Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 12:29 PM BST

PDB ID	:	3VJL
Title	:	Crystal structure of human depiptidyl peptidase IV (DPP-4) in complex with a prolylthiazolidine inhibitor #2
Authors	:	Akahoshi, F.; Kishida, H.; Miyaguchi, I.; Yoshida, T.; Ishii, S.
Deposited on	:	2011-10-24
Resolution	:	2.39 Å(reported)

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

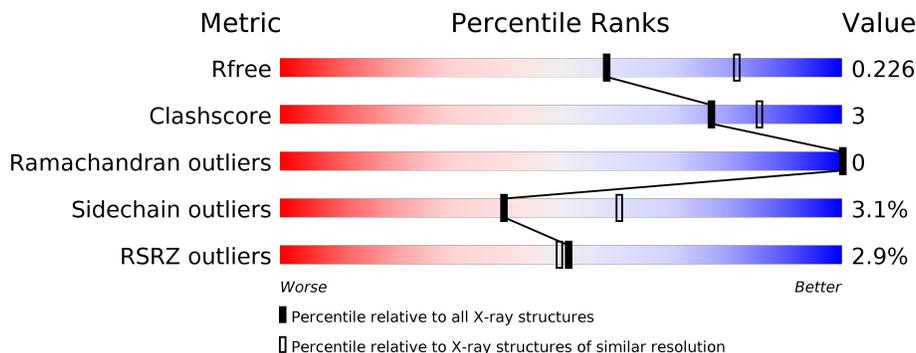|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.39 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.
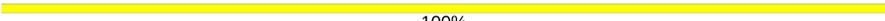
Metric	Percentile Ranks	Value
Rfree		0.226
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		3.1%
RSRZ outliers		2.9%

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 on 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	740	3% 88% 10% •
1	B	740	2% 89% 9% ••
2	C	2	50% 50%
2	D	2	100%
2	E	2	50% 50%
2	F	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	2	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13126 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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	729	Total 5971	C 3831	N 983	O 1131	S 26	0	0	0
1	B	729	Total 5971	C 3831	N 983	O 1131	S 26	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	767	HIS	-	expression tag	UNP P27487
A	768	HIS	-	expression tag	UNP P27487
A	769	HIS	-	expression tag	UNP P27487
A	770	HIS	-	expression tag	UNP P27487
A	771	HIS	-	expression tag	UNP P27487
A	772	HIS	-	expression tag	UNP P27487
B	767	HIS	-	expression tag	UNP P27487
B	768	HIS	-	expression tag	UNP P27487
B	769	HIS	-	expression tag	UNP P27487
B	770	HIS	-	expression tag	UNP P27487
B	771	HIS	-	expression tag	UNP P27487
B	772	HIS	-	expression tag	UNP P27487

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

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

Continued on next page...

Continued from previous page...

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

- Molecule 3 is [(2S,4S)-4-{4-[1-phenyl-3-(trifluoromethyl)-1H-pyrazol-5-yl]piperidin-1-yl}pyrrolidin-2-yl](1,3-thiazolidin-3-yl)methanone (three-letter code: W94) (formula: $C_{23}H_{28}F_3N_5OS$).
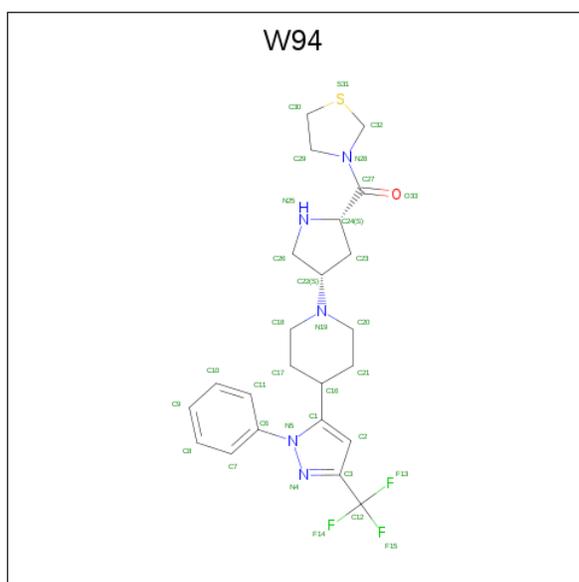
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total 33	C 23	F 3	N 5	O 1	S 1	0	0
3	B	1	Total 33	C 23	F 3	N 5	O 1	S 1	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
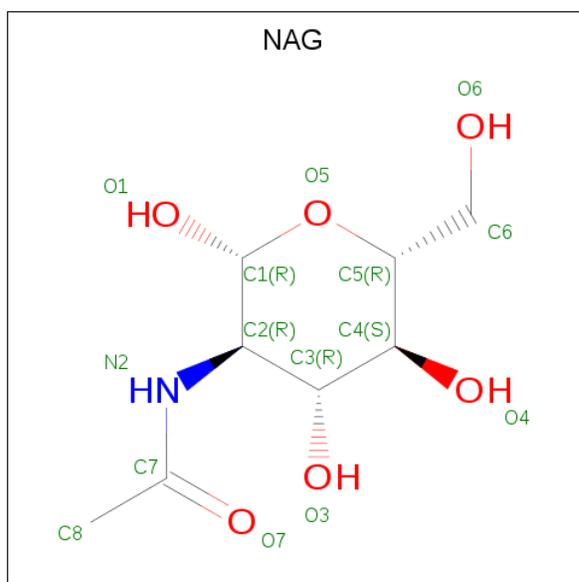
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	420	Total 420	O 420	0	0
5	B	474	Total 474	O 474	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: Dipeptidyl peptidase 4
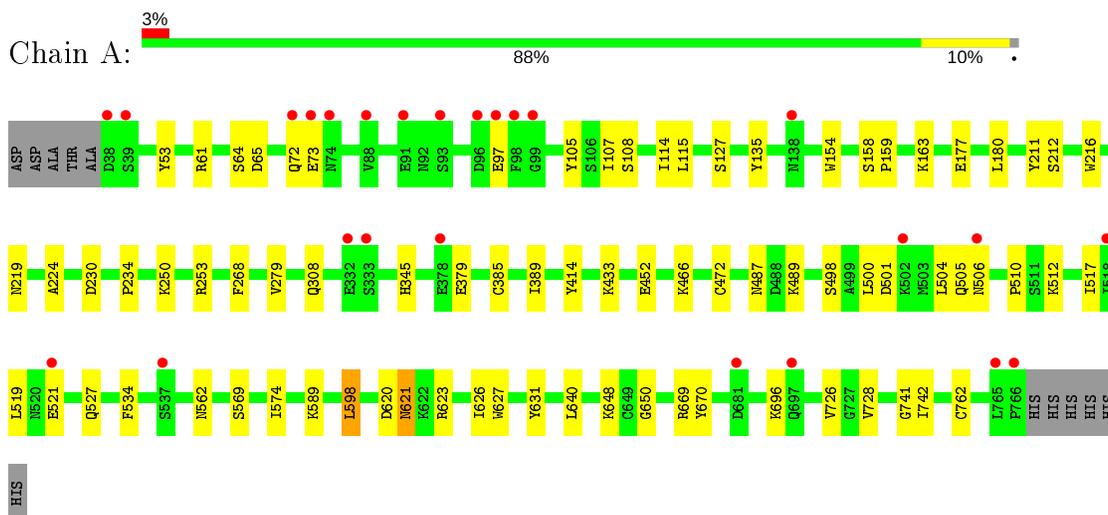
- Molecule 1: Dipeptidyl peptidase 4
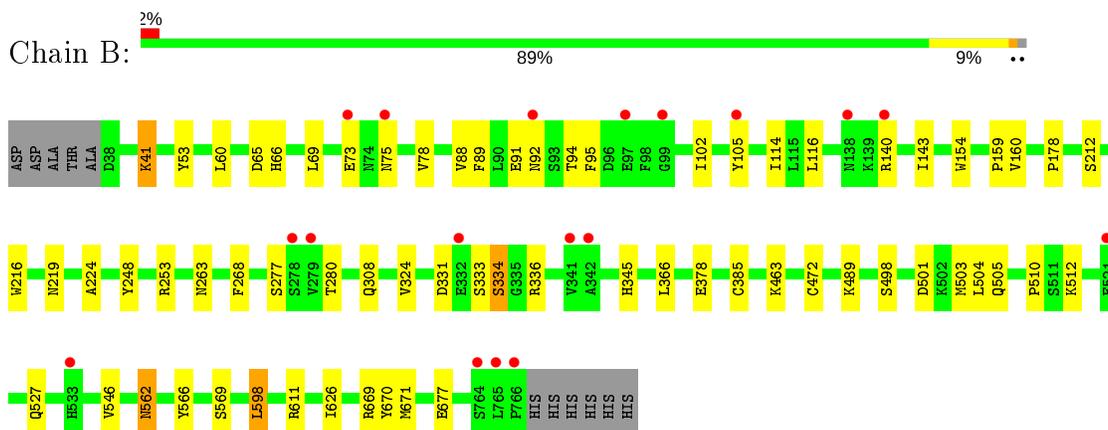
- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

