Full wwPDB X-ray Structure Validation Report (i)

May 28, 2025 – 01:31 pm BST

PDB ID	:	9HH4 / pdb_00009hh4
Title	:	Crystal structure of the family S1_19 carrageenan sulfatase ZgCgsA from Zobellia galactanivorans in complex with hybrid b-k-neocarratetraose
Authors	:	Chevenier, A.; Czjzek, M.; Michel, G.; Ficko-Blean, E.
Deposited on	:	2024-11-21
Resolution	:	1.92 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 (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:

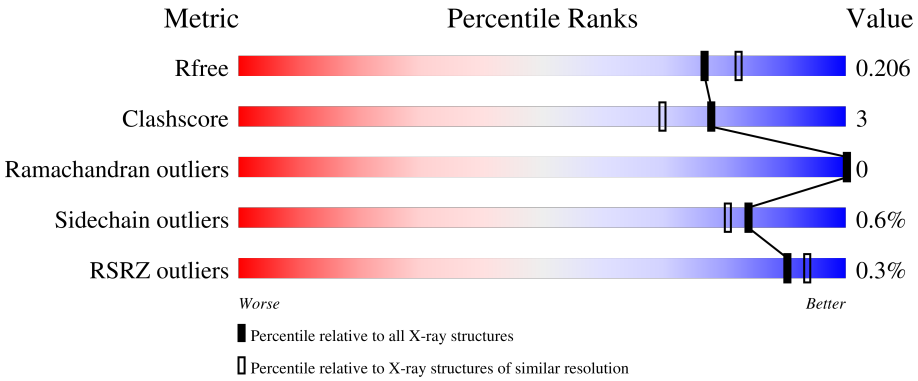
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 1.92 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}	164625	1028 (1.92-1.92)
Clashscore	180529	1100 (1.92-1.92)
Ramachandran outliers	177936	1087 (1.92-1.92)
Sidechain outliers	177891	1087 (1.92-1.92)
RSRZ outliers	164620	1028 (1.92-1.92)

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	480	91% 6% .
1	B	480	88% 9% .
1	C	480	91% 6% .
1	D	480	90% 7% .
2	I	4	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	4	50% 50%
2	K	4	50% 50%
2	L	4	75% 25%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 16711 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 Sulfatase, family S1-19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	465	Total	C	N	O	S	0	1	0
			3737	2389	634	701	13			
1	B	466	Total	C	N	O	S	0	2	0
			3740	2391	632	704	13			
1	C	464	Total	C	N	O	S	0	0	0
			3680	2351	622	694	13			
1	D	466	Total	C	N	O	S	0	0	0
			3718	2375	628	702	13			

There are 52 discrepancies between the modelled and reference sequences:

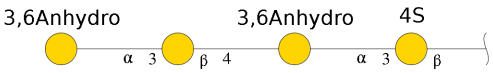
Chain	Residue	Modelled	Actual	Comment	Reference
A	32	MET	-	initiating methionine	UNP G0L000
A	33	GLY	-	expression tag	UNP G0L000
A	34	SER	-	expression tag	UNP G0L000
A	35	SER	-	expression tag	UNP G0L000
A	36	HIS	-	expression tag	UNP G0L000
A	37	HIS	-	expression tag	UNP G0L000
A	38	HIS	-	expression tag	UNP G0L000
A	39	HIS	-	expression tag	UNP G0L000
A	40	HIS	-	expression tag	UNP G0L000
A	41	HIS	-	expression tag	UNP G0L000
A	42	GLY	-	expression tag	UNP G0L000
A	43	SER	-	expression tag	UNP G0L000
A	96	OSE	CYS	conflict	UNP G0L000
B	32	MET	-	initiating methionine	UNP G0L000
B	33	GLY	-	expression tag	UNP G0L000
B	34	SER	-	expression tag	UNP G0L000
B	35	SER	-	expression tag	UNP G0L000
B	36	HIS	-	expression tag	UNP G0L000
B	37	HIS	-	expression tag	UNP G0L000
B	38	HIS	-	expression tag	UNP G0L000
B	39	HIS	-	expression tag	UNP G0L000

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	40	HIS	-	expression tag	UNP G0L000
B	41	HIS	-	expression tag	UNP G0L000
B	42	GLY	-	expression tag	UNP G0L000
B	43	SER	-	expression tag	UNP G0L000
B	96	OSE	CYS	conflict	UNP G0L000
C	32	MET	-	initiating methionine	UNP G0L000
C	33	GLY	-	expression tag	UNP G0L000
C	34	SER	-	expression tag	UNP G0L000
C	35	SER	-	expression tag	UNP G0L000
C	36	HIS	-	expression tag	UNP G0L000
C	37	HIS	-	expression tag	UNP G0L000
C	38	HIS	-	expression tag	UNP G0L000
C	39	HIS	-	expression tag	UNP G0L000
C	40	HIS	-	expression tag	UNP G0L000
C	41	HIS	-	expression tag	UNP G0L000
C	42	GLY	-	expression tag	UNP G0L000
C	43	SER	-	expression tag	UNP G0L000
C	96	OSE	CYS	conflict	UNP G0L000
D	32	MET	-	initiating methionine	UNP G0L000
D	33	GLY	-	expression tag	UNP G0L000
D	34	SER	-	expression tag	UNP G0L000
D	35	SER	-	expression tag	UNP G0L000
D	36	HIS	-	expression tag	UNP G0L000
D	37	HIS	-	expression tag	UNP G0L000
D	38	HIS	-	expression tag	UNP G0L000
D	39	HIS	-	expression tag	UNP G0L000
D	40	HIS	-	expression tag	UNP G0L000
D	41	HIS	-	expression tag	UNP G0L000
D	42	GLY	-	expression tag	UNP G0L000
D	43	SER	-	expression tag	UNP G0L000
D	96	OSE	CYS	conflict	UNP G0L000

- Molecule 2 is an oligosaccharide called 3,6-anhydro-alpha-D-galactopyranose-(1-3)-beta-D-galactopyranose-(1-4)-3,6-anhydro-alpha-D-galactopyranose-(1-3)-4-O-sulfo-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	I	4	Total C O S 47 24 22 1	0	0	0
2	J	4	Total C O S 47 24 22 1	0	0	0
2	K	4	Total C O S 47 24 22 1	0	0	0
2	L	4	Total C O S 47 24 22 1	0	0	0

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	B	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0
3	D	1	Total Ca 1 1	0	0

- Molecule 4 is BROMIDE ION (CCD ID: BR) (formula: Br).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	3	Total Br 3 3	0	0
4	B	3	Total Br 3 3	0	0
4	C	3	Total Br 3 3	0	0
4	D	3	Total Br 3 3	0	0

- Molecule 5 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

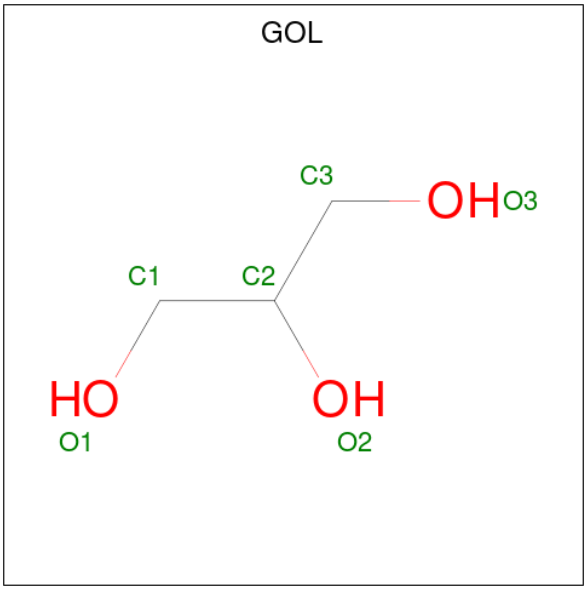
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	8	Total Cl 8 8	0	0
5	B	4	Total Cl 4 4	0	0
5	C	5	Total Cl 5 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	4	Total Cl 4 4	0	0

- Molecule 6 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	A	1	Total C O 6 3 3	0	0
6	A	1	Total C O 6 3 3	0	0
6	B	1	Total C O 6 3 3	0	0
6	B	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	425	Total O 438 438	0	13

Continued on next page...

