



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 09:32 PM JST

PDB ID : 7WR3  
Title : Crystal structure of MBP-fused OspC3 in complex with calmodulin  
Authors : Hou, Y.J.; Zeng, H.; Shao, F.; Ding, J.  
Deposited on : 2022-01-26  
Resolution : 1.87 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 : 2.36  
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

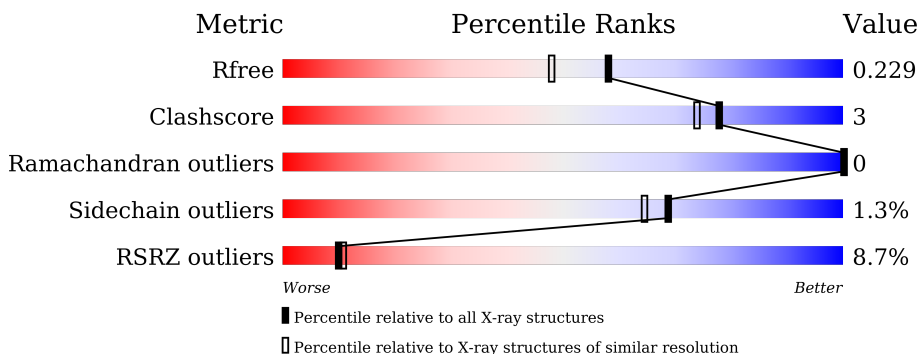
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.87 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	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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	806	
1	B	806	
2	C	153	
2	D	153	
3	F	2	
3	G	2	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 16225 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 MBP-fused OspC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	794	6296	4027	1058	1192	19	0	0	0
1	B	792	6275	4013	1056	1187	19	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P0AEX9
A	0	PRO	-	expression tag	UNP P0AEX9
A	1	MET	-	expression tag	UNP P0AEX9
A	83	ALA	ASP	engineered mutation	UNP P0AEX9
A	84	ALA	LYS	engineered mutation	UNP P0AEX9
A	173	ALA	GLU	engineered mutation	UNP P0AEX9
A	174	ALA	ASN	engineered mutation	UNP P0AEX9
A	240	ALA	LYS	engineered mutation	UNP P0AEX9
A	360	ALA	-	linker	UNP P0AEX9
A	361	ALA	-	linker	UNP P0AEX9
A	362	LEU	-	linker	UNP P0AEX9
A	363	ALA	-	linker	UNP P0AEX9
A	364	ALA	-	linker	UNP P0AEX9
A	365	ALA	-	linker	UNP P0AEX9
A	366	GLN	-	linker	UNP P0AEX9
A	367	THR	-	linker	UNP P0AEX9
A	368	ASN	-	linker	UNP P0AEX9
A	369	ALA	-	linker	UNP P0AEX9
A	370	ALA	-	linker	UNP P0AEX9
A	371	ALA	-	linker	UNP P0AEX9
B	-1	GLY	-	expression tag	UNP P0AEX9
B	0	PRO	-	expression tag	UNP P0AEX9
B	1	MET	-	expression tag	UNP P0AEX9
B	83	ALA	ASP	engineered mutation	UNP P0AEX9
B	84	ALA	LYS	engineered mutation	UNP P0AEX9

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	173	ALA	GLU	engineered mutation	UNP P0AEX9
B	174	ALA	ASN	engineered mutation	UNP P0AEX9
B	240	ALA	LYS	engineered mutation	UNP P0AEX9
B	360	ALA	-	linker	UNP P0AEX9
B	361	ALA	-	linker	UNP P0AEX9
B	362	LEU	-	linker	UNP P0AEX9
B	363	ALA	-	linker	UNP P0AEX9
B	364	ALA	-	linker	UNP P0AEX9
B	365	ALA	-	linker	UNP P0AEX9
B	366	GLN	-	linker	UNP P0AEX9
B	367	THR	-	linker	UNP P0AEX9
B	368	ASN	-	linker	UNP P0AEX9
B	369	ALA	-	linker	UNP P0AEX9
B	370	ALA	-	linker	UNP P0AEX9
B	371	ALA	-	linker	UNP P0AEX9

- Molecule 2 is a protein called Calmodulin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	138	1093	673	175	236	9	0	0	0
2	D	144	1140	701	183	247	9	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

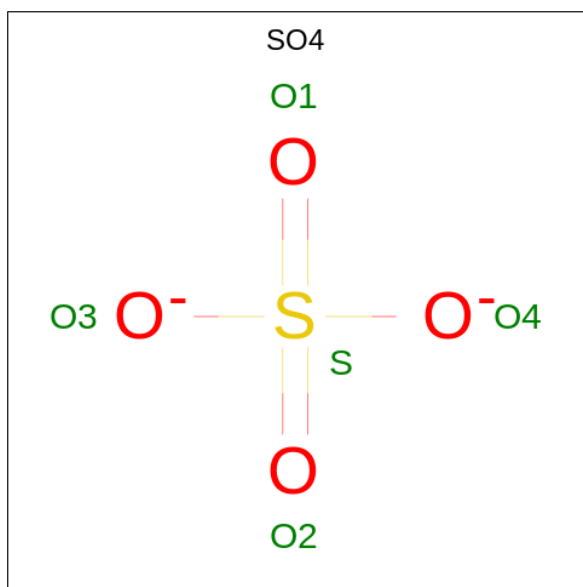
Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	SER	-	expression tag	UNP P0DP23
C	-2	GLY	-	expression tag	UNP P0DP23
C	-1	ARG	-	expression tag	UNP P0DP23
C	0	PRO	-	expression tag	UNP P0DP23
D	-3	SER	-	expression tag	UNP P0DP23
D	-2	GLY	-	expression tag	UNP P0DP23
D	-1	ARG	-	expression tag	UNP P0DP23
D	0	PRO	-	expression tag	UNP P0DP23

- Molecule 3 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



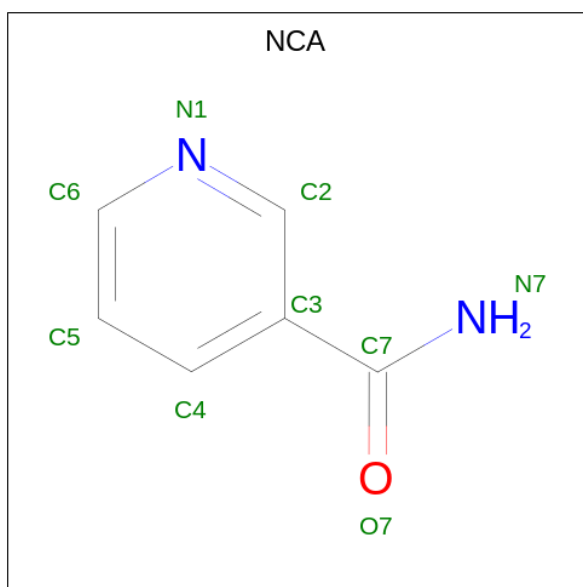
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	F	2	Total	C	O	0	0	0
			23	12	11			
3	G	2	Total	C	O	0	0	0
			23	12	11			

