wwPDB X-ray Structure Validation Summary Report (i)

Jan 29, 2024 – 03:07 PM EST

PDB ID	:	8V1F
Title	:	TMPRSS2 complexed with the noncovalent inhibitor 6-amidino-2-napthol
Authors	:	Fraser, B.J.; Dong, A.; Kutera, M.; Seitova, A.; Li, Y.; Hutchinson, A.; Edwards, A.; Benard, F.; Halabelian, L.; Arrowsmith, C.; Structural Genomics Consortium (SGC)
Deposited on	:	2023-11-20
Resolution	:	2.19 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

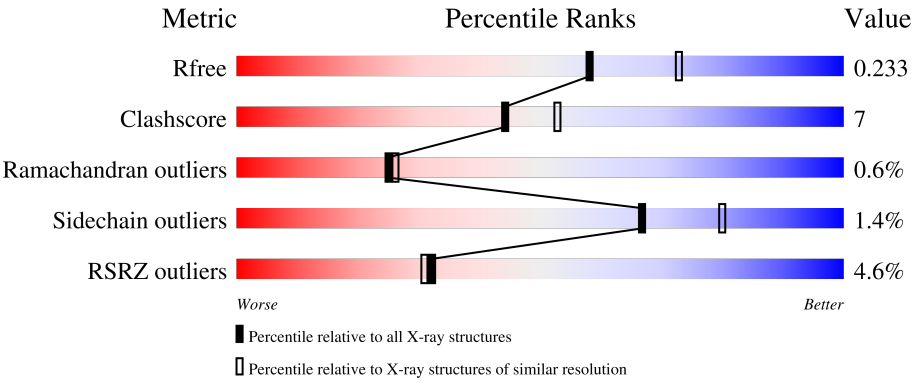
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.19 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}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	110	7% 76% 17% • 5%
1	C	110	15% 80% 15% 5%
2	B	249	2% 83% 13% • •
2	D	249	88% 9% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	2	100%
4	F	3	33% 67%

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
11	PG4	D	607	-	-	X	-
3	NAG	E	2	-	-	-	X
7	EDO	B	612	-	-	-	X

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 5926 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 Transmembrane protease serine 2 non-catalytic chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	105	Total	C	N	O	S	0	0	0
			798	497	138	155	8			
1	C	104	Total	C	N	O	S	0	0	0
			797	494	141	154	8			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	146	ALA	-	expression tag	UNP O15393
A	147	ALA	-	expression tag	UNP O15393
A	251	ASP	SER	engineered mutation	UNP O15393
A	252	ASP	ARG	engineered mutation	UNP O15393
A	253	ASP	GLN	engineered mutation	UNP O15393
A	254	ASP	SER	engineered mutation	UNP O15393
A	255	LYS	ARG	engineered mutation	UNP O15393
C	146	ALA	-	expression tag	UNP O15393
C	147	ALA	-	expression tag	UNP O15393
C	251	ASP	SER	engineered mutation	UNP O15393
C	252	ASP	ARG	engineered mutation	UNP O15393
C	253	ASP	GLN	engineered mutation	UNP O15393
C	254	ASP	SER	engineered mutation	UNP O15393
C	255	LYS	ARG	engineered mutation	UNP O15393

- Molecule 2 is a protein called Transmembrane protease serine 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	241	Total	C	N	O	S	0	1	0
			1847	1189	308	334	16			
2	D	243	Total	C	N	O	S	0	3	0
			1894	1217	319	342	16			

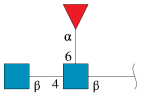
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	493	GLU	-	expression tag	UNP O15393
B	494	PHE	-	expression tag	UNP O15393
B	495	VAL	-	expression tag	UNP O15393
B	496	GLU	-	expression tag	UNP O15393
B	497	HIS	-	expression tag	UNP O15393
B	498	HIS	-	expression tag	UNP O15393
B	499	HIS	-	expression tag	UNP O15393
B	500	HIS	-	expression tag	UNP O15393
B	501	HIS	-	expression tag	UNP O15393
B	502	HIS	-	expression tag	UNP O15393
B	503	HIS	-	expression tag	UNP O15393
B	504	HIS	-	expression tag	UNP O15393
D	493	GLU	-	expression tag	UNP O15393
D	494	PHE	-	expression tag	UNP O15393
D	495	VAL	-	expression tag	UNP O15393
D	496	GLU	-	expression tag	UNP O15393
D	497	HIS	-	expression tag	UNP O15393
D	498	HIS	-	expression tag	UNP O15393
D	499	HIS	-	expression tag	UNP O15393
D	500	HIS	-	expression tag	UNP O15393
D	501	HIS	-	expression tag	UNP O15393
D	502	HIS	-	expression tag	UNP O15393
D	503	HIS	-	expression tag	UNP O15393
D	504	HIS	-	expression tag	UNP O15393

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

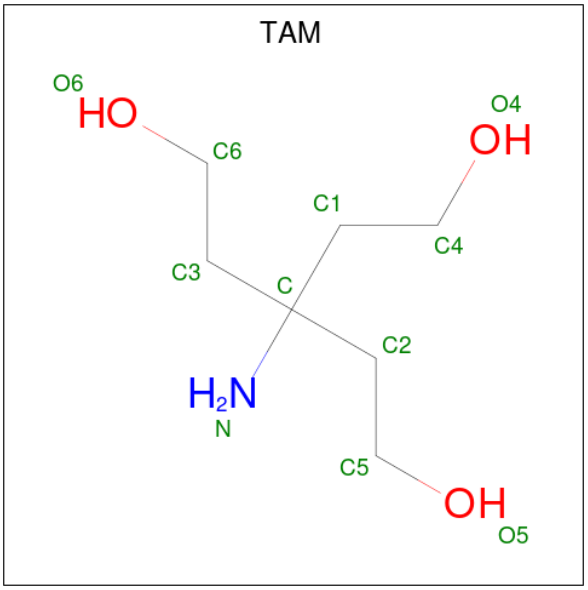
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	2	28	16	2	10	0	0	0

- Molecule 4 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
4	F	3	Total C N O 38 22 2 14	0	0	0

- Molecule 5 is TRIS(HYDROXYETHYL)AMINOMETHANE (three-letter code: TAM) (formula: $C_7H_{17}NO_3$).

