Full wwPDB X-ray Structure Validation Report ⓘ

Nov 4, 2023 – 10:11 AM EDT

PDB ID	:	7FQK
Title	:	Crystal Structure of human Legumain in complex with (2S)-N-[(3S)-5-amino-1-(1,3-oxazol-2-yl)-5-oxopent-1-yn-3-yl]-1-[1-[4-(trifluoromethoxy)phenyl]cyclopropanecarbonyl]pyrrolidine-2-carboxamide
Authors	:	Ehler, A.; Benz, J.; Bartels, B.; Rudolph, M.G.
Deposited on	:	2022-10-05
Resolution	:	1.97 Å(reported)

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

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

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

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

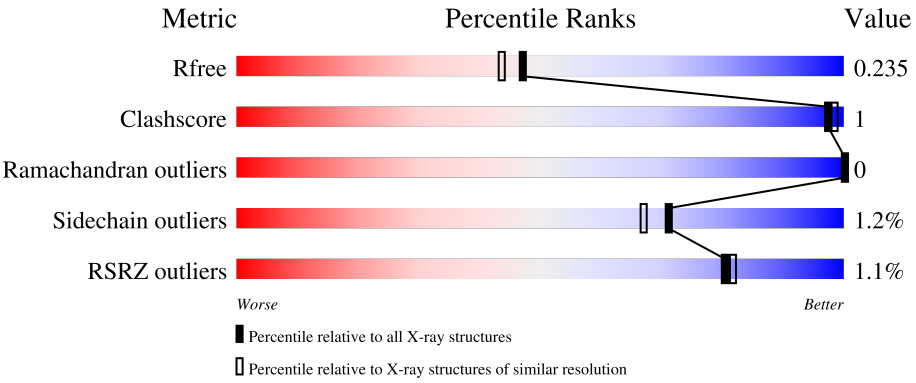
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	444	% 57% · 41%
1	B	444	% 57% · 41%
1	C	444	57% · 41%
1	D	444	56% · 41%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	2	100%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 9215 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 Legumain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	263	Total	C	N	O	S	0	0	0
			2114	1335	361	403	15			
1	B	261	Total	C	N	O	S	0	0	0
			2094	1323	355	401	15			
1	C	262	Total	C	N	O	S	0	0	0
			2103	1329	357	402	15			
1	D	262	Total	C	N	O	S	0	2	0
			2117	1338	361	403	15			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q99538
A	1	LYS	-	expression tag	UNP Q99538
A	2	LEU	-	expression tag	UNP Q99538
A	3	CYS	-	expression tag	UNP Q99538
A	4	ILE	-	expression tag	UNP Q99538
A	5	LEU	-	expression tag	UNP Q99538
A	6	LEU	-	expression tag	UNP Q99538
A	7	ALA	-	expression tag	UNP Q99538
A	8	VAL	-	expression tag	UNP Q99538
A	9	VAL	-	expression tag	UNP Q99538
A	10	ALA	-	expression tag	UNP Q99538
A	11	PHE	-	expression tag	UNP Q99538
A	12	VAL	-	expression tag	UNP Q99538
A	13	GLY	-	expression tag	UNP Q99538
A	14	LEU	-	expression tag	UNP Q99538
A	15	SER	-	expression tag	UNP Q99538
A	16	LEU	-	expression tag	UNP Q99538
A	17	GLY	-	expression tag	UNP Q99538
A	147	SNN	ASP	conflict	UNP Q99538
A	272	GLN	ASN	conflict	UNP Q99538
A	434	VAL	-	expression tag	UNP Q99538

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	435	ASP	-	expression tag	UNP Q99538
A	436	HIS	-	expression tag	UNP Q99538
A	437	HIS	-	expression tag	UNP Q99538
A	438	HIS	-	expression tag	UNP Q99538
A	439	HIS	-	expression tag	UNP Q99538
A	440	HIS	-	expression tag	UNP Q99538
A	441	HIS	-	expression tag	UNP Q99538
A	442	HIS	-	expression tag	UNP Q99538
A	443	HIS	-	expression tag	UNP Q99538
B	0	MET	-	initiating methionine	UNP Q99538
B	1	LYS	-	expression tag	UNP Q99538
B	2	LEU	-	expression tag	UNP Q99538
B	3	CYS	-	expression tag	UNP Q99538
B	4	ILE	-	expression tag	UNP Q99538
B	5	LEU	-	expression tag	UNP Q99538
B	6	LEU	-	expression tag	UNP Q99538
B	7	ALA	-	expression tag	UNP Q99538
B	8	VAL	-	expression tag	UNP Q99538
B	9	VAL	-	expression tag	UNP Q99538
B	10	ALA	-	expression tag	UNP Q99538
B	11	PHE	-	expression tag	UNP Q99538
B	12	VAL	-	expression tag	UNP Q99538
B	13	GLY	-	expression tag	UNP Q99538
B	14	LEU	-	expression tag	UNP Q99538
B	15	SER	-	expression tag	UNP Q99538
B	16	LEU	-	expression tag	UNP Q99538
B	17	GLY	-	expression tag	UNP Q99538
B	147	SNN	ASP	conflict	UNP Q99538
B	272	GLN	ASN	conflict	UNP Q99538
B	434	VAL	-	expression tag	UNP Q99538
B	435	ASP	-	expression tag	UNP Q99538
B	436	HIS	-	expression tag	UNP Q99538
B	437	HIS	-	expression tag	UNP Q99538
B	438	HIS	-	expression tag	UNP Q99538
B	439	HIS	-	expression tag	UNP Q99538
B	440	HIS	-	expression tag	UNP Q99538
B	441	HIS	-	expression tag	UNP Q99538
B	442	HIS	-	expression tag	UNP Q99538
B	443	HIS	-	expression tag	UNP Q99538
C	0	MET	-	initiating methionine	UNP Q99538
C	1	LYS	-	expression tag	UNP Q99538
C	2	LEU	-	expression tag	UNP Q99538

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	CYS	-	expression tag	UNP Q99538
C	4	ILE	-	expression tag	UNP Q99538
C	5	LEU	-	expression tag	UNP Q99538
C	6	LEU	-	expression tag	UNP Q99538
C	7	ALA	-	expression tag	UNP Q99538
C	8	VAL	-	expression tag	UNP Q99538
C	9	VAL	-	expression tag	UNP Q99538
C	10	ALA	-	expression tag	UNP Q99538
C	11	PHE	-	expression tag	UNP Q99538
C	12	VAL	-	expression tag	UNP Q99538
C	13	GLY	-	expression tag	UNP Q99538
C	14	LEU	-	expression tag	UNP Q99538
C	15	SER	-	expression tag	UNP Q99538
C	16	LEU	-	expression tag	UNP Q99538
C	17	GLY	-	expression tag	UNP Q99538
C	147	SNN	ASP	conflict	UNP Q99538
C	272	GLN	ASN	conflict	UNP Q99538
C	434	VAL	-	expression tag	UNP Q99538
C	435	ASP	-	expression tag	UNP Q99538
C	436	HIS	-	expression tag	UNP Q99538
C	437	HIS	-	expression tag	UNP Q99538
C	438	HIS	-	expression tag	UNP Q99538
C	439	HIS	-	expression tag	UNP Q99538
C	440	HIS	-	expression tag	UNP Q99538
C	441	HIS	-	expression tag	UNP Q99538
C	442	HIS	-	expression tag	UNP Q99538
C	443	HIS	-	expression tag	UNP Q99538
D	0	MET	-	initiating methionine	UNP Q99538
D	1	LYS	-	expression tag	UNP Q99538
D	2	LEU	-	expression tag	UNP Q99538
D	3	CYS	-	expression tag	UNP Q99538
D	4	ILE	-	expression tag	UNP Q99538
D	5	LEU	-	expression tag	UNP Q99538
D	6	LEU	-	expression tag	UNP Q99538
D	7	ALA	-	expression tag	UNP Q99538
D	8	VAL	-	expression tag	UNP Q99538
D	9	VAL	-	expression tag	UNP Q99538
D	10	ALA	-	expression tag	UNP Q99538
D	11	PHE	-	expression tag	UNP Q99538
D	12	VAL	-	expression tag	UNP Q99538
D	13	GLY	-	expression tag	UNP Q99538
D	14	LEU	-	expression tag	UNP Q99538

Continued on next page...