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



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	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is NICOTINAMIDE (three-letter code: NCA) (formula: C<sub>6</sub>H<sub>6</sub>N<sub>2</sub>O) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			9	6	2	1		
5	B	1	Total	C	N	O	0	0
			9	6	2	1		

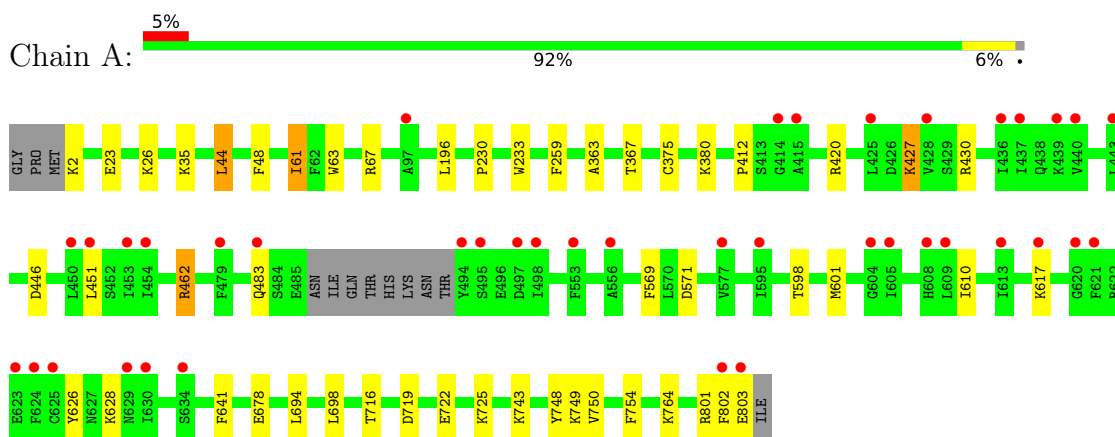
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	614	Total	O	0	0
			614	614		
6	B	680	Total	O	0	0
			680	680		
6	C	14	Total	O	0	0
			14	14		
6	D	29	Total	O	0	0
			29	29		

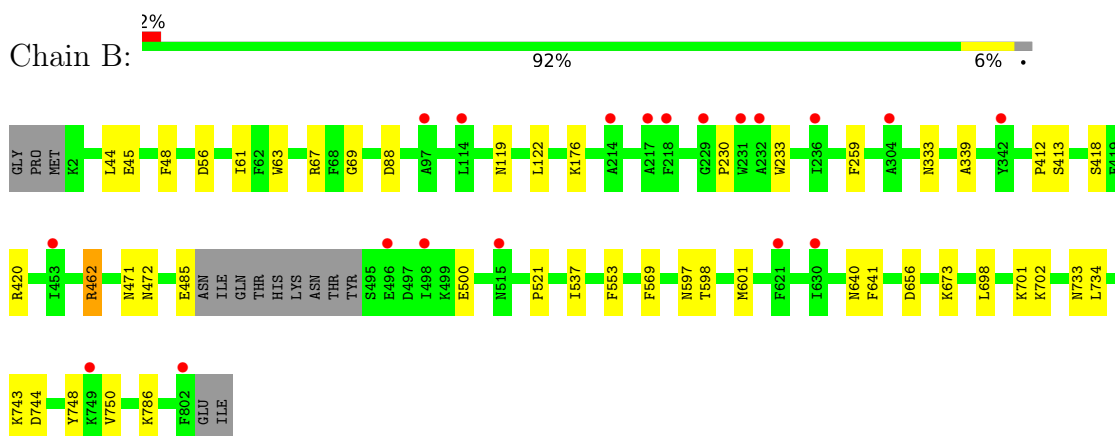
### 3 Residue-property plots [i](#)

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.

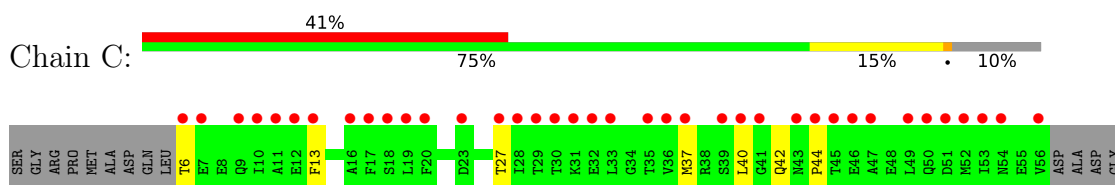
- Molecule 1: MBP-fused OspC3

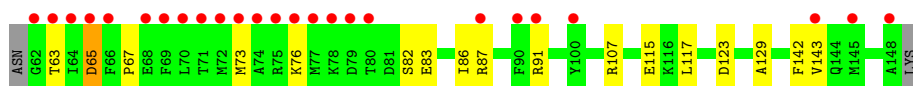


- Molecule 1: MBP-fused OspC3

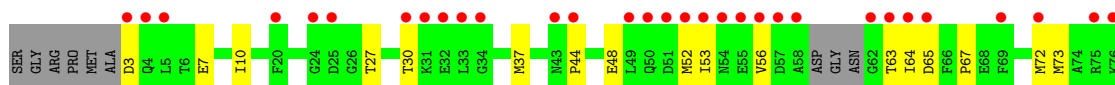
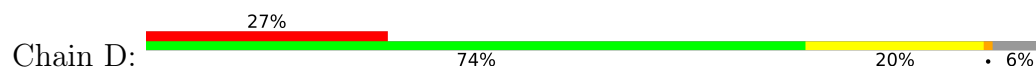


- Molecule 2: Calmodulin-1

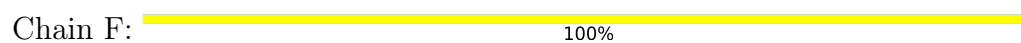




- Molecule 2: Calmodulin-1



- Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose



- Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.85Å 201.87Å 101.52Å 90.00° 102.11° 90.00°	Depositor
Resolution (Å)	49.99 – 1.87 49.99 – 1.87	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.99-1.87) 95.7 (49.99-1.87)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 1.87Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.196 , 0.228 0.196 , 0.229	Depositor DCC
$R_{free}$ test set	1990 reflections (1.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.0	Xtrriage
Anisotropy	0.106	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 48.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16225	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.27 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.8969e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, GLC, NCA

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	0.35	0/6438	0.50	1/8702 (0.0%)
1	B	0.35	0/6416	0.52	2/8672 (0.0%)
2	C	0.26	0/1104	0.45	0/1480
2	D	0.27	0/1151	0.45	0/1543
All	All	0.34	0/15109	0.50	3/20397 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	44	LEU	CA-CB-CG	6.16	129.47	115.30
1	B	44	LEU	CA-CB-CG	5.44	127.81	115.30
1	B	462	ARG	NE-CZ-NH2	-5.18	117.71	120.30

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	6296	0	6212	31	0
1	B	6275	0	6197	28	0
2	C	1093	0	1029	13	0
2	D	1140	0	1074	17	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	23	0	20	0	0
3	G	23	0	20	0	0
4	A	10	0	0	0	0
4	B	10	0	0	0	0
5	A	9	0	6	0	0
5	B	9	0	6	0	0
6	A	614	0	0	8	0
6	B	680	0	0	9	1
6	C	14	0	0	0	0
6	D	29	0	0	1	0
All	All	16225	0	14564	86	1