Continued from previous page...

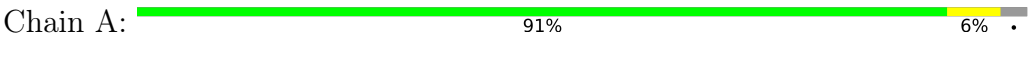
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	444	Total O 463 463	0	19
7	C	314	Total O 325 325	0	11
7	D	334	Total O 343 343	0	9

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.

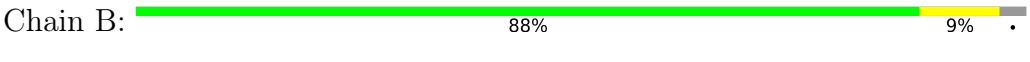
- Molecule 1: Sulfatase, family S1-19

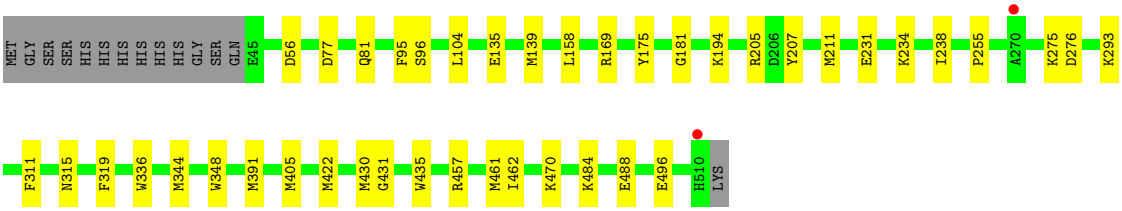
Chain A:



- Molecule 1: Sulfatase, family S1-19

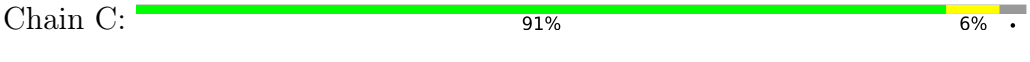
Chain B:

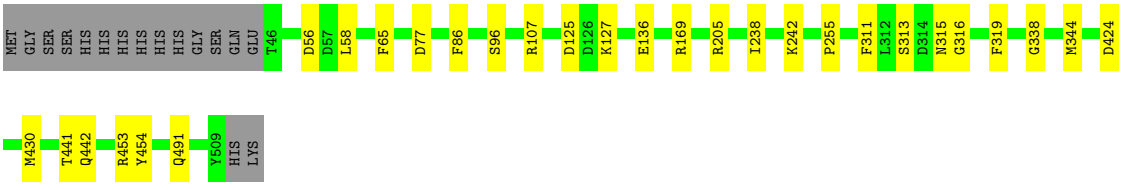




- Molecule 1: Sulfatase, family S1-19

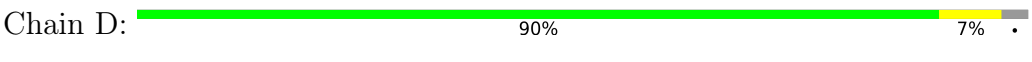
Chain C:





- Molecule 1: Sulfatase, family S1-19

Chain D:



• Molecule 2: 3,6-anhydro-alpha-D-galactopyranose-(1-3)-beta-D-galactopyranose-(1-4)-3,6-anhydro-alpha-D-galactopyranose-(1-3)-4-O-sulfo-beta-D-galactopyranose

Chain I:

• Molecule 2: 3,6-anhydro-alpha-D-galactopyranose-(1-3)-beta-D-galactopyranose-(1-4)-3,6-anhydro-alpha-D-galactopyranose-(1-3)-4-O-sulfo-beta-D-galactopyranose

Chain J:

• Molecule 2: 3,6-anhydro-alpha-D-galactopyranose-(1-3)-beta-D-galactopyranose-(1-4)-3,6-anhydro-alpha-D-galactopyranose-(1-3)-4-O-sulfo-beta-D-galactopyranose

Chain K:

• Molecule 2: 3,6-anhydro-alpha-D-galactopyranose-(1-3)-beta-D-galactopyranose-(1-4)-3,6-anhydro-alpha-D-galactopyranose-(1-3)-4-O-sulfo-beta-D-galactopyranose

Chain L:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.23Å 197.60Å 75.02Å 90.00° 99.79° 90.00°	Depositor
Resolution (Å)	36.57 – 1.92 36.57 – 1.92	Depositor EDS
% Data completeness (in resolution range)	99.4 (36.57-1.92) 93.8 (36.57-1.92)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.02 (at 1.92Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.167 , 0.207 0.165 , 0.206	Depositor DCC
R_{free} test set	8868 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	30.1	Xtriage
Anisotropy	0.371	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 48.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	16711	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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: BR, 9RN, GAL, CL, OSE, GOL, CA, G4S

