



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 4, 2023 – 03:18 PM EDT

PDB ID : 7FQJ
Title : Crystal Structure of human Legumain in complex with (2S)-N-[(1S)-3-amino-1-cyano-3-oxopropyl]-1-[1-[4-[(2,4-difluorophenyl)methoxy]phenyl]cyclopropa ncarbonyl]pyrrolidine-2-carboxamide
Authors : Ehler, A.; Benz, J.; Bartels, B.; Rudolph, M.G.
Deposited on : 2022-10-05
Resolution : 1.70 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)
Xtrriage (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

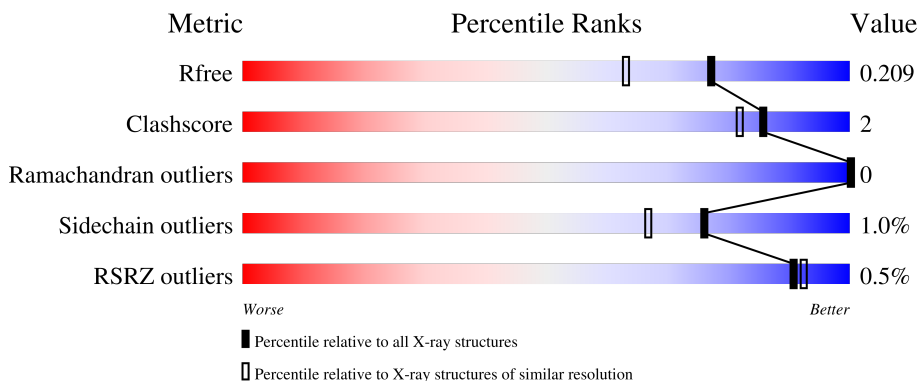
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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

Mol	Chain	Length	Quality of chain
1	A	444	
1	B	444	
1	C	444	
2	E	2	

2 Entry composition i

There are 11 unique types of molecules in this entry. The entry contains 7348 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
			Total	C	N	O	S			
1	A	265	2167	1371	372	409	15	0	6	0
1	B	264	2162	1371	366	410	15	0	6	0
1	C	260	2116	1341	360	400	15	0	4	0

There are 90 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
A	435	ASP	-	expression tag	UNP Q99538
A	436	HIS	-	expression tag	UNP Q99538

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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
C	3	CYS	-	expression tag	UNP Q99538
C	4	ILE	-	expression tag	UNP Q99538

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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

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



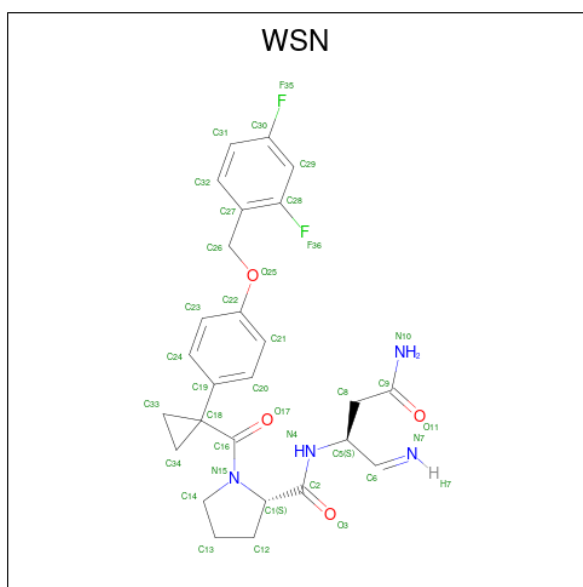
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	2	28	16	2	10	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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

- Molecule 4 is N-[(2R)-4-amino-1-imino-4-oxobutan-2-yl]-1-(1-{4-[(2,4-difluorophenyl)methoxy]phenyl}cyclopropane-1-carbonyl)-L-prolinamide (three-letter code: WSN) (formula: C₂₆H₂₈F₂N₄O₄).



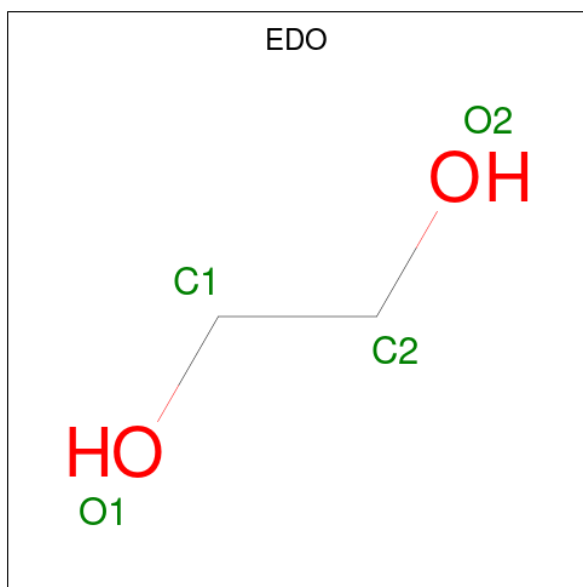
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	0	0
			36	26	2	4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
4	B	1	Total	C	F	N	O	0	1
			45	33	4	4	4		
4	C	1	Total	C	F	N	O	0	0
			36	26	2	4	4		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



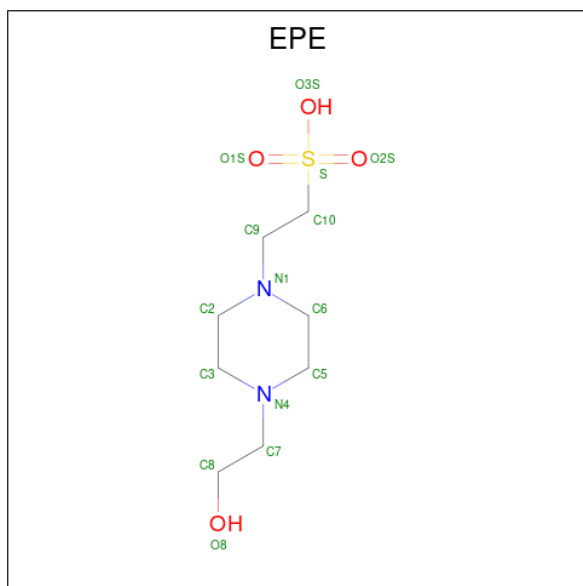
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	C O	0	0
			4	2 2		
5	C	1	Total	C O	0	0
			4	2 2		
5	C	1	Total	C O	0	0
			4	2 2		
5	C	1	Total	C O	0	0
			4	2 2		

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



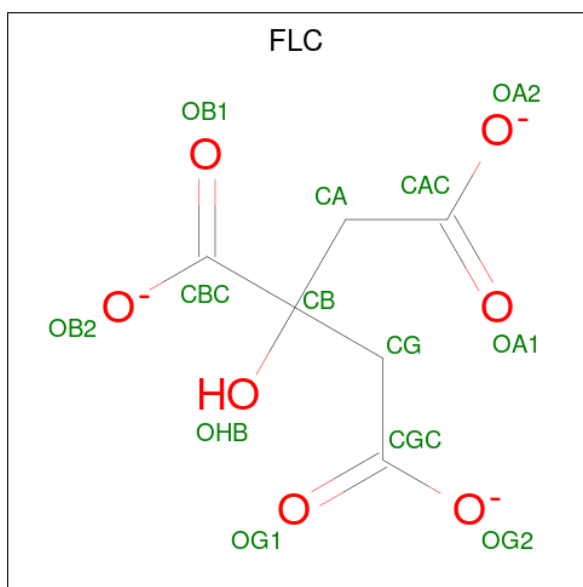
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			13	8	5		

- Molecule 7 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 8 is CITRATE ANION (three-letter code: FLC) (formula: $C_6H_5O_7$).

