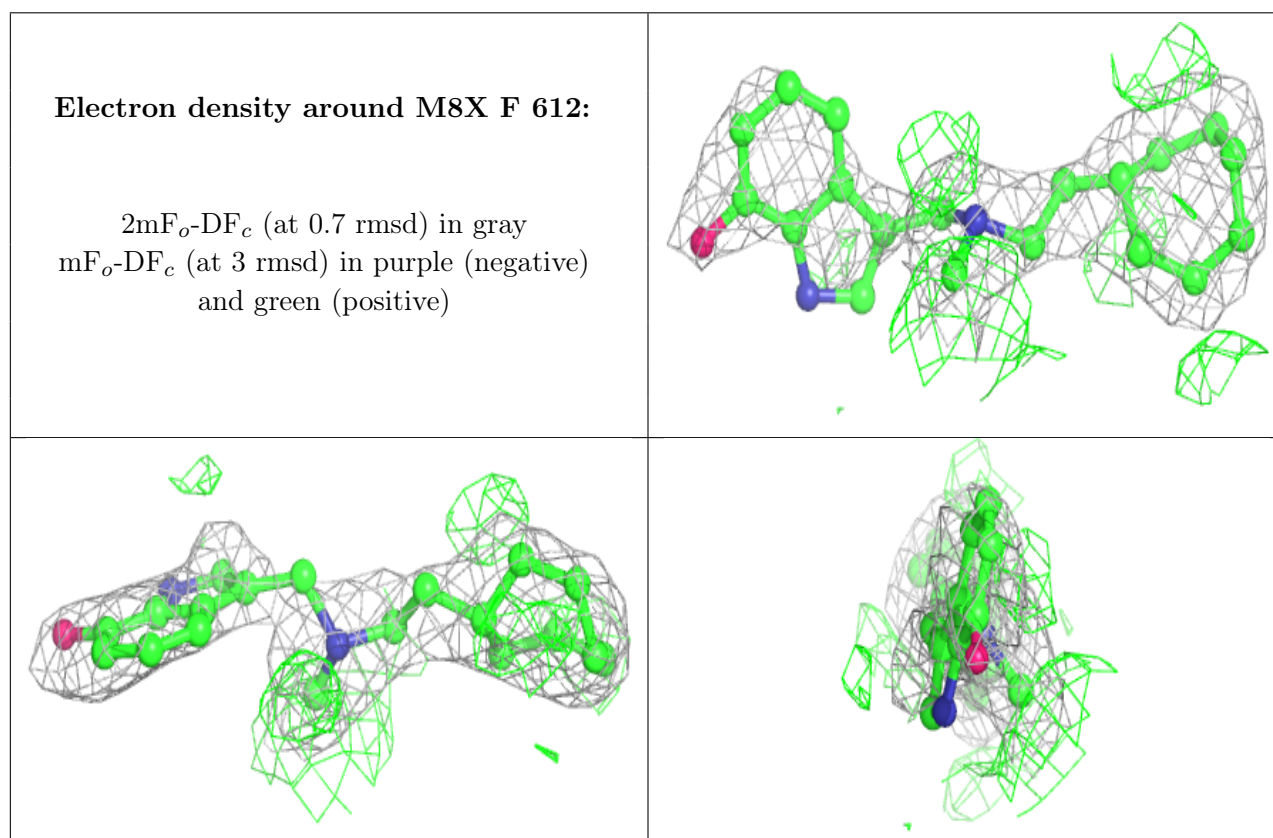
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
5	CL	F	608	1/1	0.58	0.18	98,98,98,98	0
8	NAG	F	614	14/15	0.66	0.30	128,143,155,159	0
9	GOL	F	617	6/6	0.71	0.29	113,120,124,126	0
9	GOL	F	618	6/6	0.74	0.28	99,105,107,108	0
8	NAG	F	613	14/15	0.78	0.38	141,148,153,153	0
4	SO4	F	602	5/5	0.79	0.29	113,115,121,124	0
7	M8X	F	612	22/22	0.83	0.23	59,76,84,87	8
5	CL	F	605	1/1	0.85	0.10	78,78,78,78	0
5	CL	F	607	1/1	0.85	0.16	106,106,106,106	0
9	GOL	F	615	6/6	0.86	0.28	61,67,70,70	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
6	NA	F	611	1/1	0.89	0.09	88,88,88,88	0
9	GOL	F	616	6/6	0.89	0.31	66,70,76,83	0
4	SO4	F	601	5/5	0.92	0.14	47,56,61,62	5
5	CL	F	606	1/1	0.92	0.13	96,96,96,96	0
5	CL	F	604	1/1	0.92	0.23	94,94,94,94	0
4	SO4	F	603	5/5	0.94	0.12	74,85,88,89	5
5	CL	F	609	1/1	0.94	0.11	86,86,86,86	0
5	CL	F	610	1/1	0.96	0.07	86,86,86,86	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around M8X F 612:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.