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	0.44	0/3835	0.62	1/5194 (0.0%)
1	B	0.43	0/3838	0.62	1/5202 (0.0%)
1	C	0.36	0/3775	0.56	0/5126
1	D	0.38	0/3813	0.59	1/5171 (0.0%)
All	All	0.40	0/15261	0.60	3/20693 (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	B	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	95	PHE	CA-CB-CG	7.23	121.03	113.80
1	D	95	PHE	CA-CB-CG	5.83	119.63	113.80
1	A	95	PHE	CA-CB-CG	5.82	119.62	113.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	169	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	169	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	3737	0	3592	19	0
1	B	3740	0	3577	24	0
1	C	3680	0	3477	17	0
1	D	3718	0	3537	18	0
2	I	47	0	19	0	0
2	J	47	0	19	1	0
2	K	47	0	19	1	0
2	L	47	0	19	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	3	0	0	0	0
4	B	3	0	0	1	0
4	C	3	0	0	0	0
4	D	3	0	0	1	0
5	A	8	0	0	1	0
5	B	4	0	0	0	0
5	C	5	0	0	0	0
5	D	4	0	0	0	0
6	A	18	0	24	3	0
6	B	12	0	16	0	0
6	D	12	0	16	1	0
7	A	438	0	0	2	0
7	B	463	0	0	2	0
7	C	325	0	0	1	0
7	D	343	0	0	2	0
All	All	16711	0	14315	78	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (78) 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:D:403:ASP:HA	1:D:422:MET:HE3	1.50	0.90
1:B:77:ASP:O	1:B:81:GLN:HG2	1.84	0.77
1:C:125:ASP:OD1	1:C:127:LYS:HG2	1.91	0.70
1:B:484:LYS:HE3	1:B:488:GLU:OE2	1.94	0.68
1:D:132:VAL:O	1:D:169:ARG:NH2	2.27	0.68
1:A:69:THR:H	6:A:615:GOL:H11	1.61	0.66
1:D:296:GLN:NE2	7:D:701:HOH:O	2.30	0.64
1:B:211:MET:HB2	4:B:604:BR:BR	2.55	0.62
1:A:216:LYS:NZ	7:A:705:HOH:O	2.34	0.60
1:B:496:GLU:H	1:B:496:GLU:CD	2.12	0.58
1:B:348:TRP:CZ3	1:B:391:MET:HE3	2.39	0.58
1:B:405:MET:HG2	1:B:422:MET:HE3	1.86	0.57
1:D:453:ARG:NH2	7:D:702:HOH:O	2.37	0.56
1:A:187:SER:O	1:A:191:LYS:HG3	2.06	0.56
1:A:275:LYS:HG3	1:A:278:ARG:NH2	2.20	0.55
1:B:194:LYS:NZ	7:B:704:HOH:O	2.39	0.55
1:A:474:LYS:NZ	7:A:707:HOH:O	2.38	0.55
1:C:238:ILE:O	1:C:242:LYS:HG3	2.06	0.55
1:D:169:ARG:HD2	4:D:602:BR:BR	2.64	0.53
1:B:457:ARG:O	1:B:461:MET:HG3	2.09	0.52
1:A:69:THR:H	6:A:615:GOL:C1	2.23	0.52
1:D:482:LYS:HZ1	6:D:609:GOL:H11	1.75	0.52
1:B:181:GLY:HA3	1:B:207:TYR:CZ	2.45	0.52
1:D:311:PHE:O	1:D:344:MET:HA	2.10	0.52
1:A:149:TYR:OH	1:A:235:ASN:ND2	2.43	0.51
1:C:125:ASP:CG	1:C:127:LYS:HG2	2.34	0.51
1:D:181:GLY:HA3	1:D:207:TYR:CZ	2.44	0.51
1:D:403:ASP:CA	1:D:422:MET:HE3	2.32	0.51
1:C:86:PHE:HB2	1:C:344:MET:HB3	1.92	0.51
1:B:135:GLU:H	1:B:135:GLU:CD	2.19	0.50
1:B:348:TRP:CE3	1:B:391:MET:HE3	2.46	0.50
6:A:613:GOL:H31	1:B:431:GLY:HA3	1.93	0.50
1:A:129:ASN:HD21	1:A:163:LYS:NZ	2.11	0.48
1:B:234:LYS:O	1:B:238:ILE:HG13	2.13	0.48
1:B:311:PHE:O	1:B:344:MET:HA	2.14	0.48
1:A:156:TRP:CD2	1:A:166:PRO:HG3	2.48	0.48
1:C:65:PHE:N	1:C:77:ASP:OD1	2.47	0.47
1:C:442:GLN:NE2	1:C:442:GLN:HA	2.30	0.47
1:A:139:MET:HB3	1:A:368:TYR:CD2	2.50	0.47
1:B:234:LYS:NZ	7:B:711:HOH:O	2.47	0.46
1:D:84:SER:CB	1:D:356:GLN:HG2	2.46	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:497:ALA:HA	1:D:500:GLU:HG3	1.99	0.45
1:A:191:LYS:HE3	1:A:191:LYS:HB3	1.83	0.45
1:D:361:PRO:HB2	1:D:406:ILE:HD12	1.98	0.45
1:A:205[A]:ARG:HG2	5:A:606:CL:CL	2.54	0.45
1:C:311:PHE:O	1:C:344:MET:HA	2.17	0.45
1:A:165:HIS:O	1:A:169:ARG:HG2	2.17	0.45
1:D:451:ALA:HB1	1:D:458:LEU:HD22	1.99	0.45
1:B:275:LYS:HE3	1:B:276:ASP:OD1	2.17	0.44
1:B:435:TRP:CD2	1:B:462:ILE:HD11	2.52	0.44
1:C:453:ARG:HG2	1:C:454:TYR:CZ	2.52	0.44
1:C:424:ASP:OD1	1:C:441:THR:OG1	2.34	0.44
1:C:107:ARG:NH1	1:C:136:GLU:OE2	2.47	0.43
1:C:491:GLN:NE2	7:C:712:HOH:O	2.45	0.43
1:D:84:SER:OG	1:D:356:GLN:HG2	2.17	0.43
1:B:205:ARG:NE	2:J:1:G4S:H5	2.34	0.43
1:C:430:MET:HE3	1:C:430:MET:HB3	1.79	0.43
1:D:56:ASP:O	1:D:255:PRO:HD2	2.19	0.43
1:B:470:LYS:HE3	1:B:470:LYS:HA	2.00	0.43
1:B:336:TRP:NE1	1:B:430:MET:HG3	2.33	0.42
1:B:56:ASP:O	1:B:255:PRO:HD2	2.19	0.42
1:B:104[B]:LEU:HD13	1:B:158:LEU:HD11	2.01	0.42
1:C:242:LYS:HE2	1:C:242:LYS:HB3	1.77	0.42
1:C:205:ARG:NE	2:K:1:G4S:H5	2.35	0.42
1:D:175:TYR:CD1	1:D:231:GLU:HG3	2.54	0.42
1:C:56:ASP:O	1:C:255:PRO:HD2	2.20	0.42
1:D:508:PRO:HG2	1:D:509:TYR:CE2	2.55	0.42
1:C:316:GLY:HA3	1:C:338:GLY:O	2.21	0.41
1:D:139:MET:HB3	1:D:368:TYR:CD2	2.56	0.41
1:A:368:TYR:HB3	1:A:369:PRO:CD	2.51	0.41
1:C:58:LEU:HB3	1:C:313:SER:HB2	2.02	0.41
1:B:175:TYR:CD1	1:B:231:GLU:HG3	2.56	0.41
1:B:293:LYS:HD3	1:B:293:LYS:HA	1.92	0.41
1:A:156:TRP:CE3	1:A:166:PRO:HG3	2.57	0.40
1:A:478:VAL:HB	1:A:483:ASP:HB3	2.03	0.40
1:A:484:LYS:HE3	1:A:488:GLU:OE1	2.21	0.40
1:A:311:PHE:O	1:A:344:MET:HA	2.21	0.40
1:A:484:LYS:O	1:A:488:GLU:HG3	2.21	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	463/480 (96%)	452 (98%)	11 (2%)	0	100	100
1	B	465/480 (97%)	451 (97%)	14 (3%)	0	100	100
1	C	461/480 (96%)	448 (97%)	13 (3%)	0	100	100
1	D	463/480 (96%)	449 (97%)	14 (3%)	0	100	100
All	All	1852/1920 (96%)	1800 (97%)	52 (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	390/406 (96%)	388 (100%)	2 (0%)	86	85
1	B	389/406 (96%)	386 (99%)	3 (1%)	79	74
1	C	378/406 (93%)	376 (100%)	2 (0%)	86	85
1	D	385/406 (95%)	382 (99%)	3 (1%)	79	74
All	All	1542/1624 (95%)	1532 (99%)	10 (1%)	84	80

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	319	PHE
1	A	350	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	139	MET
1	B	315	ASN
1	B	319	PHE
1	C	315	ASN
1	C	319	PHE
1	D	275	LYS
1	D	315	ASN
1	D	319	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	129	ASN
1	A	235	ASN
1	A	303	GLN
1	A	356	GLN
1	B	78	ASN
1	B	145	ASN
1	B	315	ASN
1	B	356	GLN
1	B	491	GLN
1	C	78	ASN
1	C	157	HIS
1	C	320	ASN
1	C	442	GLN
1	D	81	GLN
1	D	196	GLN
1	D	244	GLN
1	D	296	GLN
1	D	303	GLN
1	D	315	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)