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

Chain D:

MAG1 NAG2

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

Chain E:

MAG1 NAG2

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

Chain F:

MAG1 NAG2

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

Chain G:

MAG1 NAG2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	117.81Å 125.83Å 137.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.39 29.89 – 2.39	Depositor EDS
% Data completeness (in resolution range)	99.1 (30.00-2.39) 99.1 (29.89-2.39)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	10.44 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.184 , 0.230 0.181 , 0.226	Depositor DCC
R_{free} test set	4025 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	27.8	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.5	EDS
L-test for twinning[2]	$<\|L\|> = 0.51$, $<L^2> = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	13126	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% 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: W94, 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	A	0.51	0/6143	0.60	0/8355
1	B	0.51	0/6143	0.61	0/8355
All	All	0.51	0/12286	0.60	0/16710

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5971	0	5684	37	0
1	B	5971	0	5683	36	0
2	C	28	0	25	0	0
2	D	28	0	25	0	0
2	E	28	0	25	0	0
2	F	28	0	25	0	0
2	G	28	0	25	0	0
3	A	33	0	28	0	0
3	B	33	0	28	0	0
4	A	42	0	39	1	0
4	B	42	0	39	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	420	0	0	3	0
5	B	474	0	0	5	0
All	All	13126	0	11626	72	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 (72) 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:253:ARG:HH21	1:B:253:ARG:HH21	1.22	0.87
1:A:640:LEU:HD11	1:A:650:GLY:HA3	1.62	0.81
4:A:1229:NAG:H81	5:A:2134:HOH:O	1.81	0.80
1:A:219:ASN:H	1:A:308:GLN:HE22	1.38	0.71
1:B:345:HIS:HD2	5:B:2628:HOH:O	1.74	0.71
1:B:219:ASN:H	1:B:308:GLN:HE22	1.46	0.62
1:A:105:TYR:HB2	1:A:114:ILE:HD11	1.83	0.59
1:B:334:SER:HB3	1:B:336:ARG:H	1.69	0.57
1:B:331:ASP:OD1	1:B:334:SER:HB2	2.05	0.56
1:B:546:VAL:HG21	1:B:626:ILE:HD11	1.88	0.56
1:B:512:LYS:HE3	1:B:527:GLN:OE1	2.06	0.56
1:A:510:PRO:HD3	1:A:569:SER:HB2	1.89	0.54
1:A:620:ASP:OD2	1:A:623:ARG:HD3	2.08	0.54
1:B:489:LYS:HE2	5:B:2597:HOH:O	2.07	0.54
1:A:114:ILE:HG23	1:A:135:TYR:HB3	1.89	0.54
1:B:510:PRO:HD3	1:B:569:SER:HB2	1.89	0.54
1:A:219:ASN:N	1:A:308:GLN:HE22	2.04	0.53
1:A:72:GLN:HG3	1:A:73:GLU:HG3	1.92	0.50
1:B:75:ASN:ND2	1:B:92:ASN:H	2.09	0.50
1:A:696:LYS:HG3	1:A:728:VAL:HG22	1.93	0.50
1:A:648:LYS:HE2	1:A:762:CYS:O	2.12	0.50
1:B:41:LYS:HE3	1:B:53:TYR:OH	2.12	0.50
1:A:621:ASN:HD22	1:A:621:ASN:H	1.57	0.50
1:B:78:VAL:HG23	1:B:89:PHE:HB2	1.95	0.49
1:B:219:ASN:N	1:B:308:GLN:HE22	2.11	0.48
1:B:277:SER:HB3	1:B:280:THR:HG22	1.96	0.47
1:A:234:PRO:HB2	1:B:248:TYR:CZ	2.50	0.47
1:A:726:VAL:HG23	1:A:728:VAL:HG23	1.96	0.47
1:B:501:ASP:O	1:B:505:GLN:HG2	2.13	0.47
1:A:626:ILE:O	1:A:650:GLY:HA2	2.15	0.47
1:B:677:GLU:CD	1:B:677:GLU:H	2.18	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:69:LEU:HD23	1:B:78:VAL:HG22	1.97	0.47
1:A:127:SER:HB3	1:A:211:TYR:CD1	2.50	0.47
1:B:105:TYR:HB3	1:B:114:ILE:HD11	1.96	0.46
1:B:75:ASN:HD22	1:B:91:GLU:HA	1.80	0.46
1:B:562:ASN:HD22	1:B:562:ASN:C	2.19	0.46
1:A:534:PHE:HE1	1:A:574:ILE:HD11	1.82	0.45
1:B:611:ARG:NH2	5:B:2664:HOH:O	2.50	0.45
1:A:107:ILE:HG22	1:A:108:SER:O	2.17	0.45
1:B:95:PHE:CE2	1:B:116:LEU:HD11	2.53	0.44
1:A:65:ASP:OD2	1:A:466:LYS:HB2	2.18	0.44
1:A:224:ALA:HB1	1:A:268:PHE:CZ	2.52	0.44
1:B:345:HIS:CD2	5:B:2628:HOH:O	2.59	0.44
1:A:64:SER:HB2	5:A:2385:HOH:O	2.18	0.44
1:B:159:PRO:HD3	1:B:216:TRP:CB	2.47	0.43
1:A:414:TYR:CE1	1:A:433:LYS:HE2	2.53	0.43
1:A:501:ASP:O	1:A:505:GLN:HG2	2.19	0.43
1:A:154:TRP:CE2	1:A:212:SER:HB3	2.54	0.43
1:A:345:HIS:HE1	1:A:389:ILE:O	2.02	0.42
1:B:598:LEU:HD22	1:B:671:MET:HG2	2.01	0.42
1:A:669:ARG:HD2	1:A:670:TYR:CZ	2.54	0.42
1:B:143:ILE:HD13	1:B:178:PRO:HB2	2.02	0.42
1:A:53:TYR:HB3	1:A:500:LEU:HD11	2.02	0.42
1:B:41:LYS:HG2	5:B:2813:HOH:O	2.20	0.41
1:B:60:LEU:HD12	1:B:60:LEU:C	2.40	0.41
1:A:114:ILE:CG2	1:A:135:TYR:HB3	2.49	0.41
1:B:512:LYS:HE3	1:B:527:GLN:CD	2.39	0.41
1:B:669:ARG:HD2	1:B:670:TYR:CZ	2.56	0.41
1:A:345:HIS:HD2	5:A:2077:HOH:O	2.04	0.41
1:B:154:TRP:CE2	1:B:212:SER:HB2	2.56	0.41
1:B:224:ALA:HB1	1:B:268:PHE:CZ	2.56	0.41
1:A:598:LEU:HG	1:A:631:TYR:OH	2.20	0.41
1:A:105:TYR:HA	1:A:115:LEU:O	2.21	0.41
1:A:512:LYS:HE3	1:A:527:GLN:OE1	2.20	0.41
1:A:158:SER:OG	1:A:163:LYS:HB2	2.21	0.40
1:A:177:GLU:HG3	1:A:180:LEU:HD12	2.02	0.40
1:A:517:ILE:HD12	1:A:519:LEU:HG	2.03	0.40
1:B:159:PRO:HD3	1:B:216:TRP:HB3	2.01	0.40
1:B:88:VAL:HG11	1:B:91:GLU:HG2	2.03	0.40
1:A:159:PRO:HD3	1:A:216:TRP:HB3	2.03	0.40
1:A:741:GLY:O	1:A:742:ILE:C	2.60	0.40
1:B:65:ASP:HB3	1:B:66:HIS:CE1	2.57	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	727/740 (98%)	700 (96%)	27 (4%)	0	100	100
1	B	727/740 (98%)	703 (97%)	24 (3%)	0	100	100
All	All	1454/1480 (98%)	1403 (96%)	51 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	654/663 (99%)	634 (97%)	20 (3%)	40	60
1	B	654/663 (99%)	633 (97%)	21 (3%)	39	59
All	All	1308/1326 (99%)	1267 (97%)	41 (3%)	40	60