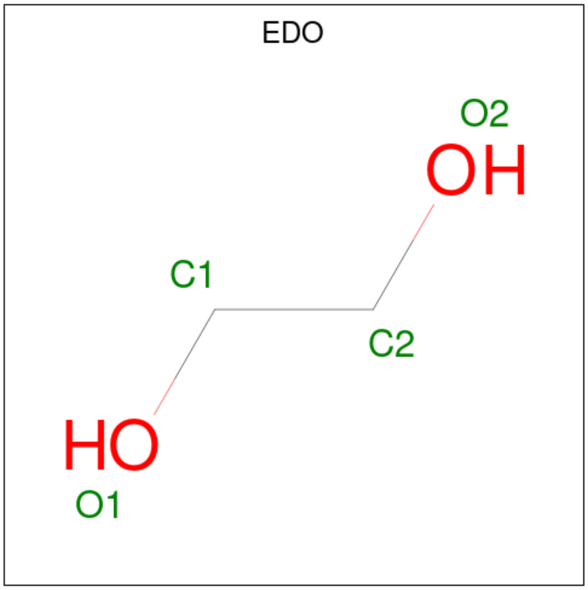


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 11 7 1 3	0	0
5	B	1	Total C N O 11 7 1 3	0	0
5	B	1	Total C N O 11 7 1 3	0	0
5	D	1	Total C N O 11 7 1 3	0	0

- Molecule 6 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total X 1 1	0	0
6	C	1	Total X 1 1	0	0
6	D	1	Total X 1 1	0	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



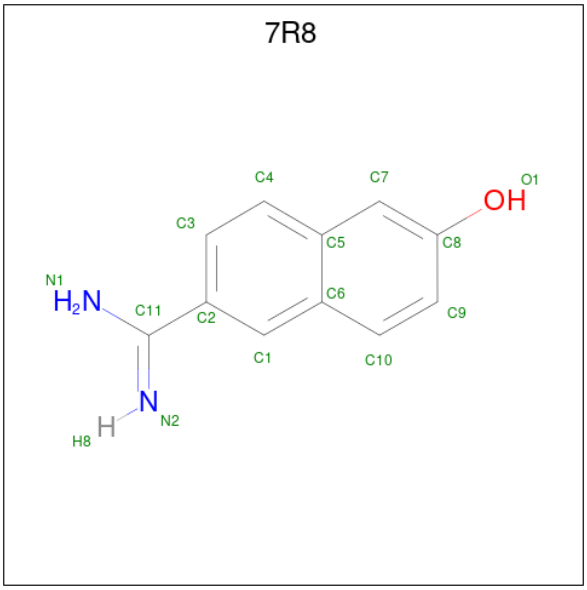
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	C	1	Total C O 4 2 2	0	0
7	D	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

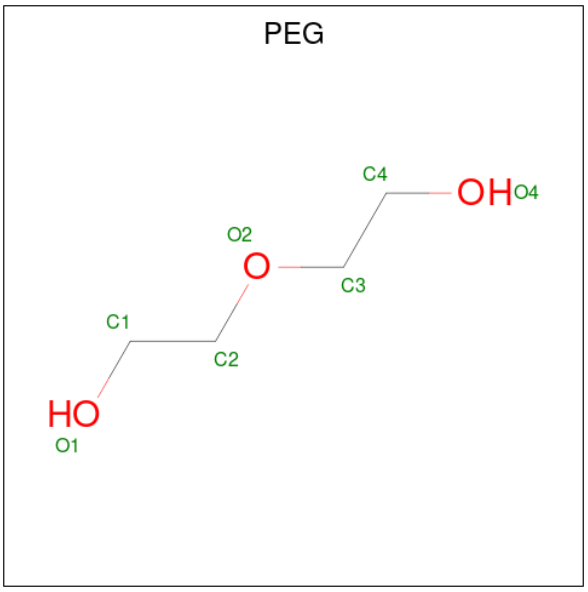
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total C O 4 2 2	0	0
7	D	1	Total C O 4 2 2	0	0
7	D	1	Total C O 4 2 2	0	0
7	D	1	Total C O 4 2 2	0	0
7	D	1	Total C O 4 2 2	0	0
7	D	1	Total C O 4 2 2	0	0

- Molecule 8 is 6-oxidanylnaphthalene-2-carboximidamide (three-letter code: 7R8) (formula: $C_{11}H_{10}N_2O$) (labeled as "Ligand of Interest" by depositor).



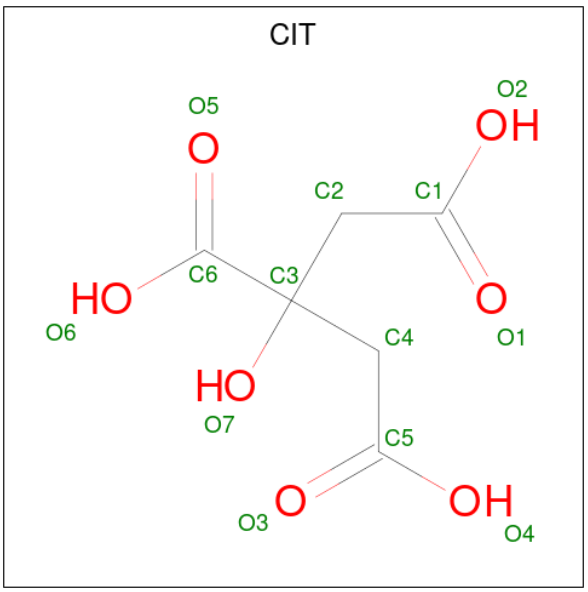
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C N O 14 11 2 1	0	0
8	D	1	Total C N O 14 11 2 1	0	0

- Molecule 9 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



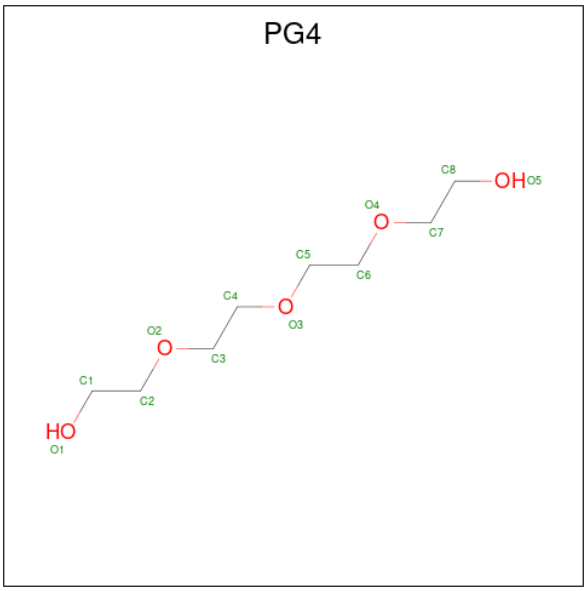
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	1	Total C O 7 4 3	0	0

- Molecule 10 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	C	1	Total C O 13 6 7	0	0
10	C	1	Total C O 13 6 7	0	0

- Molecule 11 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	C	1	Total C O 13 8 5	0	0
11	D	1	Total C O 13 8 5	0	0

- Molecule 12 is water.