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	OSE	B	96	1,3	8,9,10	0.96	1 (12%)	5,12,14	1.85	1 (20%)
1	OSE	C	96	1,3	8,9,10	0.87	0	5,12,14	1.39	1 (20%)
1	OSE	D	96	1,3	8,9,10	1.00	0	5,12,14	1.38	1 (20%)
1	OSE	A	96	1,3	8,9,10	0.87	0	5,12,14	1.65	1 (20%)

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	OSE	B	96	1,3	-	3/4/8/10	-
1	OSE	C	96	1,3	-	3/4/8/10	-
1	OSE	D	96	1,3	-	3/4/8/10	-
1	OSE	A	96	1,3	-	3/4/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	96	OSE	OG-CB	-2.04	1.38	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	96	OSE	OG-S-O2S	-2.99	97.81	106.88
1	A	96	OSE	OG-S-O2S	-2.73	98.62	106.88
1	D	96	OSE	O3S-S-O2S	2.32	116.57	108.49
1	C	96	OSE	O3S-S-O2S	2.18	116.08	108.49

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	96	OSE	CB-OG-S-O2S
1	A	96	OSE	CB-OG-S-O3S
1	B	96	OSE	CB-OG-S-O3S
1	A	96	OSE	CB-OG-S-O1S
1	B	96	OSE	CB-OG-S-O2S
1	B	96	OSE	CB-OG-S-O1S
1	C	96	OSE	CB-OG-S-O2S
1	D	96	OSE	CB-OG-S-O2S
1	C	96	OSE	CB-OG-S-O1S
1	D	96	OSE	CB-OG-S-O1S
1	C	96	OSE	CA-CB-OG-S
1	D	96	OSE	CA-CB-OG-S

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

16 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	G4S	I	1	2	16,16,16	0.57	0	19,24,24	0.79	0
2	9RN	I	2	2	11,11,12	0.38	0	15,16,18	0.97	0
2	GAL	I	3	2	11,11,12	0.77	1 (9%)	15,15,17	0.84	1 (6%)
2	9RN	I	4	2	11,11,12	0.38	0	15,16,18	1.12	2 (13%)
2	G4S	J	1	2	16,16,16	0.40	0	19,24,24	0.66	0
2	9RN	J	2	2	11,11,12	0.38	0	15,16,18	0.95	0
2	GAL	J	3	2	11,11,12	0.73	1 (9%)	15,15,17	0.70	0
2	9RN	J	4	2	11,11,12	0.38	0	15,16,18	1.02	0
2	G4S	K	1	2	16,16,16	0.42	0	19,24,24	0.60	0
2	9RN	K	2	2	11,11,12	0.39	0	15,16,18	0.90	0
2	GAL	K	3	2	11,11,12	0.75	1 (9%)	15,15,17	0.80	0
2	9RN	K	4	2	11,11,12	0.39	0	15,16,18	1.06	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	G4S	L	1	2	16,16,16	0.52	0	19,24,24	0.88	0
2	9RN	L	2	2	11,11,12	0.35	0	15,16,18	0.93	0
2	GAL	L	3	2	11,11,12	0.71	1 (9%)	15,15,17	0.80	1 (6%)
2	9RN	L	4	2	11,11,12	0.40	0	15,16,18	0.95	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	G4S	I	1	2	-	2/7/27/27	0/1/1/1
2	9RN	I	2	2	-	-	0/3/2/2
2	GAL	I	3	2	-	0/2/19/22	0/1/1/1
2	9RN	I	4	2	-	-	0/3/2/2
2	G4S	J	1	2	-	4/7/27/27	0/1/1/1
2	9RN	J	2	2	-	-	0/3/2/2
2	GAL	J	3	2	-	0/2/19/22	0/1/1/1
2	9RN	J	4	2	-	-	0/3/2/2
2	G4S	K	1	2	-	3/7/27/27	0/1/1/1
2	9RN	K	2	2	-	-	0/3/2/2
2	GAL	K	3	2	-	0/2/19/22	0/1/1/1
2	9RN	K	4	2	-	-	0/3/2/2
2	G4S	L	1	2	-	0/7/27/27	0/1/1/1
2	9RN	L	2	2	-	-	0/3/2/2
2	GAL	L	3	2	-	0/2/19/22	0/1/1/1
2	9RN	L	4	2	-	-	0/3/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	3	GAL	O4-C4	2.22	1.48	1.43
2	I	3	GAL	O4-C4	2.19	1.48	1.43
2	L	3	GAL	O4-C4	2.16	1.48	1.43
2	K	3	GAL	O4-C4	2.15	1.48	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	3	GAL	O4-C4-C3	-2.56	104.42	110.35
2	L	3	GAL	O4-C4-C3	-2.39	104.82	110.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	4	9RN	C1-O5-C5	2.34	115.36	112.19
2	I	4	9RN	C3-C4-C5	-2.15	97.32	101.99

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	1	G4S	C4-O4-S-O9
2	J	1	G4S	C4-O4-S-O7
2	J	1	G4S	C4-O4-S-O8
2	J	1	G4S	C4-O4-S-O9
2	K	1	G4S	C4-O4-S-O7
2	K	1	G4S	C4-O4-S-O8
2	K	1	G4S	C4-O4-S-O9
2	J	1	G4S	C4-C5-C6-O6
2	I	1	G4S	C4-O4-S-O8

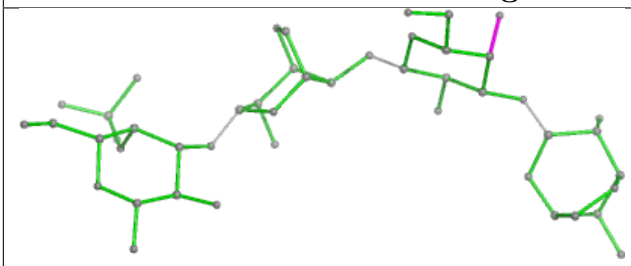
There are no ring outliers.

2 monomers are involved in 2 short contacts:

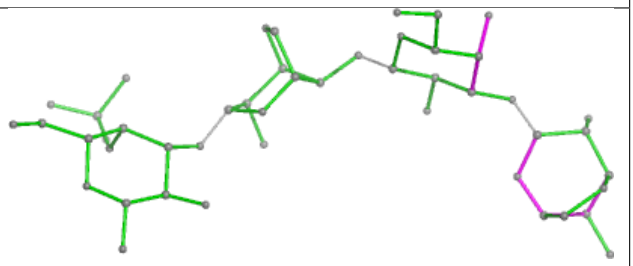
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	K	1	G4S	1	0
2	J	1	G4S	1	0