All (41) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	61	ARG
1	A	97	GLU
1	A	230	ASP
1	A	250	LYS
1	A	279	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	379	GLU
1	A	385	CYS
1	A	452	GLU
1	A	472	CYS
1	A	487	ASN
1	A	489	LYS
1	A	498	SER
1	A	504	LEU
1	A	506	ASN
1	A	521	GLU
1	A	562	ASN
1	A	589	LYS
1	A	598	LEU
1	A	621	ASN
1	A	627	TRP
1	B	41	LYS
1	B	73	GLU
1	B	94	THR
1	B	102	ILE
1	B	140	ARG
1	B	160	VAL
1	B	263	ASN
1	B	324	VAL
1	B	333	SER
1	B	334	SER
1	B	366	LEU
1	B	378	GLU
1	B	385	CYS
1	B	463	LYS
1	B	472	CYS
1	B	498	SER
1	B	503	MET
1	B	504	LEU
1	B	562	ASN
1	B	566	TYR
1	B	598	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (28) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	ASN
1	A	112	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	138	ASN
1	A	169	ASN
1	A	263	ASN
1	A	286	GLN
1	A	308	GLN
1	A	345	HIS
1	A	487	ASN
1	A	506	ASN
1	A	562	ASN
1	A	572	ASN
1	A	621	ASN
1	A	679	ASN
1	A	694	ASN
1	B	75	ASN
1	B	112	GLN
1	B	138	ASN
1	B	169	ASN
1	B	263	ASN
1	B	286	GLN
1	B	308	GLN
1	B	345	HIS
1	B	505	GLN
1	B	562	ASN
1	B	572	ASN
1	B	679	ASN
1	B	694	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)

10 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	0.59	0	17,19,21	1.00	0
2	NAG	C	2	2	14,14,15	0.52	0	17,19,21	1.24	1 (5%)
2	NAG	D	1	1,2	14,14,15	0.75	0	17,19,21	1.05	1 (5%)
2	NAG	D	2	2	14,14,15	0.51	0	17,19,21	1.01	1 (5%)
2	NAG	E	1	1,2	14,14,15	0.56	0	17,19,21	0.83	0
2	NAG	E	2	2	14,14,15	0.58	0	17,19,21	1.21	1 (5%)
2	NAG	F	1	1,2	14,14,15	0.62	0	17,19,21	1.31	1 (5%)
2	NAG	F	2	2	14,14,15	0.58	0	17,19,21	1.26	1 (5%)
2	NAG	G	1	1,2	14,14,15	0.53	0	17,19,21	0.95	1 (5%)
2	NAG	G	2	2	14,14,15	0.51	0	17,19,21	1.91	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 Torsion 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	NAG	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	1/6/23/26	0/1/1/1
2	NAG	E	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	NAG	F	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	1/6/23/26	0/1/1/1
2	NAG	G	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	2	NAG	C1-O5-C5	6.06	120.40	112.19
2	F	1	NAG	C1-O5-C5	4.68	118.53	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	NAG	C1-O5-C5	3.23	116.57	112.19
2	F	2	NAG	C1-O5-C5	3.14	116.44	112.19
2	D	1	NAG	O5-C1-C2	-2.94	106.65	111.29
2	E	2	NAG	O5-C5-C6	2.92	111.79	107.20
2	G	2	NAG	C3-C4-C5	2.63	114.94	110.24
2	D	2	NAG	C1-O5-C5	2.54	115.63	112.19
2	G	2	NAG	O5-C5-C6	2.33	110.85	107.20
2	G	2	NAG	C4-C3-C2	2.22	114.27	111.02
2	G	1	NAG	O5-C1-C2	-2.09	107.99	111.29

There are no chirality outliers.

All (5) torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
2	C	2	NAG	C4-C5-C6-O6
2	C	2	NAG	O5-C5-C6-O6
2	D	2	NAG	C4-C5-C6-O6
2	G	2	NAG	C4-C5-C6-O6
2	F	2	NAG	C4-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.