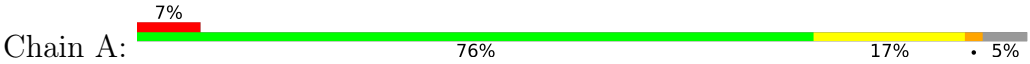
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	34	Total O 34 34	0	0
12	B	127	Total O 127 127	0	0
12	C	28	Total O 28 28	0	0
12	D	121	Total O 121 121	0	0

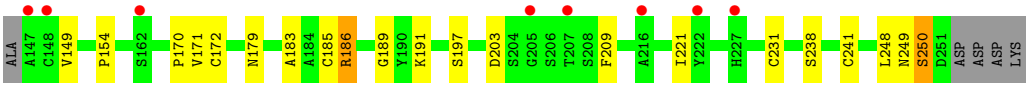
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Transmembrane protease serine 2 non-catalytic chain

Chain A:



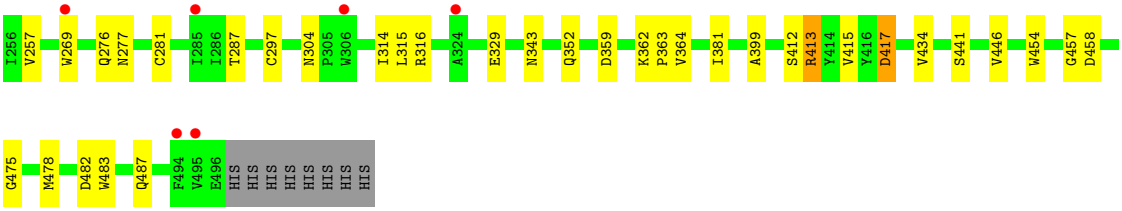


• Molecule 1: Transmembrane protease serine 2 non-catalytic chain

Chain C:

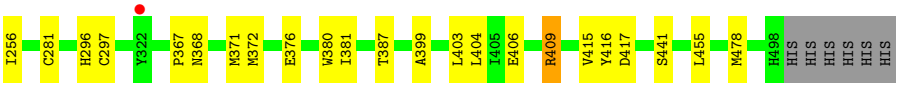
• Molecule 2: Transmembrane protease serine 2

Chain B:



• Molecule 2: Transmembrane protease serine 2

Chain D:



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

Chain E:

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

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	118.94Å 172.57Å 83.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.67 – 2.19 38.64 – 2.19	Depositor EDS
% Data completeness (in resolution range)	99.5 (38.67-2.19) 99.5 (38.64-2.19)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 2.18Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.172 , 0.223 0.181 , 0.233	Depositor DCC
R_{free} test set	462 reflections (1.04%)	wwPDB-VP
Wilson B-factor (Å^2)	37.9	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 46.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5926	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.93% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, PEG, TAM, 7R8, EDO, CIT, FUC, UNX, PG4

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	0.45	0/815	0.78	0/1105
1	C	0.42	0/814	0.78	0/1103
2	B	0.47	0/1905	0.81	0/2606
2	D	0.47	0/1960	0.81	1/2677 (0.0%)
All	All	0.46	0/5494	0.80	1/7491 (0.0%)

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
2	B	0	2
2	D	0	1
All	All	0	4

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	416	TYR	CB-CA-C	5.59	121.58	110.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	186	ARG	Sidechain
2	B	316	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	B	413	ARG	Sidechain
2	D	409	ARG	Sidechain

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	798	0	728	16	0
1	C	797	0	732	11	0
2	B	1847	0	1753	24	1
2	D	1894	0	1810	20	1
3	E	28	0	25	0	0
4	F	38	0	34	3	0
5	A	11	0	17	1	0
5	B	22	0	34	2	0
5	D	11	0	17	1	0
6	A	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
7	A	12	0	18	3	0
7	B	36	0	54	8	0
7	C	4	0	6	0	0
7	D	28	0	42	0	0
8	B	14	0	0	0	0
8	D	14	0	0	0	0
9	C	7	0	10	1	0
10	C	26	0	10	2	0
11	C	13	0	18	0	0
11	D	13	0	18	9	0
12	A	34	0	0	0	0
12	B	127	0	0	4	0
12	C	28	0	0	0	0
12	D	121	0	0	2	0
All	All	5926	0	5326	70	1

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 7.

The worst 5 of 70 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)
2:B:478:MET:HG3	12:B:732:HOH:O	1.89	0.72
2:B:482:ASP:OD1	12:B:701:HOH:O	2.08	0.71
2:D:478:MET:HG3	12:D:756:HOH:O	1.88	0.71
2:D:376:GLU:OE1	11:D:607:PG4:H42	1.91	0.71
1:A:186:ARG:CB	7:A:604:EDO:H22	2.27	0.65

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:417:ASP:OD1	2:D:417:ASP:OD2[6_554]	1.96	0.24

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	103/110 (94%)	97 (94%)	5 (5%)	1 (1%)	15	14
1	C	102/110 (93%)	99 (97%)	2 (2%)	1 (1%)	15	14
2	B	240/249 (96%)	230 (96%)	9 (4%)	1 (0%)	34	37
2	D	244/249 (98%)	235 (96%)	8 (3%)	1 (0%)	34	37
All	All	689/718 (96%)	661 (96%)	24 (4%)	4 (1%)	25	26

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	250	SER
1	C	164	GLN
2	B	415	VAL
2	D	415	VAL

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	84/95 (88%)	82 (98%)	2 (2%)	49	62
1	C	86/95 (90%)	84 (98%)	2 (2%)	50	63
2	B	192/210 (91%)	189 (98%)	3 (2%)	62	76
2	D	200/210 (95%)	199 (100%)	1 (0%)	88	94
All	All	562/610 (92%)	554 (99%)	8 (1%)	67	80

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	372	MET
1	C	214	THR
2	B	417	ASP
2	B	413	ARG
1	C	165	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	198	GLN
1	C	247	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)

5 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	$\#\|Z\| > 2$
3	NAG	E	1	3,1	14,14,15	0.40	0	17,19,21	1.05	0
3	NAG	E	2	3	14,14,15	0.41	0	17,19,21	0.86	0
4	NAG	F	1	4,1	14,14,15	0.43	0	17,19,21	0.92	1 (5%)
4	NAG	F	2	4	14,14,15	0.49	0	17,19,21	0.92	0
4	FUC	F	3	4	10,10,11	0.41	0	14,14,16	0.71	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
3	NAG	E	1	3,1	-	2/6/23/26	0/1/1/1
3	NAG	E	2	3	-	2/6/23/26	0/1/1/1
4	NAG	F	1	4,1	-	4/6/23/26	0/1/1/1
4	NAG	F	2	4	-	4/6/23/26	0/1/1/1
4	FUC	F	3	4	-	-	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	1	NAG	C4-C3-C2	2.02	113.98	111.02

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	2	NAG	C8-C7-N2-C2

Continued on next page...

Continued from previous page...