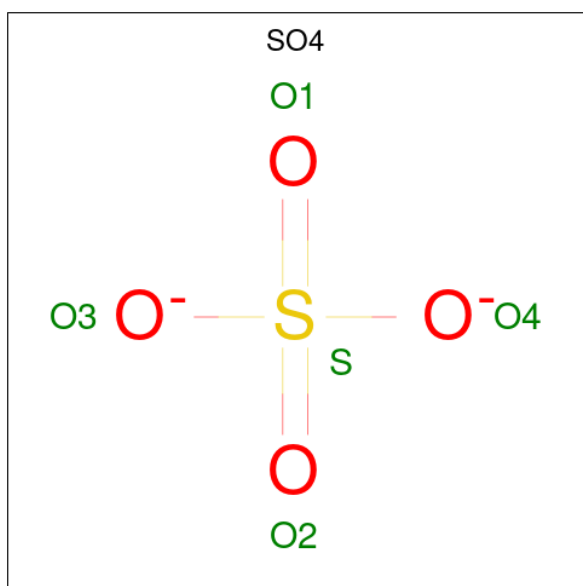


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C O 13 6 7	0	0

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	1	Total Mg 1 1	0	0

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	C	1	Total	O	S	0	0
			5	4	1		

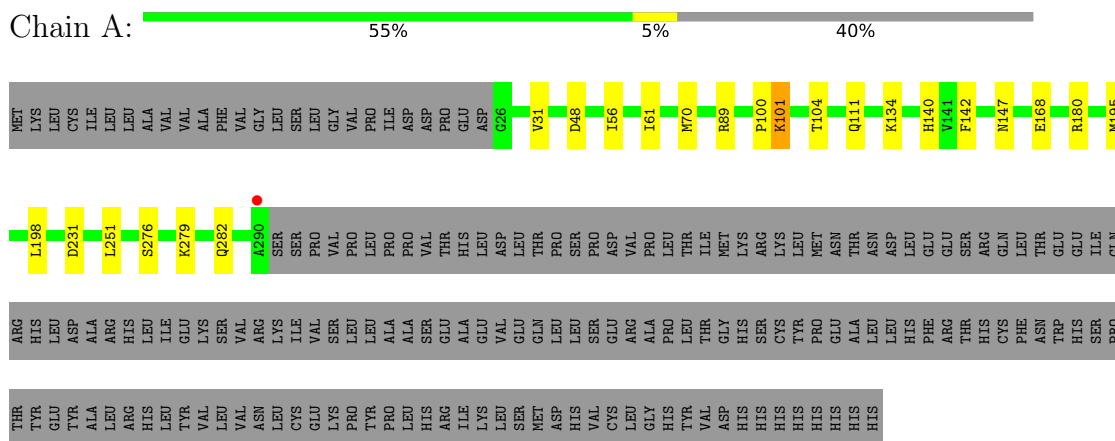
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	245	Total	O	0	2
			247	247		
11	B	240	Total	O	0	1
			241	241		
11	C	179	Total	O	0	0
			179	179		

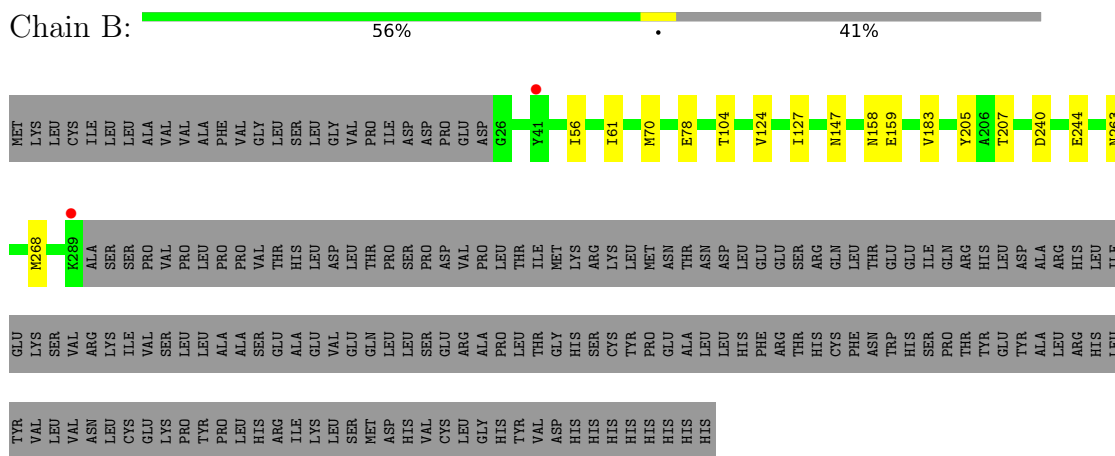
3 Residue-property plots [i](#)

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.

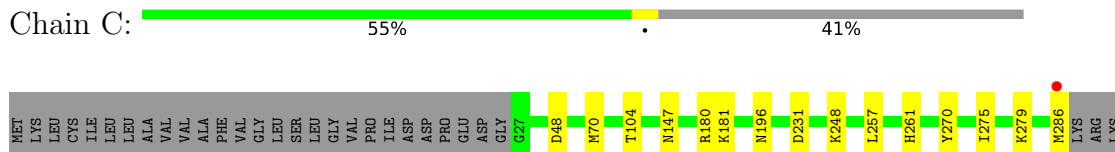
- Molecule 1: Legumain



- Molecule 1: Legumain



- Molecule 1: Legumain



ALA SER
SER SER
PRO PRO
VAL VAL
SER SER
PRO PRO
LEU LEU
PRO PRO
PRO VAL
THR THR
HIS HIS
LEU LEU
ASP ASP
LEU LEU
THR THR
PRO PRO
SER SER
PRO PRO
ASP ASP
VAL VAL
PRO PRO
LEU LEU
THR THR
ILE ILE
MET MET
LYS LYS
SER SER
ARG ARG
LEU LEU
TYR TYR
MET MET
ASN ASN
THR THR
ASN ASN
LEU LEU
ASP ASP
LEU LEU
HIS HIS
GLU GLU
GLU GLU
SER SER
ARG ARG
GLN GLN
LEU LEU
THR THR
GLU GLU
LEU LEU
ASP ASP
ALA ALA
ALA ALA
HIS HIS
HIS HIS
LEU LEU
ILE ILE
GLU GLU
TYR TYR
VAL VAL
SER SER
LEU LEU
VAL VAL

ARG LYS
ILE ILE
VAL VAL
GLU GLU
LYS LYS
PRO PRO
TYR TYR
PRO PRO
ALA ALA
PRO PRO
HIS HIS
SER SER
GLU GLU
ALA ALA
HIS HIS
LEU LEU
VAL VAL
SER SER
GLN GLN
LEU LEU
PRO PRO
SER SER
VAL VAL
CYS CYS
LEU LEU
GLY GLY
ALA ALA
PRO PRO
TYR TYR
THR THR
VAL VAL
ASP ASP
HIS HIS
LYS LYS
SER SER
ARG ARG
LEU LEU
TYR TYR
HIS HIS
HIS HIS
HIS HIS
HIS HIS
HIS HIS

ASN LEU
LEU LEU
CYS CYS
GLU GLU
LYS LYS
PRO PRO
TYR TYR
PRO PRO
ALA ALA
PRO PRO
HIS HIS
SER SER
GLU GLU
ALA ALA
HIS HIS
LEU LEU
VAL VAL
SER SER
MET MET
ASP ASP
HIS HIS
VAL VAL
CYS CYS
LEU LEU
GLY GLY
HIS HIS
TYR TYR
VAL VAL
ASP ASP
HIS HIS
HIS HIS
HIS HIS
HIS HIS
HIS HIS
HIS HIS