Oligosaccharide Chain C

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain G

Bond lengths

Bond angles

Torsions

Rings
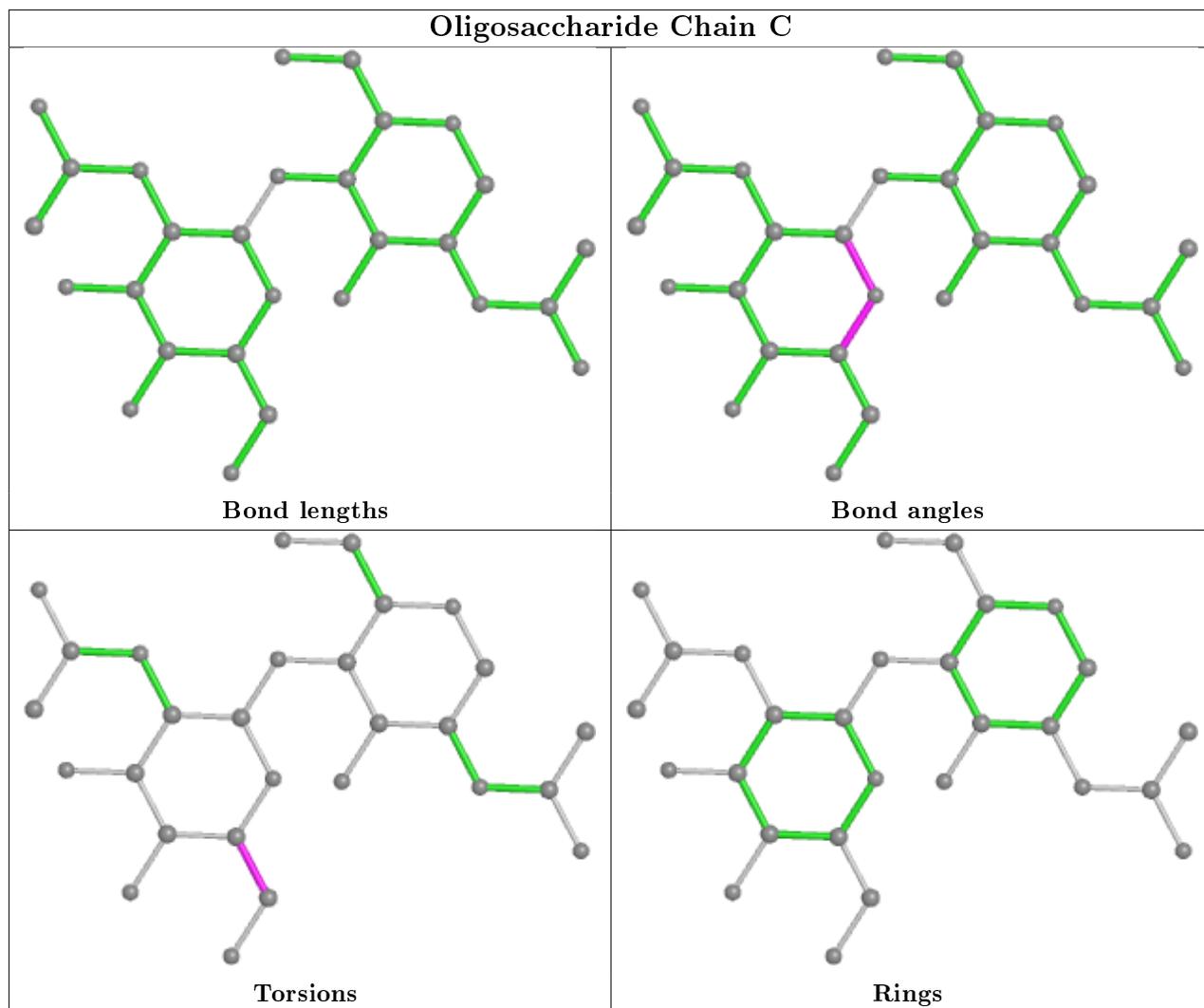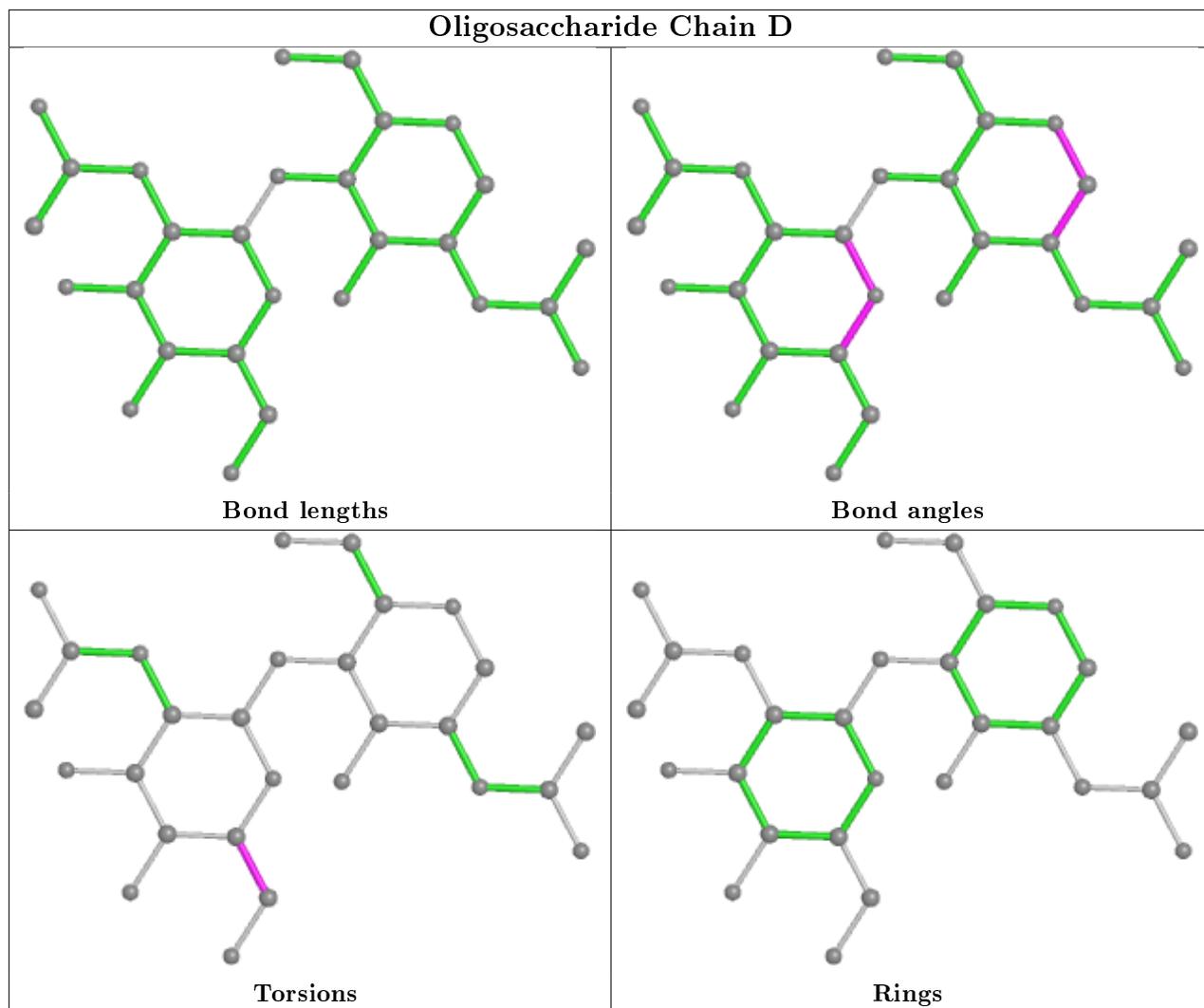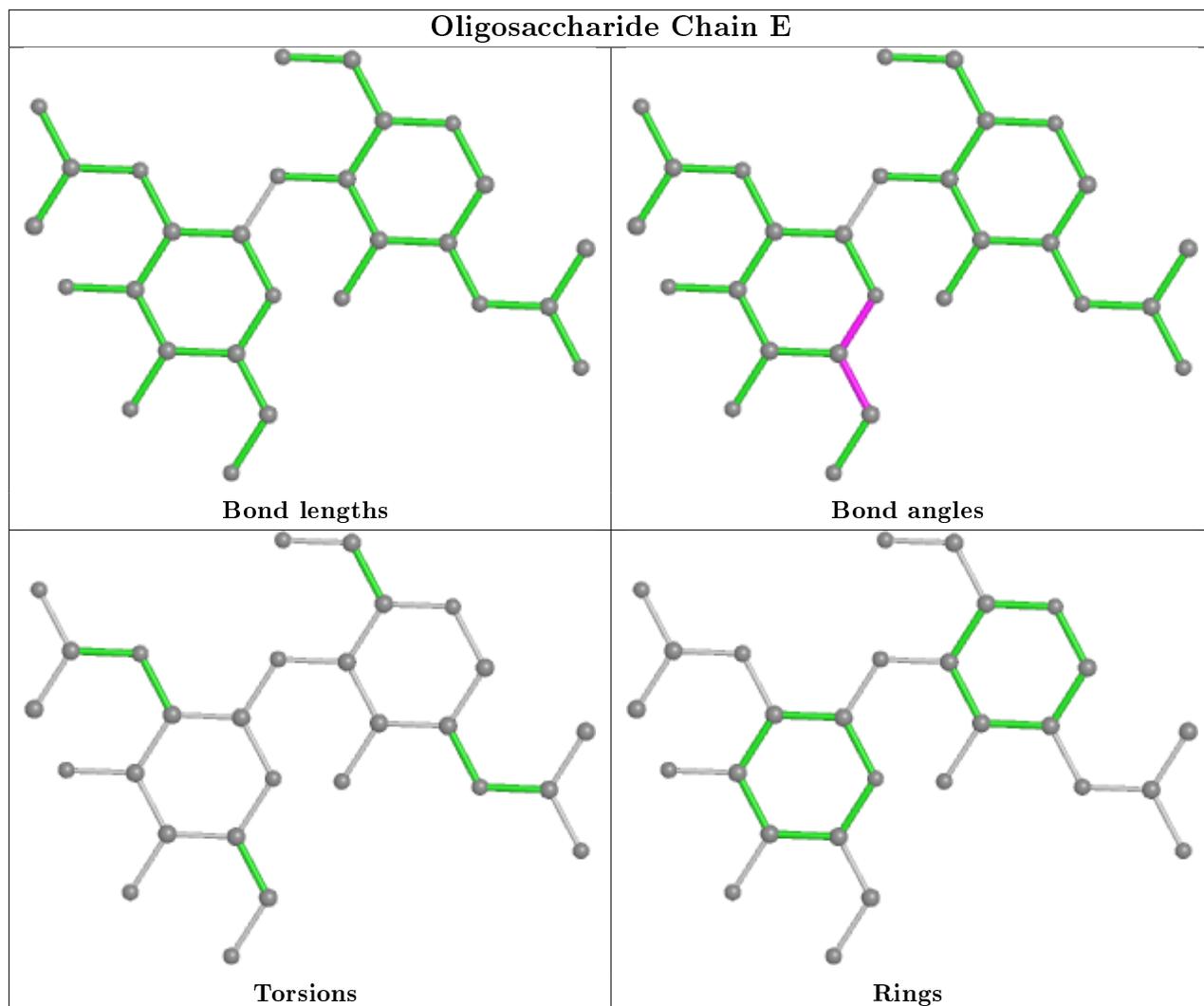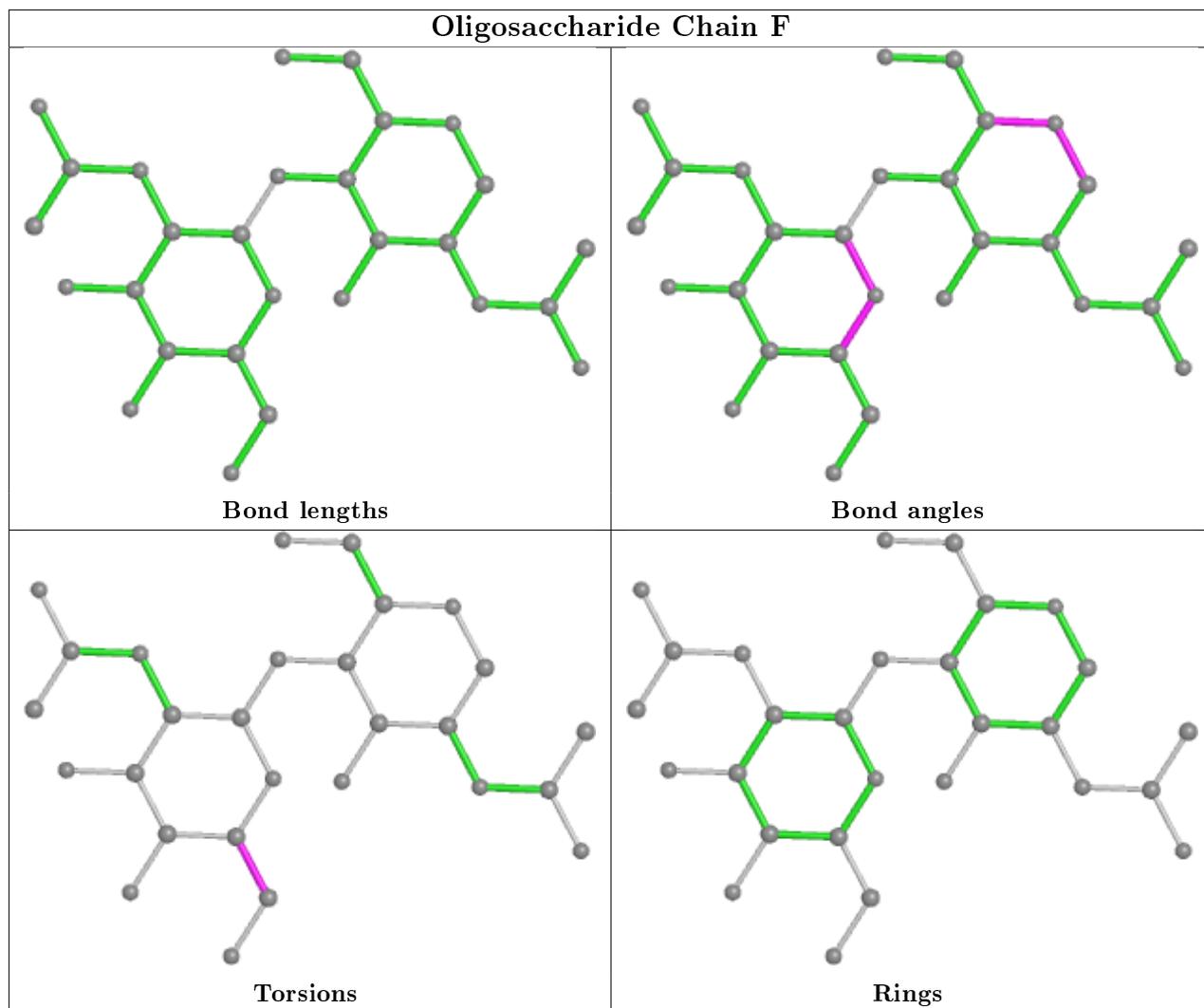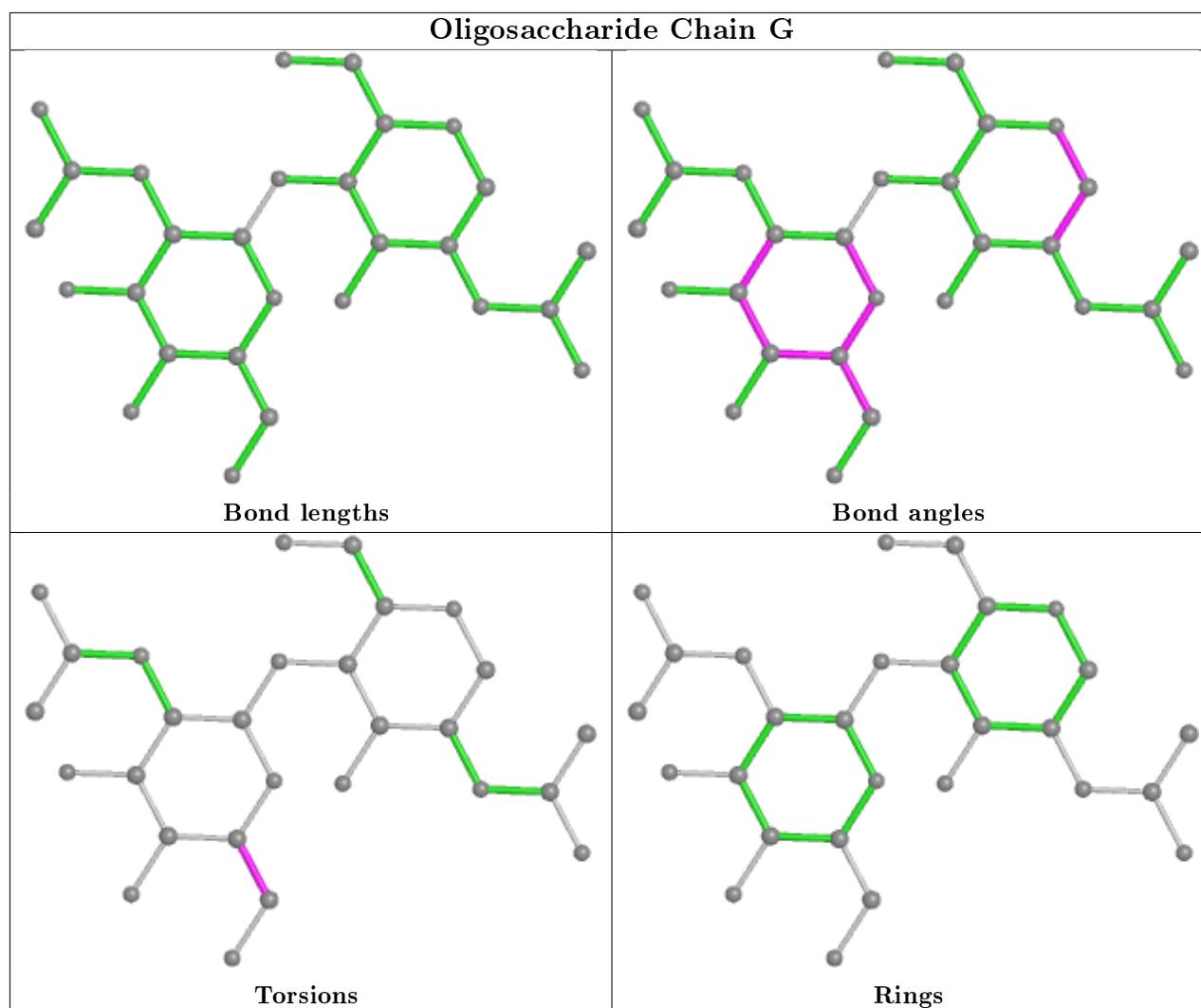
5.6 Ligand geometry ⓘ