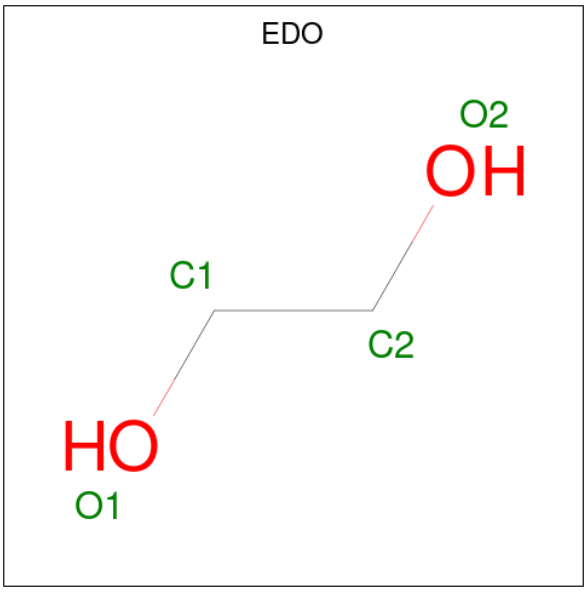
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	15	SER	-	expression tag	UNP Q99538
D	16	LEU	-	expression tag	UNP Q99538
D	17	GLY	-	expression tag	UNP Q99538
D	147	SNN	ASP	conflict	UNP Q99538
D	272	GLN	ASN	conflict	UNP Q99538
D	434	VAL	-	expression tag	UNP Q99538
D	435	ASP	-	expression tag	UNP Q99538
D	436	HIS	-	expression tag	UNP Q99538
D	437	HIS	-	expression tag	UNP Q99538
D	438	HIS	-	expression tag	UNP Q99538
D	439	HIS	-	expression tag	UNP Q99538
D	440	HIS	-	expression tag	UNP Q99538
D	441	HIS	-	expression tag	UNP Q99538
D	442	HIS	-	expression tag	UNP Q99538
D	443	HIS	-	expression tag	UNP Q99538

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

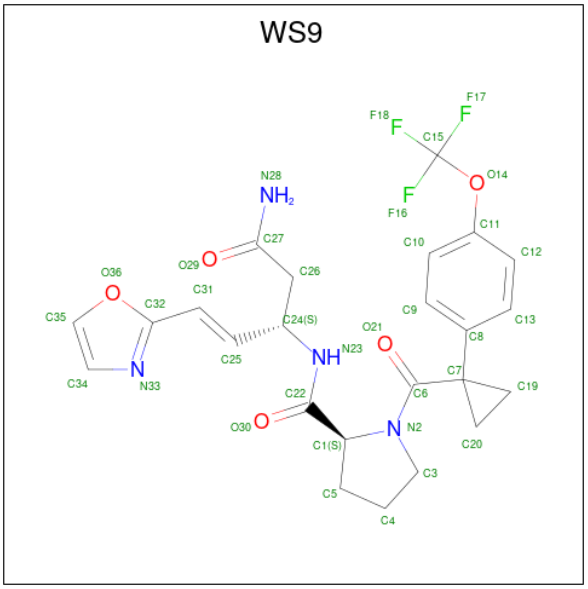
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



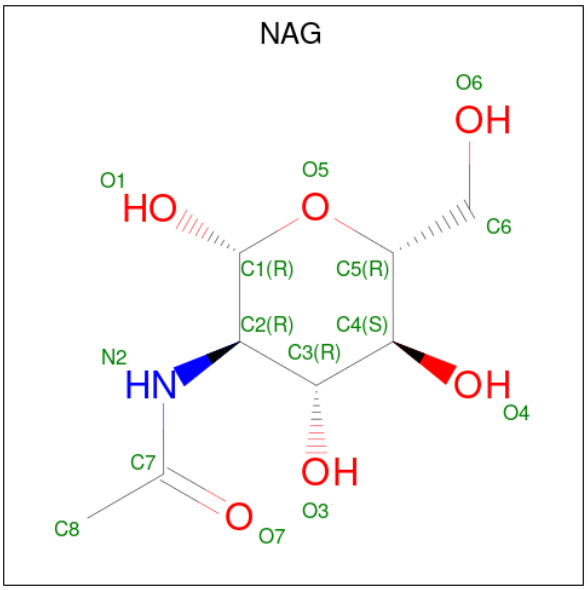
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0

- Molecule 4 is N-[(1E,3R)-5-amino-1-(1,3-oxazol-2-yl)-5-oxopent-1-en-3-yl]-1-{1-[4-(trifluoromethoxy)phenyl]cyclopropane-1-carbonyl}-L-prolinamide (three-letter code: WS9) (formula: $C_{24}H_{25}F_3N_4O_5$) (labeled as "Ligand of Interest" by depositor).



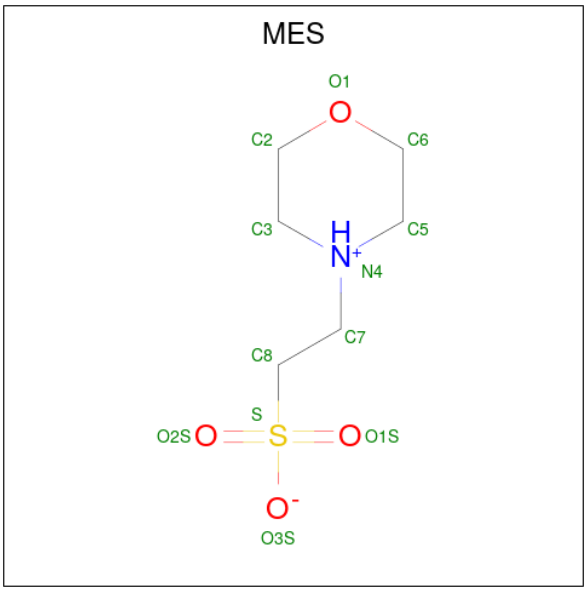
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	0	0
			36	24	3	4	5		
4	B	1	Total	C	F	N	O	0	0
			36	24	3	4	5		
4	C	1	Total	C	F	N	O	0	0
			36	24	3	4	5		
4	D	1	Total	C	F	N	O	0	0
			36	24	3	4	5		

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O		0	0
			14	8	1	5			
5	B	1	Total	C	N	O		0	0
			14	8	1	5			
5	D	1	Total	C	N	O		0	0
			14	8	1	5			

- Molecule 6 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
6	D	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 7 is water.

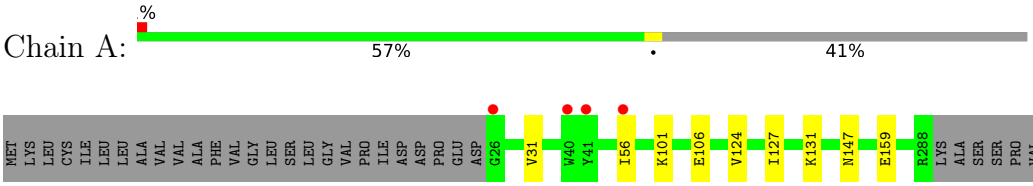
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	136	Total	O	0	0
			136	136		
7	B	126	Total	O	0	0
			126	126		
7	C	141	Total	O	0	0
			141	141		
7	D	134	Total	O	0	0
			134	134		

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.

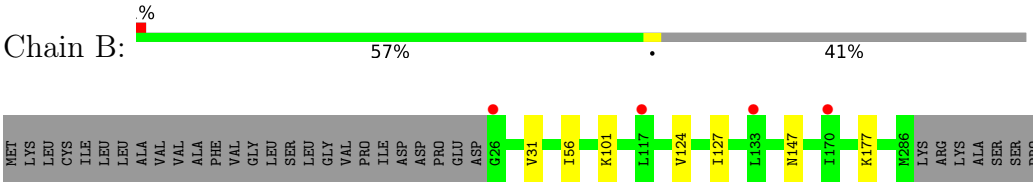
• Molecule 1: Legumain

Chain A:



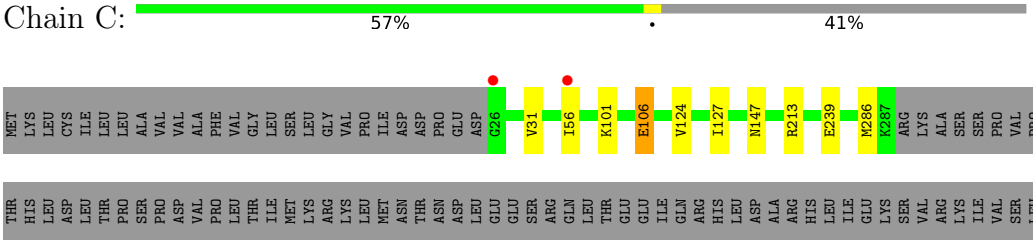
• Molecule 1: Legumain

Chain B:



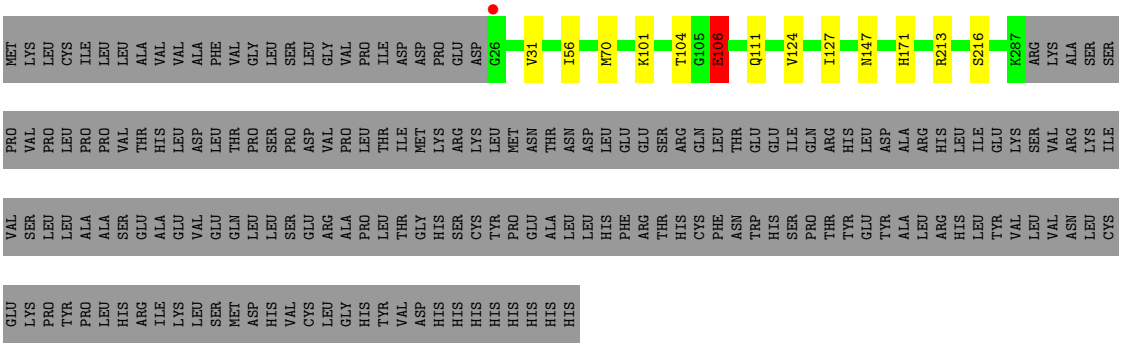
• Molecule 1: Legumain

Chain C:



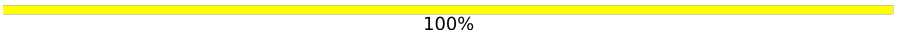
• Molecule 1: Legumain

Chain D:



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

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	102.49Å 102.49Å 330.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.13 – 1.97 87.13 – 1.97	Depositor EDS
% Data completeness (in resolution range)	41.7 (87.13-1.97) 41.7 (87.13-1.97)	Depositor EDS
R_{merge}	0.53	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.194 , 0.233 0.199 , 0.235	Depositor DCC
R_{free} test set	2736 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	23.9	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 34.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9215	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.25% 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: SNN, EDO, WS9, NAG, MES

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	0.82	0/2163	0.83	0/2932
1	B	0.79	0/2143	0.83	0/2907
1	C	0.83	2/2152 (0.1%)	0.82	1/2918 (0.0%)
1	D	0.83	1/2172 (0.0%)	0.84	1/2945 (0.0%)
All	All	0.82	3/8630 (0.0%)	0.83	2/11702 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	106	GLU	CD-OE2	8.28	1.34	1.25
1	C	106	GLU	CD-OE1	7.29	1.33	1.25
1	C	239	GLU	CD-OE1	-5.63	1.19	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	213	ARG	NE-CZ-NH1	6.33	123.46	120.30
1	C	213	ARG	NE-CZ-NH1	5.37	122.98	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2114	0	1997	3	0
1	B	2094	0	1971	2	0
1	C	2103	0	1984	2	0
1	D	2117	0	2005	6	0
2	E	28	0	25	0	0
3	A	4	0	6	0	0
3	B	4	0	6	0	0
3	D	4	0	6	0	0
4	A	36	0	0	0	0
4	B	36	0	0	0	0
4	C	36	0	0	0	0
4	D	36	0	0	0	0
5	A	14	0	13	0	0
5	B	14	0	13	0	0
5	D	14	0	13	0	0
6	B	12	0	13	0	0
6	D	12	0	13	0	0
7	A	136	0	0	1	0
7	B	126	0	0	0	0
7	C	141	0	0	0	0
7	D	134	0	0	1	0
All	All	9215	0	8065	13	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 1.

All (13) 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:124:VAL:HA	1:A:127:ILE:HD12	1.92	0.52
1:B:124:VAL:HA	1:B:127:ILE:HD12	1.94	0.50
1:B:31:VAL:HG21	1:B:56:ILE:HD12	1.92	0.50
1:C:124:VAL:HA	1:C:127:ILE:HD12	1.94	0.50
1:A:31:VAL:HG21	1:A:56:ILE:HD12	1.95	0.48
1:D:124:VAL:HA	1:D:127:ILE:HD12	1.95	0.48
1:D:111[B]:GLN:NE2	1:D:111[B]:GLN:H	2.11	0.47
1:C:31:VAL:HG21	1:C:56:ILE:HD12	1.97	0.46
1:D:31:VAL:HG21	1:D:56:ILE:HD12	1.98	0.46
1:D:106:GLU:H	1:D:106:GLU:CD	2.18	0.46
1:A:131:LYS:NZ	7:A:604:HOH:O	2.50	0.44
1:D:171:HIS:HB3	7:D:713:HOH:O	2.18	0.44
1:D:70:MET:O	1:D:104:THR:HA	2.21	0.41

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	260/444 (59%)	255 (98%)	5 (2%)	0	100	100
1	B	258/444 (58%)	254 (98%)	4 (2%)	0	100	100
1	C	259/444 (58%)	253 (98%)	6 (2%)	0	100	100
1	D	261/444 (59%)	256 (98%)	5 (2%)	0	100	100
All	All	1038/1776 (58%)	1018 (98%)	20 (2%)	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	229/396 (58%)	226 (99%)	3 (1%)	69	64
1	B	227/396 (57%)	225 (99%)	2 (1%)	78	77
1	C	228/396 (58%)	225 (99%)	3 (1%)	69	64
1	D	230/396 (58%)	227 (99%)	3 (1%)	69	64
All	All	914/1584 (58%)	903 (99%)	11 (1%)	71	67

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	101	LYS
1	A	106	GLU
1	A	159	GLU
1	B	101	LYS
1	B	177	LYS
1	C	101	LYS
1	C	106	GLU
1	C	286	MET
1	D	101	LYS
1	D	106	GLU
1	D	216	SER

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	SNN	C	147	1	5,6,8	1.08	0	3,6,11	2.43	2 (66%)
1	SNN	A	147	1	5,6,8	1.19	0	3,6,11	2.44	2 (66%)
1	SNN	D	147	1	5,6,8	1.18	0	3,6,11	2.90	2 (66%)
1	SNN	B	147	1	5,6,8	1.46	1 (20%)	3,6,11	2.46	2 (66%)

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	SNN	C	147	1	-	3/3/5/12	-
1	SNN	A	147	1	-	2/3/5/12	-
1	SNN	D	147	1	-	3/3/5/12	-
1	SNN	B	147	1	-	3/3/5/12	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	147	SNN	C4-CA	-2.53	1.48	1.54

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	147	SNN	C4-CA-C	-3.57	104.77	111.44
1	D	147	SNN	CA-C4-C5	-3.44	101.42	114.44
1	C	147	SNN	CA-C4-C5	-3.37	101.69	114.44
1	A	147	SNN	CA-C4-C5	-3.20	102.31	114.44
1	B	147	SNN	C4-CA-C	-3.16	105.54	111.44
1	B	147	SNN	CA-C4-C5	-2.84	103.69	114.44
1	A	147	SNN	C4-CA-C	-2.76	106.28	111.44
1	C	147	SNN	C4-CA-C	-2.49	106.79	111.44

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	147	SNN	C5-C4-CA-N
1	A	147	SNN	CA-C4-C5-O5
1	B	147	SNN	O-C-CA-C4
1	B	147	SNN	C5-C4-CA-N
1	B	147	SNN	CA-C4-C5-O5
1	C	147	SNN	O-C-CA-C4
1	C	147	SNN	C5-C4-CA-N
1	C	147	SNN	CA-C4-C5-O5
1	D	147	SNN	O-C-CA-C4
1	D	147	SNN	CA-C4-C5-O5
1	D	147	SNN	C5-C4-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

2 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	E	1	2,1	14,14,15	1.13	0	17,19,21	1.68	5 (29%)
2	NAG	E	2	2	14,14,15	1.80	3 (21%)	17,19,21	1.16	2 (11%)

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	E	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	2	NAG	C1-C2	3.80	1.58	1.52
2	E	2	NAG	C4-C5	2.80	1.58	1.53
2	E	2	NAG	O3-C3	2.02	1.47	1.43

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	1	NAG	C1-O5-C5	3.07	116.34	112.19
2	E	1	NAG	O4-C4-C3	-2.66	104.20	110.35
2	E	1	NAG	C1-C2-N2	-2.62	106.02	110.49
2	E	1	NAG	C2-N2-C7	-2.57	119.24	122.90
2	E	2	NAG	C2-N2-C7	-2.32	119.59	122.90
2	E	1	NAG	O5-C5-C6	-2.14	103.84	107.20
2	E	2	NAG	O4-C4-C5	2.12	114.57	109.30

