



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 09:22 AM EDT

PDB ID : 8DCF  
Title : RNA ligase RtcB from *Pyrococcus horikoshii* in complex with Cu<sup>2+</sup> and GTP  
Authors : Jacewicz, A.; Dantuluri, S.; Shuman, S.  
Deposited on : 2022-06-16  
Resolution : 2.42 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](mailto: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 : 2.35.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

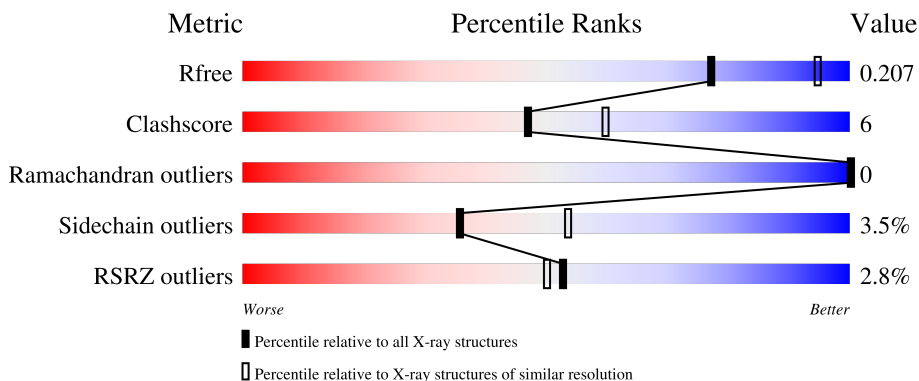
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


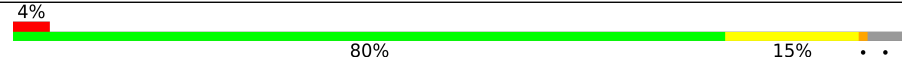
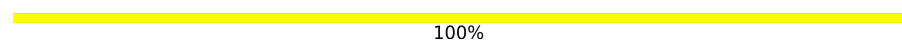
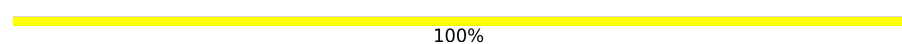
The reported resolution of this entry is 2.42 Å.

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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	501	 2% 80% 15%
1	B	501	 4% 80% 15%
2	C	2	 100%
2	D	2	 100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	GOL	B	502	-	-	X	-

## 2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 7866 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 tRNA-splicing ligase RtcB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	480	3766	2375	695	680	16	24	1	0
1	B	480	3759	2370	693	680	16	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

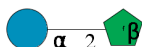
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP O59245
A	-18	GLY	-	expression tag	UNP O59245
A	-17	SER	-	expression tag	UNP O59245
A	-16	SER	-	expression tag	UNP O59245
A	-15	HIS	-	expression tag	UNP O59245
A	-14	HIS	-	expression tag	UNP O59245
A	-13	HIS	-	expression tag	UNP O59245
A	-12	HIS	-	expression tag	UNP O59245
A	-11	HIS	-	expression tag	UNP O59245
A	-10	HIS	-	expression tag	UNP O59245
A	-9	SER	-	expression tag	UNP O59245
A	-8	SER	-	expression tag	UNP O59245
A	-7	GLY	-	expression tag	UNP O59245
A	-6	LEU	-	expression tag	UNP O59245
A	-5	VAL	-	expression tag	UNP O59245
A	-4	PRO	-	expression tag	UNP O59245
A	-3	ARG	-	expression tag	UNP O59245
A	-2	GLY	-	expression tag	UNP O59245
A	-1	SER	-	expression tag	UNP O59245
A	0	HIS	-	expression tag	UNP O59245
B	-19	MET	-	initiating methionine	UNP O59245
B	-18	GLY	-	expression tag	UNP O59245
B	-17	SER	-	expression tag	UNP O59245
B	-16	SER	-	expression tag	UNP O59245
B	-15	HIS	-	expression tag	UNP O59245

*Continued on next page...*

Continued from previous page...

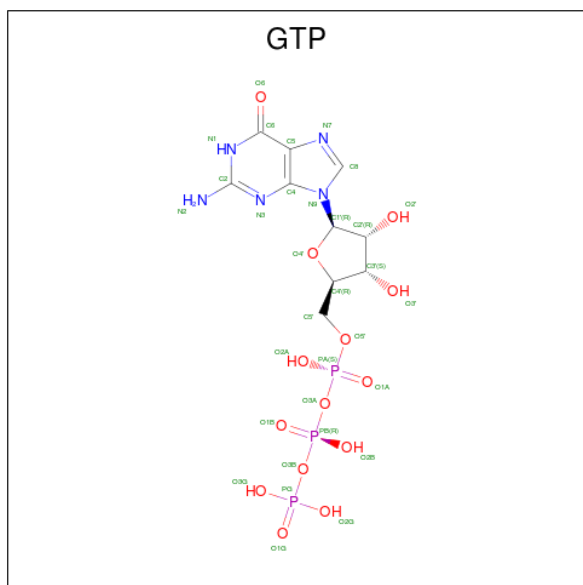
Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	expression tag	UNP O59245
B	-13	HIS	-	expression tag	UNP O59245
B	-12	HIS	-	expression tag	UNP O59245
B	-11	HIS	-	expression tag	UNP O59245
B	-10	HIS	-	expression tag	UNP O59245
B	-9	SER	-	expression tag	UNP O59245
B	-8	SER	-	expression tag	UNP O59245
B	-7	GLY	-	expression tag	UNP O59245
B	-6	LEU	-	expression tag	UNP O59245
B	-5	VAL	-	expression tag	UNP O59245
B	-4	PRO	-	expression tag	UNP O59245
B	-3	ARG	-	expression tag	UNP O59245
B	-2	GLY	-	expression tag	UNP O59245
B	-1	SER	-	expression tag	UNP O59245
B	0	HIS	-	expression tag	UNP O59245

- Molecule 2 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



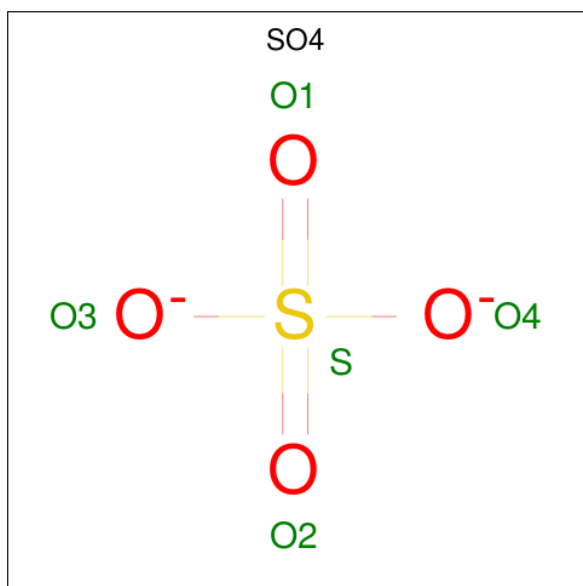
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	2	Total	C	O	0	0	0
			23	12	11			
2	D	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
3	A	1	32	10	5	14	3	1	0
3	B	1	32	10	5	14	3	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	1	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		

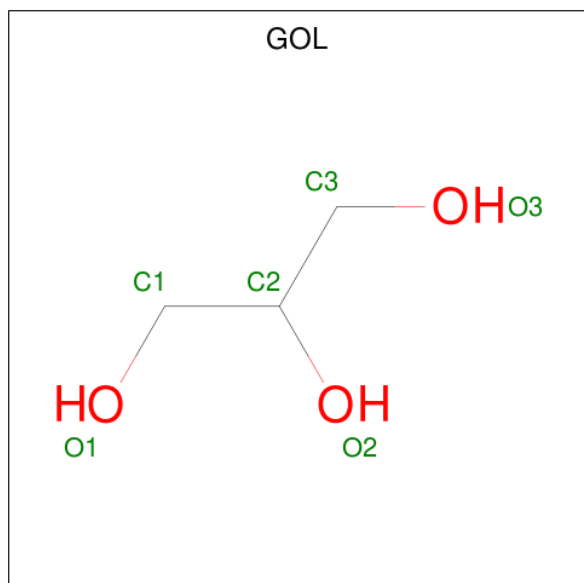
- Molecule 5 is COPPER (II) ION (three-letter code: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	3	Total	Cu	0	0
			3	3		
5	B	3	Total	Cu	0	0
			3	3		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cl 1 1	0	0
6	B	1	Total Cl 1 1	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 8 is water.

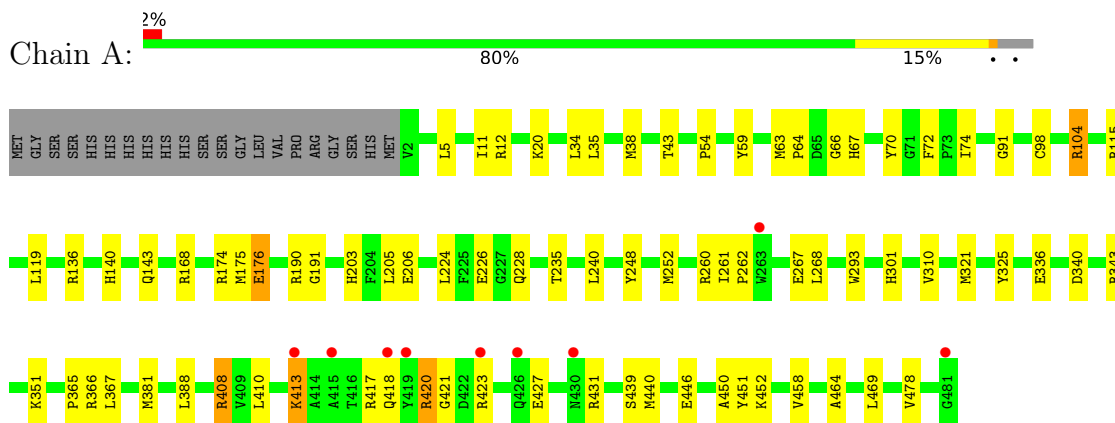
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	62	Total O 62 62	0	0
8	B	54	Total O 54 54	0	0