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

All (86) 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:375:CYS:SG	6:A:1521:HOH:O	2.40	0.78
1:A:462:ARG:NH2	6:A:1002:HOH:O	2.16	0.77
1:A:802:PHE:O	6:A:1001:HOH:O	2.09	0.70
1:B:176:LYS:NZ	6:B:1006:HOH:O	2.23	0.66
1:A:380:LYS:NZ	2:C:115:GLU:O	2.30	0.63
1:A:63:TRP:CD1	1:A:67:ARG:HG3	2.35	0.62
1:A:617:LYS:NZ	6:A:1007:HOH:O	2.33	0.61
1:B:462:ARG:NH2	6:B:1003:HOH:O	2.22	0.60
2:C:107:ARG:NH1	2:C:123:ASP:OD1	2.35	0.60
1:A:716:THR:HG23	1:A:719:ASP:H	1.69	0.58
1:A:412:PRO:HA	1:A:420:ARG:HH21	1.68	0.58
1:A:446:ASP:OD1	1:A:628:LYS:NZ	2.26	0.57
2:C:37:MET:HE3	2:C:44:PRO:HG3	1.85	0.57
1:B:418:SER:OG	1:B:420:ARG:NH1	2.37	0.57
1:A:678:GLU:OE2	6:A:1003:HOH:O	2.18	0.56
2:D:87:ARG:NH2	6:D:205:HOH:O	2.38	0.56
1:B:45:GLU:HB2	1:B:67:ARG:HD2	1.88	0.54
1:B:48:PHE:CG	1:B:61:ILE:HD12	2.43	0.54
1:B:56:ASP:OD1	6:B:1001:HOH:O	2.18	0.54
1:B:69:GLY:HA3	1:B:333:ASN:O	2.08	0.54
1:B:733:ASN:OD1	6:B:1002:HOH:O	2.19	0.53
2:D:30:THR:HA	2:D:64:ILE:HD11	1.89	0.53
1:B:673:LYS:HA	1:B:702:LYS:HE3	1.91	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:748:TYR:HD2	1:A:749:LYS:HD2	1.74	0.52
1:B:701:LYS:HD3	1:B:744:ASP:CG	2.29	0.52
1:A:678:GLU:OE2	1:A:678:GLU:N	2.34	0.52
2:D:99:GLY:O	2:D:138:ASN:HA	2.10	0.52
2:D:107:ARG:NH1	2:D:123:ASP:OD1	2.43	0.51
1:A:694:LEU:HD21	1:A:754:PHE:HZ	1.76	0.51
2:C:129:ALA:HA	2:C:142:PHE:CE1	2.47	0.50
1:B:420:ARG:NH1	6:B:1030:HOH:O	2.42	0.50
1:A:743:LYS:NZ	6:A:1015:HOH:O	2.40	0.49
1:B:598:THR:HA	1:B:601:MET:HG3	1.95	0.49
1:B:412:PRO:HA	1:B:420:ARG:HH21	1.78	0.49
1:B:748:TYR:CE2	1:B:786:LYS:HE3	2.47	0.48
1:A:451:LEU:HB2	1:A:610:ILE:HD11	1.95	0.48
1:A:801:ARG:NH1	6:A:1027:HOH:O	2.46	0.48
1:B:597:ASN:O	1:B:601:MET:HG3	2.14	0.48
1:A:48:PHE:CG	1:A:61:ILE:HD12	2.49	0.47
1:A:598:THR:HA	1:A:601:MET:HG3	1.96	0.47
2:D:65:ASP:CG	2:D:67:PRO:HD2	2.34	0.47
2:D:73:MET:O	2:D:77:MET:HG2	2.15	0.47
1:B:230:PRO:HA	1:B:233:TRP:CE2	2.50	0.47
1:B:640:ASN:HB2	6:B:1184:HOH:O	2.15	0.46
2:D:123:ASP:HB3	2:D:127:ARG:NH1	2.29	0.46
1:A:722:GLU:HA	1:A:725:LYS:HE2	1.98	0.46
2:C:82:SER:O	2:C:86:ILE:HG12	2.16	0.46
1:A:230:PRO:HA	1:A:233:TRP:CE2	2.51	0.46
2:C:13:PHE:HD2	2:C:40:LEU:HD22	1.81	0.45
2:C:65:ASP:HB2	2:C:67:PRO:HD2	1.98	0.44
1:A:363:ALA:O	1:A:367:THR:HG23	2.18	0.44
2:C:37:MET:HB3	2:C:42:GLN:HB2	1.99	0.44
2:C:83:GLU:HG3	2:C:143:VAL:HG12	1.97	0.44
2:D:7:GLU:O	2:D:10:ILE:HG13	2.18	0.44
1:A:427:LYS:N	1:A:427:LYS:HD2	2.32	0.44
1:A:427:LYS:O	1:A:430:ARG:HB3	2.17	0.44
1:A:569:PHE:CZ	1:A:641:PHE:HB2	2.53	0.44
1:A:698:LEU:HD21	1:A:750:VAL:HG11	2.00	0.44
2:D:27:THR:HG21	2:D:63:THR:HB	2.00	0.44
2:D:37:MET:HE1	2:D:44:PRO:HG3	2.00	0.44
2:D:89:ALA:O	2:D:92:VAL:HG22	2.18	0.44
2:D:27:THR:CG2	2:D:63:THR:HB	2.48	0.43
1:B:119:ASN:CG	1:B:122:LEU:HD13	2.38	0.43
2:D:48:GLU:O	2:D:52:MET:HG3	2.19	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:521:PRO:HB3	1:B:537:ILE:HD11	1.99	0.43
1:B:471:ASN:HB3	1:B:472:ASN:H	1.64	0.43
2:C:76:LYS:HE3	2:C:76:LYS:HB3	1.71	0.43
2:D:52:MET:HG2	2:D:72:MET:HE3	1.99	0.43
2:C:87:ARG:HG2	2:C:91:ARG:CZ	2.48	0.43
1:A:446:ASP:HB3	1:A:626:TYR:CG	2.54	0.43
1:A:571:ASP:OD2	6:A:1004:HOH:O	2.21	0.42
1:B:500:GLU:HG3	1:B:553:PHE:HE1	1.84	0.42
1:B:569:PHE:CZ	1:B:641:PHE:HB2	2.53	0.42
1:B:743:LYS:HD2	6:B:1392:HOH:O	2.20	0.42
2:C:27:THR:HA	2:C:65:ASP:HA	2.01	0.42
1:A:2:LYS:HA	1:A:2:LYS:HD3	1.87	0.42
1:B:698:LEU:HD21	1:B:750:VAL:HG11	2.02	0.41
1:B:63:TRP:CD1	1:B:67:ARG:HG3	2.54	0.41
1:B:339:ALA:HB3	6:B:1104:HOH:O	2.19	0.41
1:A:35:LYS:HB3	1:A:35:LYS:HE2	1.70	0.41
1:B:88:ASP:OD1	6:B:1004:HOH:O	2.22	0.41
1:A:23:GLU:O	1:A:26:LYS:HG2	2.20	0.41
2:D:129:ALA:HA	2:D:142:PHE:CE1	2.56	0.40
1:A:380:LYS:HD3	2:C:117:LEU:HD22	2.04	0.40
1:B:656:ASP:OD2	2:D:116:LYS:NZ	2.55	0.40
2:D:53:ILE:HA	2:D:56:VAL:HG22	2.04	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:B:1002:HOH:O	6:B:1459:HOH:O[2_455]	2.16	0.04