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

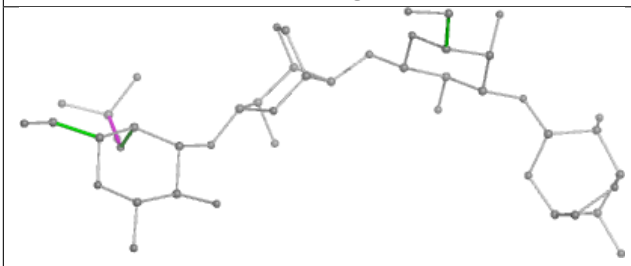
Oligosaccharide Chain I



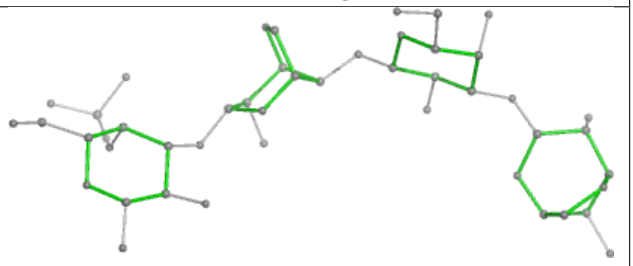
Bond lengths



Bond angles

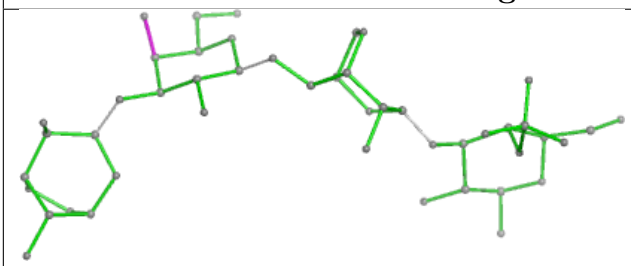


Torsions

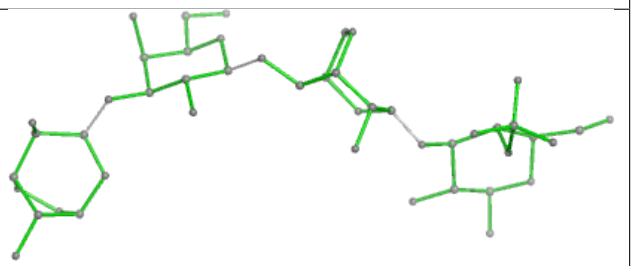


Rings

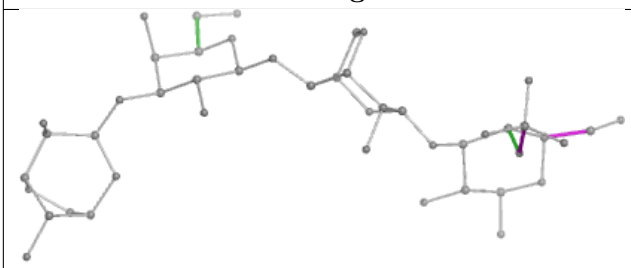
Oligosaccharide Chain J



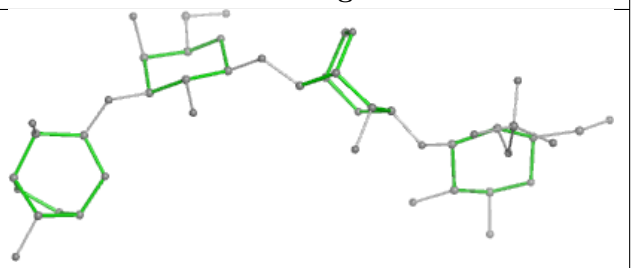
Bond lengths



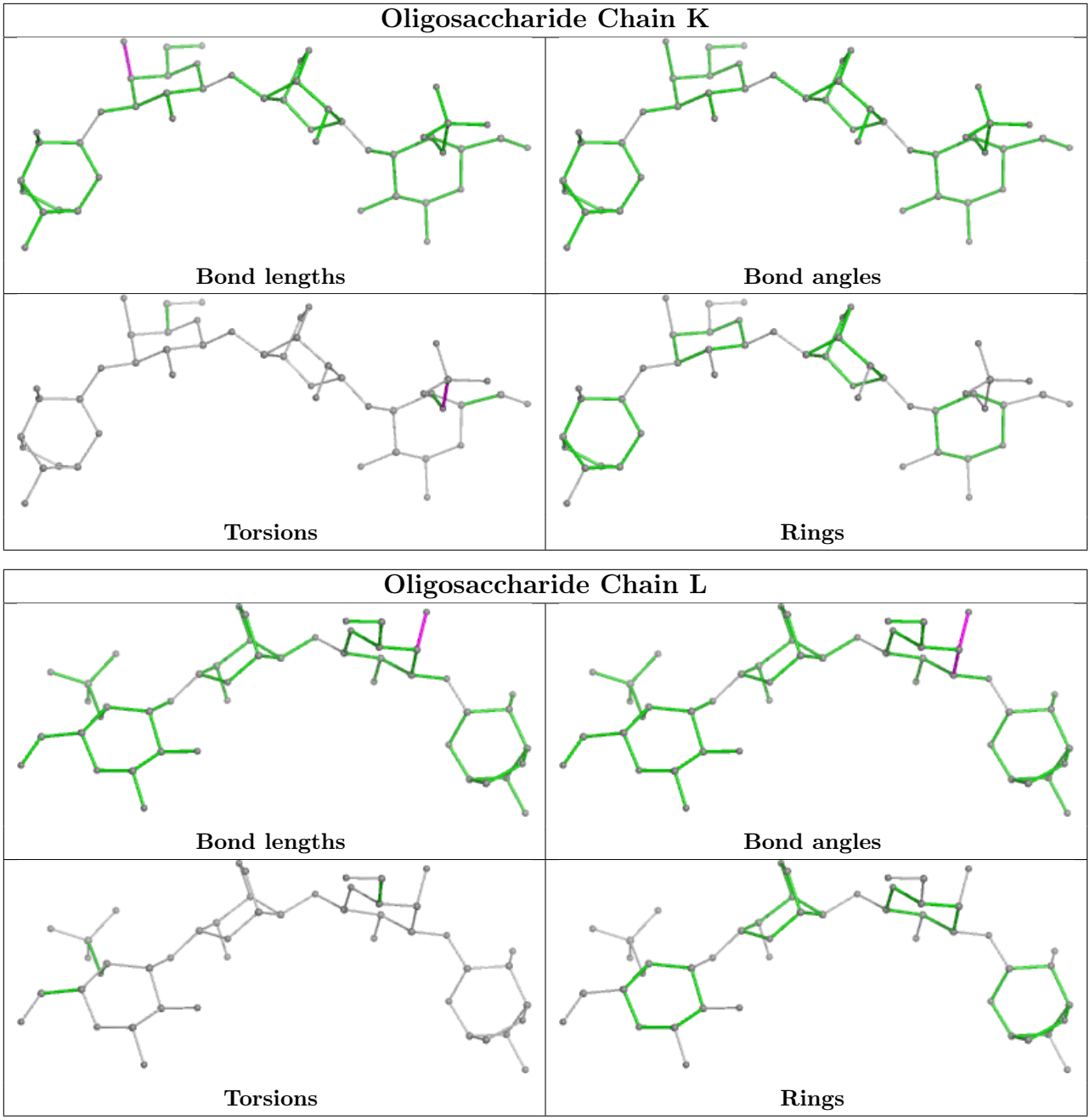
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