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

Chain E:  50% 50%

MAG1
MAG2

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	229.39Å 66.57Å 79.11Å 90.00° 100.93° 90.00°	Depositor
Resolution (Å)	112.62 – 1.70 112.62 – 1.70	Depositor EDS
% Data completeness (in resolution range)	70.2 (112.62-1.70) 70.2 (112.62-1.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 1.70Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.169 , 0.201 0.179 , 0.209	Depositor DCC
R_{free} test set	4751 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	18.2	Xtrriage
Anisotropy	0.038	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 42.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7348	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: EPE, NAG, PG4, FLC, SNN, SO4, EDO, WSN, MG

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	0.66	0/2234	0.77	0/3026
1	B	0.64	0/2230	0.77	0/3019
1	C	0.68	0/2177	0.76	0/2953
All	All	0.66	0/6641	0.76	0/8998

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	2167	0	2074	15	0
1	B	2162	0	2065	6	0
1	C	2116	0	2016	8	0
2	E	28	0	25	0	0
3	A	14	0	13	0	0
3	C	14	0	13	0	0
4	A	36	0	0	0	0
4	B	45	0	0	0	0
4	C	36	0	0	0	0
5	A	4	0	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	12	0	18	2	0
6	A	13	0	18	0	0
7	A	15	0	17	0	0
8	B	13	0	4	0	0
9	B	1	0	0	0	0
10	C	5	0	0	0	0
11	A	247	0	0	6	0
11	B	241	0	0	0	0
11	C	179	0	0	0	0
All	All	7348	0	6269	29	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 2.

All (29) 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:282[B]:GLN:OE1	11:A:601:HOH:O	1.77	1.00
1:A:89[A]:ARG:HD2	11:A:817[A]:HOH:O	1.92	0.69
1:A:111[B]:GLN:HG3	11:A:749:HOH:O	1.99	0.62
1:A:168[A]:GLU:HG2	11:A:635:HOH:O	2.05	0.56
1:B:240:ASP:OD1	1:B:244[B]:GLU:HG3	2.09	0.52
1:C:181:LYS:NZ	5:C:504:EDO:H22	2.25	0.51
1:A:168[A]:GLU:CG	11:A:635:HOH:O	2.58	0.51
1:C:275:ILE:HA	5:C:504:EDO:H12	1.95	0.48
1:A:140:HIS:NE2	1:A:180[A]:ARG:NH2	2.61	0.48
1:A:195:MET:O	1:A:198:LEU:HB2	2.14	0.47
1:C:180:ARG:NH1	1:C:181:LYS:HE2	2.30	0.47
1:A:48:ASP:OD2	1:A:231:ASP:OD2	2.33	0.47
1:B:56:ILE:HG23	1:B:61:ILE:HD12	1.97	0.47
1:C:257[B]:LEU:HD11	1:C:261:HIS:NE2	2.30	0.46
1:B:70:MET:O	1:B:104:THR:HA	2.16	0.45
1:C:70:MET:O	1:C:104:THR:HA	2.18	0.44
1:C:48:ASP:OD2	1:C:231:ASP:OD2	2.36	0.44
1:C:248:LYS:O	1:C:279:LYS:HE3	2.17	0.44
1:A:100:PRO:C	1:A:101:LYS:HD2	2.38	0.44
1:A:251:LEU:HB2	1:A:276:SER:HA	2.00	0.43
1:B:124:VAL:HA	1:B:127:ILE:HD12	2.00	0.42
1:A:31:VAL:HA	1:A:142:PHE:O	2.20	0.42
1:B:183:VAL:HA	1:B:205:TYR:O	2.20	0.41
1:B:207:THR:HA	1:B:268:MET:O	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:279:LYS:HB2	1:A:282[B]:GLN:HG3	2.02	0.41
1:C:196:ASN:HA	1:C:270:TYR:CD1	2.55	0.41
1:A:70:MET:O	1:A:104:THR:HA	2.20	0.41
1:A:56:ILE:HG23	1:A:61:ILE:HD12	2.02	0.41
1:A:111[A]:GLN:CG	11:A:803:HOH:O	2.69	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	268/444 (60%)	262 (98%)	6 (2%)	0	100	100
1	B	267/444 (60%)	257 (96%)	10 (4%)	0	100	100
1	C	261/444 (59%)	254 (97%)	7 (3%)	0	100	100
All	All	796/1332 (60%)	773 (97%)	23 (3%)	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	236/396 (60%)	234 (99%)	2 (1%)	81	74
1	B	236/396 (60%)	232 (98%)	4 (2%)	60	46

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	231/396 (58%)	230 (100%)	1 (0%)	91	87
All	All	703/1188 (59%)	696 (99%)	7 (1%)	76	67

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

Mol	Chain	Res	Type
1	A	101	LYS
1	A	134	LYS
1	B	78	GLU
1	B	158	ASN
1	B	159	GLU
1	B	263	ASN
1	C	286	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	88	ASN
1	B	65	GLN
1	B	263	ASN
1	C	88	ASN
1	C	138	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

3 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	A	147	1	5,6,8	0.94	0	3,6,11	3.01	2 (66%)
1	SNN	C	147	1	5,6,8	0.85	0	3,6,11	3.28	2 (66%)
1	SNN	B	147	1	5,6,8	0.84	0	3,6,11	3.30	3 (100%)

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

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	147	SNN	C4-CA-C	-4.01	103.96	111.44
1	B	147	SNN	C4-CA-C	-3.96	104.04	111.44
1	C	147	SNN	CA-C4-C5	-3.89	99.72	114.44
1	A	147	SNN	C4-CA-C	-3.61	104.70	111.44
1	B	147	SNN	CA-C4-C5	-3.39	101.59	114.44
1	A	147	SNN	CA-C4-C5	-3.34	101.78	114.44
1	B	147	SNN	O5-C5-C4	-2.33	118.64	125.43

There are no chirality outliers.

All (9) torsion outliers are listed below:

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

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	1,2	14,14,15	0.59	0	17,19,21	1.01	0
2	NAG	E	2	2	14,14,15	0.52	0	17,19,21	1.62	1 (5%)

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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	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(°)	Ideal(°)
2	E	2	NAG	C1-O5-C5	5.74	119.97	112.19