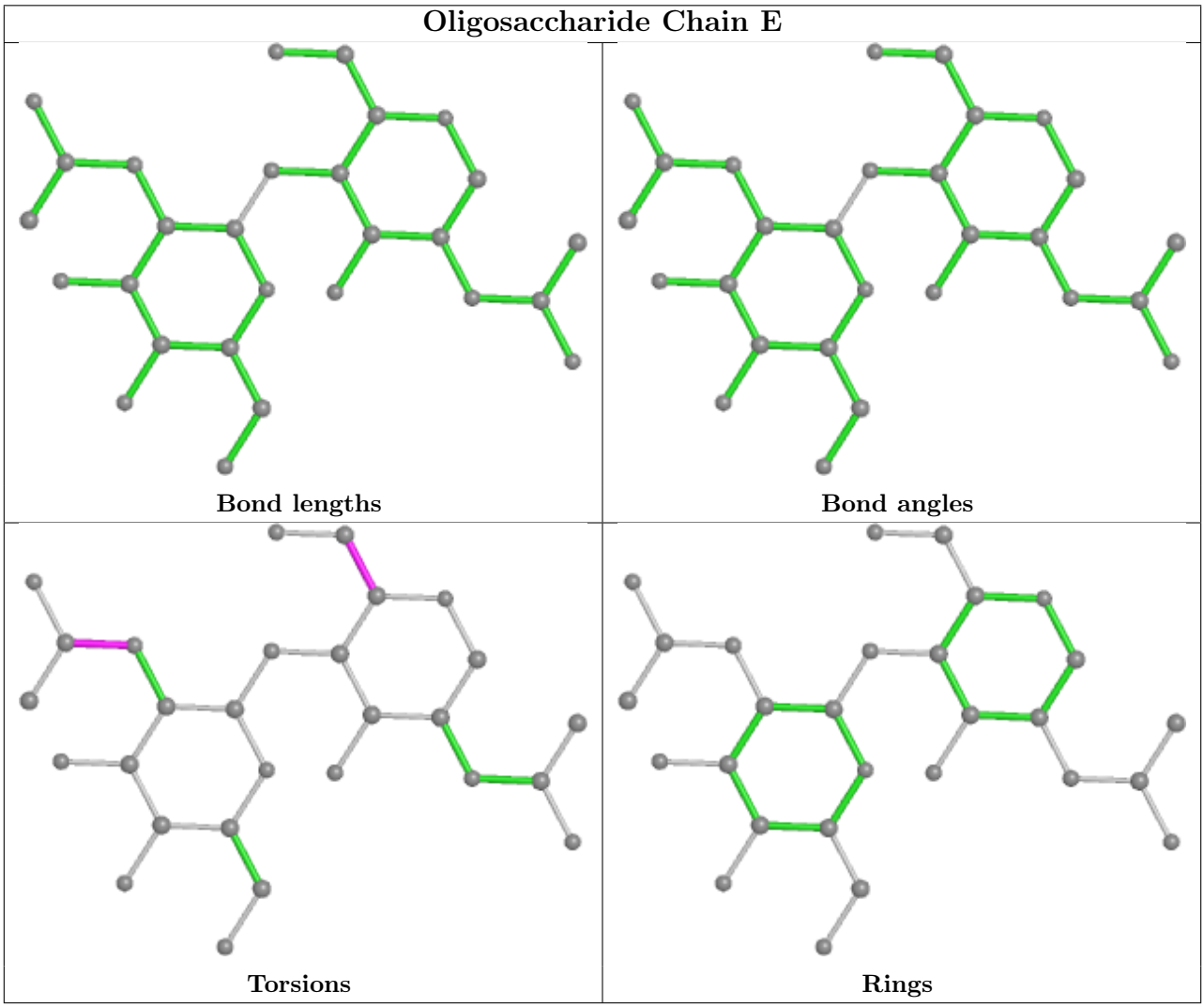
Mol	Chain	Res	Type	Atoms
3	E	2	NAG	O7-C7-N2-C2
4	F	2	NAG	C3-C2-N2-C7
4	F	2	NAG	C8-C7-N2-C2
4	F	2	NAG	O7-C7-N2-C2

There are no ring outliers.

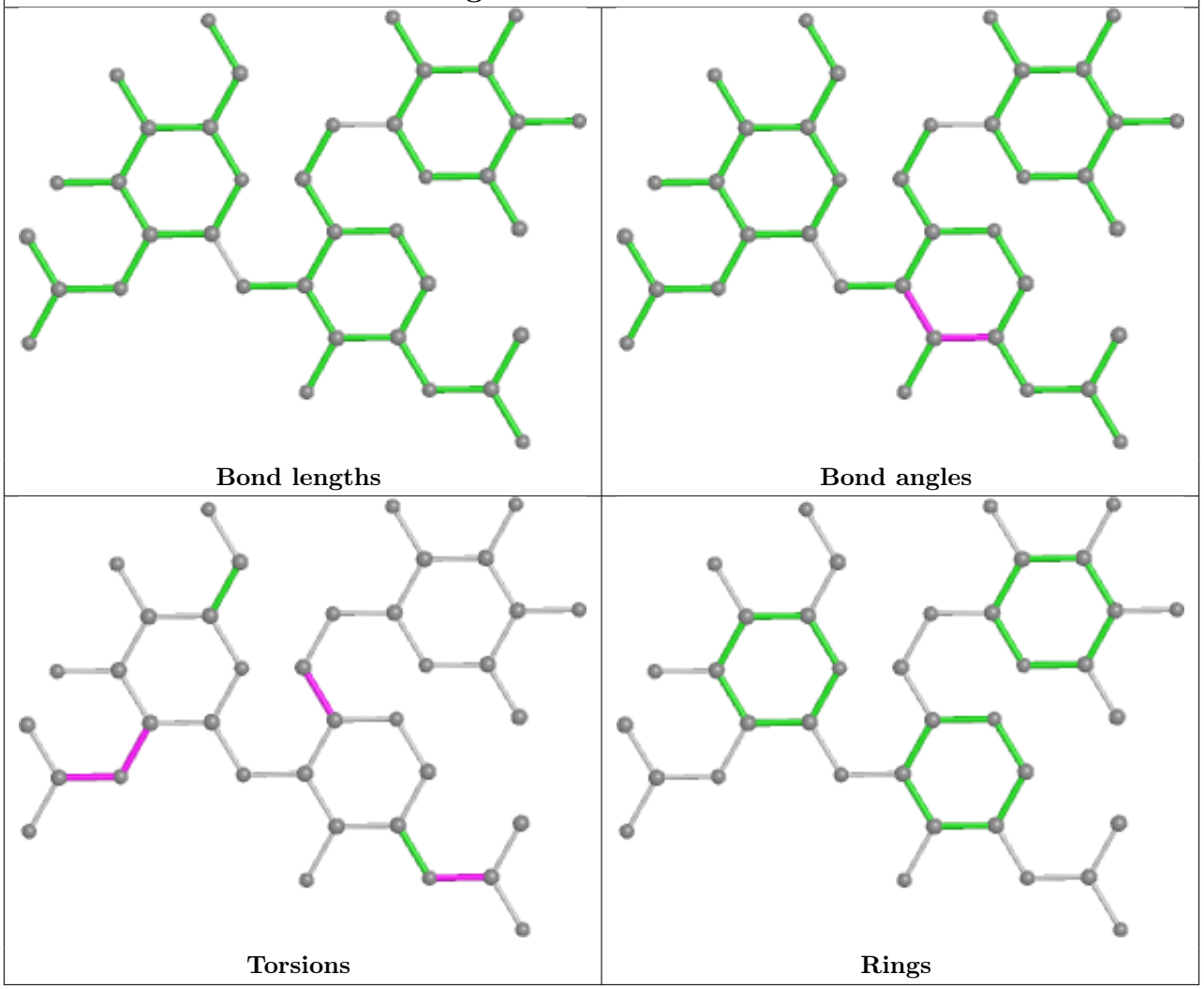
1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	3	FUC	3	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 F