## 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	790/806 (98%)	772 (98%)	18 (2%)	0	100	100
1	B	788/806 (98%)	775 (98%)	13 (2%)	0	100	100
2	C	134/153 (88%)	131 (98%)	3 (2%)	0	100	100
2	D	140/153 (92%)	136 (97%)	4 (3%)	0	100	100
All	All	1852/1918 (97%)	1814 (98%)	38 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	675/686 (98%)	666 (99%)	9 (1%)	69	64
1	B	673/686 (98%)	669 (99%)	4 (1%)	86	86
2	C	119/130 (92%)	115 (97%)	4 (3%)	37	25
2	D	124/130 (95%)	120 (97%)	4 (3%)	39	27
All	All	1591/1632 (98%)	1570 (99%)	21 (1%)	69	64

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	LEU
1	A	61	ILE
1	A	196	LEU
1	A	259	PHE
1	A	427	LYS
1	A	462	ARG
1	A	483	GLN
1	A	764	LYS
1	A	803	GLU
1	B	259	PHE
1	B	413	SER
1	B	485	GLU
1	B	734	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	6	THR
2	C	63	THR
2	C	65	ASP
2	C	73	MET
2	D	3	ASP
2	D	79	ASP
2	D	91	ARG
2	D	116	LYS

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	$\# Z  > 2$
3	GLC	F	1	3	12,12,12	1.28	1 (8%)	17,17,17	1.18	1 (5%)
3	GLC	F	2	3	11,11,12	1.43	1 (9%)	15,15,17	1.01	1 (6%)
3	GLC	G	1	3	12,12,12	1.16	1 (8%)	17,17,17	1.17	2 (11%)
3	GLC	G	2	3	11,11,12	1.49	3 (27%)	15,15,17	0.78	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	GLC	F	1	3	-	0/2/22/22	0/1/1/1
3	GLC	F	2	3	-	0/2/19/22	0/1/1/1
3	GLC	G	1	3	-	0/2/22/22	0/1/1/1
3	GLC	G	2	3	-	0/2/19/22	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	1	GLC	O5-C1	3.61	1.51	1.42
3	F	2	GLC	O5-C1	3.48	1.49	1.43
3	G	1	GLC	O5-C1	3.26	1.51	1.42
3	G	2	GLC	O5-C1	3.04	1.48	1.43
3	G	2	GLC	O5-C5	2.12	1.47	1.43
3	G	2	GLC	C2-C3	-2.05	1.49	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	2	GLC	O5-C1-C2	-2.72	106.57	110.77
3	F	1	GLC	O3-C3-C2	-2.40	104.80	110.35
3	G	1	GLC	C1-C2-C3	2.32	115.14	110.31
3	G	1	GLC	C4-C3-C2	2.30	114.84	110.82

There are no chirality outliers.

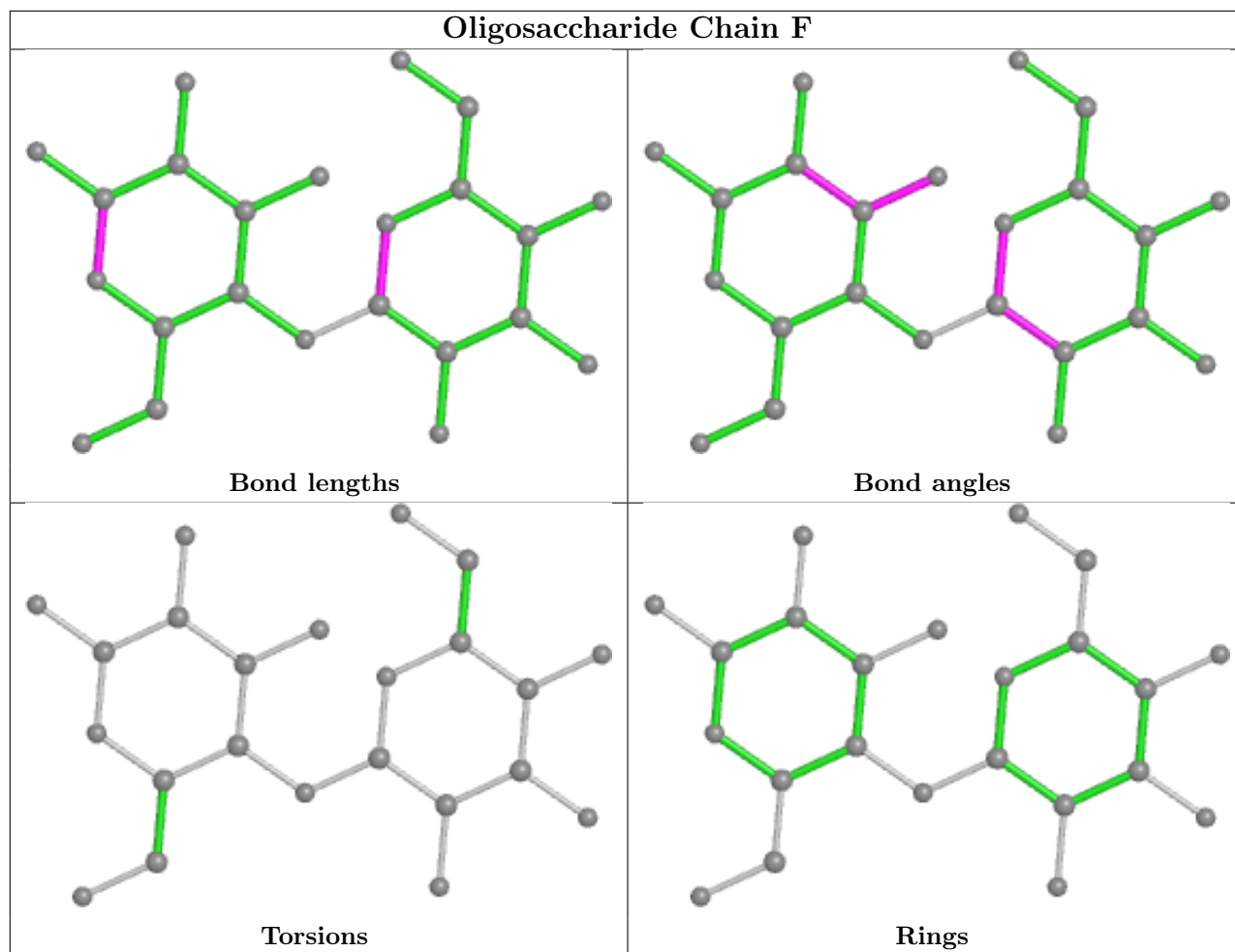
There are no torsion outliers.