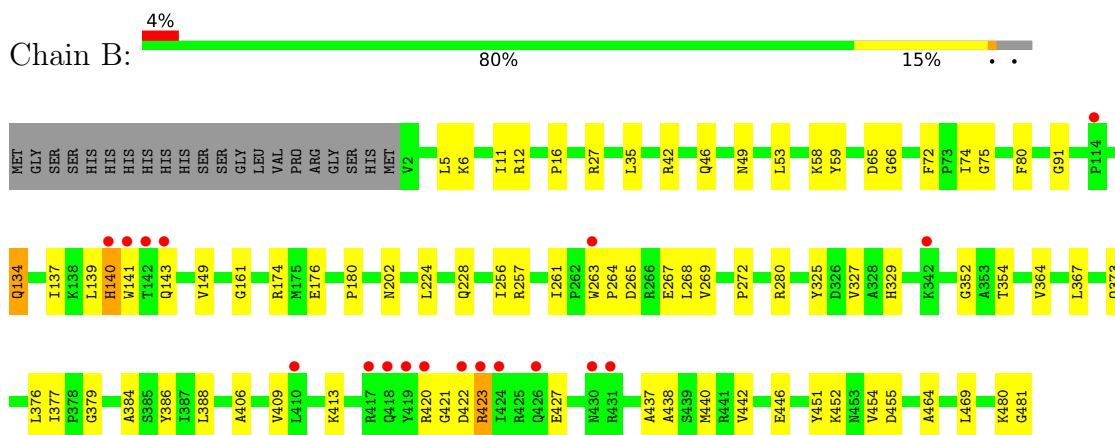
### 3 Residue-property plots [i](#)

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

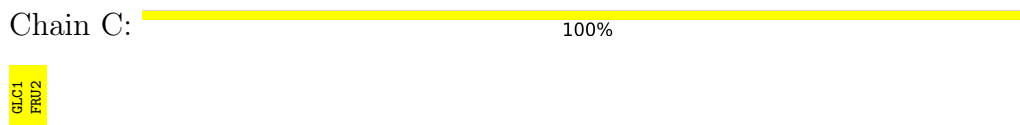
- Molecule 1: tRNA-splicing ligase RtcB



- Molecule 1: tRNA-splicing ligase RtcB



- Molecule 2: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



- Molecule 2: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



GLC1  
FRU2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.11Å 137.83Å 149.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.56 – 2.42 49.56 – 2.42	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.56-2.42) 99.9 (49.56-2.42)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 2.42Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.165 , 0.204 0.166 , 0.207	Depositor DCC
$R_{free}$ test set	3286 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.5	Xtrriage
Anisotropy	0.237	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 54.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7866	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FRU, SO4, CL, GLC, CU, GTP, GOL

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	0.39	0/3847	0.63	0/5190
1	B	0.42	0/3836	0.65	0/5175
All	All	0.40	0/7683	0.64	0/10365

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

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	3766	0	3789	46	0
1	B	3759	0	3782	49	0
2	C	23	0	21	0	0
2	D	23	0	21	1	0
3	A	32	0	12	2	0
3	B	32	0	12	2	0
4	A	50	0	0	0	0
4	B	45	0	0	0	0
5	A	3	0	0	1	0
5	B	3	0	0	0	0
6	A	1	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	1	0	0	1	0
7	B	12	0	16	6	0
8	A	62	0	0	2	0
8	B	54	0	0	0	0
All	All	7866	0	7653	95	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 6.

All (95) 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:190:ARG:HH21	7:B:502:GOL:H11	1.33	0.93
3:A:601:GTP:O1A	8:A:701:HOH:O	1.98	0.81
1:A:140:HIS:H	1:A:143:GLN:HE21	1.36	0.74
1:B:420:ARG:HH11	1:B:421:GLY:H	1.38	0.70
1:A:248:TYR:O	1:A:252:MET:HG3	1.94	0.67
1:B:140:HIS:CD2	1:B:143:GLN:HB2	2.32	0.65
1:A:98:CYS:SG	5:A:613:CU:CU	1.85	0.64
1:B:202:ASN:HB2	3:B:501:GTP:O2B	1.98	0.64
1:A:421:GLY:HA3	1:A:440:MET:HG3	1.79	0.64
1:A:408:ARG:NH1	1:A:410:LEU:O	2.24	0.62
1:A:174:ARG:HH12	1:A:176:GLU:HG2	1.64	0.62
1:A:423:ARG:O	1:A:427:GLU:HG3	1.99	0.62
1:B:174:ARG:HH12	1:B:176:GLU:HG2	1.66	0.60
1:B:388:LEU:HD23	1:B:469:LEU:HA	1.83	0.59
1:A:140:HIS:N	1:A:143:GLN:HE21	2.01	0.57
1:A:11:ILE:HD12	1:A:12:ARG:HG3	1.87	0.56
1:A:301:HIS:ND1	6:A:615:CL:CL	2.69	0.56
1:B:329:HIS:HB3	1:B:354:THR:HG22	1.88	0.56
1:B:423:ARG:O	1:B:427:GLU:HG3	2.07	0.54
1:A:11:ILE:HD13	1:A:464:ALA:HB2	1.90	0.54
1:B:58:LYS:HA	2:D:2:FRU:H11	1.89	0.54
1:B:174:ARG:NH1	1:B:176:GLU:HG2	2.24	0.53
1:A:417:ARG:O	1:A:418:GLN:HB3	2.08	0.53
1:B:388:LEU:CD2	1:B:469:LEU:HD23	2.39	0.53
1:A:365:PRO:HB2	1:A:367:LEU:HD23	1.92	0.52
1:B:6:LYS:HG3	1:B:16:PRO:HG3	1.92	0.52
1:B:442:VAL:O	1:B:446:GLU:HG2	2.09	0.52
1:A:261:ILE:HG22	1:A:262:PRO:O	2.11	0.51
1:B:264:PRO:HG2	1:B:268:LEU:HD12	1.93	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:149:VAL:HG23	1:B:180:PRO:HG3	1.93	0.50
1:B:11:ILE:HD13	1:B:464:ALA:HB2	1.93	0.50
1:A:63:MET:HB3	1:A:64:PRO:HD2	1.94	0.49
1:A:336:GLU:HG2	1:A:343:ARG:NE	2.27	0.49
1:A:54:PRO:HD3	1:A:262:PRO:HD2	1.95	0.49
1:A:104:ARG:O	1:A:321:MET:HA	2.12	0.49
1:B:11:ILE:HD12	1:B:12:ARG:HG3	1.95	0.49
1:B:72:PHE:CG	1:B:91:GLY:HA2	2.48	0.48
1:B:261:ILE:HD11	1:B:272:PRO:HG3	1.95	0.48
1:A:66:GLY:HA2	1:A:74:ILE:O	2.13	0.47
1:A:140:HIS:H	1:A:143:GLN:NE2	2.10	0.47
1:B:440:MET:H	7:B:502:GOL:H32	1.78	0.47
1:A:413:LYS:HB3	1:A:413:LYS:HE3	1.72	0.47
1:B:438:ALA:HB3	1:B:481:GLY:HA2	1.97	0.47
1:A:5:LEU:HD13	1:A:35:LEU:HD21	1.97	0.47
1:A:388:LEU:CD2	1:A:469:LEU:HD23	2.44	0.47
1:A:410:LEU:HG	1:A:450:ALA:HA	1.96	0.47
1:A:190:ARG:NH2	7:B:502:GOL:H11	2.14	0.46
1:B:27:ARG:HH12	7:B:503:GOL:H31	1.81	0.46
1:A:268:LEU:HD23	1:A:268:LEU:HA	1.77	0.46
1:B:27:ARG:HH22	7:B:503:GOL:H11	1.80	0.46
1:B:364:VAL:HG22	1:B:373:GLN:HB2	1.97	0.46
1:A:446:GLU:OE1	1:A:451:TYR:OH	2.32	0.46
1:B:161:GLY:HA2	6:B:516:CL:CL	2.52	0.46
1:B:437:ALA:HA	1:B:480:LYS:O	2.15	0.46
1:A:381:MET:HB3	3:A:601:GTP:O6	2.16	0.46
1:B:66:GLY:HA2	1:B:74:ILE:O	2.14	0.46
1:B:451:TYR:CE1	3:B:501:GTP:H2'	2.50	0.46
1:A:38:MET:HG2	1:A:43:THR:HG21	1.98	0.45
1:B:265:ASP:OD1	1:B:267:GLU:HG3	2.16	0.45
1:A:224:LEU:HA	1:A:228:GLN:OE1	2.17	0.45
1:B:49:ASN:HB3	1:B:268:LEU:CD1	2.47	0.45
1:B:75:GLY:HA2	1:B:352:GLY:O	2.16	0.45
1:A:34:LEU:HD22	1:A:452:LYS:HD3	2.00	0.44
1:B:379:GLY:HA3	1:B:384:ALA:O	2.17	0.44
1:B:446:GLU:OE1	1:B:451:TYR:OH	2.34	0.44
1:B:388:LEU:HD21	1:B:469:LEU:HD23	2.00	0.44
1:A:458:VAL:CG1	1:A:469:LEU:HD21	2.48	0.44
1:A:203:HIS:CD2	1:A:235:THR:HA	2.52	0.43
1:B:386:TYR:OH	1:B:455:ASP:OD1	2.17	0.43
1:A:67:HIS:HD2	8:A:743:HOH:O	2.01	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:191:GLY:HA3	1:A:240:LEU:HB2	2.01	0.43
1:A:119:LEU:HA	1:A:310:VAL:HG21	2.00	0.43
1:A:136:ARG:NE	1:B:134:GLN:OE1	2.52	0.43
1:A:206:GLU:HG3	1:A:478:VAL:HG22	2.01	0.42
1:A:70:TYR:CZ	1:A:267:GLU:HB3	2.55	0.42
1:A:70:TYR:O	1:A:351:LYS:NZ	2.53	0.42
1:B:406:ALA:O	1:B:452:LYS:HD2	2.20	0.42
1:A:72:PHE:CG	1:A:91:GLY:HA2	2.54	0.42
1:B:256:ILE:HG23	1:B:257:ARG:H	1.83	0.42
1:B:376:LEU:O	1:B:377:ILE:HD13	2.20	0.42
1:A:175:MET:HG3	1:A:293:TRP:HZ3	1.85	0.41
1:A:420:ARG:HA	1:A:420:ARG:CZ	2.50	0.41
1:B:139:LEU:HD11	1:B:143:GLN:OE1	2.20	0.41
1:B:224:LEU:HA	1:B:228:GLN:OE1	2.21	0.41
1:B:325:TYR:HD2	1:B:327:VAL:HG13	1.85	0.41
1:B:53:LEU:HB2	1:B:80:PHE:HE2	1.86	0.41
1:B:139:LEU:HA	1:B:139:LEU:HD12	1.65	0.41
1:B:440:MET:N	7:B:502:GOL:H32	2.36	0.41
1:A:427:GLU:O	1:A:431:ARG:HG3	2.20	0.41
1:B:5:LEU:HD13	1:B:35:LEU:HD21	2.03	0.41
1:B:406:ALA:HB2	1:B:454:VAL:HB	2.03	0.40
1:A:168:ARG:O	1:A:325:TYR:HA	2.21	0.40
1:B:49:ASN:HB3	1:B:268:LEU:HD11	2.02	0.40
1:B:42:ARG:O	1:B:46:GLN:HG3	2.21	0.40
1:B:65:ASP:OD1	1:B:65:ASP:N	2.55	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	479/501 (96%)	464 (97%)	15 (3%)	0	<a href="#">100</a> <a href="#">100</a>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	478/501 (95%)	462 (97%)	16 (3%)	0	100	100
All	All	957/1002 (96%)	926 (97%)	31 (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	390/407 (96%)	376 (96%)	14 (4%)	35	52
1	B	389/407 (96%)	376 (97%)	13 (3%)	38	56
All	All	779/814 (96%)	752 (96%)	27 (4%)	36	53