8 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	$\#|Z| > 2$
4	NAG	B	1150	1	14,14,15	0.50	0	17,19,21	1.34	2 (11%)
4	NAG	A	1085	1	14,14,15	0.42	0	17,19,21	1.61	1 (5%)
4	NAG	B	1085	1	14,14,15	0.51	0	17,19,21	1.58	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	NAG	A	1321	1	14,14,15	0.58	0	17,19,21	1.66	1 (5%)
4	NAG	A	1229	1	14,14,15	0.54	0	17,19,21	1.54	3 (17%)
3	W94	B	1000	-	35,37,37	0.91	1 (2%)	45,54,54	2.12	15 (33%)
3	W94	A	1000	-	35,37,37	0.82	1 (2%)	45,54,54	2.17	14 (31%)
4	NAG	B	1281	1	14,14,15	0.58	0	17,19,21	1.59	3 (17%)

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	B	1150	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1085	1	-	1/6/23/26	0/1/1/1
4	NAG	B	1085	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1321	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1229	1	-	2/6/23/26	0/1/1/1
3	W94	B	1000	-	-	0/26/52/52	0/5/5/5
3	W94	A	1000	-	-	0/26/52/52	0/5/5/5
4	NAG	B	1281	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1000	W94	C6-N5	-3.55	1.33	1.44
3	B	1000	W94	C6-N5	-3.30	1.34	1.44

All (40) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1000	W94	C2-C1-C16	-7.16	117.50	128.71
3	B	1000	W94	C2-C1-C16	-6.57	118.42	128.71
4	A	1085	NAG	C1-O5-C5	6.04	120.37	112.19
4	B	1085	NAG	C1-O5-C5	5.64	119.84	112.19
4	A	1321	NAG	C1-O5-C5	5.54	119.70	112.19
3	B	1000	W94	C20-N19-C18	5.11	118.42	109.08
3	A	1000	W94	C3-N4-N5	-4.79	100.09	105.66
4	B	1150	NAG	C1-O5-C5	4.31	118.03	112.19
3	B	1000	W94	C3-N4-N5	-4.28	100.69	105.66
3	A	1000	W94	C20-N19-C18	4.10	116.58	109.08

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1000	W94	C1-C2-C3	-3.88	102.87	106.08
4	B	1281	NAG	O5-C5-C6	3.82	113.19	107.20
3	B	1000	W94	C11-C6-C7	-3.79	115.72	121.33
3	B	1000	W94	C10-C11-C6	3.61	123.31	118.63
3	A	1000	W94	C8-C7-C6	3.53	123.21	118.63
4	A	1229	NAG	O5-C1-C2	-3.49	105.78	111.29
4	B	1281	NAG	C1-O5-C5	3.36	116.75	112.19
3	A	1000	W94	C2-C3-N4	3.31	116.21	111.41
3	A	1000	W94	C11-C6-C7	-3.15	116.67	121.33
4	A	1229	NAG	C1-C2-N2	3.08	115.74	110.49
4	A	1229	NAG	C1-O5-C5	3.05	116.32	112.19
3	A	1000	W94	C2-C3-C12	-2.84	124.90	127.93
3	A	1000	W94	F15-C12-C3	-2.79	107.70	112.47
3	B	1000	W94	C1-C2-C3	-2.78	103.78	106.08
3	A	1000	W94	F13-C12-C3	-2.76	107.75	112.47
3	B	1000	W94	C7-C6-N5	2.65	124.53	119.50
3	B	1000	W94	C8-C7-C6	2.59	121.99	118.63
3	B	1000	W94	F13-C12-C3	-2.54	108.13	112.47
3	B	1000	W94	C2-C3-N4	2.53	115.08	111.41
4	B	1281	NAG	C1-C2-N2	2.53	114.80	110.49
3	B	1000	W94	C2-C3-C12	-2.49	125.28	127.93
3	A	1000	W94	C11-C6-N5	2.46	124.15	119.50
3	B	1000	W94	C21-C16-C1	2.40	115.07	112.06
3	B	1000	W94	C30-C29-N28	2.37	112.07	106.85
3	B	1000	W94	C30-S31-C32	2.30	97.19	91.48
3	A	1000	W94	C30-S31-C32	2.14	96.80	91.48
3	A	1000	W94	C30-C29-N28	2.10	111.49	106.85
4	B	1150	NAG	O5-C5-C6	2.09	110.48	107.20
3	A	1000	W94	C17-C18-N19	-2.05	108.29	111.43
3	B	1000	W94	F15-C12-C3	-2.04	108.98	112.47