5.6 Ligand geometry (i)

Of 34 ligands modelled in this entry, 3 are unknown - leaving 31 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	$\#\|Z\| > 2$
7	EDO	B	608	-	3,3,3	0.13	0	2,2,2	0.34	0
8	7R8	B	604	-	15,15,15	0.85	1 (6%)	18,21,21	0.43	0
7	EDO	D	605	-	3,3,3	0.23	0	2,2,2	0.22	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	EDO	B	609	-	3,3,3	0.23	0	2,2,2	0.06	0
7	EDO	C	304	-	3,3,3	0.48	0	2,2,2	0.70	0
5	TAM	A	601	-	7,10,10	0.40	0	9,12,12	1.17	1 (11%)
11	PG4	D	607	-	12,12,12	0.92	0	11,11,11	0.72	0
7	EDO	D	606	-	3,3,3	0.29	0	2,2,2	0.52	0
5	TAM	D	602	-	7,10,10	0.78	0	9,12,12	1.54	2 (22%)
7	EDO	D	609	-	3,3,3	0.21	0	2,2,2	0.33	0
7	EDO	B	612	-	3,3,3	0.35	0	2,2,2	0.41	0
7	EDO	D	610	-	3,3,3	0.18	0	2,2,2	0.24	0
7	EDO	B	610	-	3,3,3	0.26	0	2,2,2	0.37	0
7	EDO	B	601	-	3,3,3	0.22	0	2,2,2	0.27	0
7	EDO	A	604	-	3,3,3	0.27	0	2,2,2	0.48	0
7	EDO	D	604	-	3,3,3	0.24	0	2,2,2	0.42	0
8	7R8	D	603	-	15,15,15	0.81	0	18,21,21	0.92	1 (5%)
7	EDO	B	602	-	3,3,3	0.26	0	2,2,2	0.56	0
10	CIT	C	302	-	12,12,12	1.14	1 (8%)	17,17,17	1.38	2 (11%)
7	EDO	D	611	-	3,3,3	0.23	0	2,2,2	0.21	0
7	EDO	A	603	-	3,3,3	0.10	0	2,2,2	0.08	0
7	EDO	D	608	-	3,3,3	0.24	0	2,2,2	0.38	0
11	PG4	C	305	-	12,12,12	0.50	0	11,11,11	0.37	0
5	TAM	B	603	-	7,10,10	0.43	0	9,12,12	1.70	2 (22%)
9	PEG	C	301	-	6,6,6	0.44	0	5,5,5	0.74	0
5	TAM	B	605	-	7,10,10	0.43	0	9,12,12	1.47	1 (11%)
7	EDO	A	605	-	3,3,3	0.18	0	2,2,2	0.21	0
7	EDO	B	606	-	3,3,3	0.33	0	2,2,2	0.41	0
10	CIT	C	306	-	12,12,12	1.28	1 (8%)	17,17,17	1.51	2 (11%)
7	EDO	B	611	-	3,3,3	0.07	0	2,2,2	0.21	0
7	EDO	B	607	-	3,3,3	0.24	0	2,2,2	0.51	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
7	EDO	B	608	-	-	1/1/1/1	-
8	7R8	B	604	-	-	0/4/4/4	0/2/2/2
7	EDO	D	605	-	-	1/1/1/1	-
7	EDO	B	609	-	-	1/1/1/1	-
7	EDO	C	304	-	-	1/1/1/1	-
5	TAM	A	601	-	-	8/12/12/12	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
11	PG4	D	607	-	-	7/10/10/10	-
7	EDO	D	606	-	-	1/1/1/1	-
5	TAM	D	602	-	-	6/12/12/12	-
7	EDO	D	609	-	-	1/1/1/1	-
7	EDO	B	612	-	-	1/1/1/1	-
7	EDO	D	610	-	-	1/1/1/1	-
7	EDO	B	610	-	-	0/1/1/1	-
7	EDO	B	601	-	-	0/1/1/1	-
7	EDO	A	604	-	-	1/1/1/1	-
7	EDO	D	604	-	-	1/1/1/1	-
8	7R8	D	603	-	-	0/4/4/4	0/2/2/2
7	EDO	B	602	-	-	1/1/1/1	-
10	CIT	C	302	-	-	0/16/16/16	-
7	EDO	D	611	-	-	0/1/1/1	-
7	EDO	A	603	-	-	1/1/1/1	-
7	EDO	D	608	-	-	1/1/1/1	-
11	PG4	C	305	-	-	5/10/10/10	-
5	TAM	B	603	-	-	4/12/12/12	-
9	PEG	C	301	-	-	3/4/4/4	-
5	TAM	B	605	-	-	5/12/12/12	-
7	EDO	A	605	-	-	1/1/1/1	-
7	EDO	B	606	-	-	0/1/1/1	-
10	CIT	C	306	-	-	9/16/16/16	-
7	EDO	B	611	-	-	1/1/1/1	-
7	EDO	B	607	-	-	1/1/1/1	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	C	306	CIT	C3-C6	2.83	1.56	1.53
10	C	302	CIT	O2-C1	-2.21	1.23	1.30
8	B	604	7R8	C3-C2	2.19	1.43	1.39

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	C	306	CIT	O5-C6-C3	-3.90	116.72	122.25
5	B	603	TAM	O6-C6-C3	-3.89	101.01	111.39
10	C	302	CIT	O5-C6-C3	-3.68	117.04	122.25
8	D	603	7R8	C2-C11-N1	3.59	123.46	118.05
10	C	306	CIT	O6-C6-C3	3.45	119.03	113.05

There are no chirality outliers.

