



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 23, 2023 – 09:30 pm GMT

PDB ID : 8AXS  
Title : Sialidases and Fucosidases of Akkermansia muciniphila are key for rapid growth on colonic mucin and nutrient sharing amongst mucin-associated human gut microbiota  
Authors : Sakanaka, H.; Nielsen, T.S.; Pichler, M.J.; Nordberg Karlsson, E.; Abou Hachem, M.; Morth, J.P.  
Deposited on : 2022-08-31  
Resolution : 1.30 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](#) ①) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.32.1  
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) : 2.32.1

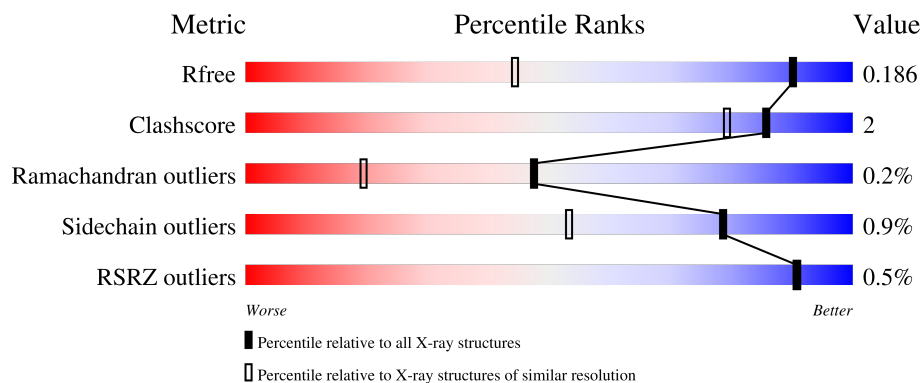
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.30 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	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (1.30-1.30)

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	578	
2	B	576	

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
6	CL	B	607	-	-	X	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 19936 atoms, of which 9196 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 Exo-alpha-sialidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	573	9169	2897	4552	848	858	14	0	27	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	MET	-	initiating methionine	UNP A0A7G6DVF6
A	477	VAL	ALA	conflict	UNP A0A7G6DVF6
A	516	LYS	ARG	conflict	UNP A0A7G6DVF6
A	598	ALA	-	expression tag	UNP A0A7G6DVF6
A	599	ALA	-	expression tag	UNP A0A7G6DVF6
A	600	ALA	-	expression tag	UNP A0A7G6DVF6
A	601	ALA	-	expression tag	UNP A0A7G6DVF6

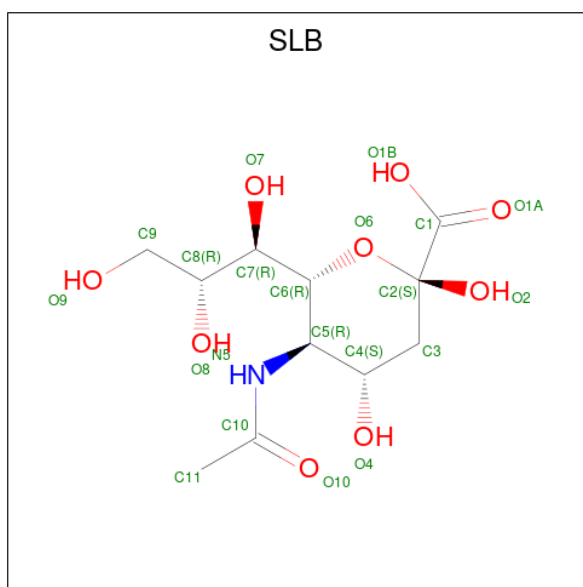
- Molecule 2 is a protein called Exo-alpha-sialidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	572	9126	2883	4539	844	848	12	0	23	0

There are 5 discrepancies between the modelled and reference sequences:

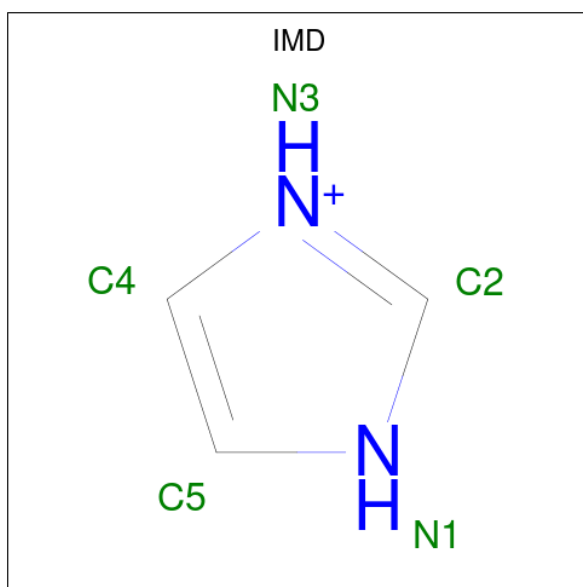
Chain	Residue	Modelled	Actual	Comment	Reference
B	22	MET	-	initiating methionine	UNP A0A7G6DVF6
B	477	VAL	ALA	conflict	UNP A0A7G6DVF6
B	516	LYS	ARG	conflict	UNP A0A7G6DVF6
B	596	ALA	-	expression tag	UNP A0A7G6DVF6
B	597	GLU	-	expression tag	UNP A0A7G6DVF6

- Molecule 3 is N-acetyl-beta-neuraminic acid (three-letter code: SLB) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>) (labeled as "Ligand of Interest" by depositor).



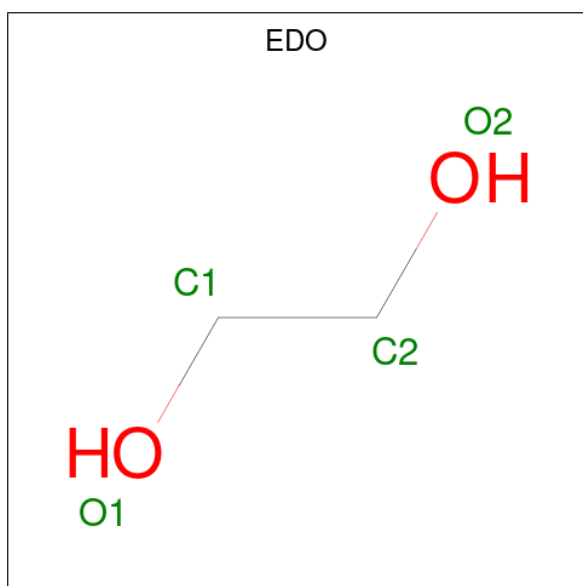
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
3	A	1	Total	C	H	N	O	0	0
			36	11	16	1	8		
3	A	1	Total	C	H	N	O	0	0
			36	11	16	1	8		
3	B	1	Total	C	H	N	O	0	0
			36	11	16	1	8		

- Molecule 4 is IMIDAZOLE (three-letter code: IMD) (formula:  $C_3H_5N_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	N		
4	A	1	Total	C	H	N	0	0
			10	3	5	2		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			10	2	6	2		
5	A	1	Total	C	H	O	0	0
			10	2	6	2		
5	A	1	Total	C	H	O	0	0
			10	2	6	2		
5	B	1	Total	C	H	O	0	0
			10	2	6	2		
5	B	1	Total	C	H	O	0	0
			10	2	6	2		
5	B	1	Total	C	H	O	0	0
			10	2	6	2		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Cl	0	0
			2	2		
6	B	2	Total	Cl	0	0
			2	2		

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

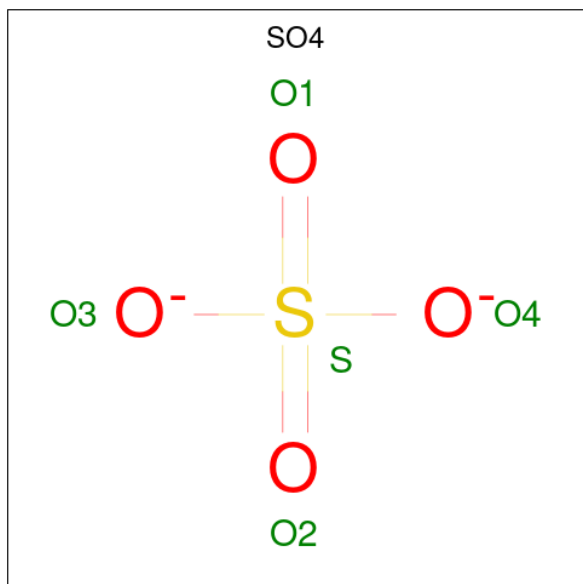
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Ca	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