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	LYS
1	A	59	TYR
1	A	104	ARG
1	A	115	ARG
1	A	176	GLU
1	A	205	LEU
1	A	226	GLU
1	A	260	ARG
1	A	340	ASP
1	A	366	ARG
1	A	408	ARG
1	A	413	LYS
1	A	420	ARG
1	A	439	SER
1	B	59	TYR
1	B	134	GLN
1	B	137	ILE
1	B	140	HIS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	141	TRP
1	B	263	TRP
1	B	269	VAL
1	B	280	ARG
1	B	367	LEU
1	B	409	VAL
1	B	413	LYS
1	B	422	ASP
1	B	423	ARG

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

Mol	Chain	Res	Type
1	A	143	GLN
1	A	426	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](#)

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

## 5.5 Carbohydrates [i](#)

4 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	GLC	C	1	2	11,11,12	0.56	0	15,15,17	1.18	2 (13%)
2	FRU	C	2	2	11,12,12	0.50	0	10,18,18	0.83	1 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	D	1	2	11,11,12	0.56	0	15,15,17	1.39	2 (13%)
2	FRU	D	2	2	11,12,12	0.79	0	10,18,18	0.73	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	GLC	C	1	2	-	2/2/19/22	0/1/1/1
2	FRU	C	2	2	-	3/5/24/24	0/1/1/1
2	GLC	D	1	2	-	2/2/19/22	0/1/1/1
2	FRU	D	2	2	-	5/5/24/24	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1	GLC	C1-O5-C5	3.18	116.50	112.19
2	D	1	GLC	O5-C5-C6	2.91	111.77	107.20
2	C	1	GLC	O5-C5-C6	2.49	111.11	107.20
2	C	1	GLC	C1-O5-C5	2.17	115.14	112.19
2	C	2	FRU	O1-C1-C2	-2.13	107.34	111.86

There are no chirality outliers.

All (12) torsion outliers are listed below:

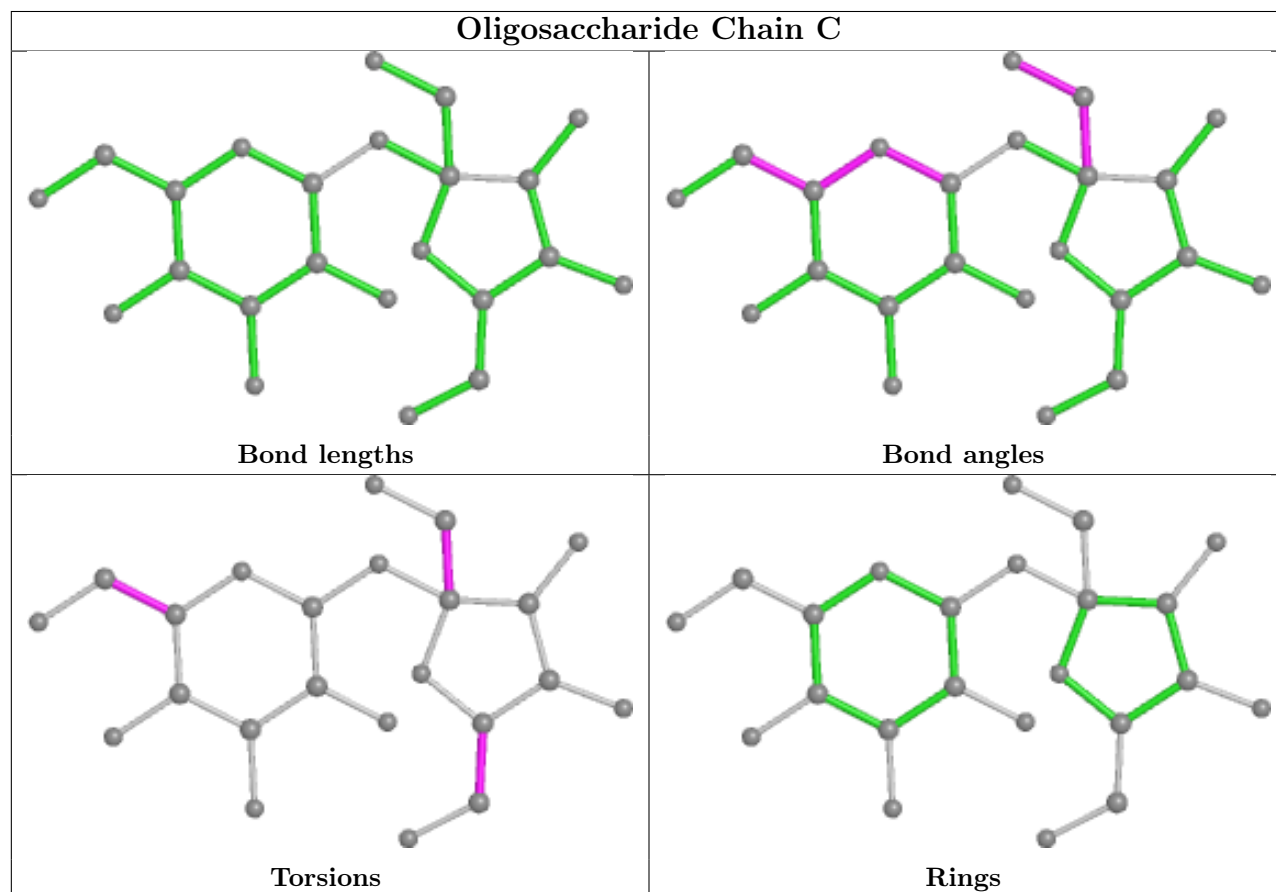
Mol	Chain	Res	Type	Atoms
2	D	2	FRU	O1-C1-C2-C3
2	D	2	FRU	O1-C1-C2-O2
2	D	2	FRU	O1-C1-C2-O5
2	D	2	FRU	C4-C5-C6-O6
2	C	1	GLC	O5-C5-C6-O6
2	D	1	GLC	O5-C5-C6-O6
2	C	1	GLC	C4-C5-C6-O6
2	D	1	GLC	C4-C5-C6-O6
2	D	2	FRU	O5-C5-C6-O6
2	C	2	FRU	O5-C5-C6-O6
2	C	2	FRU	C4-C5-C6-O6
2	C	2	FRU	O1-C1-C2-O2

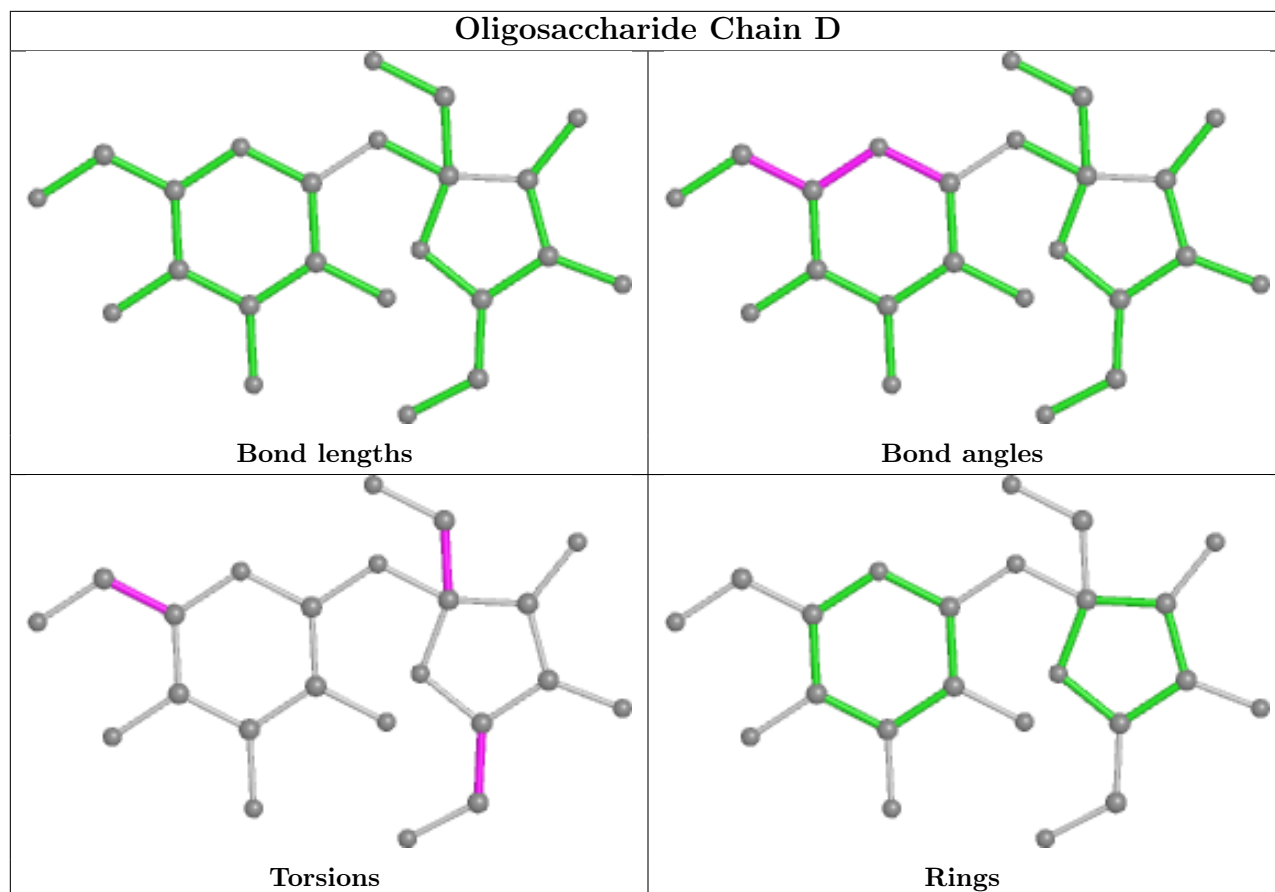
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	2	FRU	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.