5 of 63 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	601	TAM	N-C-C1-C4
5	A	601	TAM	C1-C-C2-C5
5	A	601	TAM	C3-C-C2-C5
5	A	601	TAM	N-C-C2-C5
5	A	601	TAM	C-C1-C4-O4

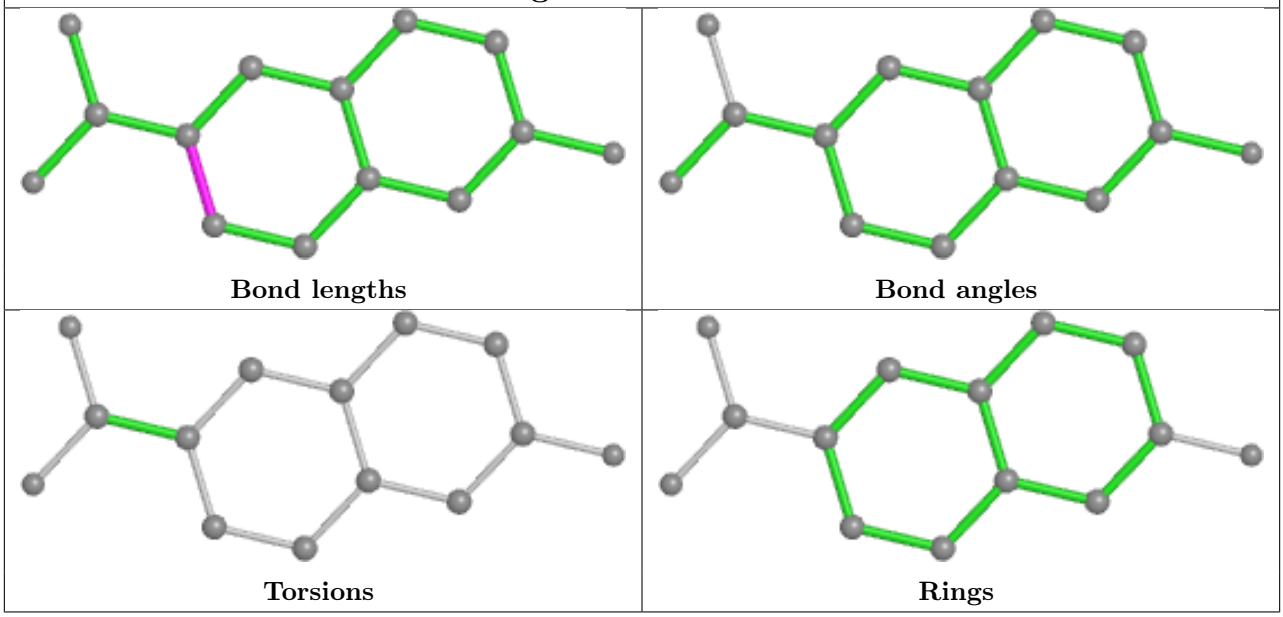
There are no ring outliers.

11 monomers are involved in 27 short contacts:

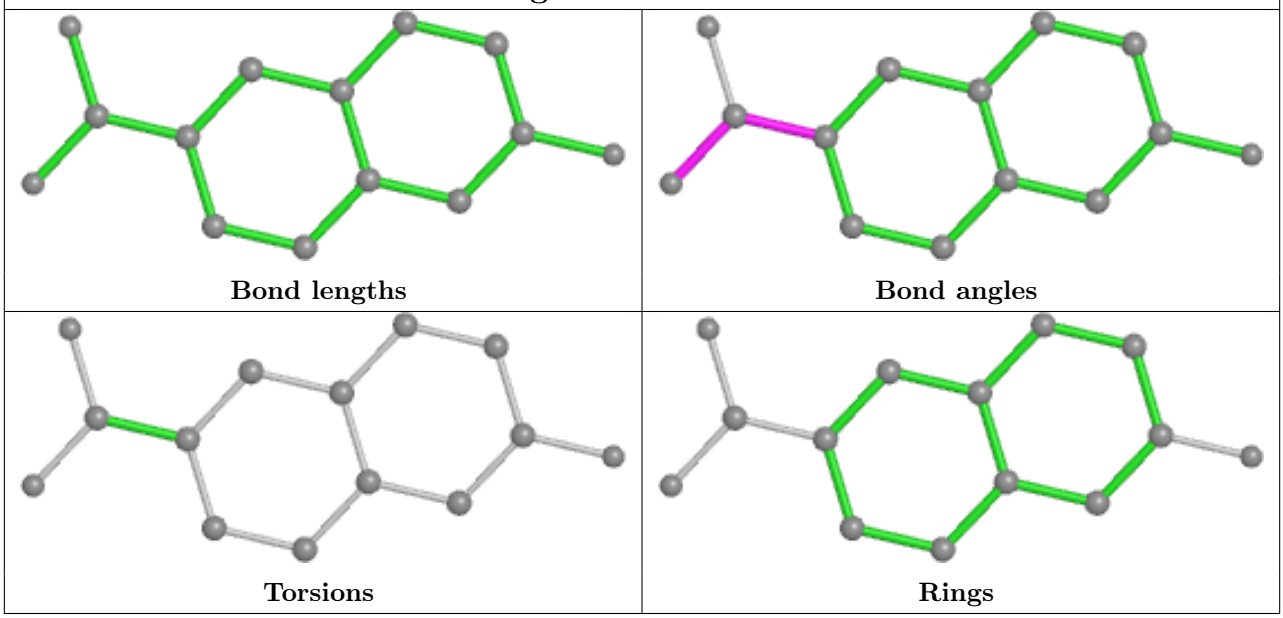
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	601	TAM	1	0
11	D	607	PG4	9	0
5	D	602	TAM	1	0
7	B	612	EDO	2	0
7	B	601	EDO	1	0
7	A	604	EDO	3	0
7	B	602	EDO	2	0
9	C	301	PEG	1	0
5	B	605	TAM	2	0
10	C	306	CIT	2	0
7	B	607	EDO	3	0

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 7R8 B 604



Ligand 7R8 D 603



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

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, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	105/110 (95%)	0.34	8 (7%) 13 12	29, 47, 79, 94	0
1	C	104/110 (94%)	0.68	17 (16%) 1 1	32, 56, 85, 97	0
2	B	241/249 (96%)	0.01	6 (2%) 57 55	26, 38, 60, 84	0
2	D	243/249 (97%)	-0.16	1 (0%) 92 91	25, 35, 55, 87	0
All	All	693/718 (96%)	0.10	32 (4%) 32 31	25, 39, 72, 97	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	148	CYS	5.9
1	C	163	SER	4.2
1	A	147	ALA	3.9
1	C	209	PHE	3.4
1	C	232	SER	3.4

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

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

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, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	E	2	14/15	0.63	0.46	82,109,130,135	0
3	NAG	E	1	14/15	0.81	0.31	59,92,107,125	0

Continued on next page...