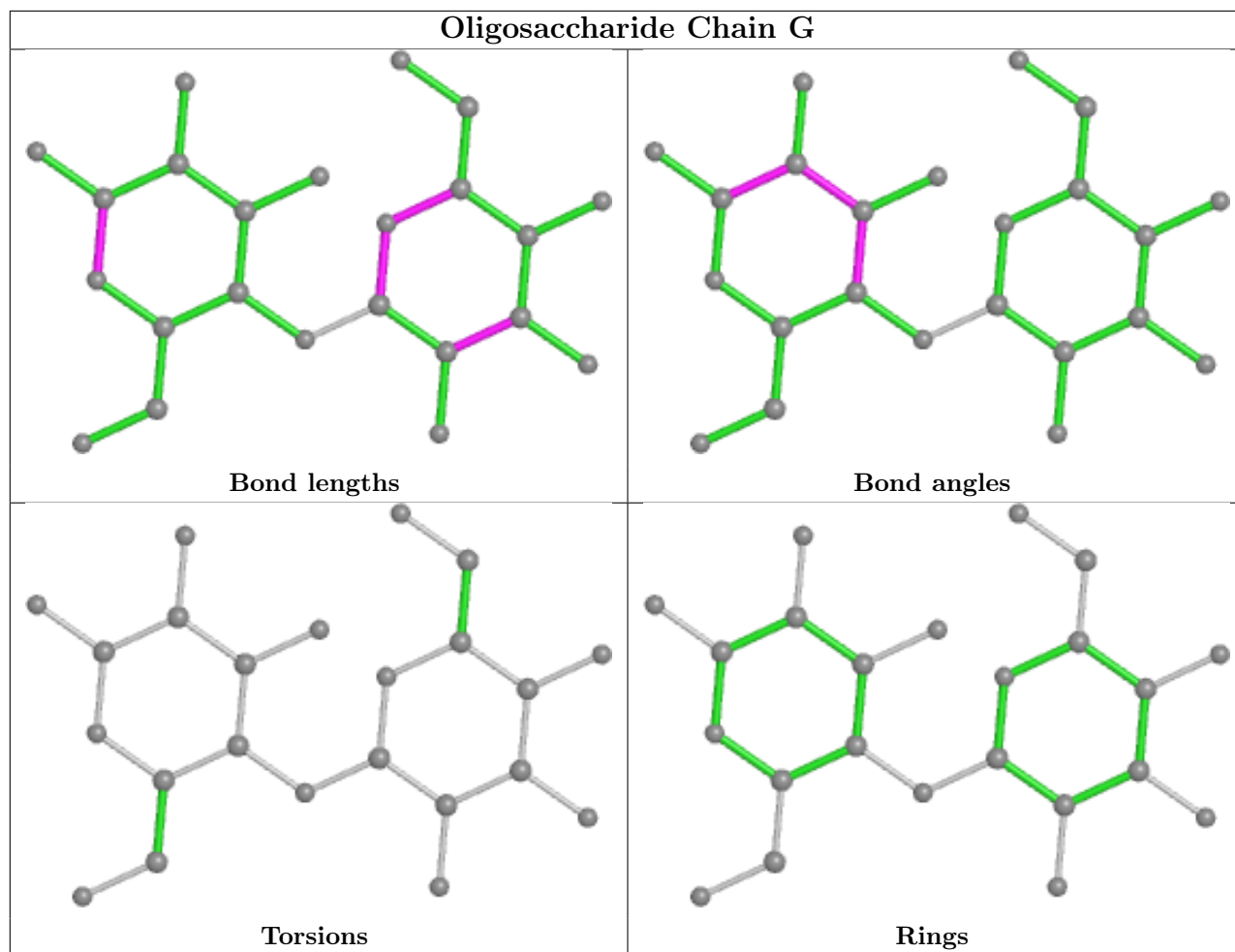
There are no ring outliers.

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](#)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	SO4	A	901	-	4,4,4	0.18	0	6,6,6	0.14	0
5	NCA	B	903	-	9,9,9	2.66	2 (22%)	11,11,11	1.57	2 (18%)
4	SO4	B	902	-	4,4,4	0.16	0	6,6,6	0.27	0
5	NCA	A	903	-	9,9,9	2.58	2 (22%)	11,11,11	1.51	3 (27%)
4	SO4	B	901	-	4,4,4	0.16	0	6,6,6	0.21	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	A	902	-	4,4,4	0.15	0	6,6,6	0.06	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
5	NCA	A	903	-	-	0/4/4/4	0/1/1/1
5	NCA	B	903	-	-	0/4/4/4	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	903	NCA	C7-N7	6.92	1.46	1.33
5	A	903	NCA	C7-N7	6.61	1.45	1.33
5	B	903	NCA	O7-C7	-2.85	1.18	1.24
5	A	903	NCA	O7-C7	-2.73	1.18	1.24

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	903	NCA	C4-C3-C2	3.12	121.17	117.63
5	B	903	NCA	C3-C2-N1	-2.85	119.27	123.49
5	A	903	NCA	C3-C7-N7	2.41	120.64	117.75
5	A	903	NCA	C3-C2-N1	-2.37	119.98	123.49
5	A	903	NCA	C4-C3-C2	2.18	120.10	117.63

There are no chirality outliers.