## 5.6 Ligand geometry [i](#)

Of 31 ligands modelled in this entry, 8 are monoatomic - leaving 23 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	$\# Z  > 2$
4	SO4	A	610	-	4,4,4	0.13	0	6,6,6	0.14	0
7	GOL	B	502	-	5,5,5	0.82	0	5,5,5	1.09	0
4	SO4	B	507	-	4,4,4	0.16	0	6,6,6	0.47	0
7	GOL	B	503	-	5,5,5	0.83	0	5,5,5	0.93	0
4	SO4	A	607	-	4,4,4	0.13	0	6,6,6	0.23	0
4	SO4	A	605	-	4,4,4	0.14	0	6,6,6	0.23	0
4	SO4	A	611	-	4,4,4	0.13	0	6,6,6	0.13	0
4	SO4	A	602	-	4,4,4	0.16	0	6,6,6	0.34	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	B	512	-	4,4,4	0.15	0	6,6,6	0.18	0
4	SO4	B	509	-	4,4,4	0.14	0	6,6,6	0.10	0
4	SO4	A	606	-	4,4,4	0.08	0	6,6,6	0.37	0
3	GTP	A	601	5	26,34,34	1.17	2 (7%)	32,54,54	1.63	5 (15%)
4	SO4	B	510	-	4,4,4	0.11	0	6,6,6	0.27	0
4	SO4	B	506	-	4,4,4	0.18	0	6,6,6	0.40	0
4	SO4	B	505	-	4,4,4	0.19	0	6,6,6	0.35	0
4	SO4	A	604	-	4,4,4	0.12	0	6,6,6	0.37	0
3	GTP	B	501	5	26,34,34	1.12	2 (7%)	32,54,54	1.45	6 (18%)
4	SO4	A	608	-	4,4,4	0.15	0	6,6,6	0.15	0
4	SO4	B	511	-	4,4,4	0.15	0	6,6,6	0.47	0
4	SO4	A	609	-	4,4,4	0.15	0	6,6,6	0.11	0
4	SO4	B	508	-	4,4,4	0.16	0	6,6,6	0.18	0
4	SO4	B	504	-	4,4,4	0.13	0	6,6,6	0.09	0
4	SO4	A	603	-	4,4,4	0.20	0	6,6,6	0.41	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GTP	B	501	5	-	5/18/38/38	0/3/3/3
7	GOL	B	503	-	-	0/4/4/4	-
7	GOL	B	502	-	-	1/4/4/4	-
3	GTP	A	601	5	-	2/18/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	GTP	C5-C6	-4.02	1.39	1.47
3	B	501	GTP	C5-C6	-3.78	1.39	1.47
3	B	501	GTP	C2-N3	2.36	1.38	1.33
3	A	601	GTP	C2-N3	2.31	1.38	1.33

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	GTP	PB-O3B-PG	-4.33	117.97	132.83
3	B	501	GTP	PB-O3B-PG	-4.05	118.93	132.83
3	A	601	GTP	C5-C6-N1	3.60	120.31	113.95

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	GTP	PA-O3A-PB	-3.33	121.41	132.83
3	B	501	GTP	C8-N7-C5	3.10	108.89	102.99
3	B	501	GTP	C5-C6-N1	3.05	119.34	113.95
3	A	601	GTP	C8-N7-C5	2.90	108.51	102.99
3	A	601	GTP	C2-N1-C6	-2.71	120.10	125.10
3	B	501	GTP	C3'-C2'-C1'	2.59	104.88	100.98
3	B	501	GTP	C2-N1-C6	-2.56	120.39	125.10
3	B	501	GTP	PA-O3A-PB	-2.11	125.59	132.83

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	501	GTP	C5'-O5'-PA-O1A
7	B	502	GOL	O2-C2-C3-O3
3	A	601	GTP	PB-O3A-PA-O2A
3	B	501	GTP	PB-O3A-PA-O1A
3	B	501	GTP	O4'-C4'-C5'-O5'
3	B	501	GTP	C3'-C4'-C5'-O5'
3	B	501	GTP	C5'-O5'-PA-O3A
3	A	601	GTP	C5'-O5'-PA-O1A

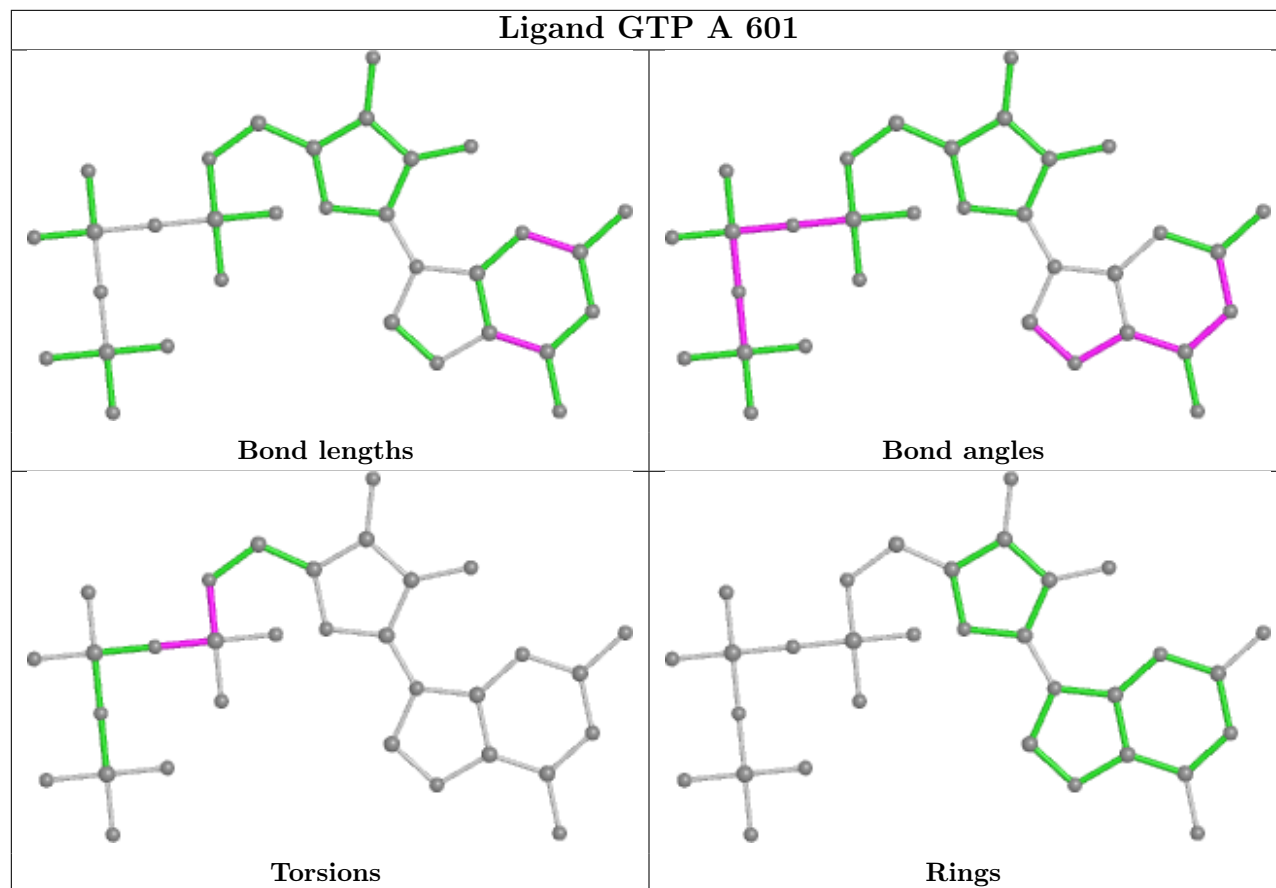
There are no ring outliers.