Of 44 ligands modelled in this entry, 37 are monoatomic - leaving 7 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	$\#\|Z\|>2$
6	GOL	D	609	-	5,5,5	0.96	0	5,5,5	0.91	0
6	GOL	B	610	-	5,5,5	0.84	0	5,5,5	1.02	0
6	GOL	B	609	-	5,5,5	0.71	0	5,5,5	1.61	1 (20%)
6	GOL	D	610	-	5,5,5	0.61	0	5,5,5	1.11	1 (20%)
6	GOL	A	614	-	5,5,5	0.82	0	5,5,5	1.12	1 (20%)
6	GOL	A	613	-	5,5,5	0.82	0	5,5,5	1.16	0
6	GOL	A	615	-	5,5,5	0.94	0	5,5,5	0.99	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	GOL	D	609	-	-	4/4/4/4	-
6	GOL	B	610	-	-	1/4/4/4	-
6	GOL	B	609	-	-	4/4/4/4	-
6	GOL	D	610	-	-	4/4/4/4	-
6	GOL	A	614	-	-	0/4/4/4	-
6	GOL	A	613	-	-	2/4/4/4	-
6	GOL	A	615	-	-	4/4/4/4	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	609	GOL	C3-C2-C1	-2.91	100.39	111.70
6	D	610	GOL	C3-C2-C1	-2.07	103.64	111.70
6	A	614	GOL	C3-C2-C1	-2.02	103.86	111.70

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	615	GOL	C1-C2-C3-O3
6	D	609	GOL	O1-C1-C2-C3
6	D	609	GOL	C1-C2-C3-O3
6	D	610	GOL	C1-C2-C3-O3
6	A	613	GOL	C1-C2-C3-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	A	615	GOL	O1-C1-C2-C3
6	B	609	GOL	O1-C1-C2-C3
6	B	609	GOL	C1-C2-C3-O3
6	D	610	GOL	O1-C1-C2-C3
6	A	613	GOL	O2-C2-C3-O3
6	A	615	GOL	O2-C2-C3-O3
6	D	609	GOL	O1-C1-C2-O2
6	D	609	GOL	O2-C2-C3-O3
6	D	610	GOL	O1-C1-C2-O2
6	B	609	GOL	O2-C2-C3-O3
6	D	610	GOL	O2-C2-C3-O3
6	A	615	GOL	O1-C1-C2-O2
6	B	609	GOL	O1-C1-C2-O2
6	B	610	GOL	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	609	GOL	1	0
6	A	613	GOL	1	0
6	A	615	GOL	2	0

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	464/480 (96%)	-0.40	1 (0%)	92	94	21, 31, 43, 61	1 (0%)
1	B	465/480 (96%)	-0.39	2 (0%)	89	92	17, 31, 44, 54	2 (0%)
1	C	463/480 (96%)	-0.06	0	100	100	27, 37, 50, 63	0
1	D	465/480 (96%)	-0.20	2 (0%)	89	92	26, 36, 49, 59	0
All	All	1857/1920 (96%)	-0.26	5 (0%)	90	93	17, 34, 47, 63	3 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	270	ALA	2.9
1	D	510	HIS	2.5
1	A	46	THR	2.3
1	B	510	HIS	2.2
1	B	270	ALA	2.1

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	OSE	D	96	10/11	0.89	0.13	30,41,44,45	0
1	OSE	C	96	10/11	0.90	0.14	32,41,46,46	0
1	OSE	A	96	10/11	0.91	0.13	24,35,40,42	0
1	OSE	B	96	10/11	0.91	0.13	27,36,40,44	0

6.3 Carbohydrates (i)

SUGAR-RSR INFOmissingINFO

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
6	GOL	A	613	6/6	0.82	0.14	37,42,45,47	0
6	GOL	D	609	6/6	0.82	0.13	47,55,59,62	0
6	GOL	A	614	6/6	0.85	0.11	45,48,51,55	0
6	GOL	A	615	6/6	0.86	0.10	38,42,48,51	0
6	GOL	B	610	6/6	0.87	0.10	38,40,49,52	0
6	GOL	B	609	6/6	0.87	0.12	37,39,41,44	0
5	CL	D	608	1/1	0.89	0.17	73,73,73,73	0
5	CL	C	609	1/1	0.90	0.11	70,70,70,70	0
3	CA	C	601	1/1	0.90	0.18	43,43,43,43	0
6	GOL	D	610	6/6	0.90	0.10	49,52,54,55	0
5	CL	C	605	1/1	0.91	0.11	55,55,55,55	0
5	CL	B	608	1/1	0.91	0.19	74,74,74,74	0
5	CL	B	605	1/1	0.93	0.10	48,48,48,48	0
5	CL	A	607	1/1	0.94	0.08	44,44,44,44	0
5	CL	A	612	1/1	0.94	0.15	61,61,61,61	0
5	CL	C	608	1/1	0.94	0.08	45,45,45,45	0
5	CL	A	605	1/1	0.94	0.13	70,70,70,70	0
3	CA	D	601	1/1	0.95	0.15	39,39,39,39	0
4	BR	D	603	1/1	0.95	0.10	78,78,78,78	0
3	CA	B	601	1/1	0.95	0.12	35,35,35,35	0
5	CL	A	609	1/1	0.96	0.06	42,42,42,42	0
4	BR	C	603	1/1	0.96	0.07	81,81,81,81	0
5	CL	C	606	1/1	0.96	0.07	41,41,41,41	0
5	CL	C	607	1/1	0.96	0.07	48,48,48,48	0
4	BR	B	602	1/1	0.96	0.07	73,73,73,73	0
5	CL	B	606	1/1	0.96	0.06	46,46,46,46	0
5	CL	B	607	1/1	0.96	0.06	39,39,39,39	0
5	CL	D	607	1/1	0.97	0.06	43,43,43,43	0
5	CL	D	605	1/1	0.97	0.05	40,40,40,40	0
5	CL	D	606	1/1	0.97	0.07	41,41,41,41	0
5	CL	A	608	1/1	0.98	0.05	36,36,36,36	0
5	CL	A	606	1/1	0.98	0.08	32,32,32,32	0

Continued on next page...