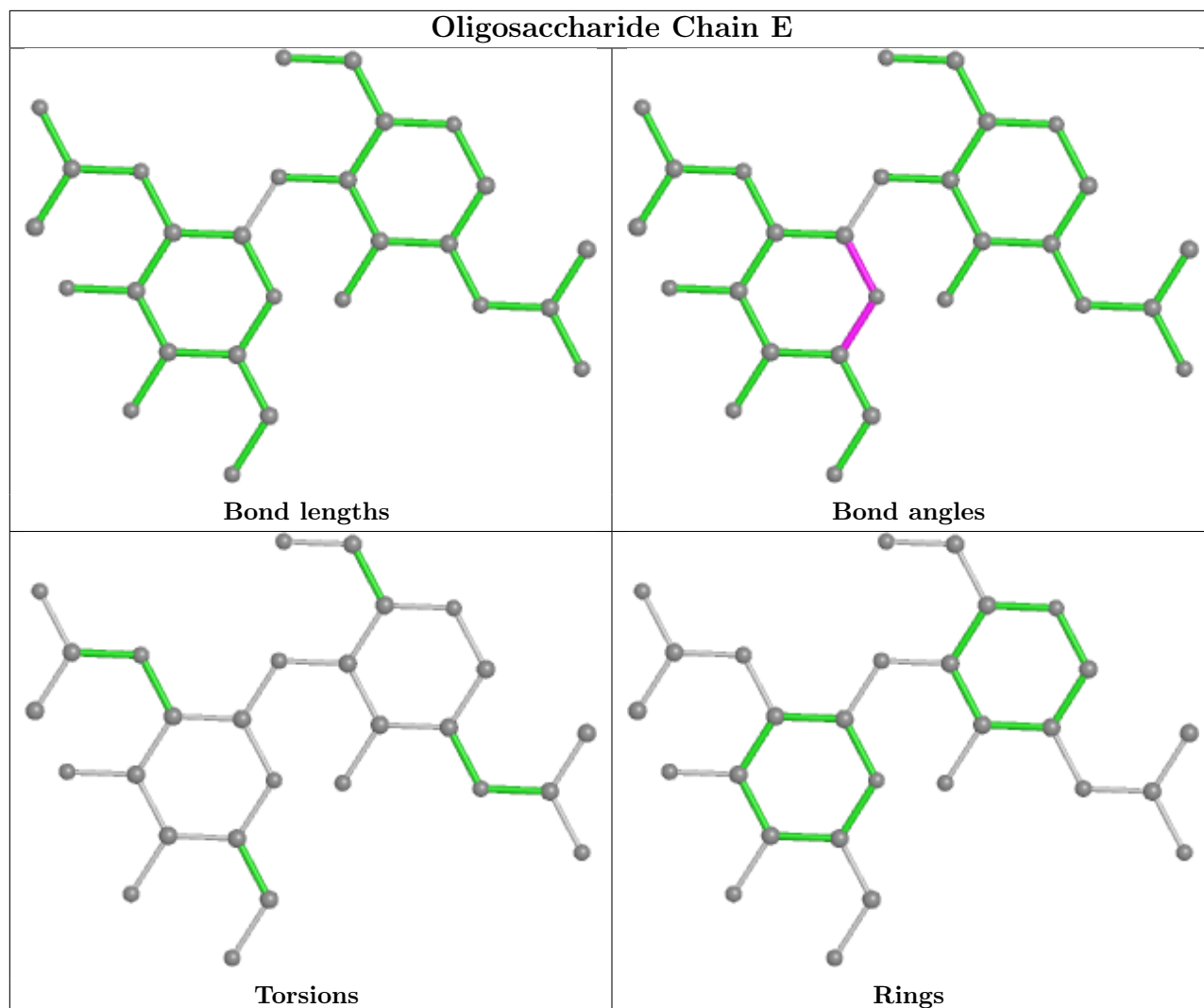
There are no chirality outliers.

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 [i](#)

Of 15 ligands modelled in this entry, 1 is monoatomic - leaving 14 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$
3	NAG	A	501	1	14,14,15	0.57	0	17,19,21	0.76	0
8	FLC	B	502	9	12,12,12	1.28	1 (8%)	17,17,17	1.53	2 (11%)
5	EDO	C	504	-	3,3,3	0.11	0	2,2,2	0.29	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	WSN	A	502	1	36,39,39	0.36	0	49,56,56	0.68	1 (2%)
10	SO4	C	506	-	4,4,4	0.28	0	6,6,6	0.14	0
4	WSN	C	502	1	36,39,39	0.27	0	49,56,56	0.73	0
4	WSN	B	501[A]	-	36,39,39	0.30	0	49,56,56	0.49	0
5	EDO	C	505	-	3,3,3	0.08	0	2,2,2	0.24	0
6	PG4	A	504	-	12,12,12	0.19	0	11,11,11	0.09	0
7	EPE	A	505	-	15,15,15	0.93	1 (6%)	18,20,20	1.76	2 (11%)
3	NAG	C	501	1	14,14,15	0.35	0	17,19,21	0.82	1 (5%)
5	EDO	A	503	-	3,3,3	0.19	0	2,2,2	0.27	0
5	EDO	C	503	-	3,3,3	0.17	0	2,2,2	0.35	0
4	WSN	B	501[B]	-	36,39,39	0.32	0	49,56,56	0.58	1 (2%)

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	A	501	1	-	0/6/23/26	0/1/1/1
8	FLC	B	502	9	-	0/16/16/16	-
5	EDO	C	504	-	-	0/1/1/1	-
4	WSN	A	502	1	-	0/33/49/49	0/4/4/4
4	WSN	C	502	1	-	1/33/49/49	0/4/4/4
4	WSN	B	501[A]	-	-	4/33/49/49	0/4/4/4
5	EDO	C	505	-	-	1/1/1/1	-
6	PG4	A	504	-	-	3/10/10/10	-
7	EPE	A	505	-	-	3/9/19/19	0/1/1/1
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
5	EDO	A	503	-	-	0/1/1/1	-
5	EDO	C	503	-	-	0/1/1/1	-
4	WSN	B	501[B]	-	-	4/33/49/49	0/4/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	505	EPE	O1S-S	3.42	1.55	1.45
8	B	502	FLC	CB-CBC	2.09	1.55	1.53

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	505	EPE	O1S-S-C10	-6.03	99.65	106.92
8	B	502	FLC	OB1-CBC-CB	-4.45	115.95	122.25
7	A	505	EPE	O3S-S-O2S	4.00	121.05	111.27
4	A	502	WSN	O25-C26-C27	2.73	117.45	109.38
8	B	502	FLC	OB2-CBC-CB	2.65	117.65	113.05
3	C	501	NAG	C1-O5-C5	2.41	115.46	112.19
4	B	501[B]	WSN	O25-C26-C27	2.30	116.18	109.38

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	505	EPE	C9-C10-S-O2S
7	A	505	EPE	C9-C10-S-O3S
4	B	501[A]	WSN	C23-C22-O25-C26
4	B	501[A]	WSN	C21-C22-O25-C26
4	B	501[B]	WSN	C23-C22-O25-C26
6	A	504	PG4	C5-C6-O4-C7
7	A	505	EPE	C9-C10-S-O1S
6	A	504	PG4	C1-C2-O2-C3
4	B	501[A]	WSN	C33-C18-C19-C20
4	B	501[A]	WSN	C33-C18-C19-C24
4	B	501[B]	WSN	C33-C18-C19-C20
4	B	501[B]	WSN	C33-C18-C19-C24
4	C	502	WSN	C33-C18-C19-C24
4	B	501[B]	WSN	C21-C22-O25-C26
6	A	504	PG4	O4-C7-C8-O5
5	C	505	EDO	O1-C1-C2-O2