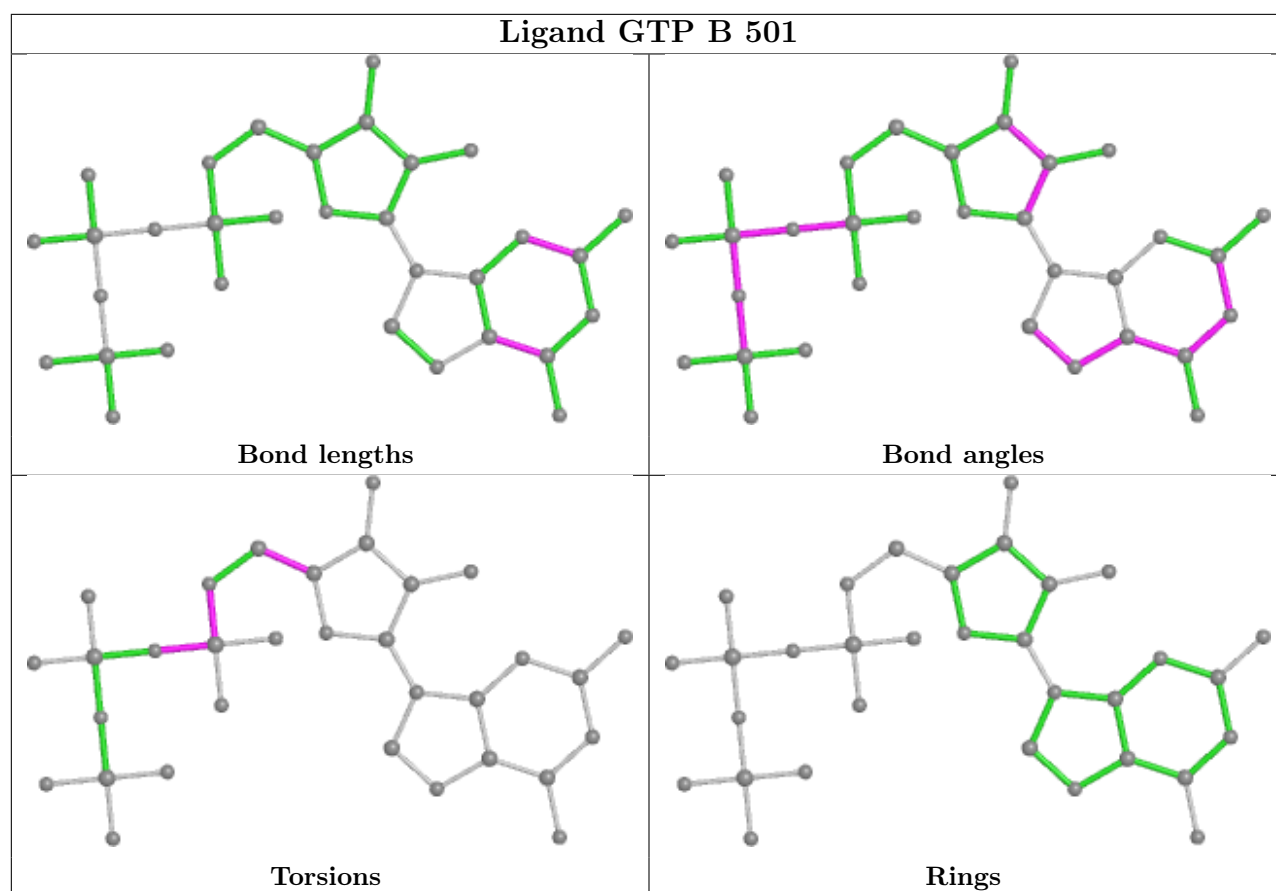
4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	502	GOL	4	0
7	B	503	GOL	2	0
3	A	601	GTP	2	0
3	B	501	GTP	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

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	480/501 (95%)	-0.13	9 (1%) 66 64	38, 52, 85, 118	110 (22%)
1	B	480/501 (95%)	-0.04	18 (3%) 40 38	41, 55, 92, 142	1 (0%)
All	All	960/1002 (95%)	-0.08	27 (2%) 53 50	38, 54, 90, 142	111 (11%)

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	263	TRP	5.2
1	B	141	TRP	4.7
1	B	140	HIS	4.5
1	B	417	ARG	4.2
1	B	423	ARG	4.1
1	A	419	TYR	3.7
1	B	418	GLN	3.6
1	B	143	GLN	3.5
1	B	419	TYR	3.2
1	A	430	ASN	3.2
1	A	423	ARG	3.1
1	B	420	ARG	3.1
1	A	481	GLY	3.0
1	A	418	GLN	3.0
1	B	342	LYS	3.0
1	B	410	LEU	2.8
1	A	415	ALA	2.7
1	B	430	ASN	2.7
1	A	263	TRP	2.6
1	A	426	GLN	2.6
1	B	424	ILE	2.4
1	A	413	LYS	2.2
1	B	431	ARG	2.1
1	B	422	ASP	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	114	PRO	2.0
1	B	142	THR	2.0
1	B	426	GLN	2.0

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

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

## 6.3 Carbohydrates [i](#)

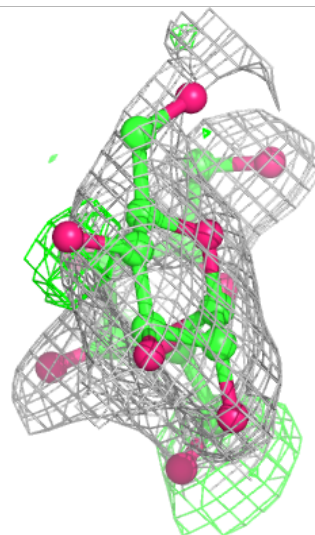
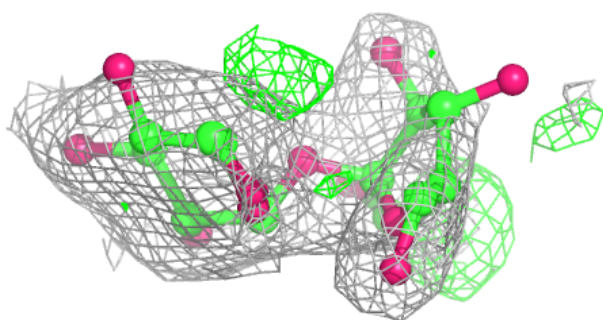
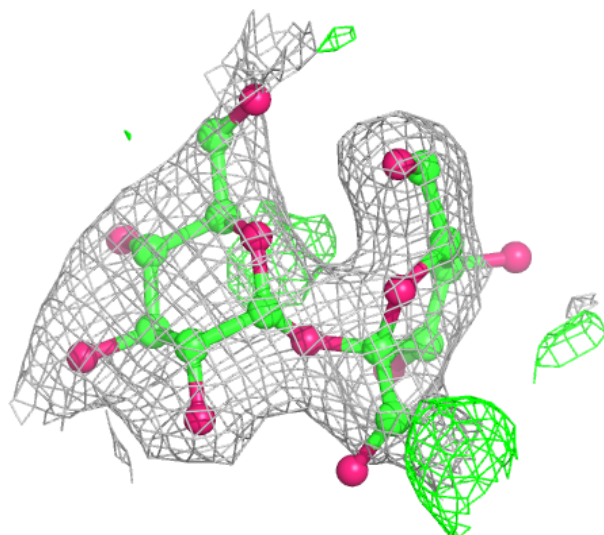
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

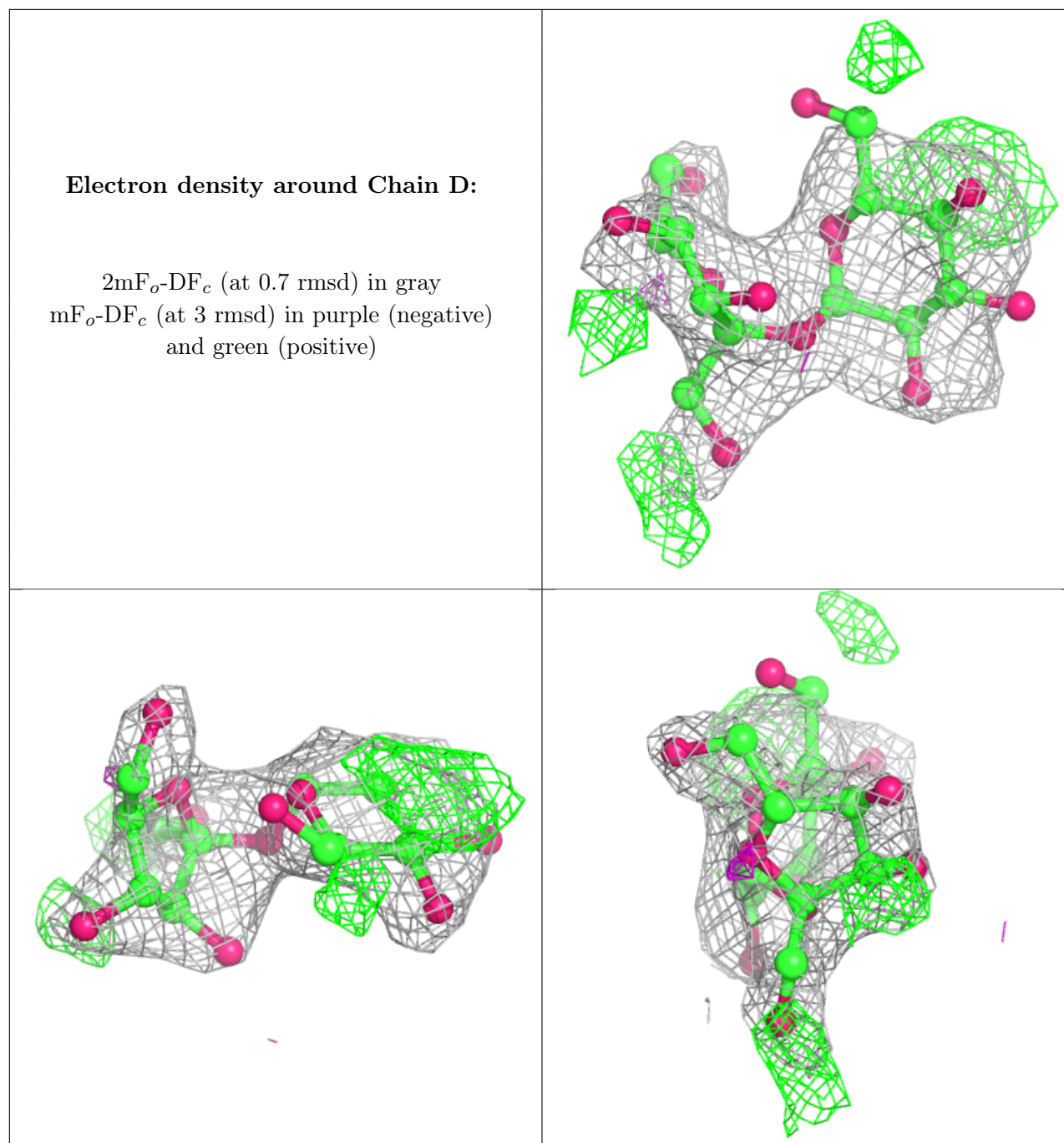
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	FRU	D	2	12/12	0.82	0.40	62,77,81,85	12
2	GLC	D	1	11/12	0.83	0.28	69,78,85,88	11
2	FRU	C	2	12/12	0.84	0.30	54,78,82,88	12
2	GLC	C	1	11/12	0.88	0.21	61,77,81,86	11

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain C:**

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





## 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<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
4	SO4	B	512	5/5	0.54	0.36	125,130,138,143	5