There are no chirality outliers.
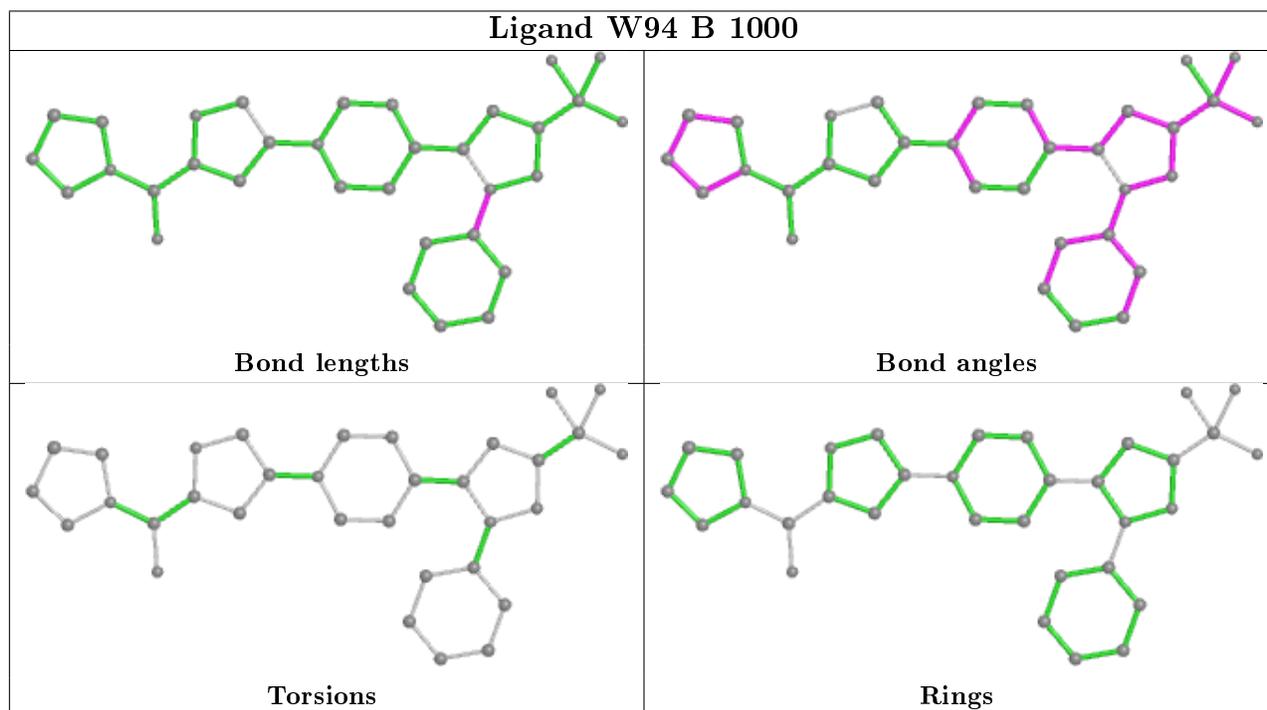
All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1229	NAG	C8-C7-N2-C2
4	A	1229	NAG	O7-C7-N2-C2
4	B	1281	NAG	C4-C5-C6-O6
4	B	1281	NAG	O5-C5-C6-O6
4	A	1085	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
4	A	1229	NAG	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 W94 B 1000

Bond lengths

Bond angles

Torsions

Rings

Ligand W94 A 1000

Bond lengths

Bond angles

Torsions

Rings
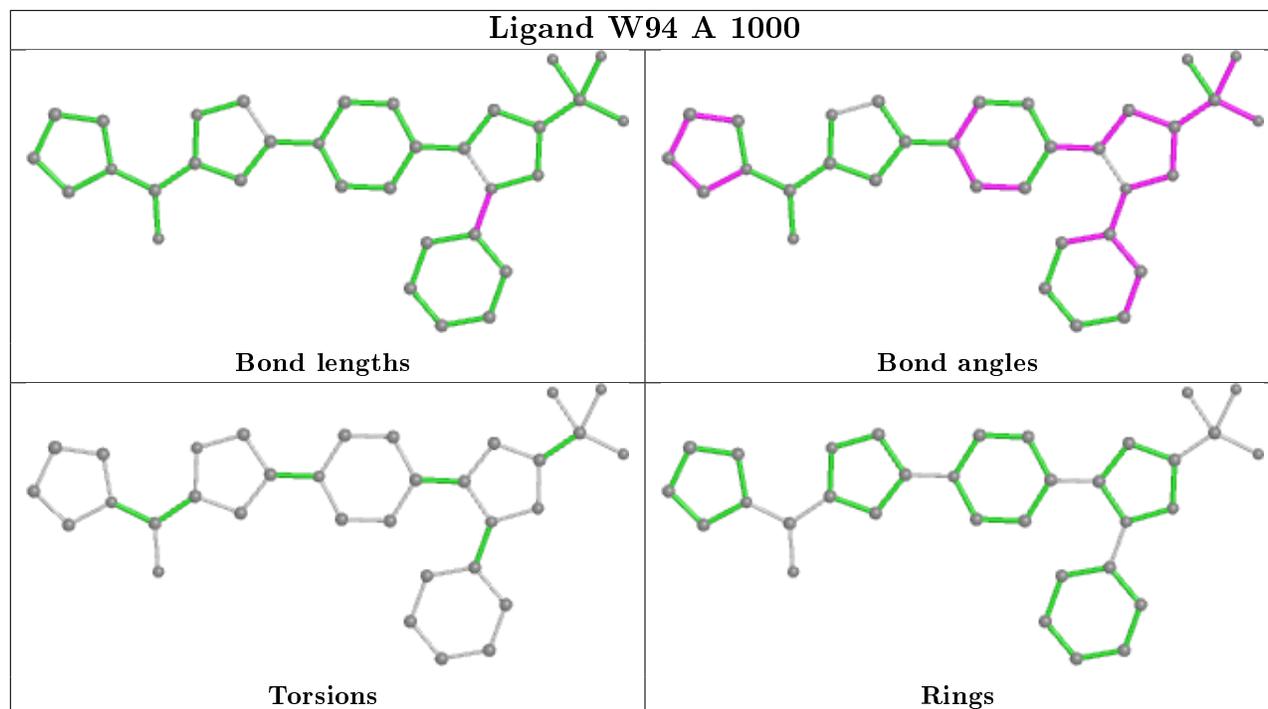
5.7 Other polymers (i)

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(Å²)	Q<0.9
1	A	729/740 (98%)	-0.05	25 (3%)	45	44	14, 23, 46, 66	0
1	B	729/740 (98%)	-0.15	18 (2%)	57	55	13, 22, 42, 56	0
All	All	1458/1480 (98%)	-0.10	43 (2%)	51	50	13, 23, 44, 66	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	766	PRO	11.8
1	B	766	PRO	10.8
1	A	765	LEU	8.7
1	B	765	LEU	5.8
1	A	39	SER	4.6
1	B	279	VAL	4.2
1	A	73	GLU	3.8
1	A	333	SER	3.6
1	A	96	ASP	3.6
1	A	97	GLU	3.5
1	B	73	GLU	3.5
1	B	97	GLU	3.4
1	A	506	ASN	3.3
1	A	38	ASP	3.2
1	A	138	ASN	3.1
1	B	342	ALA	2.8
1	B	138	ASN	2.7
1	B	99	GLY	2.7
1	B	105	TYR	2.7
1	A	72	GLN	2.6
1	B	140	ARG	2.6
1	A	93	SER	2.6
1	B	533	HIS	2.5
1	A	88	VAL	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	92	ASN	2.4
1	A	332	GLU	2.4
1	A	521	GLU	2.4
1	A	537	SER	2.4
1	A	91	GLU	2.3
1	A	378	GLU	2.3
1	B	341	VAL	2.3
1	A	518	ILE	2.3
1	B	278	SER	2.3
1	A	681	ASP	2.2
1	B	75	ASN	2.2
1	B	332	GLU	2.2
1	A	697	GLN	2.1
1	A	502	LYS	2.1
1	A	74	ASN	2.1
1	A	99	GLY	2.1
1	B	521	GLU	2.1
1	B	764	SER	2.1
1	A	98	PHE	2.0