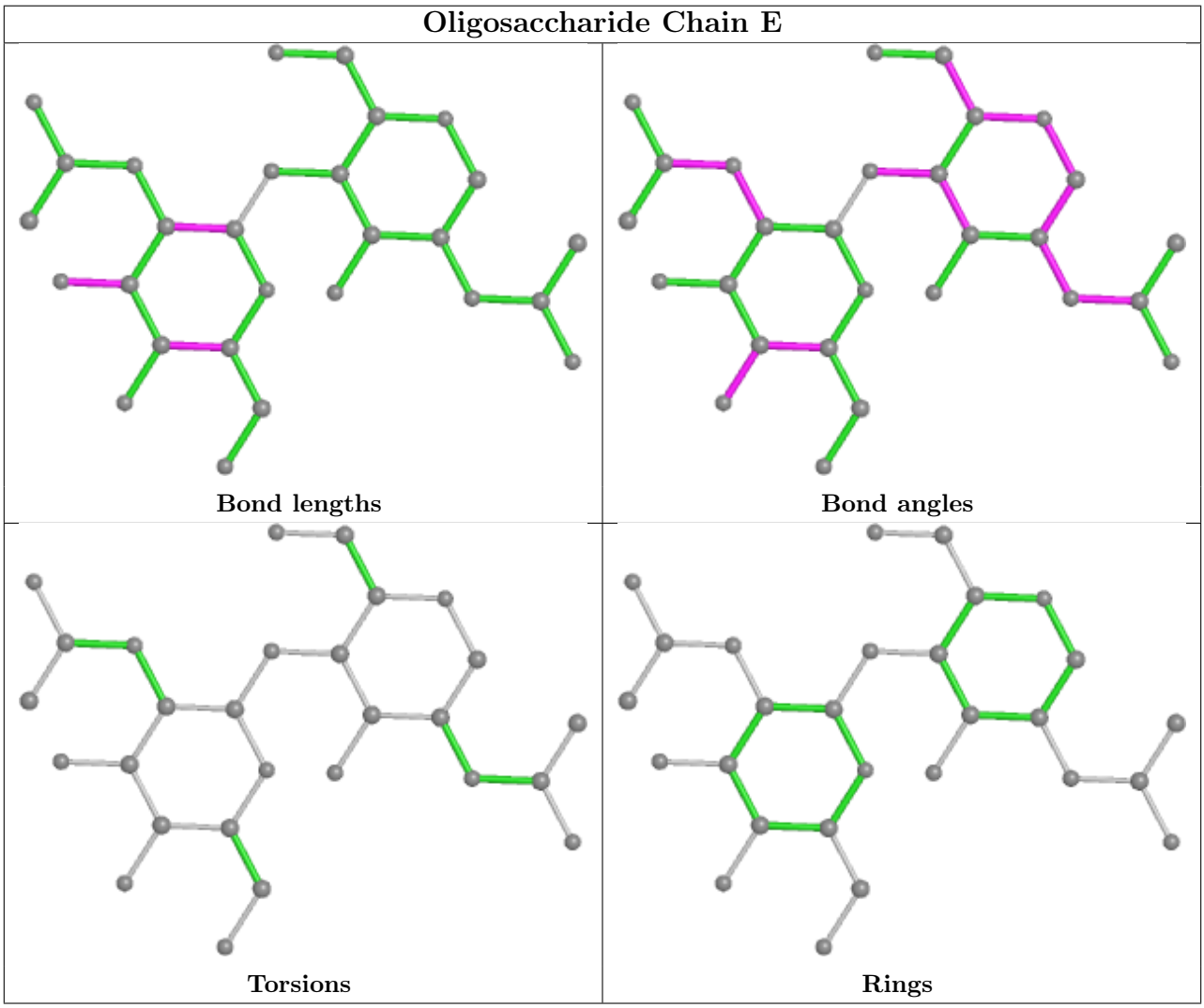
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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



5.6 Ligand geometry

12 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	$\#\|Z\|>2$
6	MES	D	501	-	12,12,12	0.58	0	14,16,16	0.62	0
5	NAG	D	504	1	14,14,15	1.28	1 (7%)	17,19,21	1.78	3 (17%)
4	WS9	C	501	1	36,39,39	0.71	1 (2%)	45,57,57	1.42	3 (6%)
3	EDO	B	501	-	3,3,3	0.51	0	2,2,2	0.42	0
3	EDO	A	501	-	3,3,3	0.44	0	2,2,2	0.34	0
4	WS9	B	502	1	36,39,39	0.56	0	45,57,57	1.35	2 (4%)
4	WS9	D	503	1	36,39,39	0.50	0	45,57,57	1.19	2 (4%)
5	NAG	B	503	1	14,14,15	1.05	1 (7%)	17,19,21	1.45	3 (17%)
3	EDO	D	502	-	3,3,3	0.88	0	2,2,2	0.78	0
5	NAG	A	503	1	14,14,15	0.83	0	17,19,21	1.88	2 (11%)
4	WS9	A	502	1	36,39,39	0.66	1 (2%)	45,57,57	1.37	4 (8%)
6	MES	B	504	-	12,12,12	0.65	0	14,16,16	0.66	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
6	MES	D	501	-	-	3/6/14/14	0/1/1/1
5	NAG	D	504	1	-	4/6/23/26	0/1/1/1
4	WS9	C	501	1	-	0/36/52/52	0/4/4/4
3	EDO	B	501	-	-	0/1/1/1	-
3	EDO	A	501	-	-	0/1/1/1	-
4	WS9	B	502	1	-	2/36/52/52	0/4/4/4
4	WS9	D	503	1	-	2/36/52/52	0/4/4/4
5	NAG	B	503	1	-	4/6/23/26	0/1/1/1
3	EDO	D	502	-	-	1/1/1/1	-
5	NAG	A	503	1	-	2/6/23/26	0/1/1/1
4	WS9	A	502	1	-	0/36/52/52	0/4/4/4
6	MES	B	504	-	-	3/6/14/14	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	501	WS9	F16-C15	2.61	1.44	1.32
4	A	502	WS9	F16-C15	2.55	1.44	1.32
5	D	504	NAG	C3-C2	2.31	1.57	1.52
5	B	503	NAG	O4-C4	2.01	1.47	1.43

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	502	WS9	C25-C24-N23	-7.25	96.87	110.53
4	C	501	WS9	C25-C24-N23	-6.91	97.51	110.53
5	A	503	NAG	C1-O5-C5	6.35	120.80	112.19
4	A	502	WS9	C25-C24-N23	-5.79	99.62	110.53
5	D	504	NAG	C1-O5-C5	5.73	119.95	112.19
4	D	503	WS9	C25-C24-N23	-5.71	99.77	110.53
4	C	501	WS9	C24-C25-C31	-4.34	108.92	126.00
4	B	502	WS9	C24-C25-C31	-3.72	111.35	126.00
4	D	503	WS9	C24-C25-C31	-3.68	111.53	126.00
4	A	502	WS9	C24-C25-C31	-3.50	112.24	126.00
4	C	501	WS9	C34-N33-C32	3.07	108.40	104.34
5	B	503	NAG	C4-C3-C2	3.05	115.49	111.02
5	B	503	NAG	C1-O5-C5	2.92	116.15	112.19
5	B	503	NAG	C1-C2-N2	-2.81	105.68	110.49
4	A	502	WS9	F17-C15-O14	-2.71	99.64	111.71
5	D	504	NAG	C4-C3-C2	2.70	114.97	111.02
5	D	504	NAG	O5-C1-C2	2.51	115.26	111.29
5	A	503	NAG	C2-N2-C7	-2.28	119.66	122.90
4	A	502	WS9	F16-C15-O14	2.21	121.56	111.71

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	504	MES	C7-C8-S-O2S
6	D	501	MES	C7-C8-S-O2S
5	B	503	NAG	C4-C5-C6-O6
5	B	503	NAG	O5-C5-C6-O6
5	A	503	NAG	O5-C5-C6-O6
6	B	504	MES	C7-C8-S-O3S
6	D	501	MES	C7-C8-S-O3S
5	B	503	NAG	C8-C7-N2-C2
5	A	503	NAG	C4-C5-C6-O6
5	D	504	NAG	C4-C5-C6-O6
5	D	504	NAG	C8-C7-N2-C2
5	D	504	NAG	O5-C5-C6-O6
6	B	504	MES	C7-C8-S-O1S
6	D	501	MES	C7-C8-S-O1S
4	B	502	WS9	C19-C7-C8-C13
4	B	502	WS9	C19-C7-C8-C9
4	D	503	WS9	C19-C7-C8-C13

Continued on next page...

Continued from previous page...