*Continued on next page...*

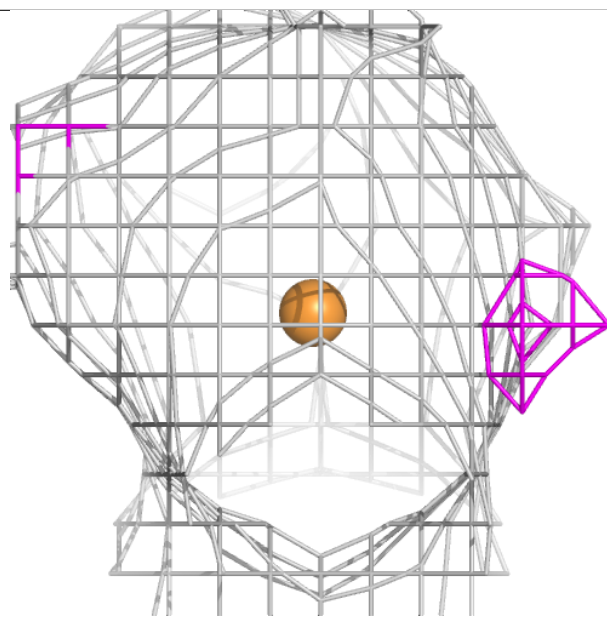
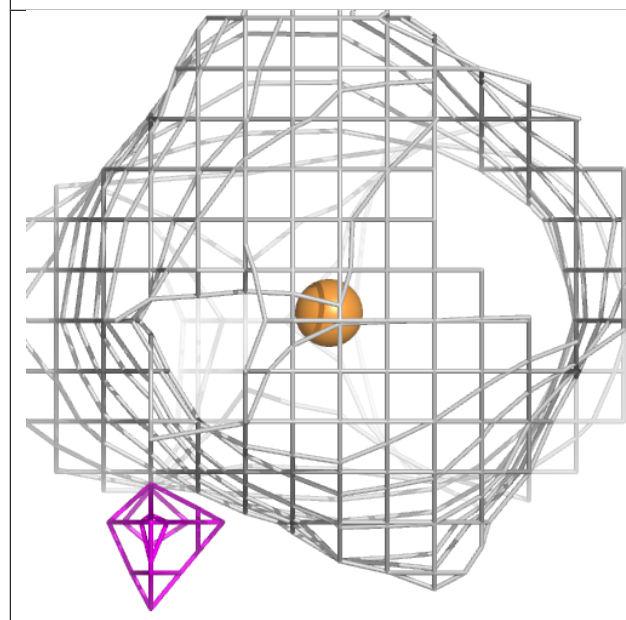
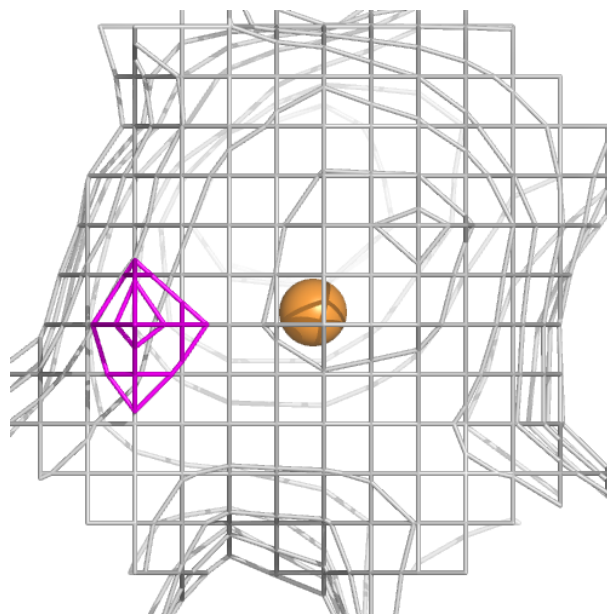
*Continued from previous page...*

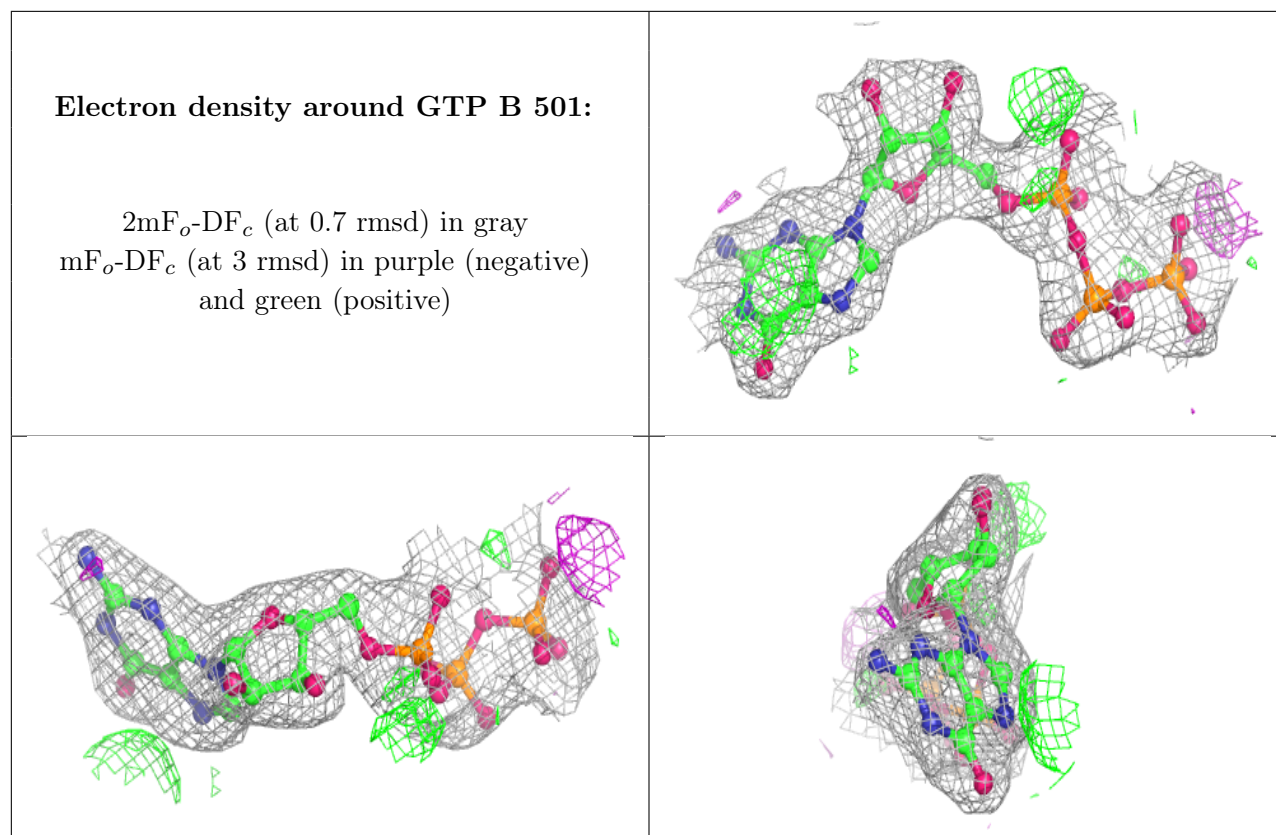
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	B	511	5/5	0.73	0.23	59,70,88,101	5
7	GOL	B	503	6/6	0.74	0.19	69,85,98,103	0
4	SO4	A	608	5/5	0.76	0.22	100,102,114,115	2
4	SO4	A	604	5/5	0.79	0.22	65,75,98,101	5
5	CU	B	515	1/1	0.89	0.06	65,65,65,65	1
4	SO4	B	508	5/5	0.90	0.26	81,91,98,101	1
7	GOL	B	502	6/6	0.91	0.18	66,68,76,83	0
4	SO4	B	510	5/5	0.91	0.12	60,69,72,83	2
4	SO4	A	609	5/5	0.92	0.13	61,62,74,75	4
4	SO4	B	506	5/5	0.93	0.15	60,63,67,69	2
3	GTP	B	501	32/32	0.93	0.24	51,61,75,78	22
4	SO4	A	607	5/5	0.93	0.16	69,72,75,96	1
4	SO4	A	611	5/5	0.93	0.17	74,90,94,102	1
5	CU	A	612	1/1	0.94	0.04	70,70,70,70	1
4	SO4	A	610	5/5	0.94	0.17	79,83,92,95	2
6	CL	A	615	1/1	0.94	0.12	63,63,63,63	1
4	SO4	A	605	5/5	0.94	0.13	64,76,84,88	2
4	SO4	B	509	5/5	0.94	0.20	72,79,83,87	2
3	GTP	A	601	32/32	0.95	0.28	48,56,65,79	28
4	SO4	B	507	5/5	0.95	0.37	43,51,57,69	3
4	SO4	A	602	5/5	0.97	0.30	44,46,58,59	4
6	CL	B	516	1/1	0.97	0.14	55,55,55,55	1
5	CU	B	514	1/1	0.98	0.15	65,65,65,65	1
4	SO4	A	603	5/5	0.98	0.20	43,47,51,53	4
4	SO4	B	505	5/5	0.99	0.12	64,67,84,88	1
4	SO4	A	606	5/5	0.99	0.12	55,57,64,68	2
4	SO4	B	504	5/5	0.99	0.33	51,60,64,68	2
5	CU	A	613	1/1	0.99	0.13	45,45,45,45	1
5	CU	A	614	1/1	0.99	0.12	63,63,63,63	1
5	CU	B	513	1/1	0.99	0.14	53,53,53,53	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 CU B 515:**