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($Å^2$)	Q<0.9
2	NAG	G	2	14/15	0.79	0.40	53,55,58,59	0
2	NAG	F	2	14/15	0.80	0.35	50,53,56,57	0
2	NAG	E	2	14/15	0.82	0.34	53,54,57,57	0
2	NAG	C	2	14/15	0.87	0.29	50,52,55,55	0
2	NAG	E	1	14/15	0.90	0.25	40,45,49,50	0
2	NAG	G	1	14/15	0.91	0.25	40,45,48,51	0
2	NAG	C	1	14/15	0.91	0.26	33,42,47,47	0
2	NAG	F	1	14/15	0.94	0.18	34,38,42,46	0
2	NAG	D	2	14/15	0.94	0.12	27,33,35,35	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	NAG	D	1	14/15	0.97	0.11	23,24,29,31	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.
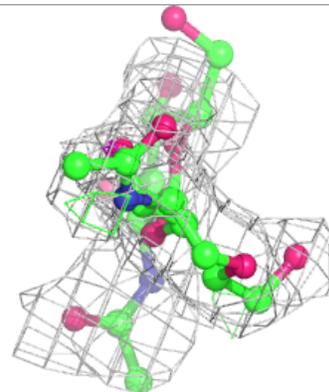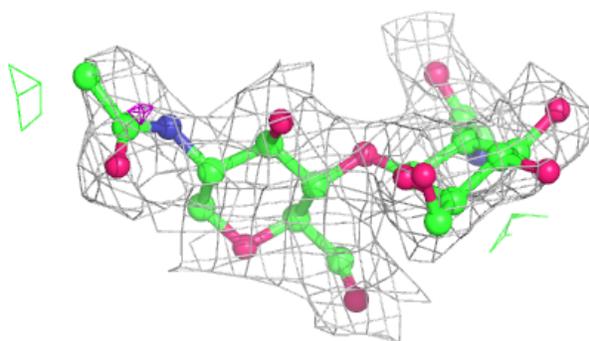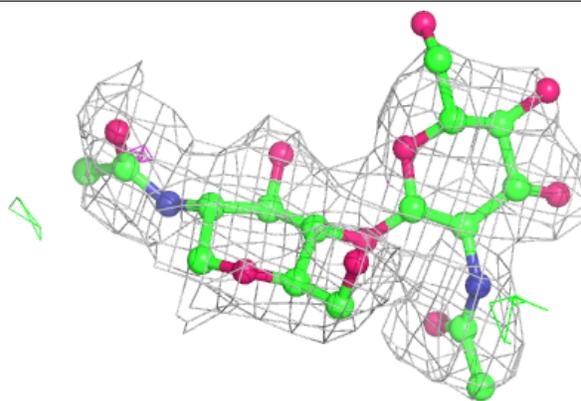

Electron density around Chain C:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
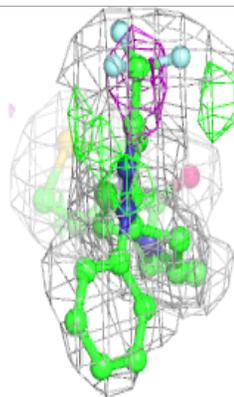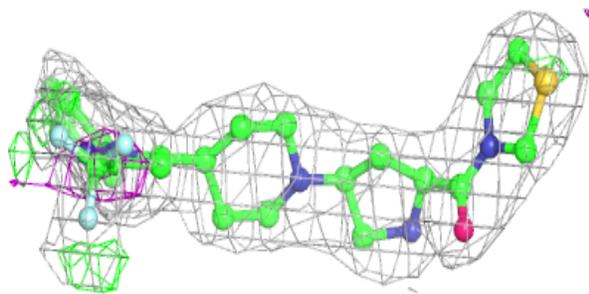mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
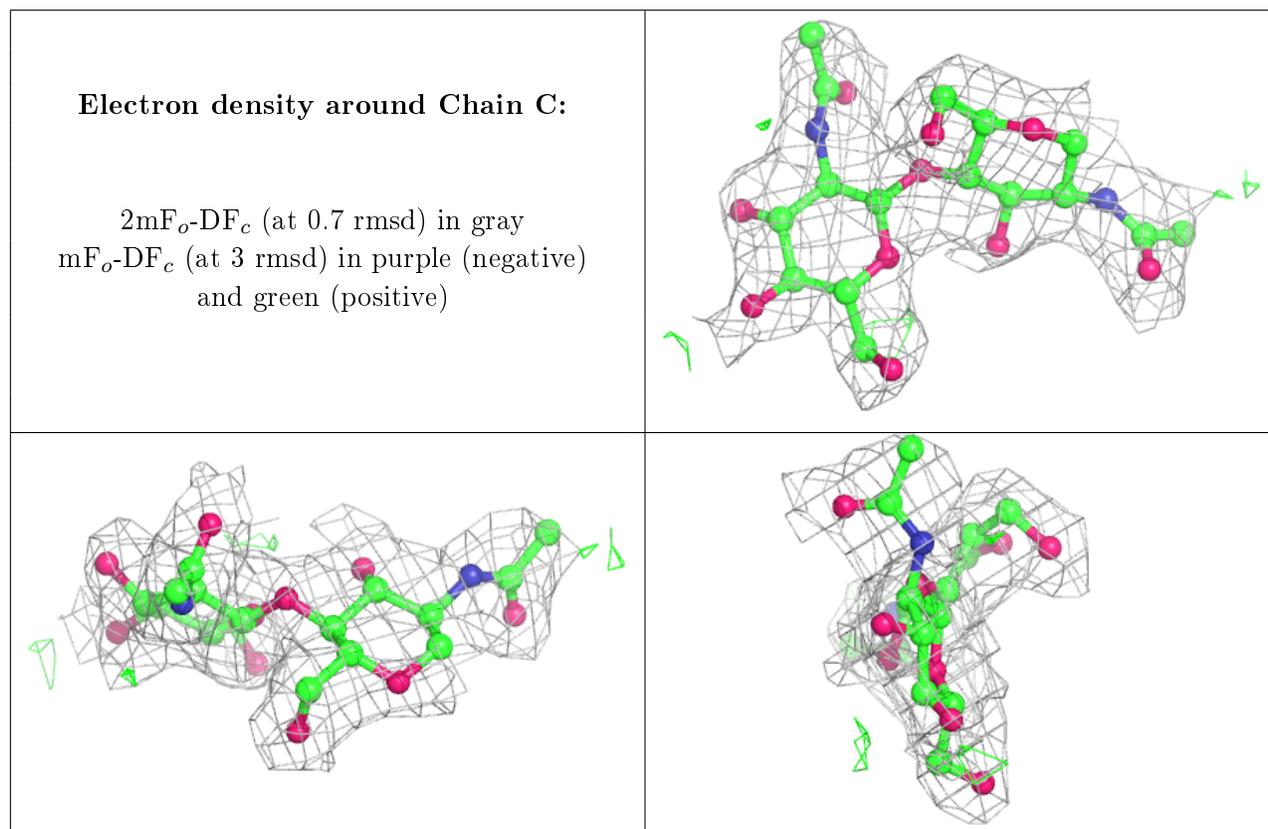
Electron density around Chain D:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
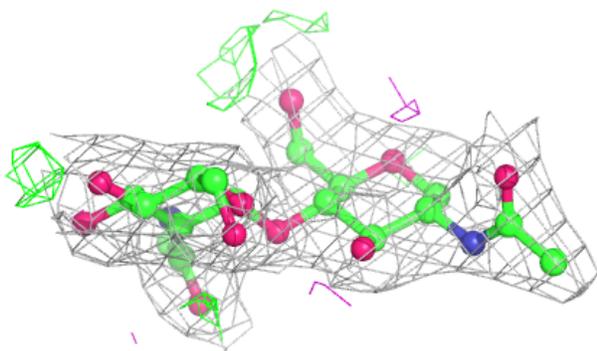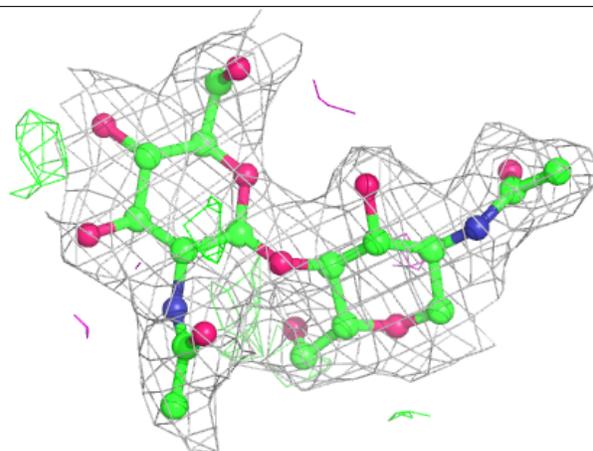
Electron density around Chain E:

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

Electron density around Chain F:

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 G:

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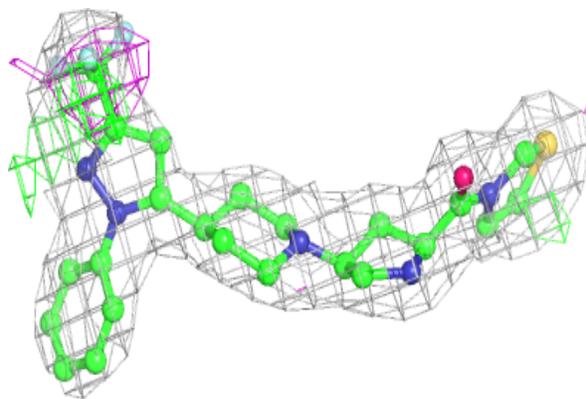and green (positive)
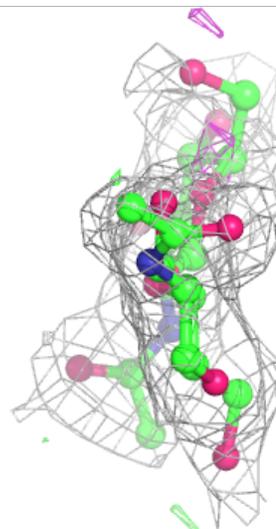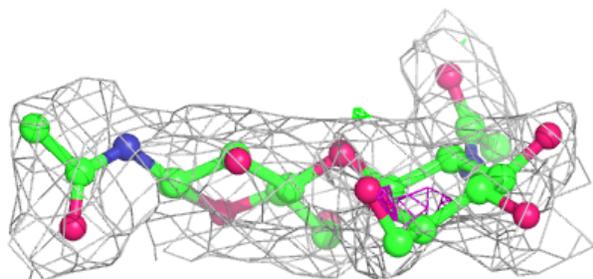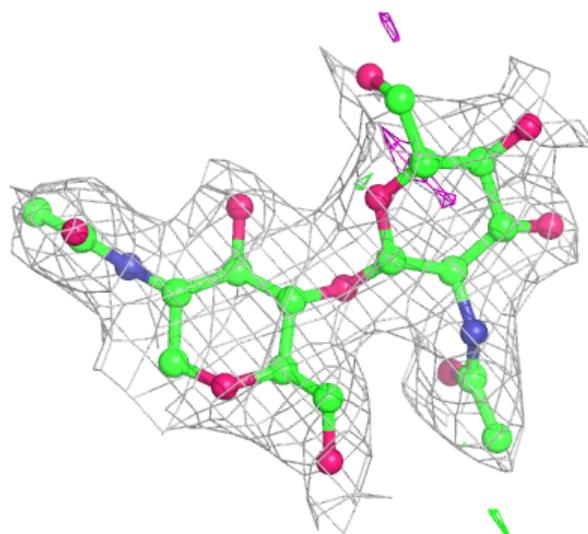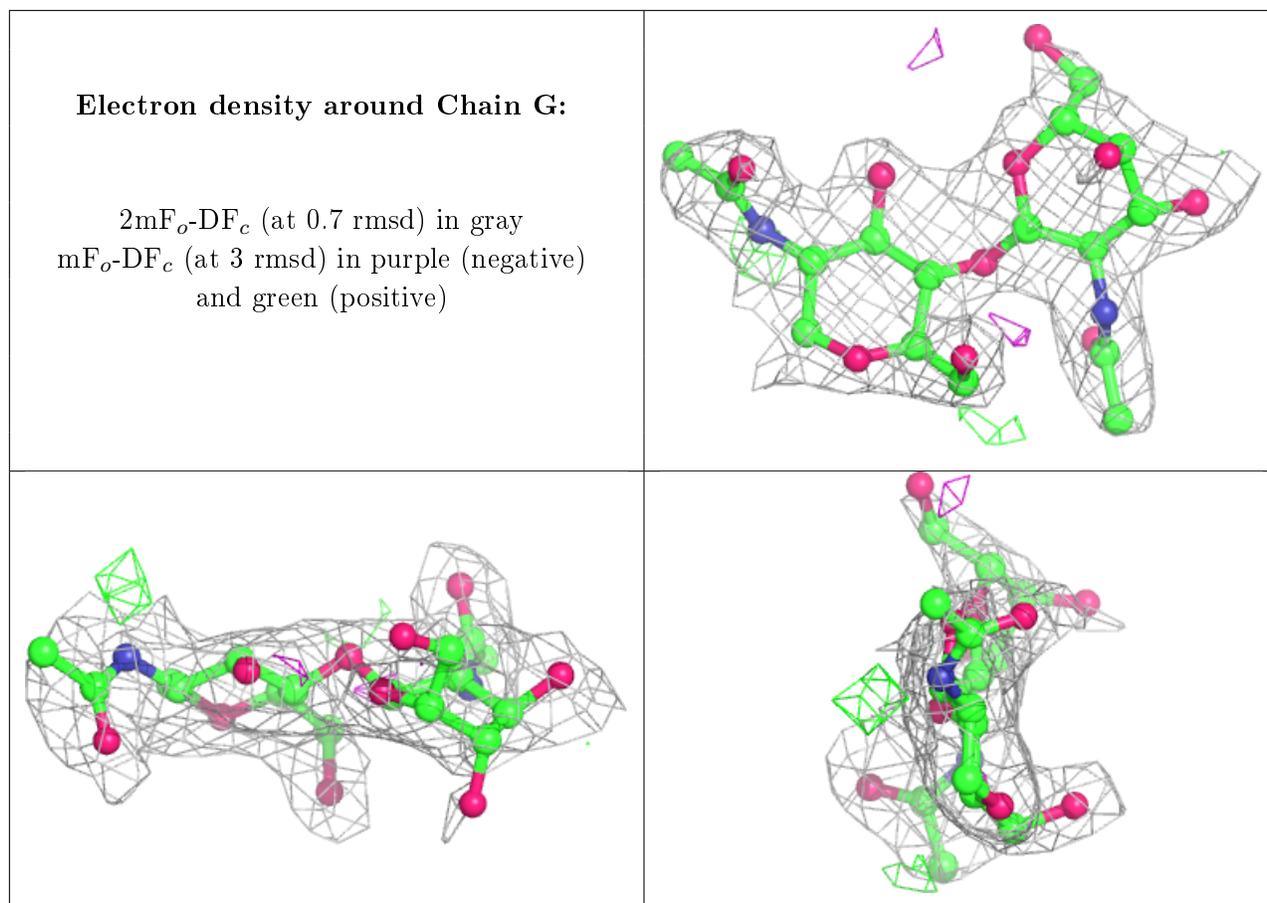
6.4 Ligands (i)

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.

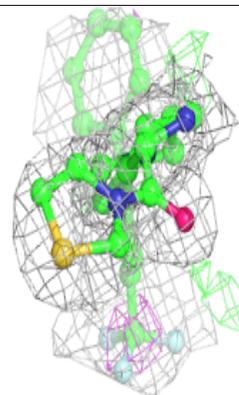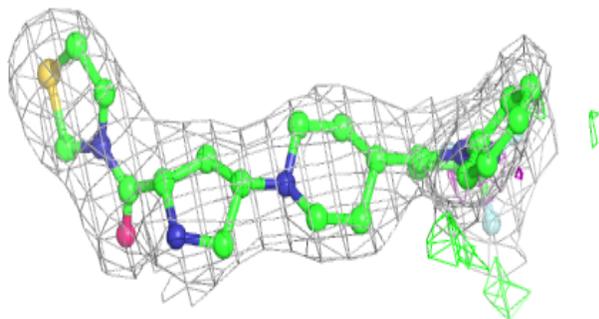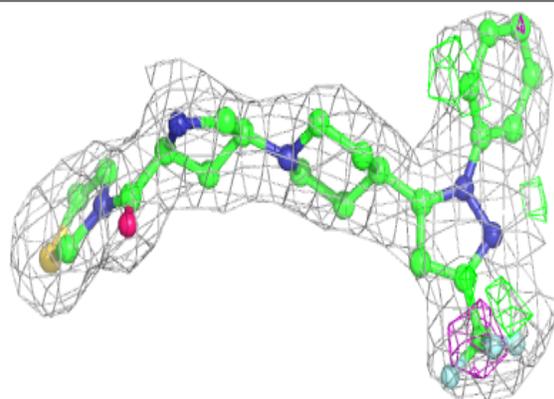| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	A	1085	14/15	0.78	0.19	53,55,57,58	0
4	NAG	B	1150	14/15	0.83	0.31	45,49,51,51	0
4	NAG	A	1229	14/15	0.85	0.22	29,33,35,38	0
4	NAG	B	1281	14/15	0.85	0.23	46,48,49,50	0
3	W94	B	1000	33/33	0.86	0.19	19,34,53,54	0
4	NAG	A	1321	14/15	0.86	0.23	35,38,40,41	0
3	W94	A	1000	33/33	0.87	0.18	16,29,52,53	0
4	NAG	B	1085	14/15	0.90	0.14	39,41,44,44	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 W94 B 1000:

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

Electron density around W94 A 1000:

$2mF_o$-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 (i)

There are no such residues in this entry.