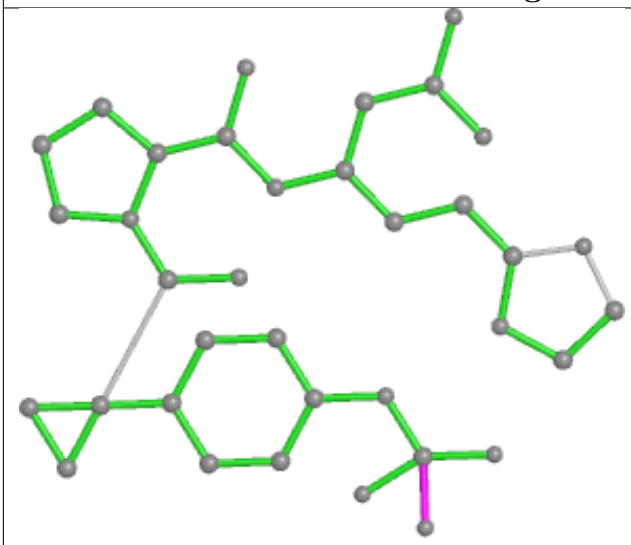
Mol	Chain	Res	Type	Atoms
4	D	503	WS9	C19-C7-C8-C9
3	D	502	EDO	O1-C1-C2-O2
5	B	503	NAG	O7-C7-N2-C2
5	D	504	NAG	O7-C7-N2-C2

There are no ring outliers.

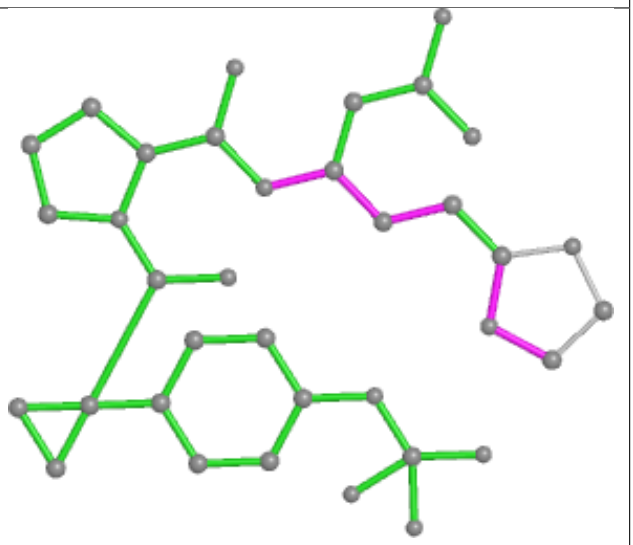
No monomer is involved in short contacts.