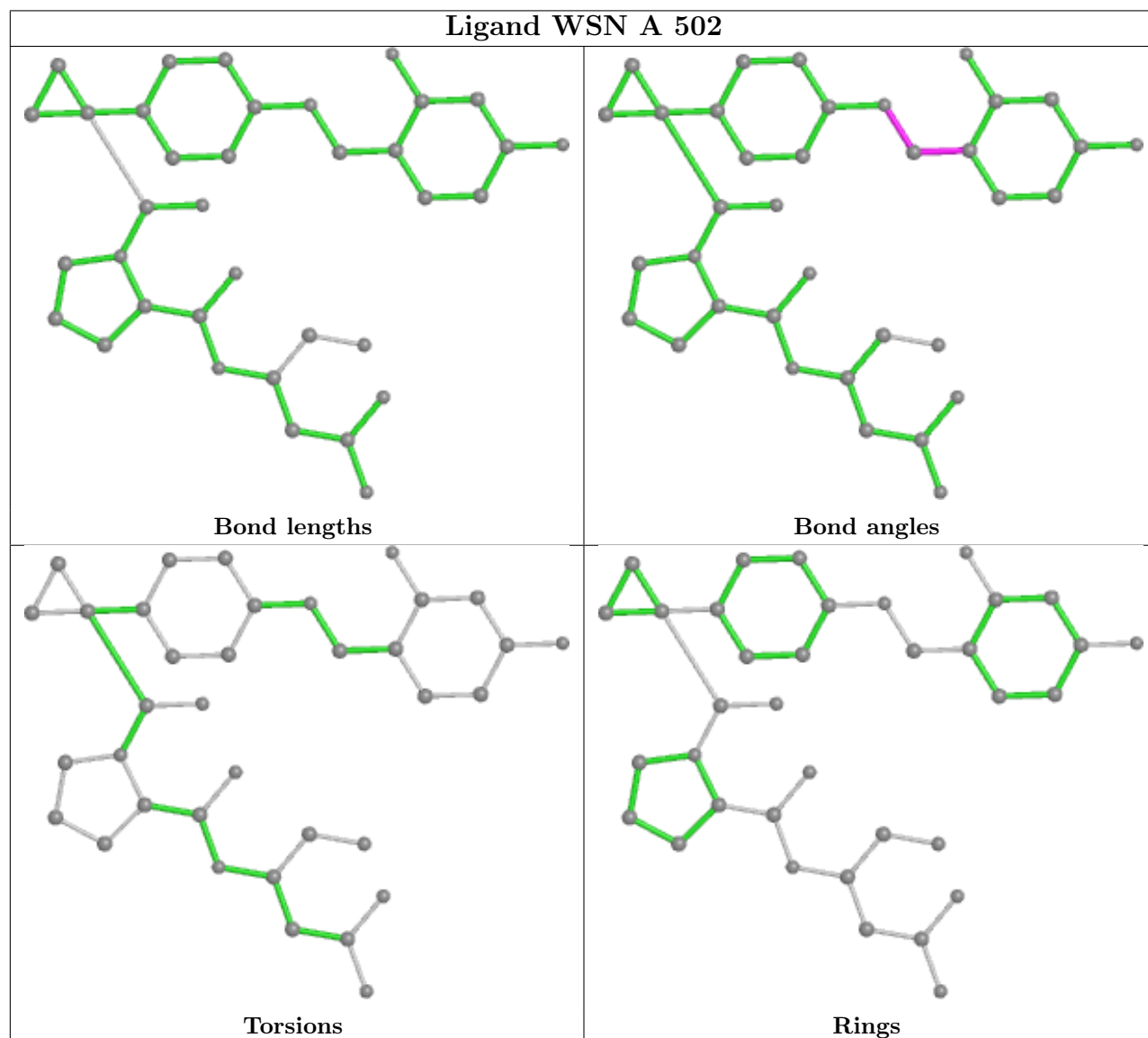
There are no ring outliers.

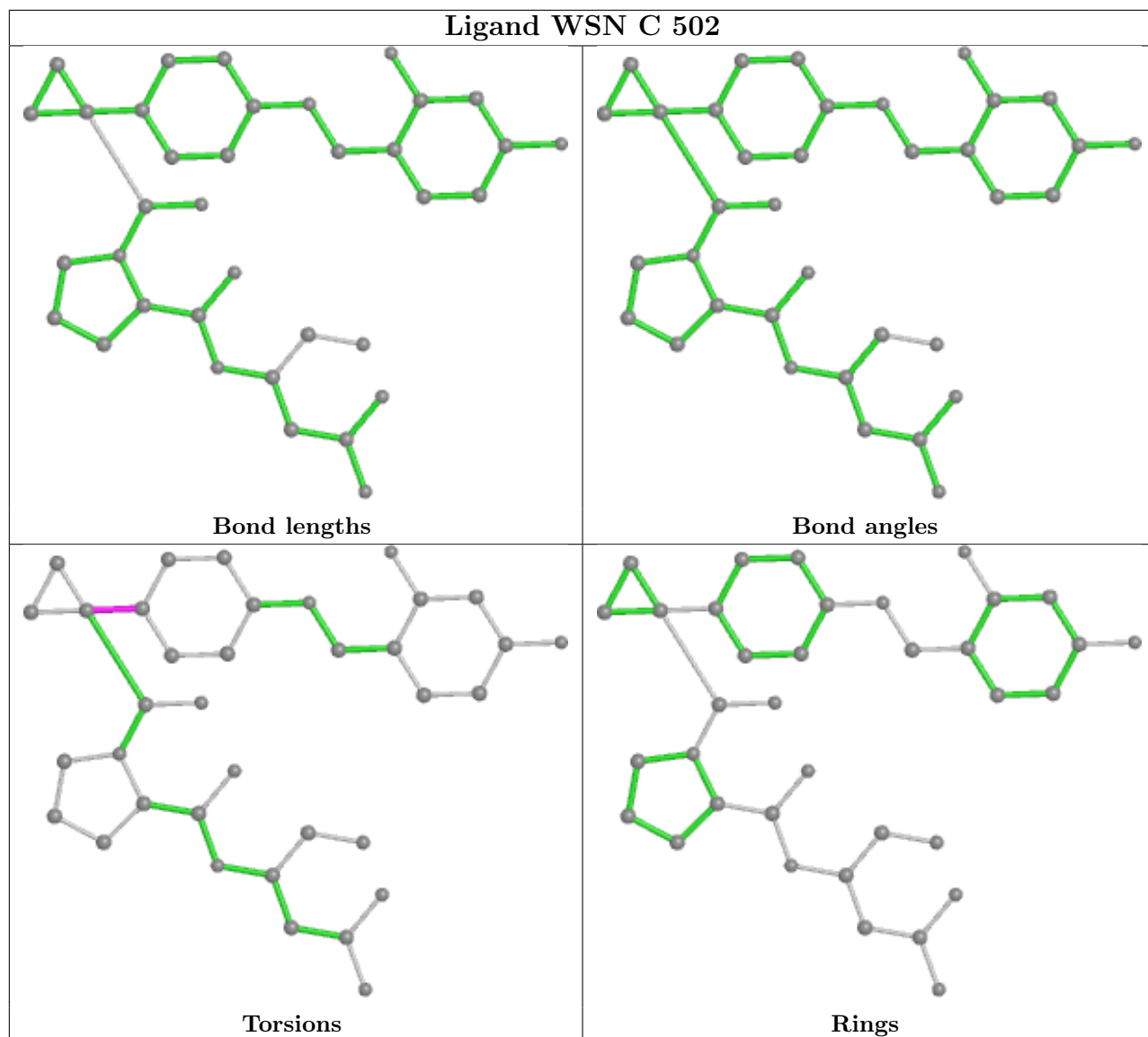
1 monomer is involved in 2 short contacts:

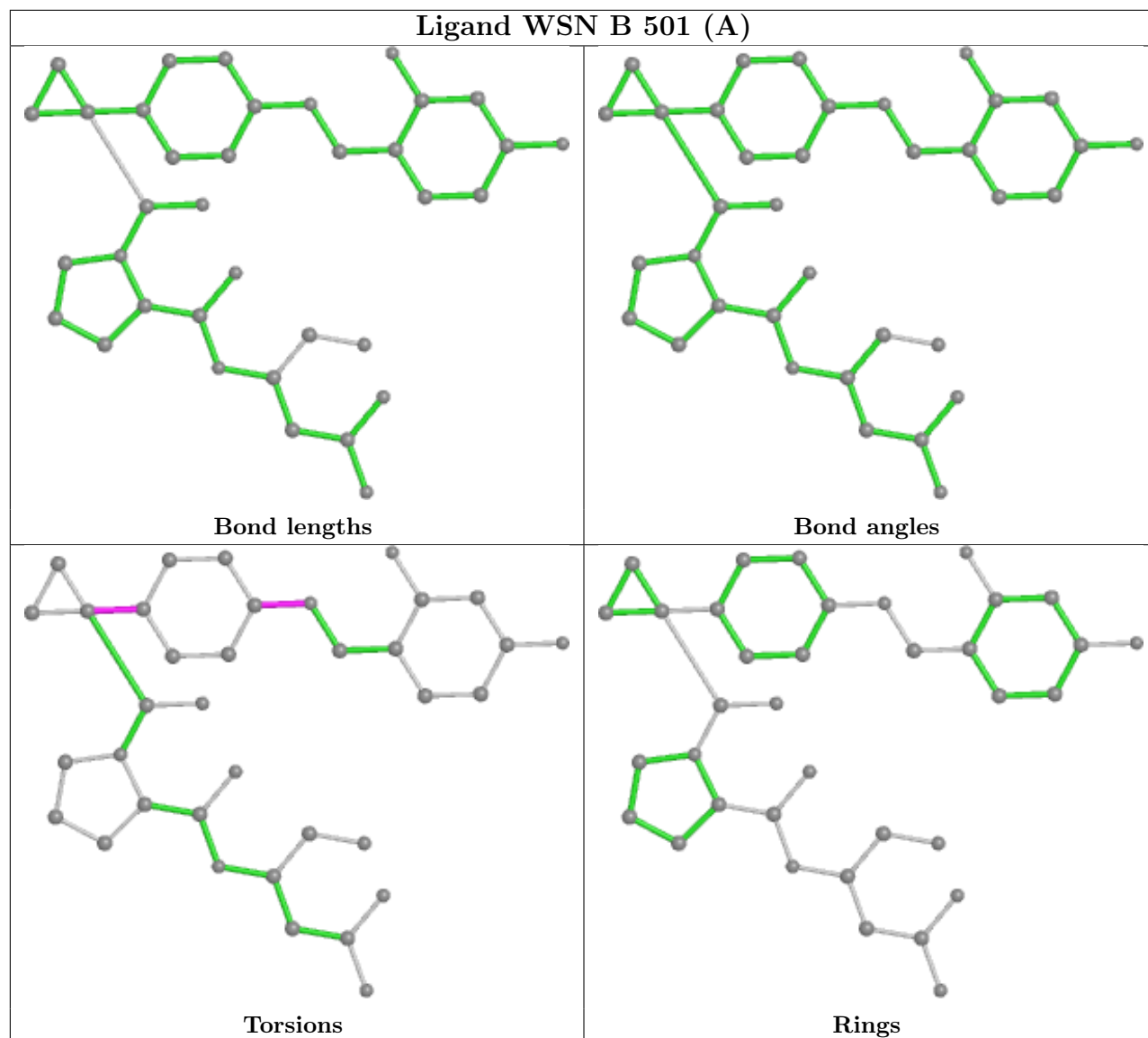
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	504	EDO	2	0

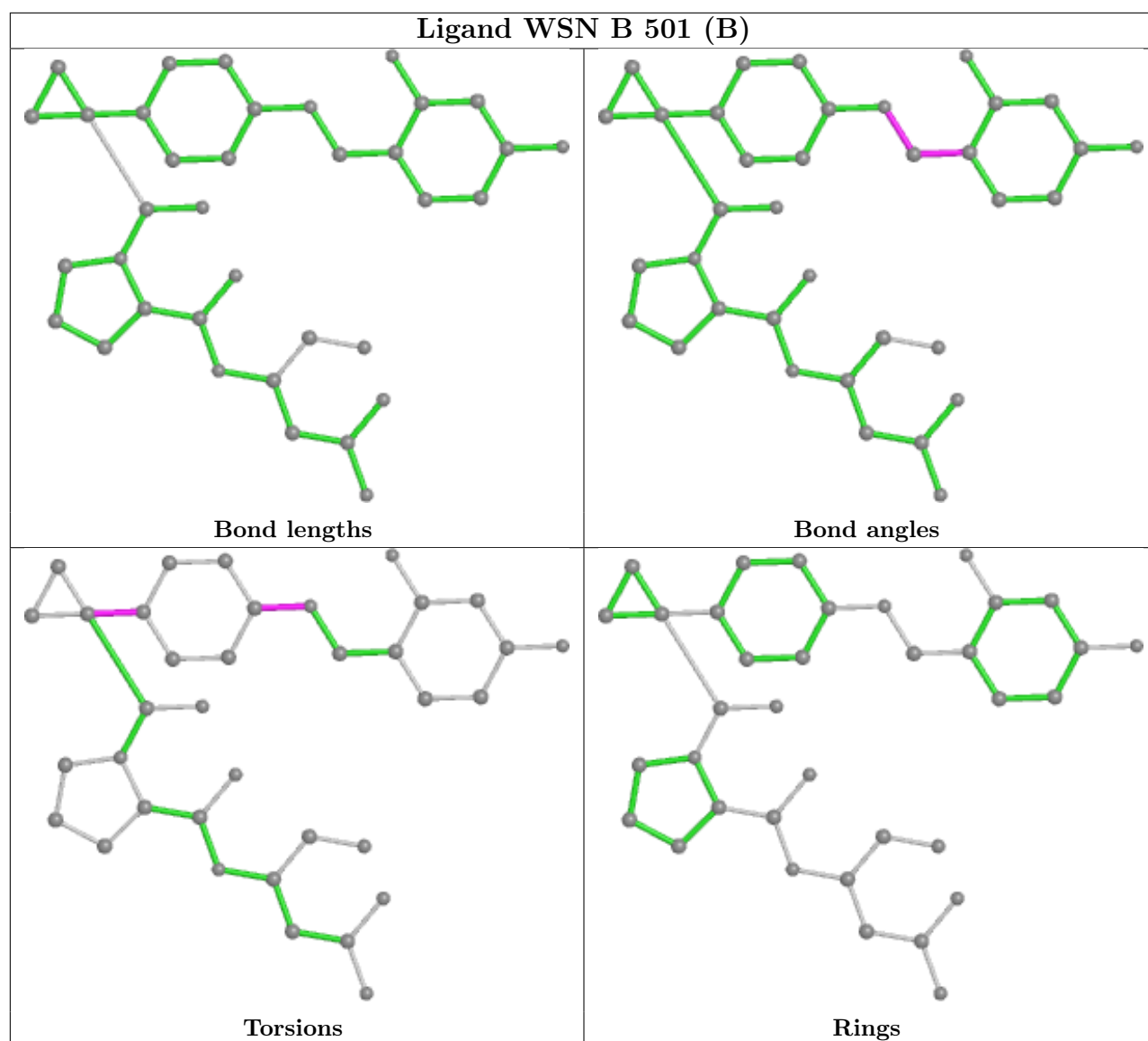
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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









5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	264/444 (59%)	-0.33	1 (0%) 92 93	12, 18, 33, 49	0
1	B	263/444 (59%)	-0.37	2 (0%) 86 88	13, 18, 31, 62	0
1	C	259/444 (58%)	-0.30	1 (0%) 92 93	13, 21, 40, 57	0
All	All	786/1332 (59%)	-0.33	4 (0%) 91 92	12, 19, 34, 62	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	290	ALA	2.7
1	B	289	LYS	2.4
1	C	286	MET	2.3
1	B	41[A]	TYR	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	SNN	A	147	7/8	0.98	0.07	11,11,12,13	0
1	SNN	B	147	7/8	0.98	0.06	12,13,15,15	0
1	SNN	C	147	7/8	0.98	0.07	13,14,15,16	0