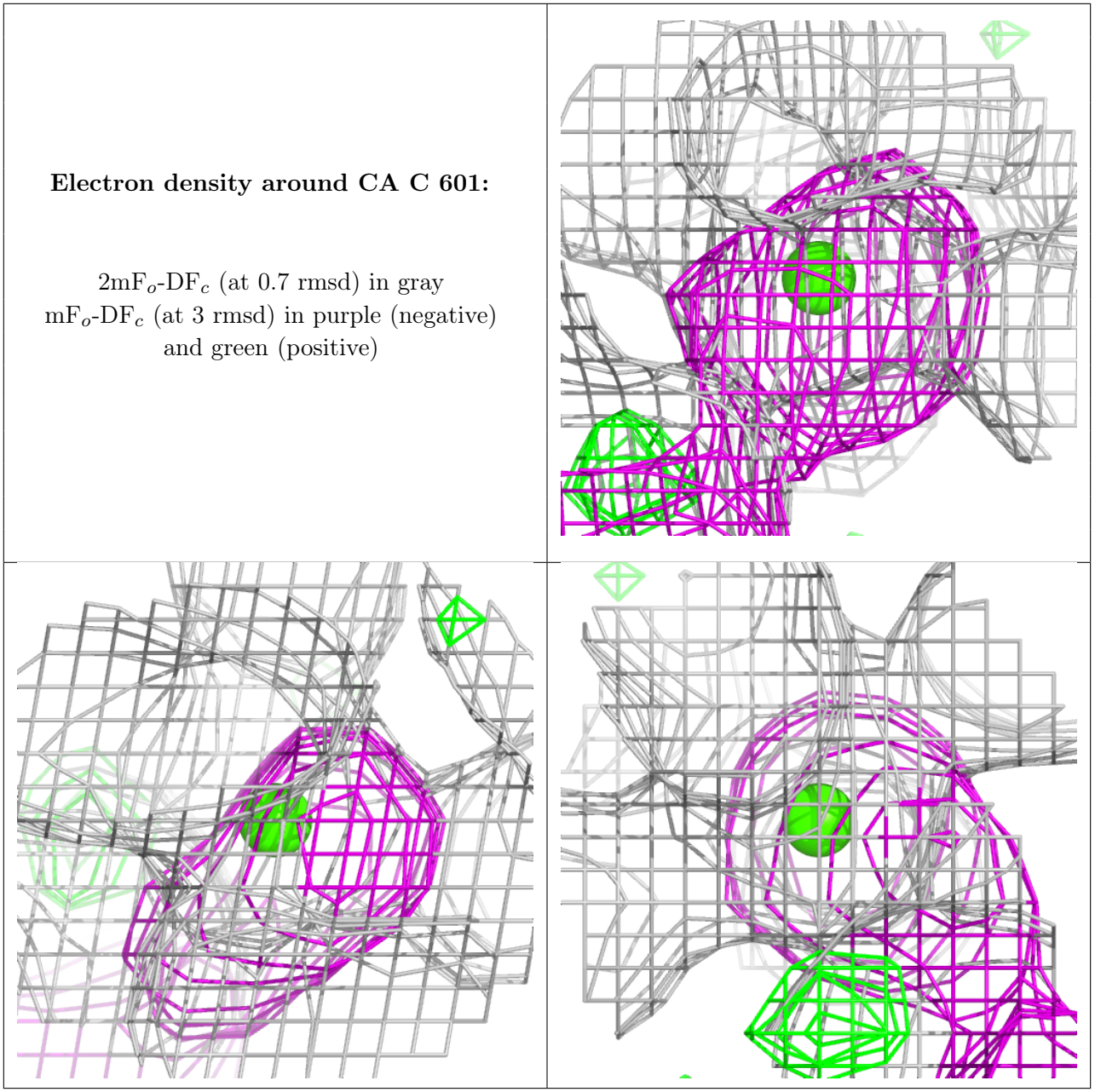
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CL	A	610	1/1	0.98	0.05	35,35,35,35	0
5	CL	A	611	1/1	0.98	0.04	38,38,38,38	0
4	BR	A	604	1/1	0.98	0.07	66,66,66,66	0
4	BR	A	602	1/1	0.99	0.03	37,37,37,37	1
4	BR	B	603	1/1	0.99	0.06	34,34,34,34	1
4	BR	C	602	1/1	0.99	0.04	36,36,36,36	1
4	BR	A	603	1/1	0.99	0.06	34,34,34,34	1
4	BR	D	602	1/1	0.99	0.04	37,37,37,37	1
3	CA	A	601	1/1	0.99	0.11	34,34,34,34	0
4	BR	D	604	1/1	0.99	0.05	36,36,36,36	1
4	BR	C	604	1/1	1.00	0.04	35,35,35,35	1
4	BR	B	604	1/1	1.00	0.03	32,32,32,32	1

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.

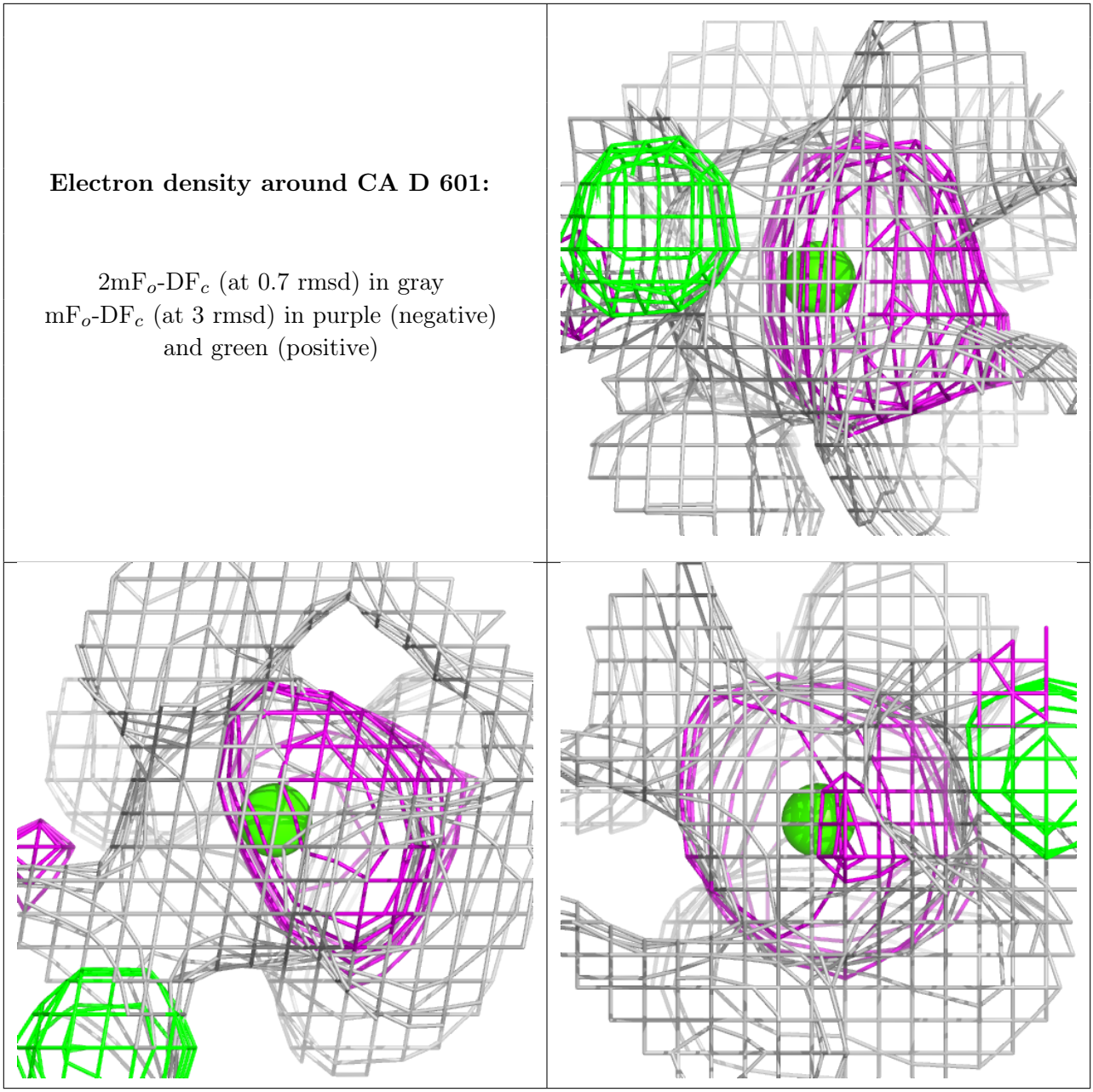
Electron density around CA C 601:

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



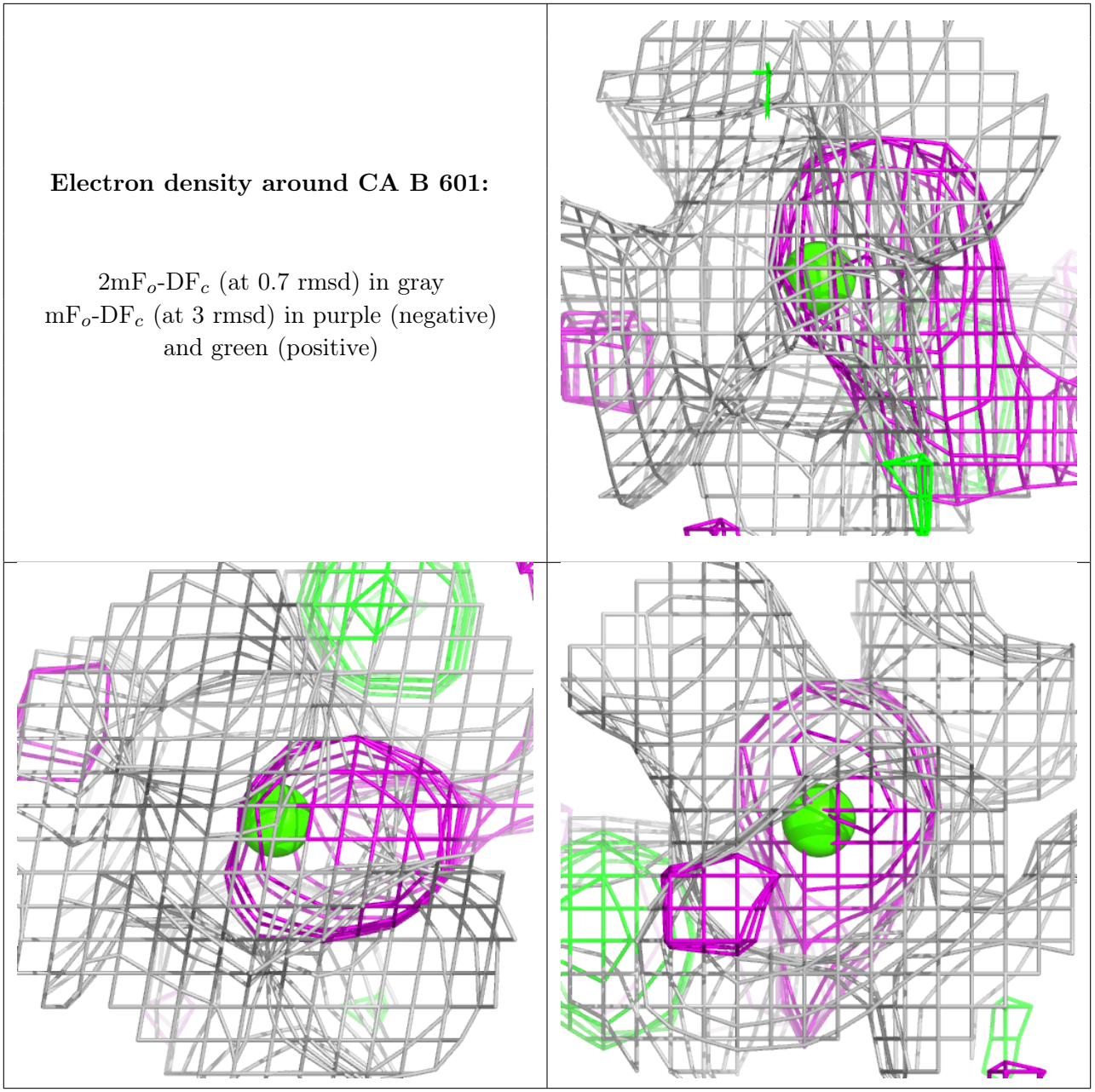
Electron density around CA D 601:

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



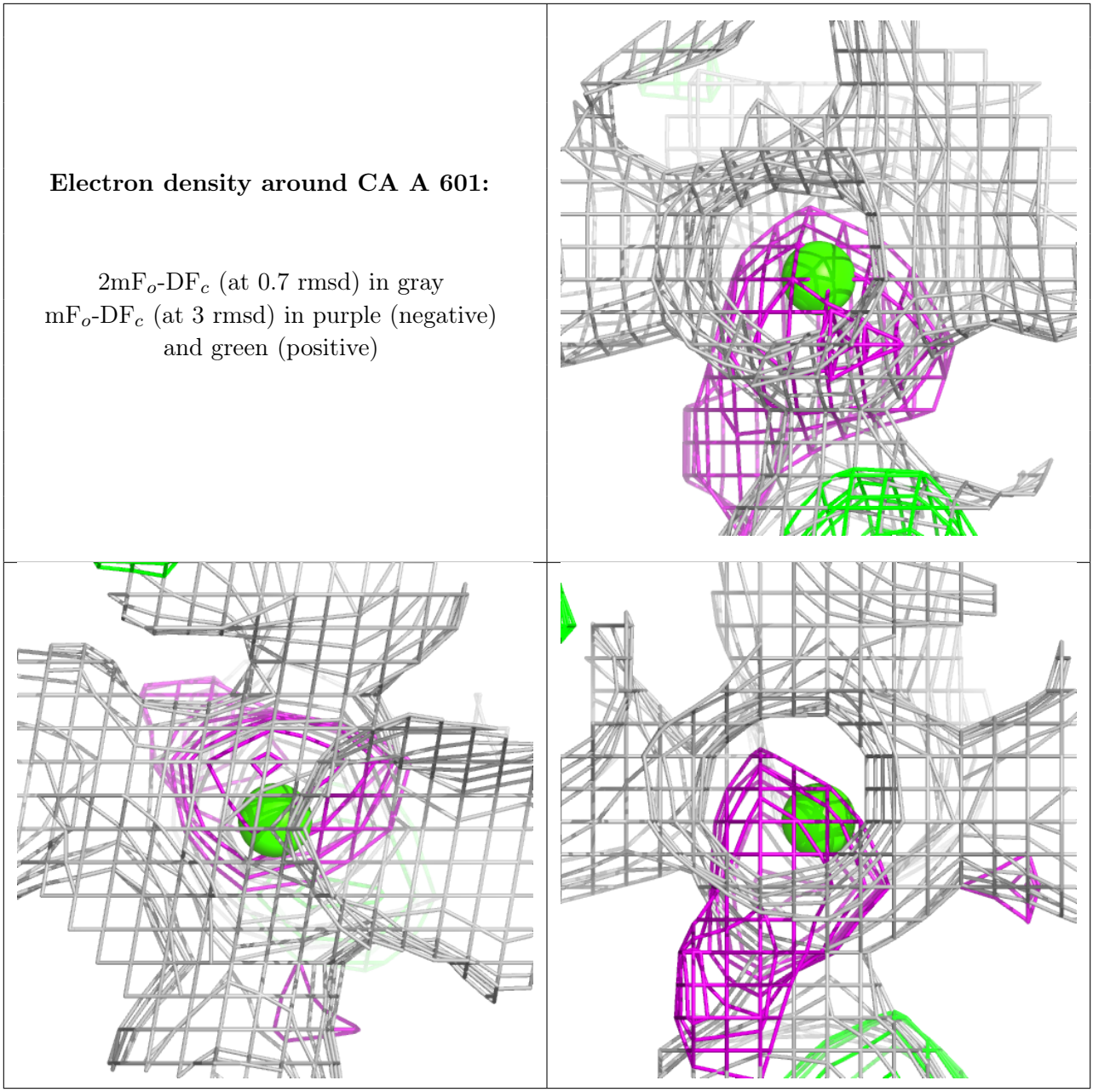
Electron density around CA B 601:

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



Electron density around CA A 601:

$2mF_o$-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.