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

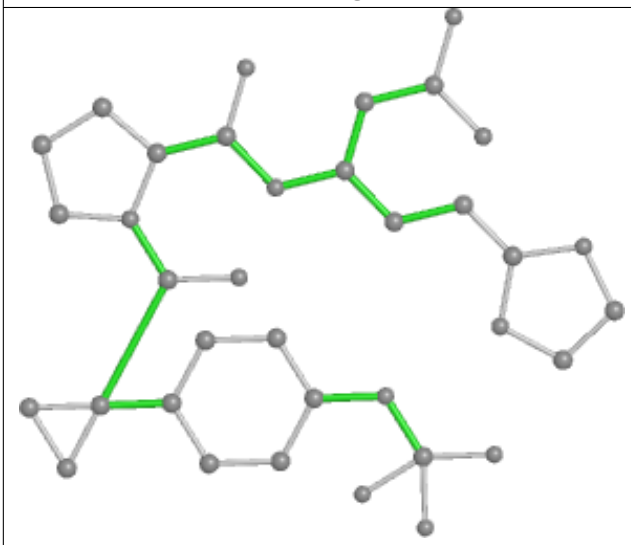
Ligand WS9 C 501



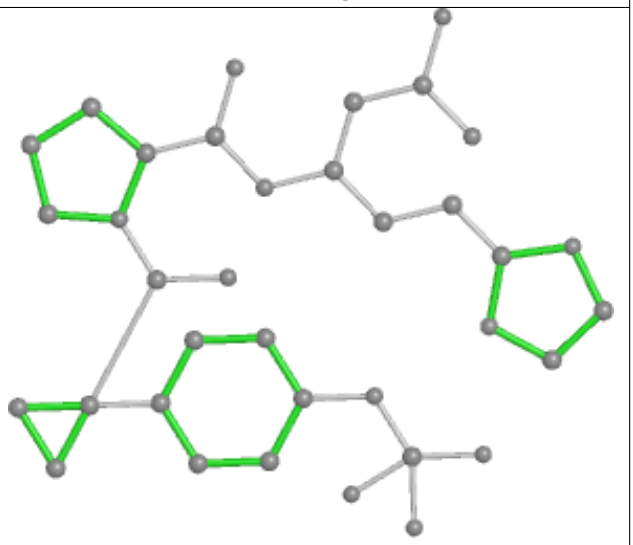
Bond lengths



Bond angles

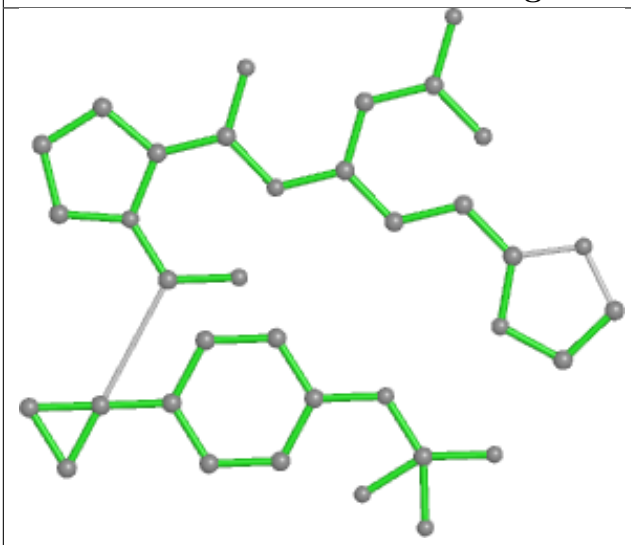


Torsions

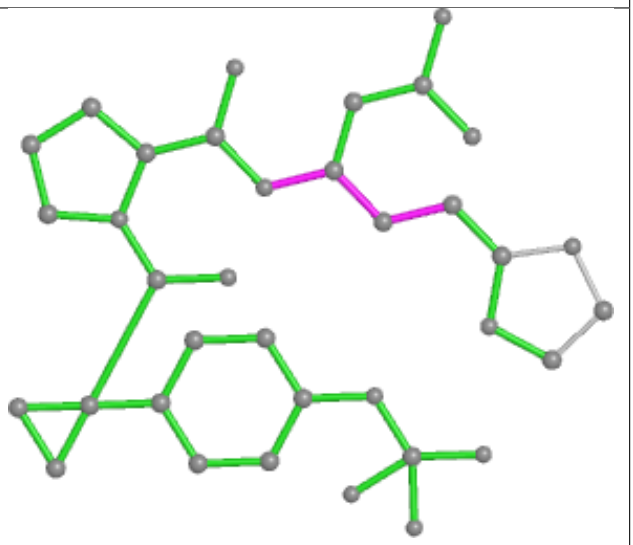


Rings

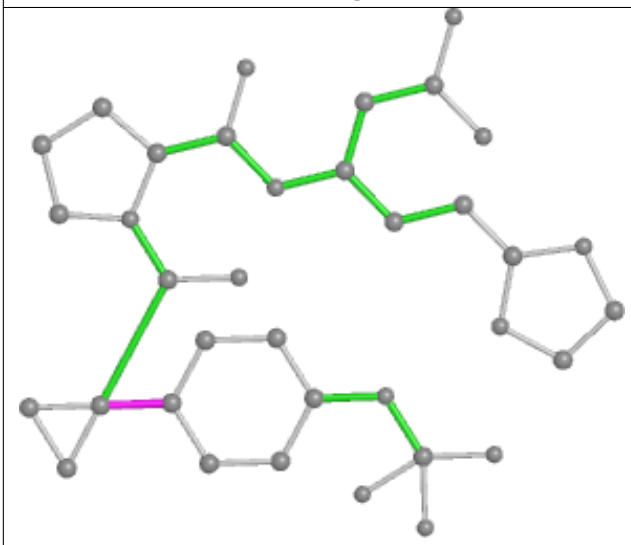
Ligand WS9 B 502



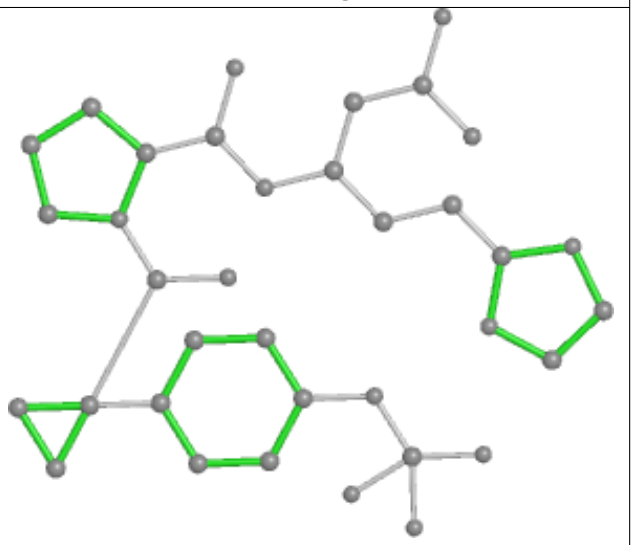
Bond lengths



Bond angles

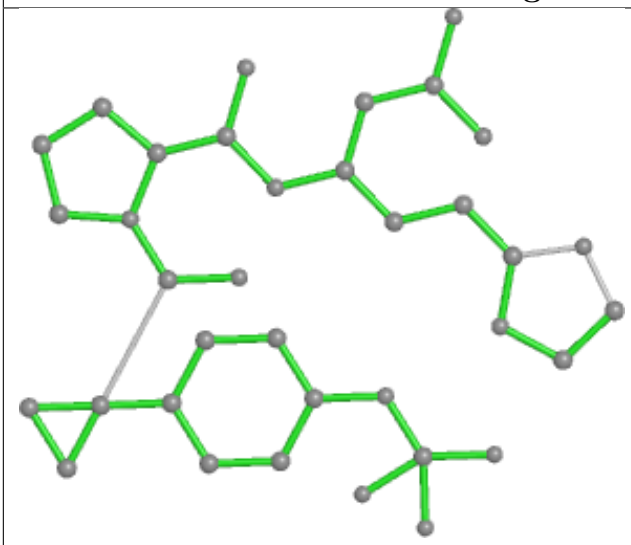


Torsions

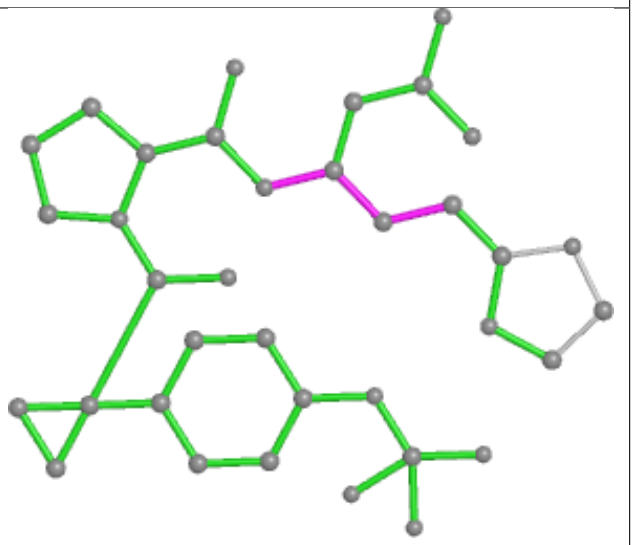


Rings

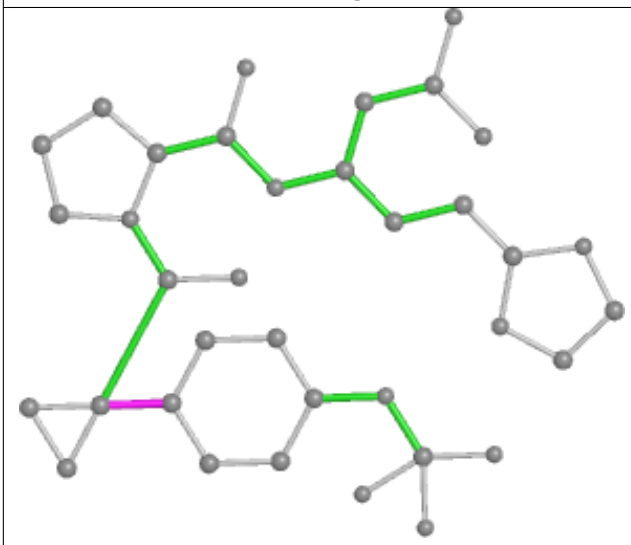
Ligand WS9 D 503



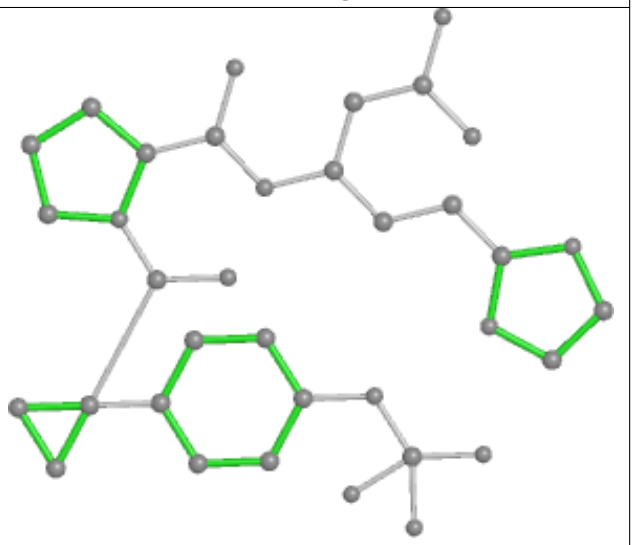
Bond lengths



Bond angles

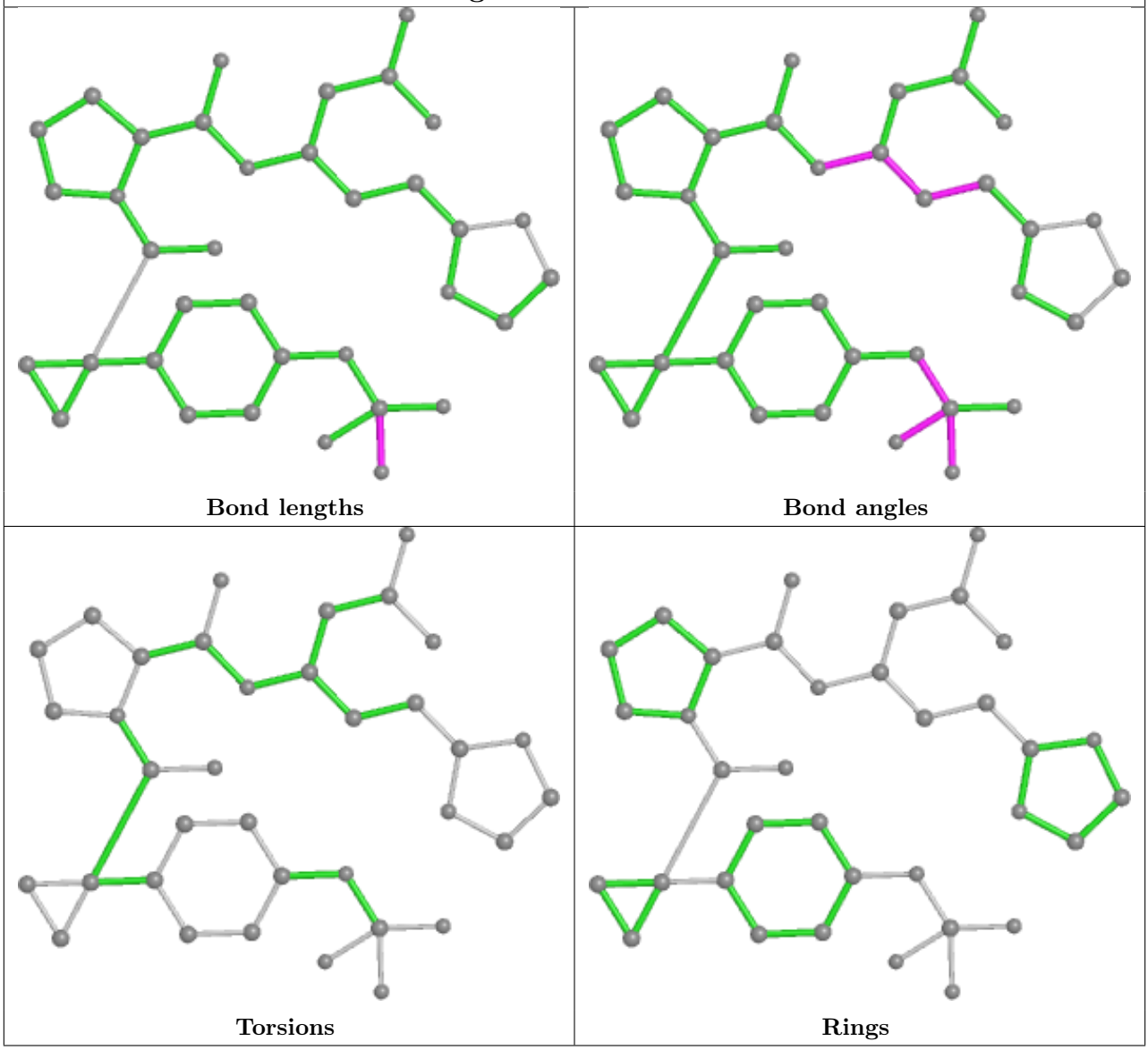


Torsions



Rings

Ligand WS9 A 502



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	146:THR	C	147:SNN	N	1.61

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	262/444 (59%)	0.43	4 (1%)	73	75	33, 38, 63, 93	0
1	B	260/444 (58%)	0.46	4 (1%)	73	75	33, 40, 70, 89	0
1	C	261/444 (58%)	0.32	2 (0%)	86	87	32, 39, 62, 94	0
1	D	261/444 (58%)	0.36	1 (0%)	92	93	32, 39, 64, 90	0
All	All	1044/1776 (58%)	0.39	11 (1%)	80	82	32, 39, 66, 94	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	26	GLY	4.9
1	B	26	GLY	4.4
1	C	26	GLY	3.8
1	D	26	GLY	3.6
1	C	56	ILE	2.6
1	A	56	ILE	2.4
1	A	41	TYR	2.2