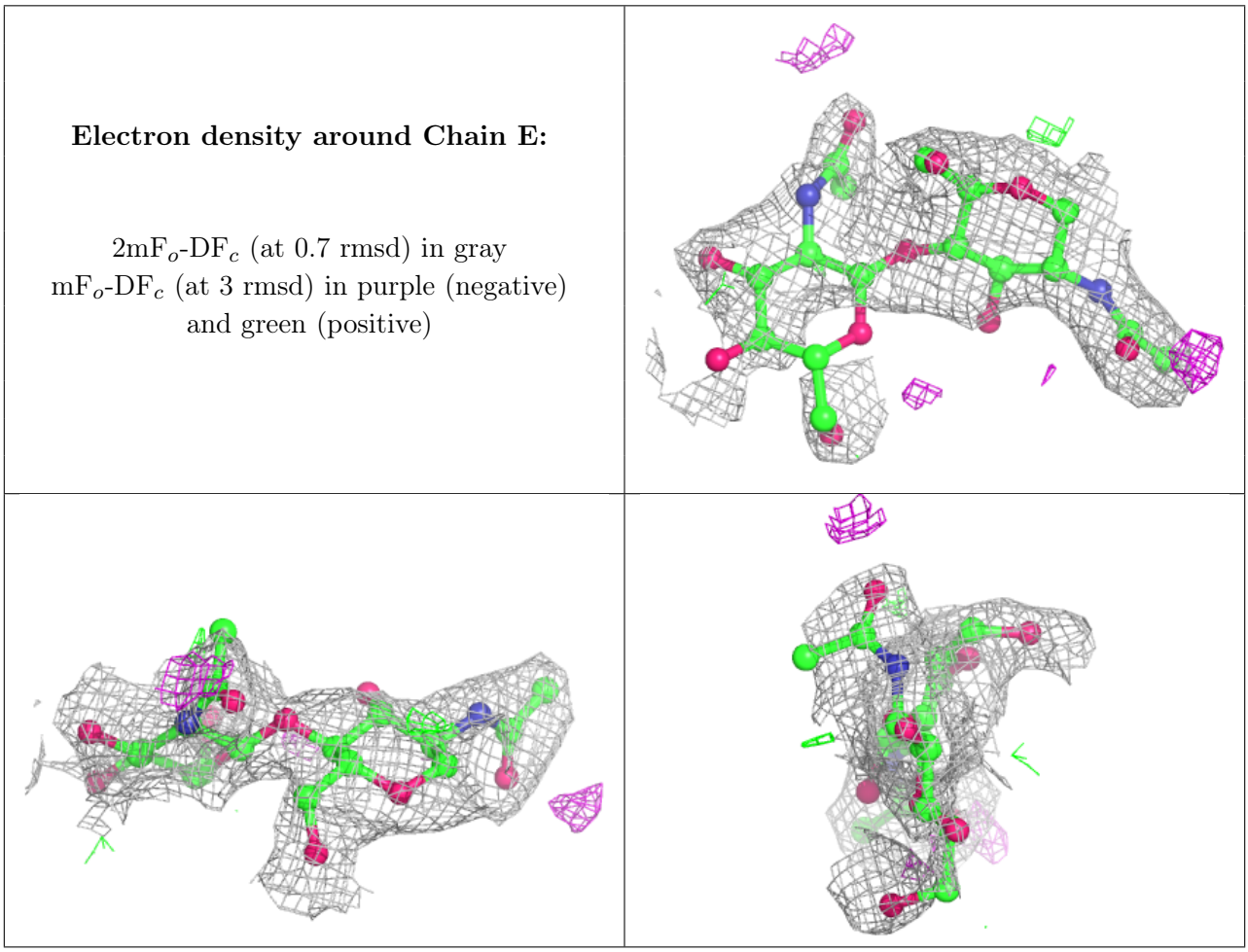
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	F	2	14/15	0.85	0.25	73,98,108,111	0
4	FUC	F	3	10/11	0.91	0.17	76,93,97,103	0
4	NAG	F	1	14/15	0.92	0.13	70,80,89,92	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.

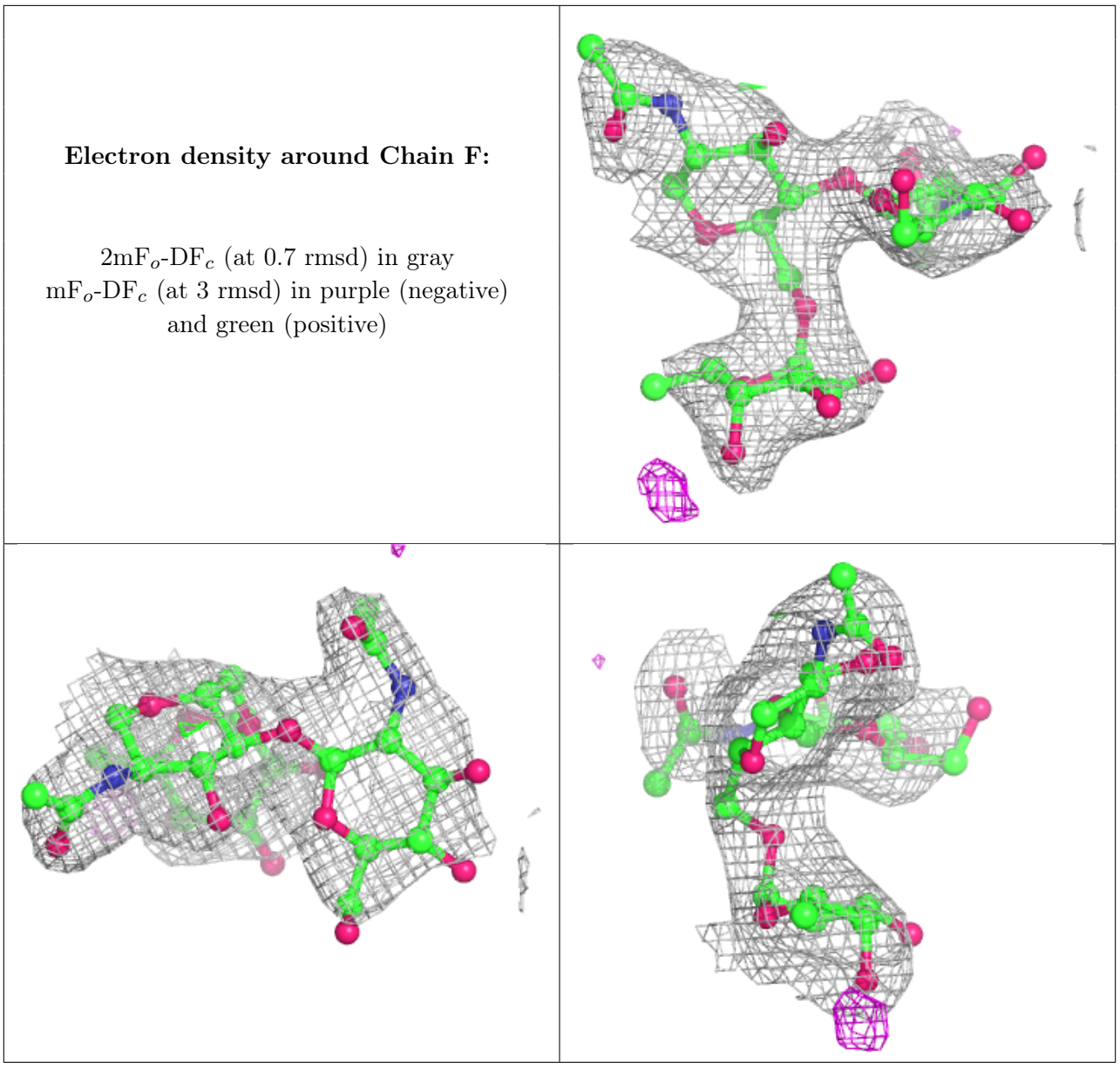
Electron density around Chain E:

$2mF_o$-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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	EDO	D	611	4/4	0.65	0.24	70,70,72,73	0
7	EDO	A	603	4/4	0.67	0.20	76,77,78,79	0
7	EDO	B	610	4/4	0.77	0.33	65,75,76,80	0
11	PG4	C	305	13/13	0.77	0.37	65,70,78,82	0
7	EDO	B	612	4/4	0.78	0.47	63,65,67,68	0

Continued on next page...