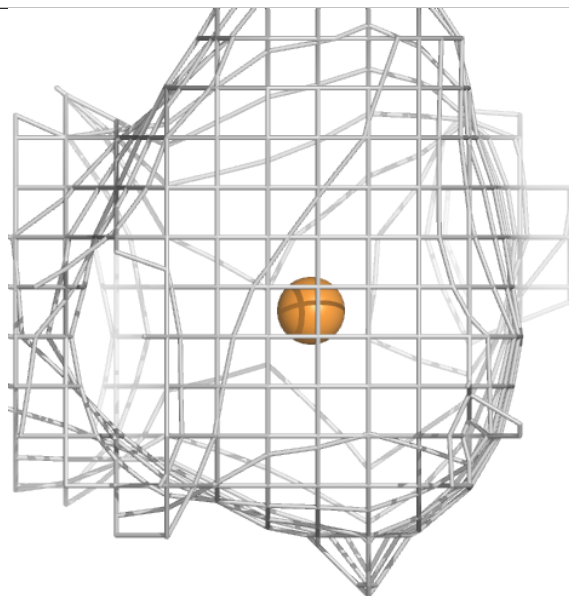
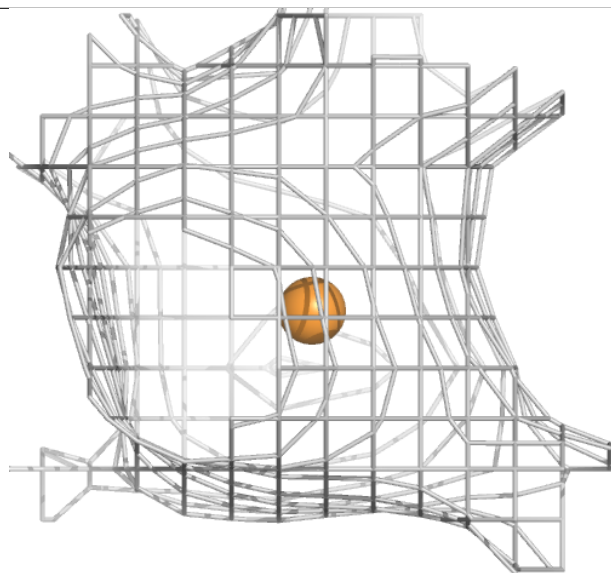
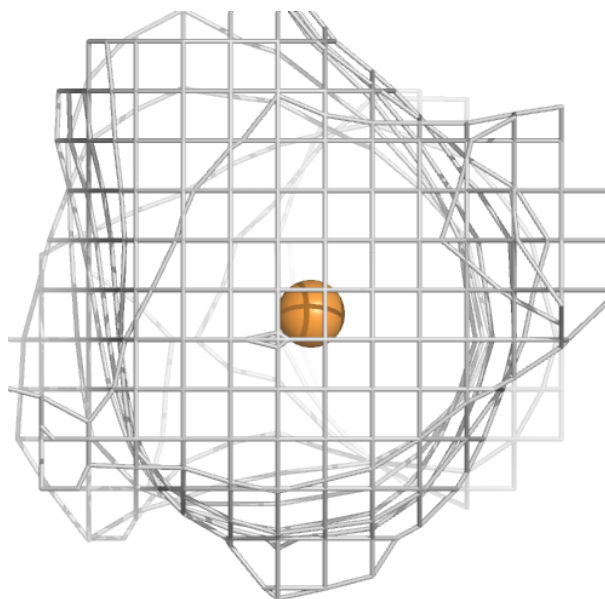
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



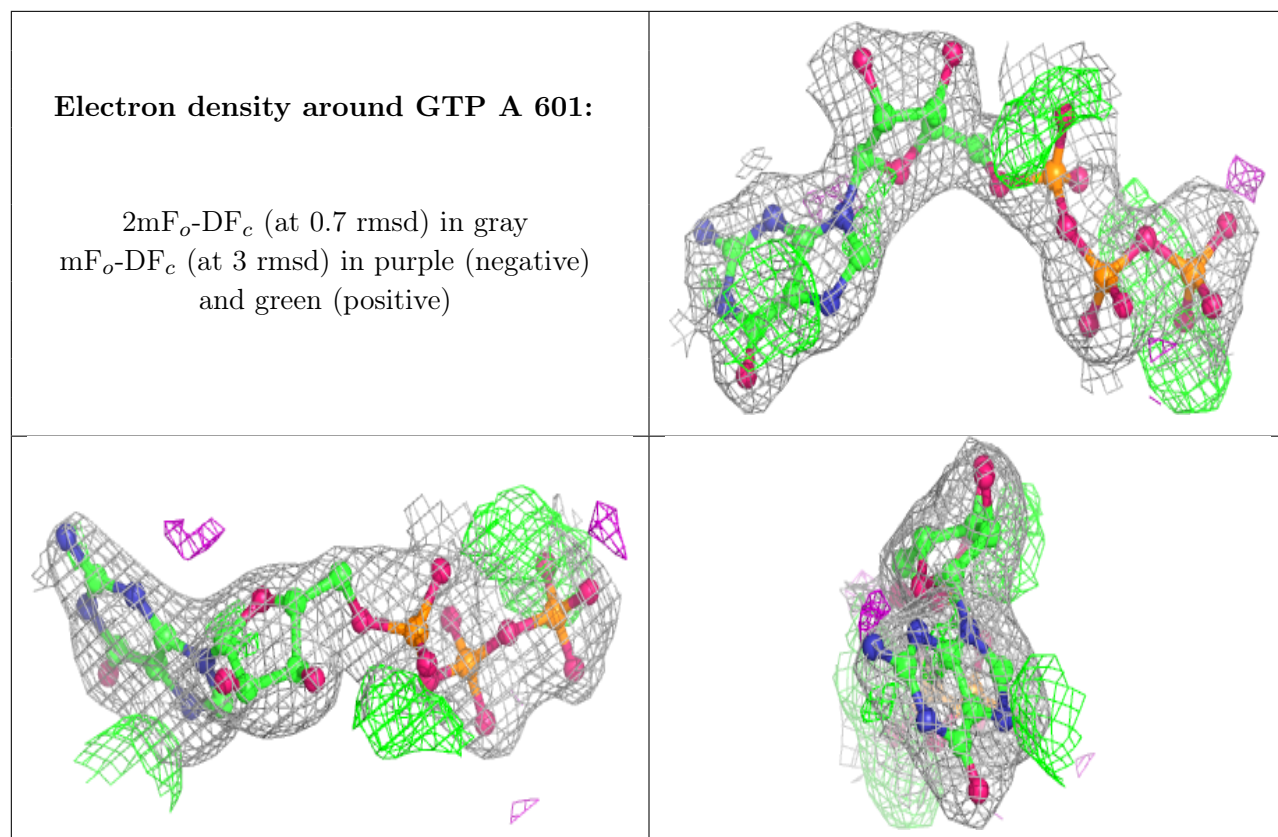


**Electron density around CU A 612:**

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

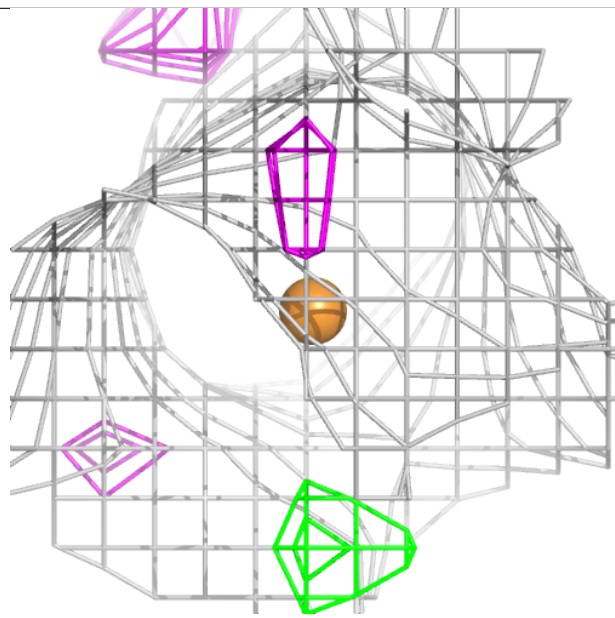
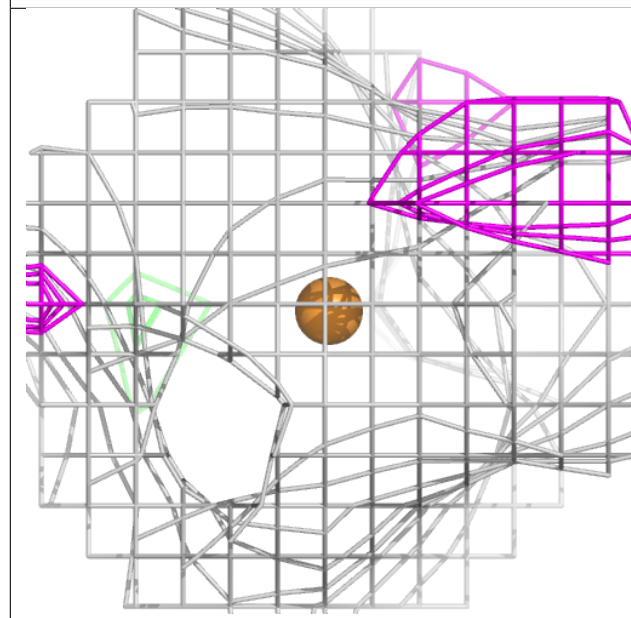
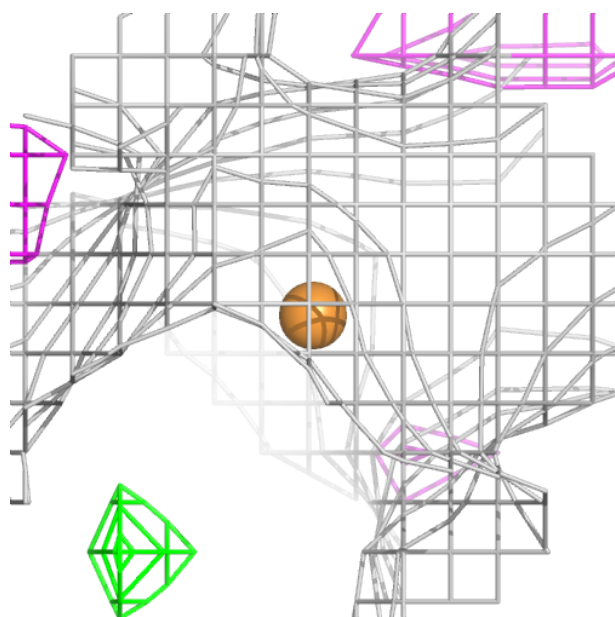