1	B	117	LEU	2.2
1	B	133	LEU	2.2
1	B	170	ILE	2.1
1	A	40	TRP	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, 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
1	SNN	D	147	7/8	0.95	0.13	30,31,31,33	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	SNN	B	147	7/8	0.97	0.10	31,34,35,35	0
1	SNN	C	147	7/8	0.98	0.13	23,24,26,27	0
1	SNN	A	147	7/8	0.98	0.12	21,24,25,25	0

6.3 Carbohydrates (i)

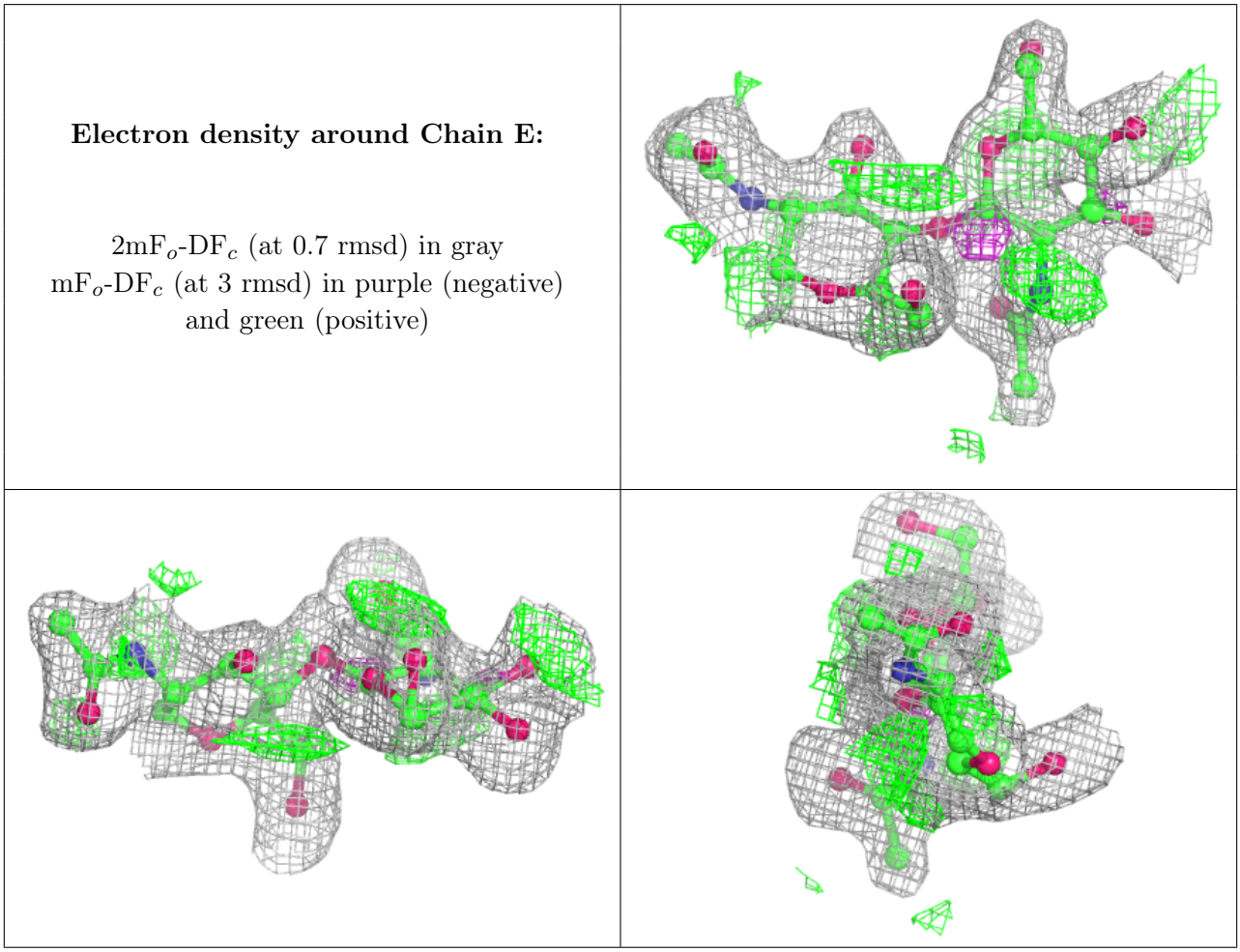
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	NAG	E	2	14/15	0.69	0.19	50,55,64,65	0
2	NAG	E	1	14/15	0.89	0.10	47,52,55,56	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)



6.4 Ligands (i)

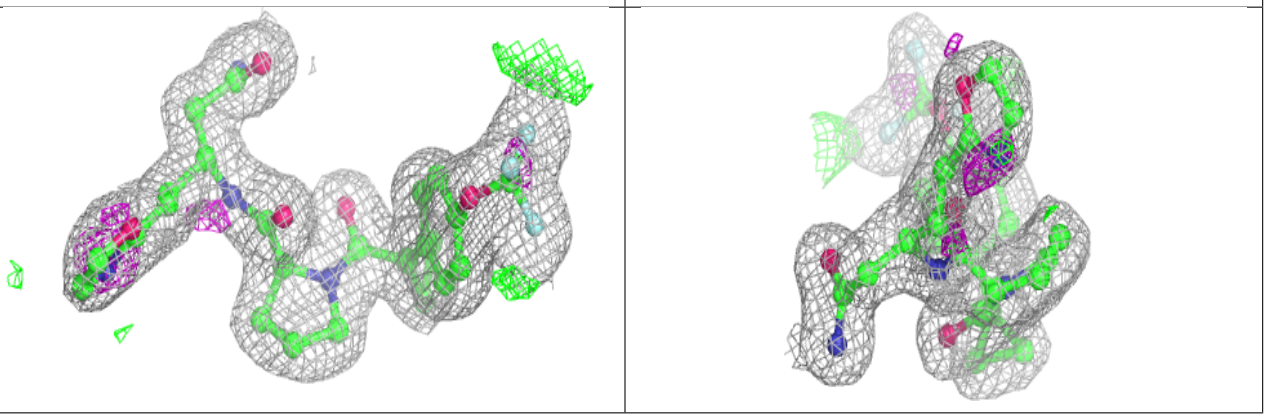
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
5	NAG	D	504	14/15	0.72	0.16	69,77,83,85	0
3	EDO	D	502	4/4	0.79	0.19	53,58,58,61	0
5	NAG	B	503	14/15	0.82	0.19	73,78,83,83	0
3	EDO	B	501	4/4	0.88	0.27	52,55,57,58	0
5	NAG	A	503	14/15	0.90	0.13	48,50,53,54	0
6	MES	B	504	12/12	0.93	0.14	70,73,76,76	0
6	MES	D	501	12/12	0.93	0.12	61,63,70,77	0
4	WS9	A	502	36/36	0.95	0.13	25,30,47,52	0
4	WS9	C	501	36/36	0.96	0.14	25,28,42,45	0
3	EDO	A	501	4/4	0.96	0.11	42,42,43,45	0
4	WS9	B	502	36/36	0.96	0.14	22,28,44,46	0
4	WS9	D	503	36/36	0.97	0.14	21,27,43,45	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.