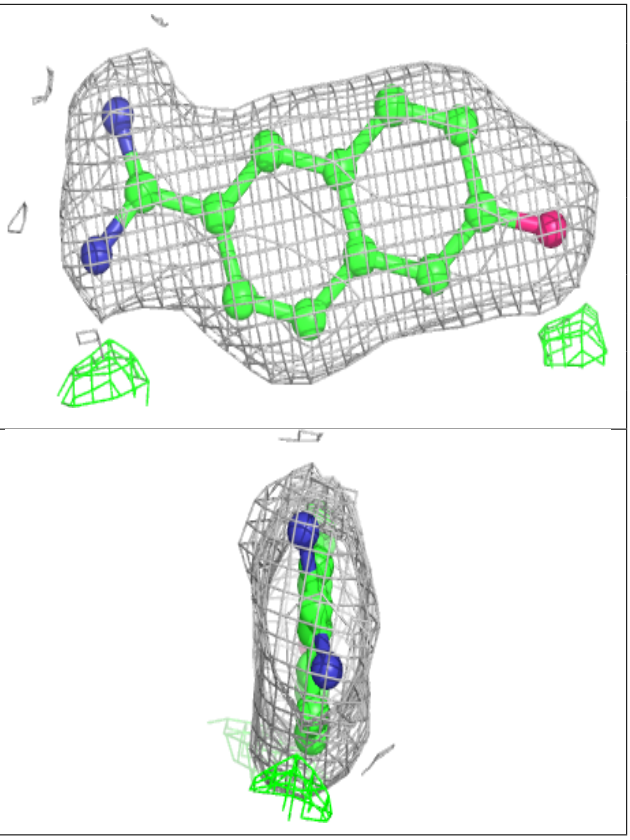
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
9	PEG	C	301	7/7	0.80	0.25	57,65,73,74	0
11	PG4	D	607	13/13	0.80	0.27	49,56,68,71	0
7	EDO	B	609	4/4	0.82	0.26	71,79,85,86	0
10	CIT	C	302	13/13	0.83	0.37	69,88,95,95	0
7	EDO	D	605	4/4	0.84	0.16	58,58,58,61	0
10	CIT	C	306	13/13	0.84	0.43	76,86,104,108	0
7	EDO	B	606	4/4	0.84	0.20	63,66,68,69	0
5	TAM	B	605	11/11	0.84	0.17	47,66,81,81	0
6	UNX	C	303	1/1	0.85	0.30	62,62,62,62	0
5	TAM	D	602	11/11	0.86	0.24	45,57,82,94	0
7	EDO	D	610	4/4	0.86	0.14	61,64,68,74	0
7	EDO	D	608	4/4	0.87	0.13	57,59,62,65	0
7	EDO	A	605	4/4	0.87	0.26	59,65,69,72	0
5	TAM	A	601	11/11	0.87	0.13	69,82,91,93	0
5	TAM	B	603	11/11	0.87	0.19	46,58,73,82	0
7	EDO	B	611	4/4	0.88	0.20	60,62,67,79	0
7	EDO	B	602	4/4	0.88	0.28	54,57,63,64	0
7	EDO	C	304	4/4	0.88	0.36	67,68,71,71	0
7	EDO	B	601	4/4	0.88	0.39	50,58,60,63	0
7	EDO	D	604	4/4	0.89	0.28	60,61,63,64	0
7	EDO	D	606	4/4	0.89	0.32	48,62,63,66	0
7	EDO	B	608	4/4	0.90	0.17	66,67,70,75	0
7	EDO	D	609	4/4	0.91	0.20	59,61,63,64	0
7	EDO	A	604	4/4	0.91	0.39	58,66,69,71	0
7	EDO	B	607	4/4	0.92	0.24	58,71,72,74	0
8	7R8	D	603	14/14	0.96	0.11	32,34,39,41	0
6	UNX	A	602	1/1	0.97	0.11	29,29,29,29	0
8	7R8	B	604	14/14	0.98	0.10	36,38,40,42	0
6	UNX	D	601	1/1	0.99	0.12	26,26,26,26	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.

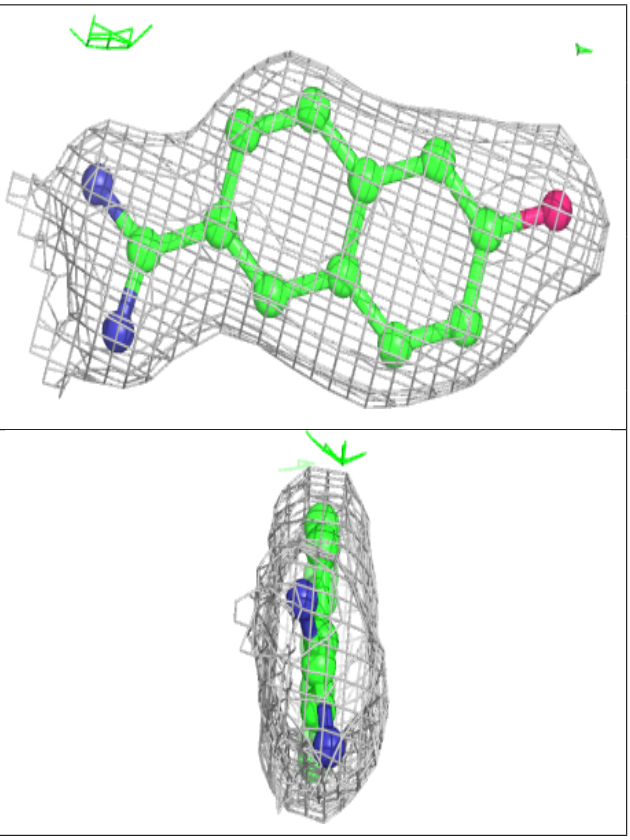
Electron density around 7R8 D 603:

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



Electron density around 7R8 B 604:

$2mF_o$-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.