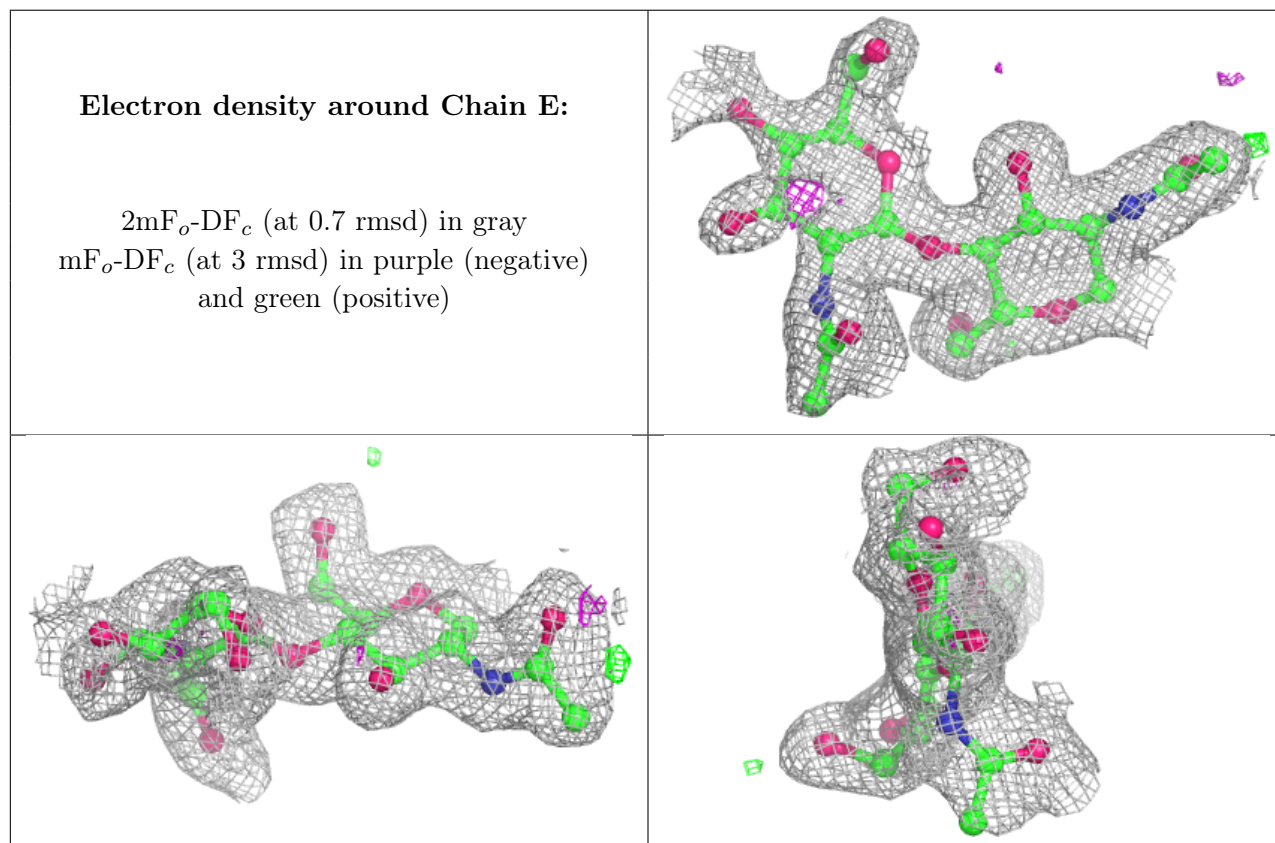
6.3 Carbohydrates [i](#)

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, 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	E	2	14/15	0.77	0.18	46,56,59,63	0
2	NAG	E	1	14/15	0.96	0.07	24,30,33,36	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.



6.4 Ligands [i](#)

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

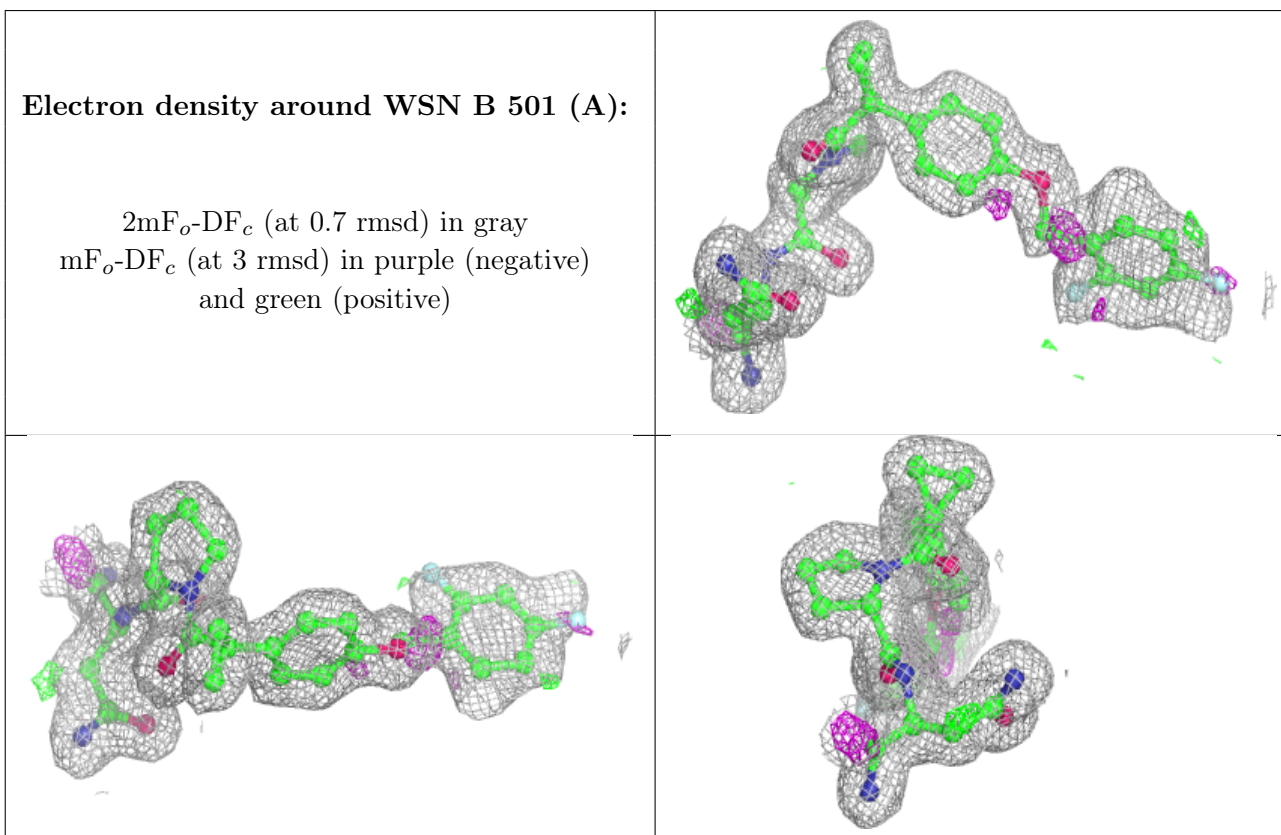
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	EDO	C	505	4/4	0.78	0.13	44,45,47,49	0
7	EPE	A	505	15/15	0.81	0.18	48,55,63,63	0
5	EDO	A	503	4/4	0.83	0.16	38,41,41,41	0
3	NAG	A	501	14/15	0.85	0.17	38,41,42,49	0

Continued on next page...

Continued from previous page...