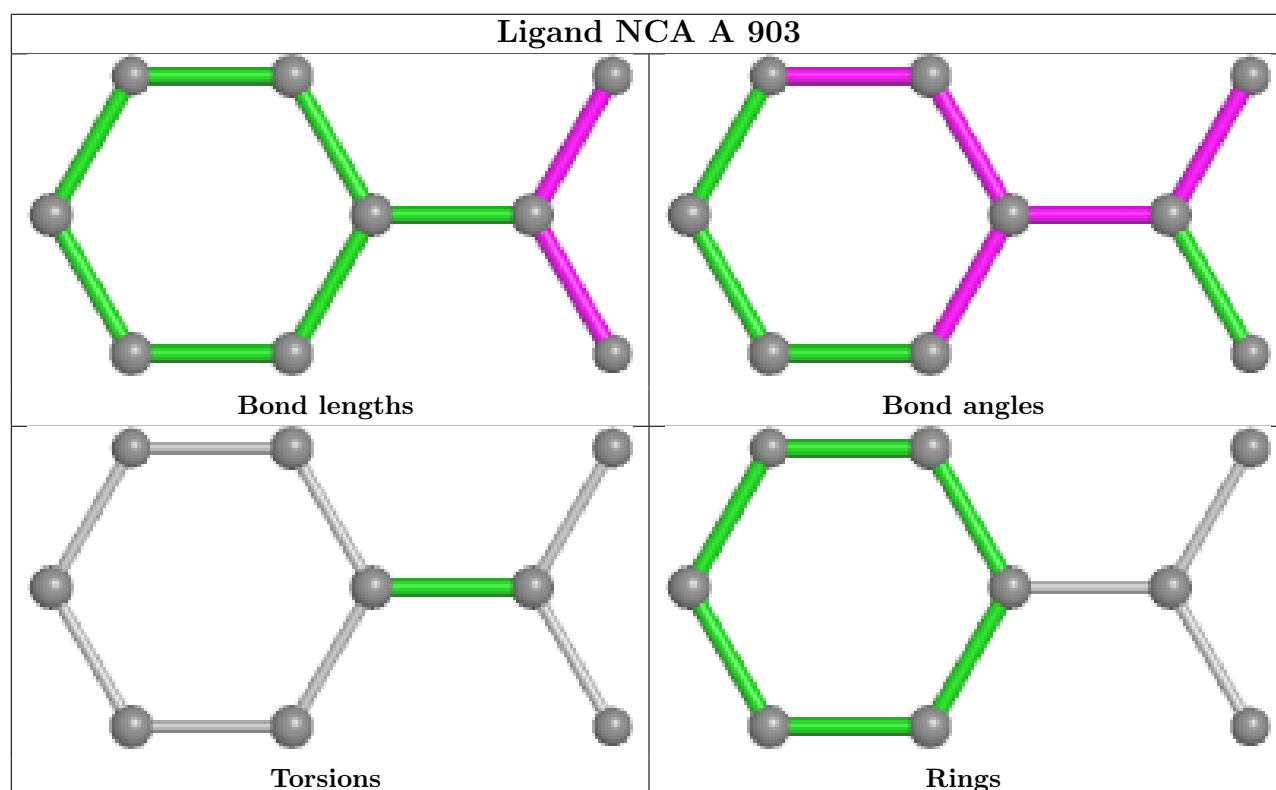
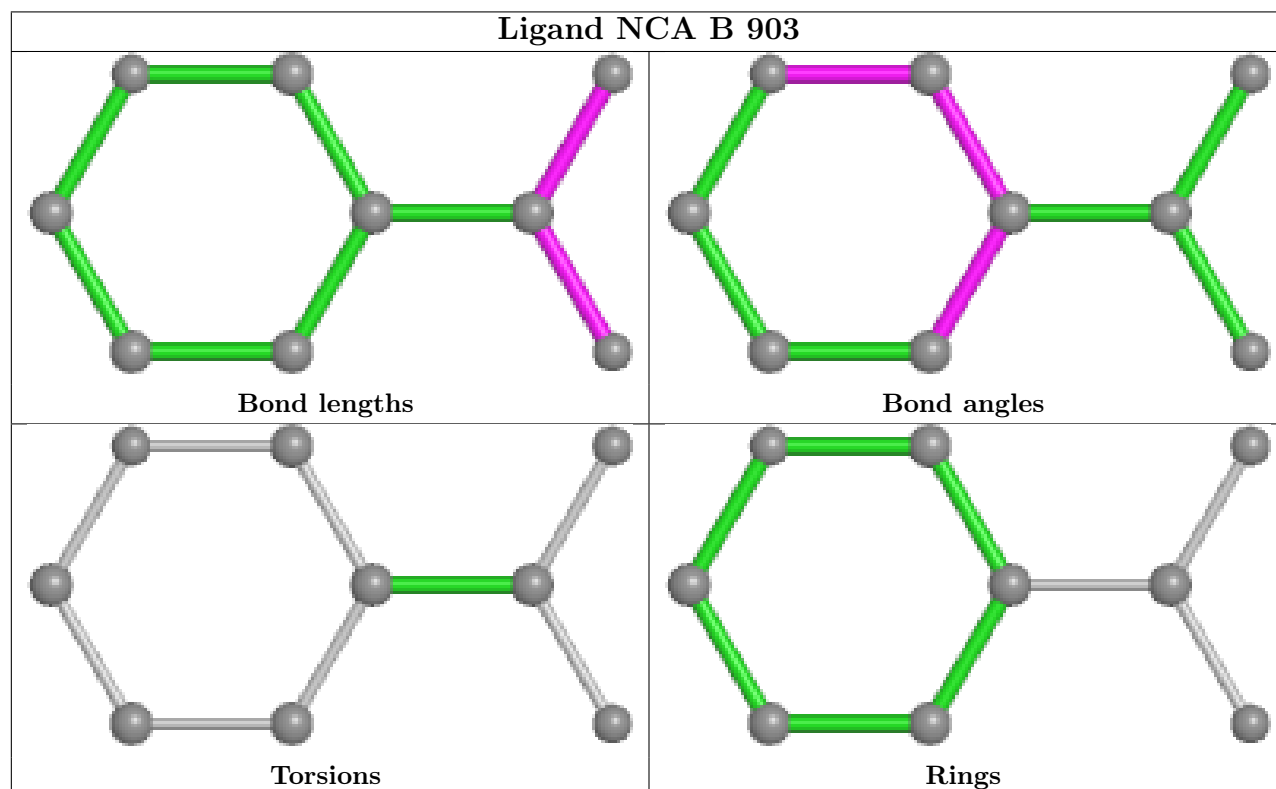
There are no torsion outliers.

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 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	794/806 (98%)	0.31	40 (5%) 28 30	17, 34, 63, 90	0
1	B	792/806 (98%)	0.10	19 (2%) 59 60	17, 32, 51, 88	0
2	C	138/153 (90%)	2.34	63 (45%) 0 0	35, 75, 110, 117	0
2	D	144/153 (94%)	1.28	41 (28%) 0 0	36, 56, 87, 97	0
All	All	1868/1918 (97%)	0.44	163 (8%) 10 11	17, 35, 78, 117	0

All (163) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	69	PHE	10.7
2	C	53	ILE	10.4
1	A	494	TYR	9.1
2	C	50	GLN	8.4
2	C	49	LEU	8.1
2	D	53	ILE	8.0
1	A	498	ILE	7.8
2	C	72	MET	7.6
1	A	450	LEU	7.4
2	C	78	LYS	7.3
2	C	70	LEU	7.2
2	C	76	LYS	7.1
2	C	28	ILE	6.6
2	C	54	ASN	6.0
2	C	19	LEU	5.9
2	C	12	GLU	5.8
1	A	497	ASP	5.7
1	A	454	ILE	5.7
2	C	7	GLU	5.6
2	C	17	PHE	5.5
2	C	31	LYS	5.5