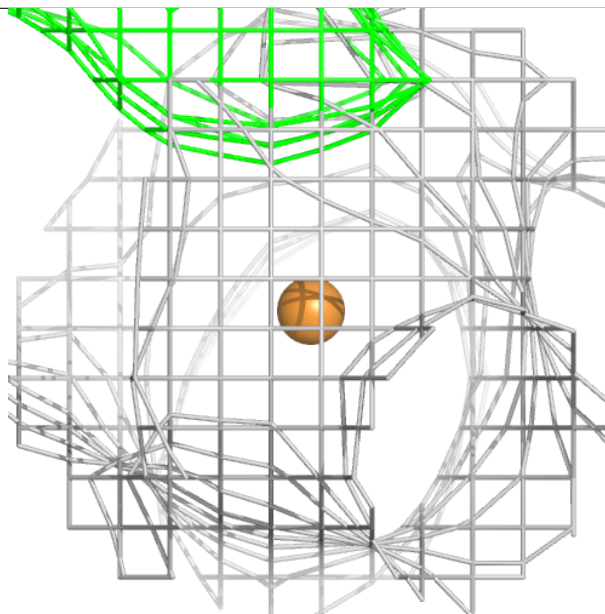
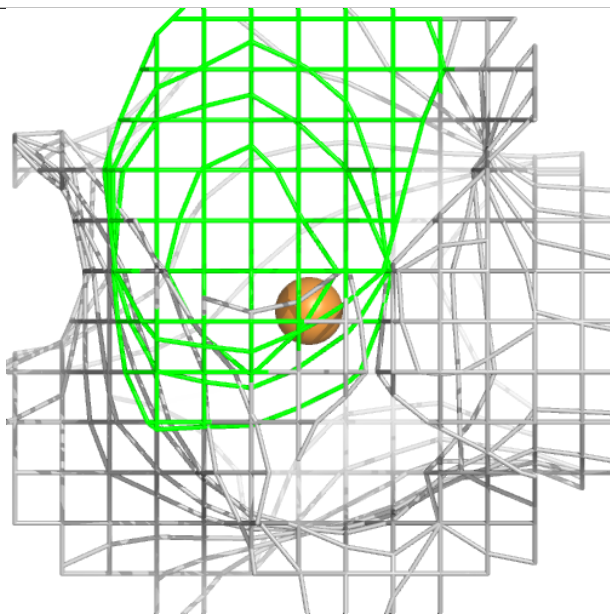
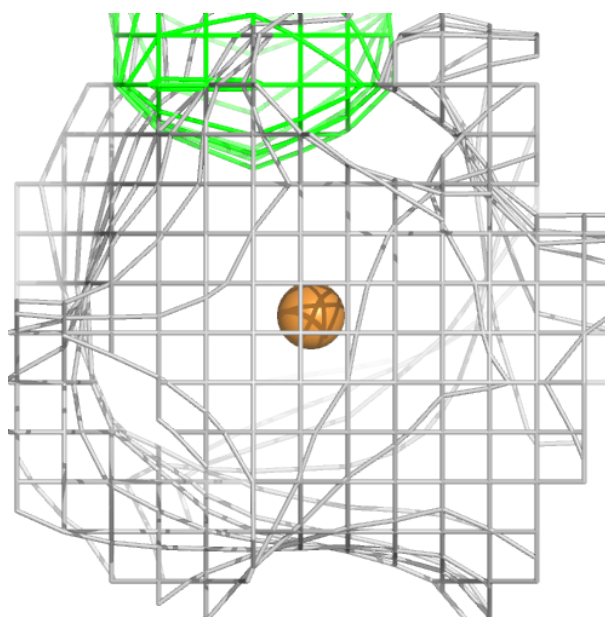
**Electron density around CU B 514:**

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



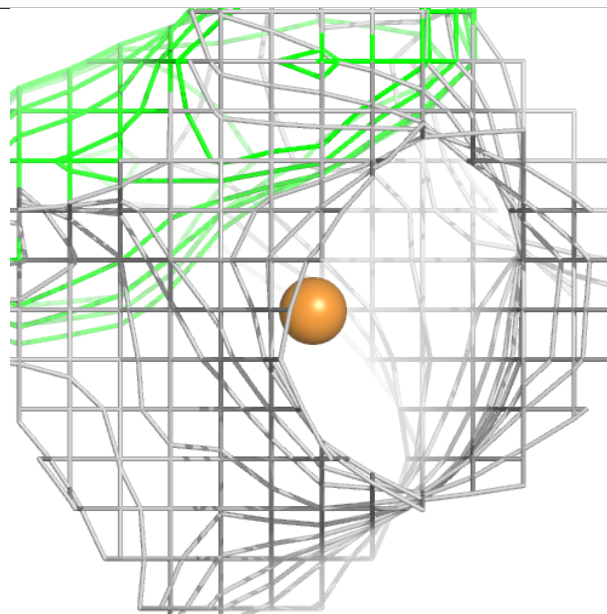
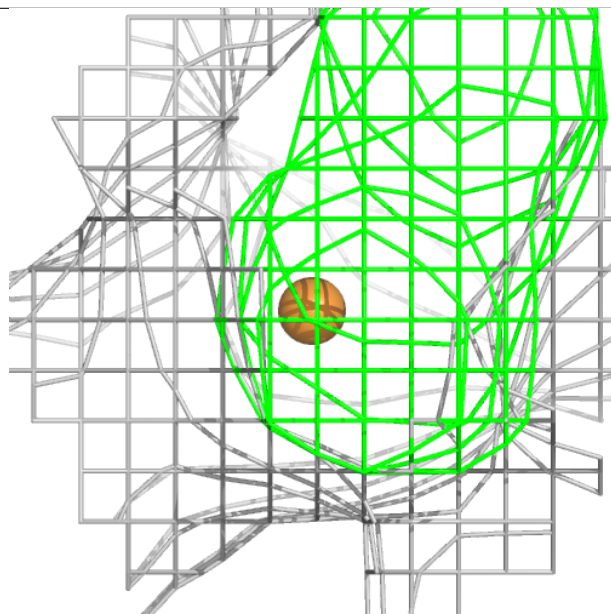
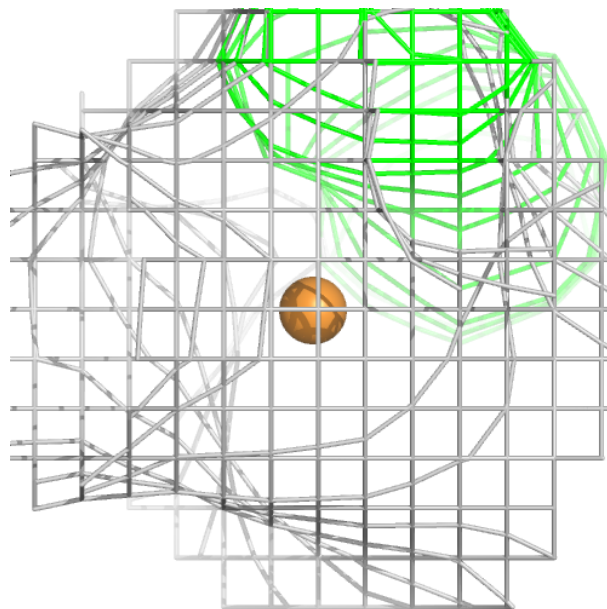
**Electron density around CU A 613:**

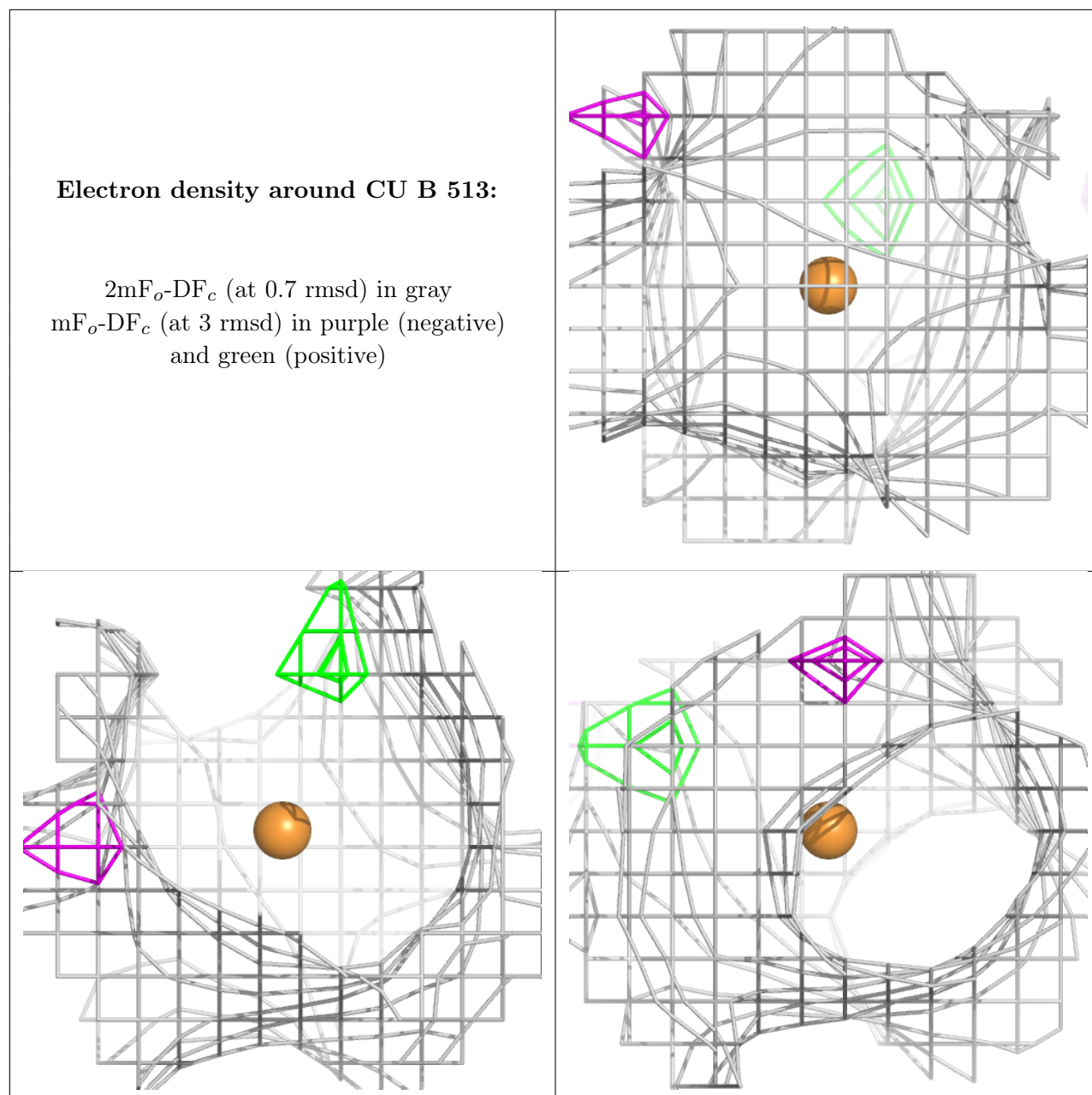
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around CU A 614:**

$2mF_o-DF_c$  (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.