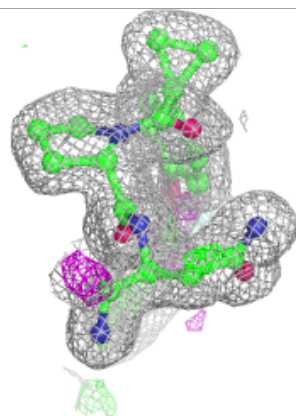
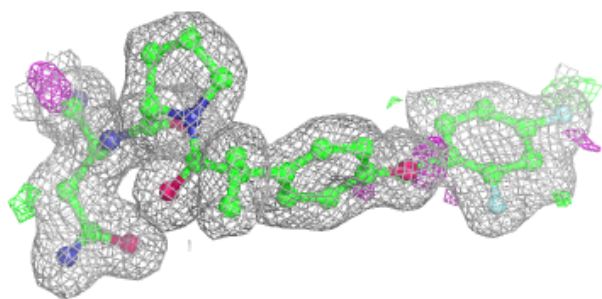
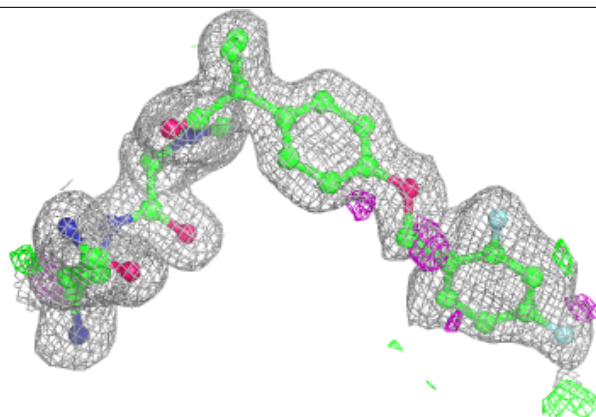
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	FLC	B	502	13/13	0.85	0.18	29,33,37,38	0
10	SO4	C	506	5/5	0.86	0.17	54,64,66,70	0
6	PG4	A	504	13/13	0.89	0.15	43,44,48,49	13
5	EDO	C	504	4/4	0.90	0.11	38,40,41,49	0
3	NAG	C	501	14/15	0.90	0.09	35,37,41,48	0
5	EDO	C	503	4/4	0.94	0.12	32,35,38,41	0
4	WSN	B	501[A]	36/36	0.96	0.09	16,19,39,43	9
9	MG	B	503	1/1	0.96	0.28	32,32,32,32	0
4	WSN	B	501[B]	36/36	0.96	0.09	16,19,27,29	9
4	WSN	C	502	36/36	0.97	0.07	14,17,25,27	0
4	WSN	A	502	36/36	0.97	0.08	12,15,25,31	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.

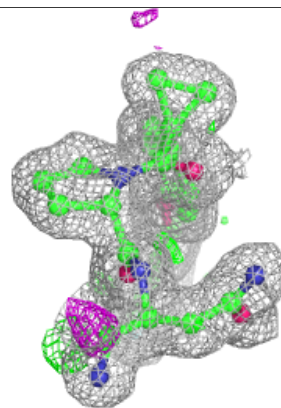
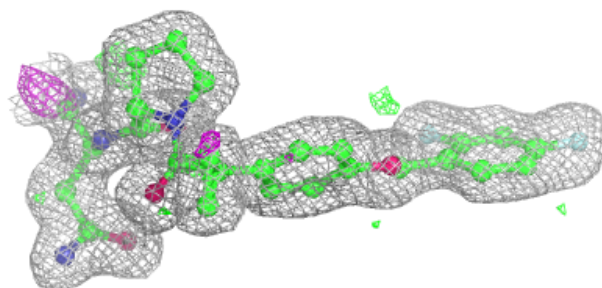
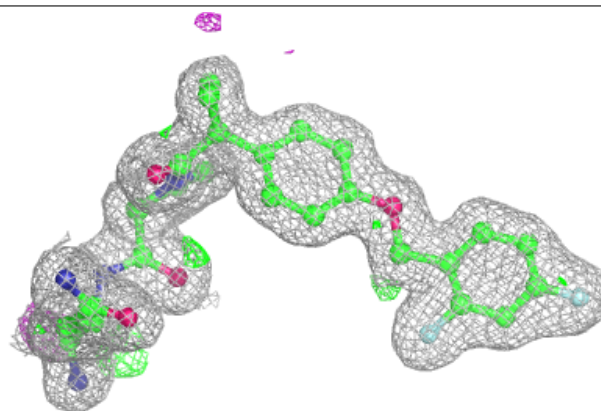


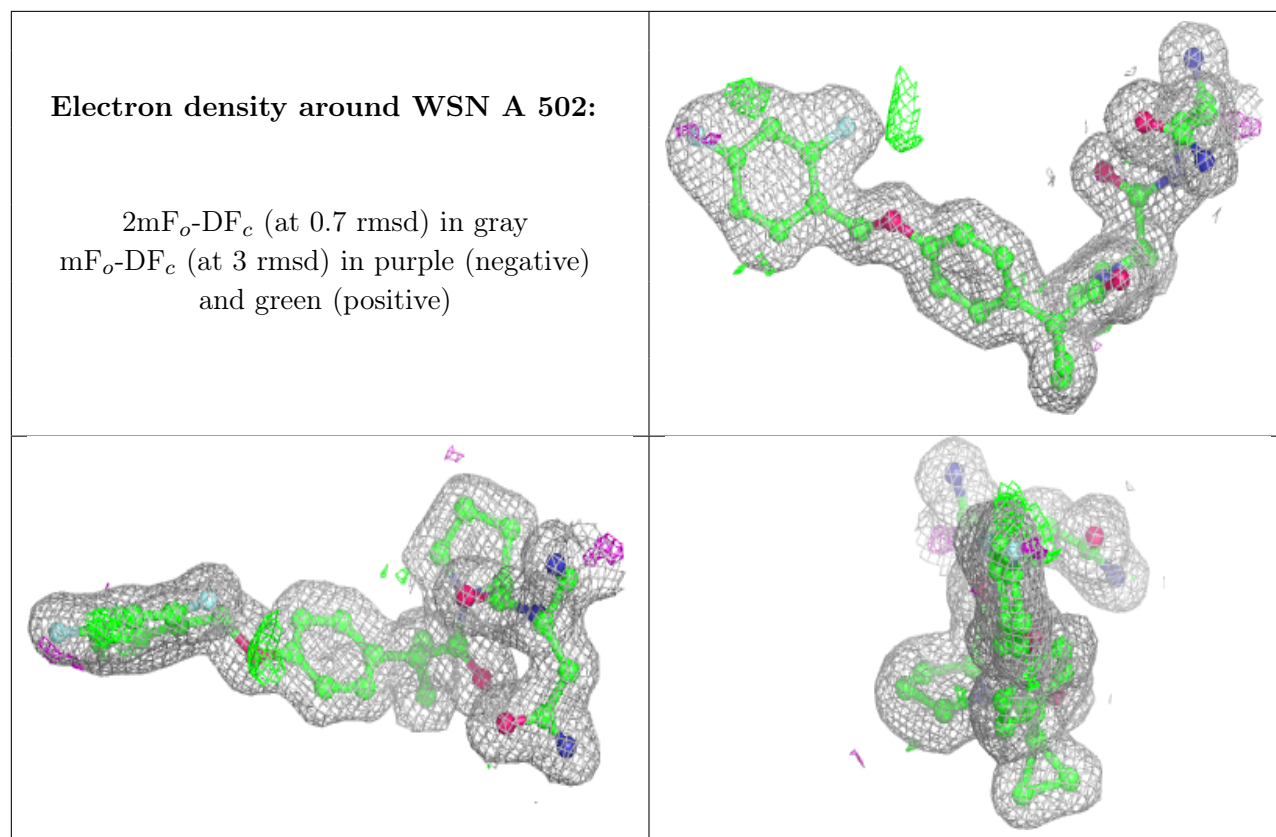
Electron density around WSN B 501 (B):

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

**Electron density around WSN C 502:**

$2mF_o-DF_c$ (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.