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	C	13	PHE	5.5
2	C	29	THR	5.3
2	C	75	ARG	5.2
1	A	802	PHE	5.2
2	C	52	MET	5.2
2	D	56	VAL	5.2
2	C	62	GLY	5.2
1	B	498	ILE	5.0
2	C	10	ILE	5.0
2	C	79	ASP	4.9
2	C	35	THR	4.9
1	A	553	PHE	4.8
2	D	72	MET	4.8
2	C	73	MET	4.7
2	C	20	PHE	4.6
1	A	453	ILE	4.5
2	C	100	TYR	4.5
2	D	98	ASN	4.5
1	A	605	ILE	4.5
2	D	3	ASP	4.4
2	D	54	ASN	4.4
1	A	436	ILE	4.4
2	C	64	ILE	4.4
2	C	39	SER	4.3
2	C	77	MET	4.3
2	C	33	LEU	4.2
2	C	56	VAL	4.2
2	C	40	LEU	4.2
1	A	425	LEU	4.1
1	A	451	LEU	4.1
1	A	613	ILE	4.1
2	C	90	PHE	4.0
2	D	49	LEU	3.9
1	A	414	GLY	3.9
1	A	439	LYS	3.9
2	C	36	VAL	3.9
2	C	16	ALA	3.7
2	D	33	LEU	3.7
2	C	66	PHE	3.7
2	D	57	ASP	3.6
2	C	47	ALA	3.6
2	C	32	GLU	3.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	C	9	GLN	3.5
2	D	64	ILE	3.5
2	D	50	GLN	3.5
1	A	443	LEU	3.5
2	C	11	ALA	3.4
2	C	45	THR	3.4
2	C	6	THR	3.4
2	C	46	GLU	3.4
2	C	41	GLY	3.4
1	A	629	ASN	3.4
1	A	630	ILE	3.4
2	C	37	MET	3.4
2	C	74	ALA	3.3
1	B	496	GLU	3.3
1	A	479	PHE	3.2
2	D	100	TYR	3.2
2	C	65	ASP	3.2
2	C	23	ASP	3.1
2	C	63	THR	3.1
1	A	624	PHE	3.1
2	D	4	GLN	3.1
1	A	483	GLN	3.0
1	B	217	ALA	3.0
2	D	149	LYS	3.0
1	A	556	ALA	3.0
1	A	495	SER	2.9
2	D	43	ASN	2.9
1	A	440	VAL	2.9
2	D	20	PHE	2.9
2	C	87	ARG	2.9
1	A	625	CYS	2.9
2	D	97	GLY	2.9
2	D	63	THR	2.8
1	A	415	ALA	2.8
2	D	58	ALA	2.8
1	A	617	LYS	2.8
1	A	428	VAL	2.8
2	D	30	THR	2.7
2	D	55	GLU	2.7
2	C	91	ARG	2.7
1	B	802	PHE	2.7
2	D	32	GLU	2.6

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	515	ASN	2.6
2	D	96	ASP	2.6
2	C	18	SER	2.6
1	B	214	ALA	2.5
2	D	52	MET	2.5
1	B	453	ILE	2.5
2	D	148	ALA	2.5
2	D	25	ASP	2.5
2	D	65	ASP	2.5
1	B	97	ALA	2.5
1	A	621	PHE	2.5
2	D	69	PHE	2.5
2	D	147	THR	2.5
1	A	623	GLU	2.5
2	D	99	GLY	2.4
2	C	71	THR	2.4
2	C	30	THR	2.4
1	A	803	GLU	2.4
1	A	604	GLY	2.3
1	B	630	ILE	2.3
1	B	749	LYS	2.3
1	B	342	TYR	2.3
1	B	231	TRP	2.3
1	A	577	VAL	2.3
2	C	145	MET	2.3
2	D	31	LYS	2.3
1	A	609	LEU	2.3
2	C	43	ASN	2.3
1	B	621	PHE	2.3
2	C	51	ASP	2.3
1	A	634	SER	2.3
2	D	144	GLN	2.2
1	A	608	HIS	2.2
1	B	229	GLY	2.2
2	C	148	ALA	2.2
1	B	232	ALA	2.2
2	D	76	LYS	2.2
1	A	437	ILE	2.2
2	D	62	GLY	2.2
1	B	236	ILE	2.2
2	C	80	THR	2.2
1	B	304	ALA	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	75	ARG	2.2
1	B	114	LEU	2.1
2	C	27	THR	2.1
2	C	68	GLU	2.1
2	D	44	PRO	2.1
1	A	595	ILE	2.1
1	A	97	ALA	2.1
2	D	78	LYS	2.1
2	C	44	PRO	2.1
1	A	620	GLY	2.1
2	D	5	LEU	2.1
1	B	218	PHE	2.1
2	D	51	ASP	2.0
2	D	24	GLY	2.0
2	D	34	GLY	2.0
2	C	143	VAL	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](#)

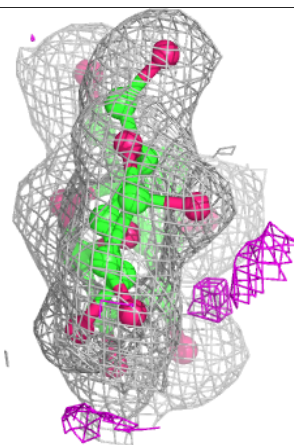
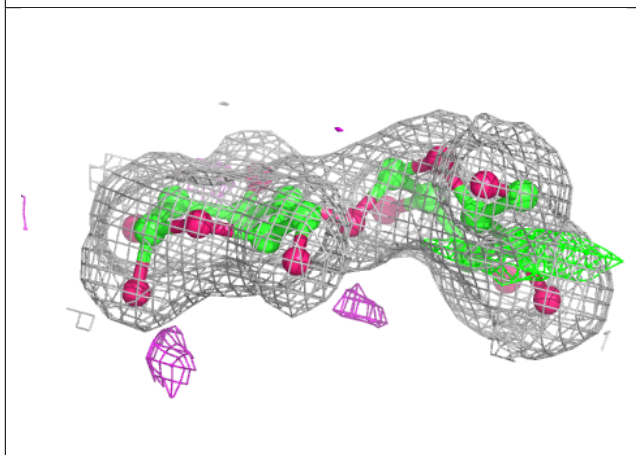
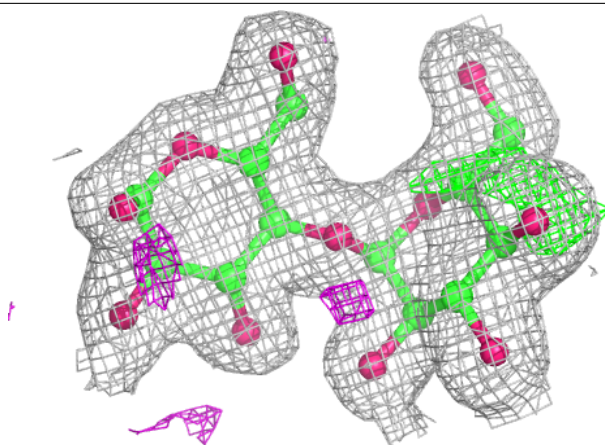
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