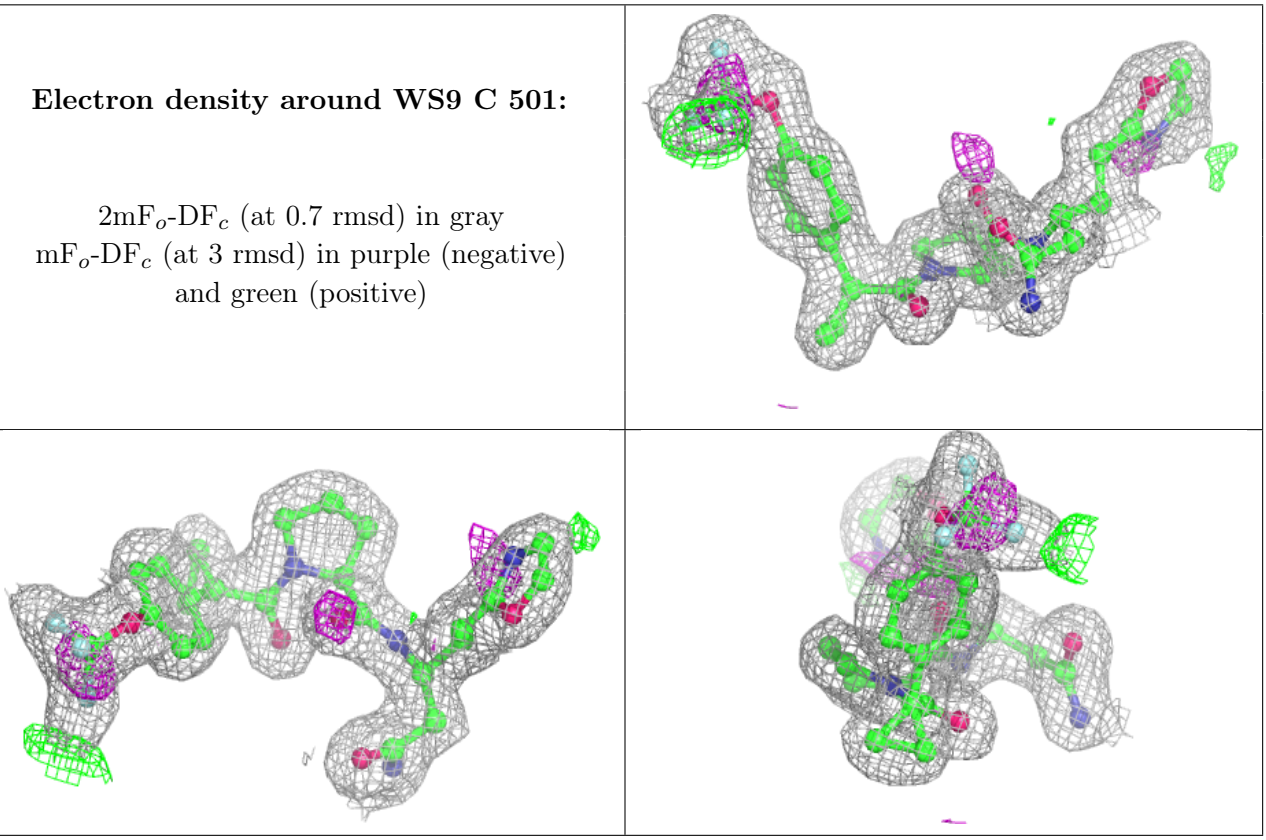
Electron density around WS9 A 502:

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



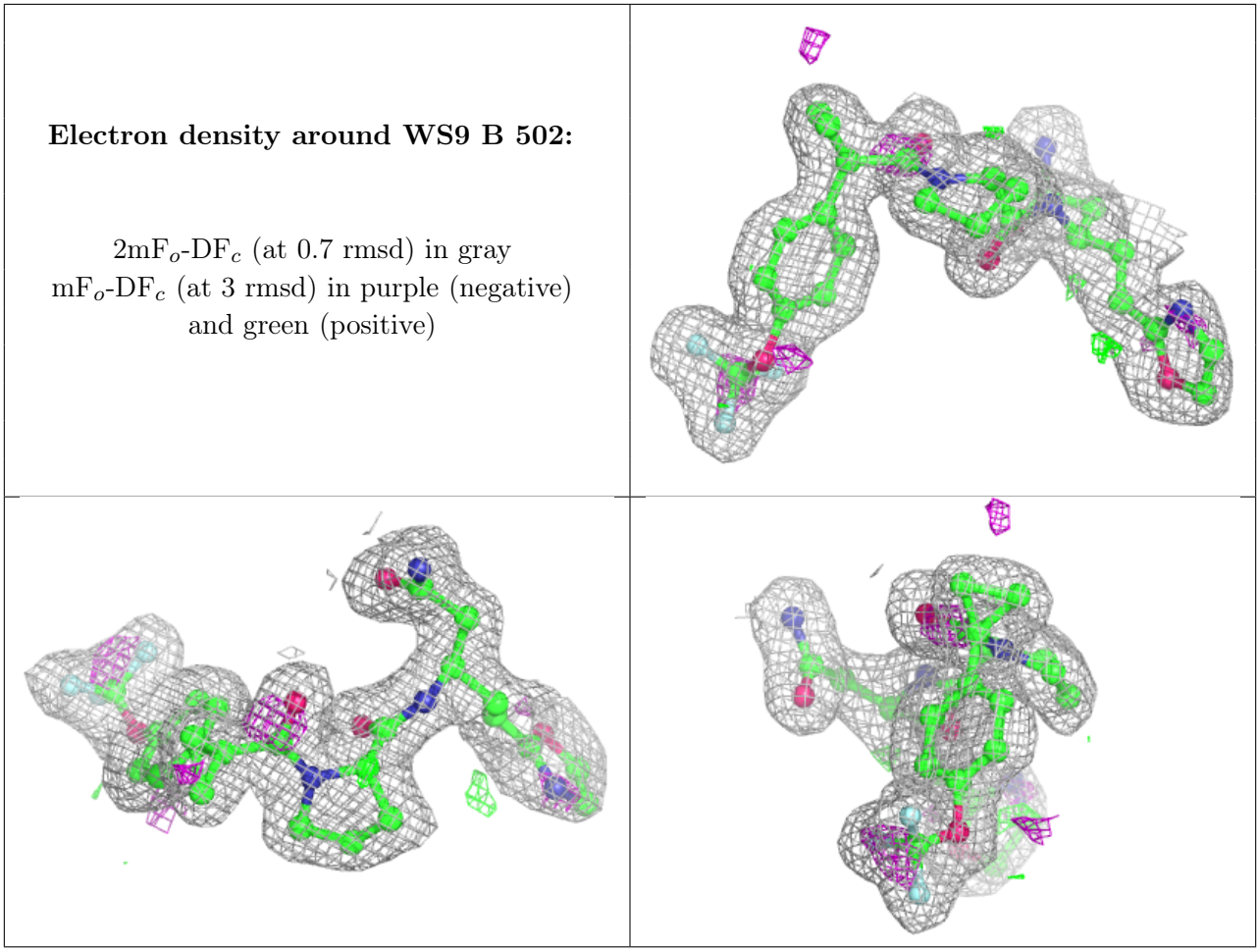
Electron density around WS9 C 501:

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



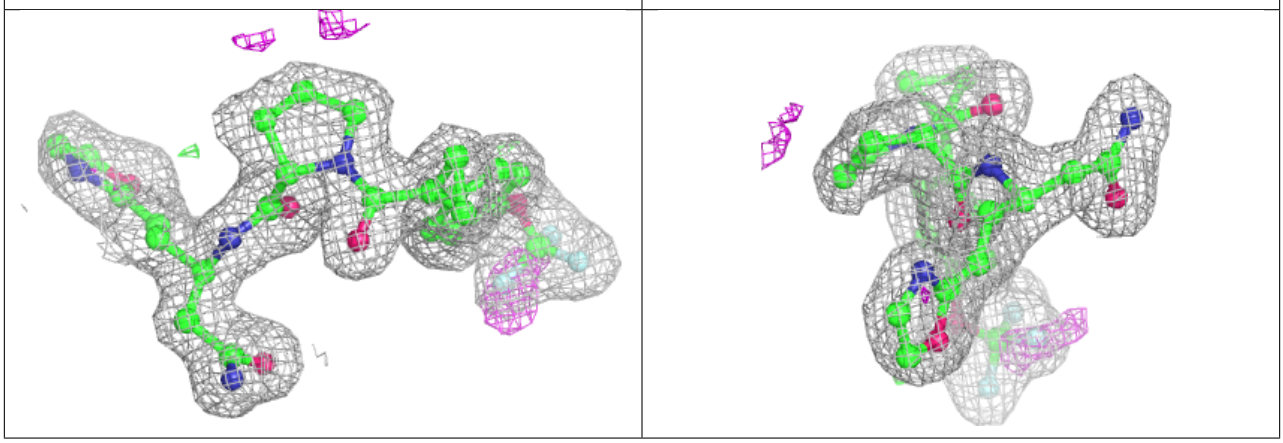
Electron density around WS9 B 502:

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



Electron density around WS9 D 503:

$2mF_o$-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.