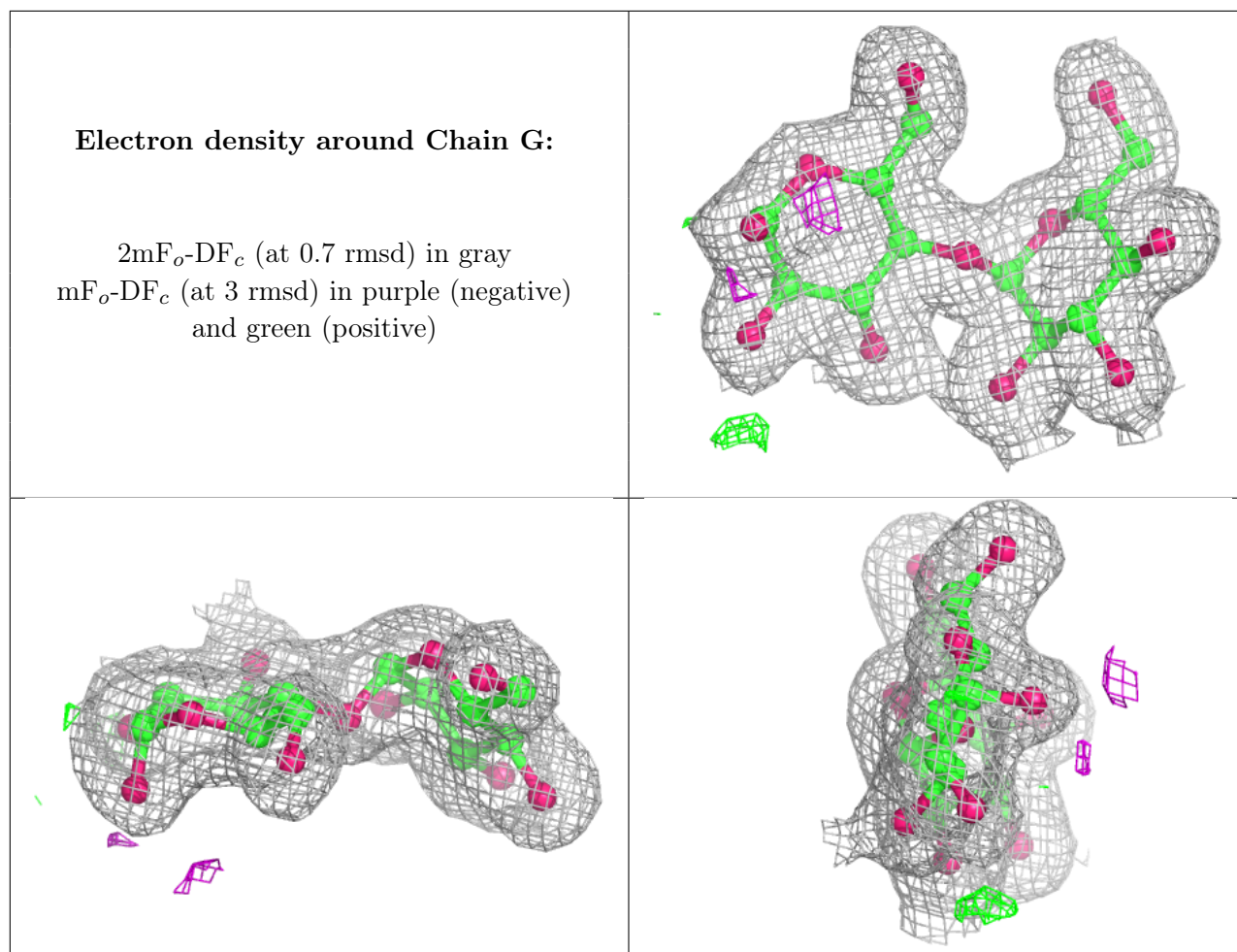
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	GLC	F	2	11/12	0.94	0.16	17,20,22,22	0
3	GLC	F	1	12/12	0.96	0.16	18,22,32,32	0
3	GLC	G	2	11/12	0.96	0.15	17,22,23,26	0
3	GLC	G	1	12/12	0.98	0.13	16,22,27,29	0

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

**Electron density around Chain F:**

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





## 6.4 Ligands [i](#)

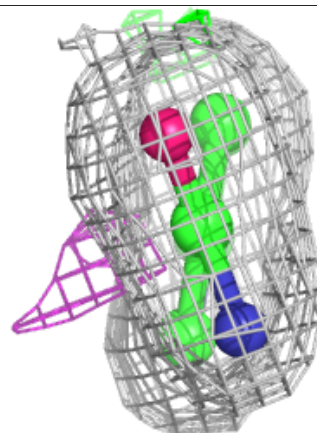
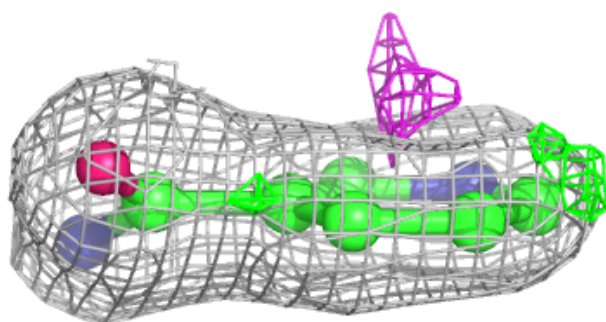
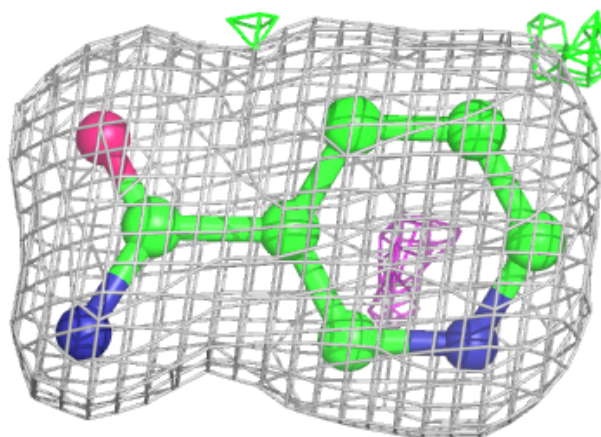
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 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	901	5/5	0.85	0.23	41,43,50,69	0
5	NCA	B	903	9/9	0.92	0.09	28,32,36,36	0
5	NCA	A	903	9/9	0.94	0.09	30,33,38,40	0
4	SO4	B	902	5/5	0.95	0.11	27,33,48,53	0
4	SO4	A	902	5/5	0.95	0.18	40,52,57,65	0
4	SO4	A	901	5/5	0.95	0.14	43,48,55,66	0

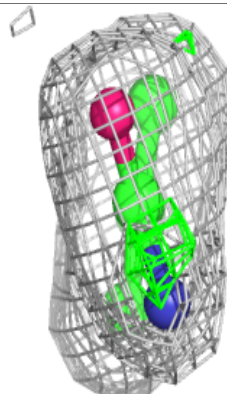
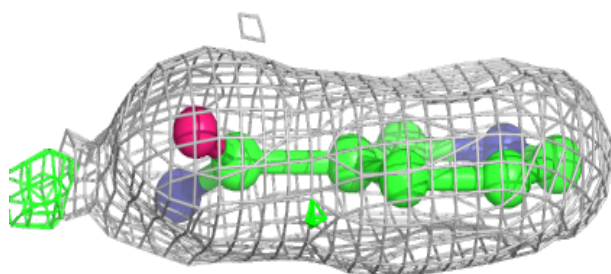
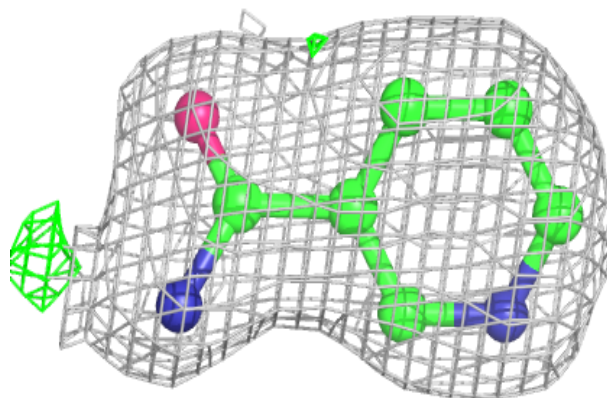
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around NCA B 903:**

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

**Electron density around NCA A 903:**

$2mF_o-DF_c$  (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.