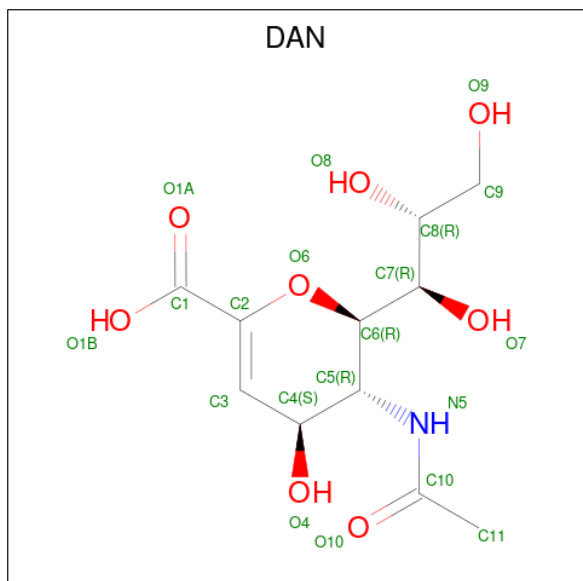
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	1	Total	Ca	0	0
			1	1		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 9 is 2-DEOXY-2,3-DEHYDRO-N-ACETYL-NEURAMINIC ACID (three-letter code: DAN) (formula: C<sub>11</sub>H<sub>17</sub>NO<sub>8</sub>).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
9	B	1	36	11	16	1	8	0	0

- Molecule 10 is water.

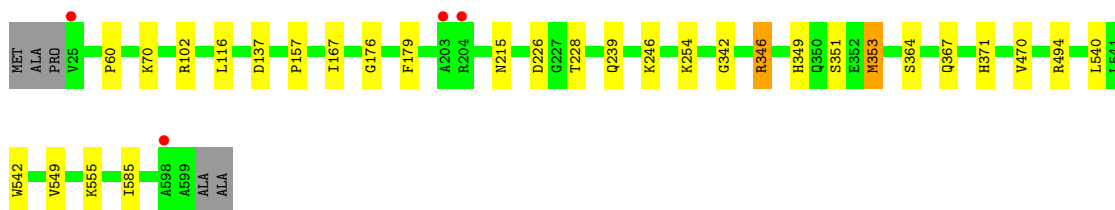
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	701	Total 701	O 701	0	0
10	B	715	Total 715	O 715	0	0

### 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.

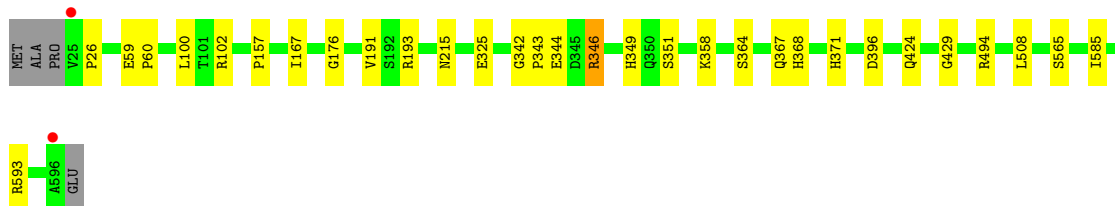
- Molecule 1: Exo-alpha-sialidase

Chain A:  94% 5%



- Molecule 2: Exo-alpha-sialidase

Chain B:  94% 5%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.07Å 51.78Å 118.44Å 90.00° 93.07° 90.00°	Depositor
Resolution (Å)	50.25 – 1.30 59.14 – 1.30	Depositor EDS
% Data completeness (in resolution range)	97.7 (50.25-1.30) 99.5 (59.14-1.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.14 (at 1.30Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.166 , 0.188 0.164 , 0.186	Depositor DCC
$R_{free}$ test set	15510 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.5	Xtrriage
Anisotropy	0.491	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 47.0	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	0.98	EDS
Total number of atoms	19936	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.25% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: DAN, CA, IMD, SO4, EDO, CL, SLB

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	0.58	0/4834	0.78	1/6552 (0.0%)
2	B	0.56	1/4799 (0.0%)	0.78	1/6506 (0.0%)
All	All	0.57	1/9633 (0.0%)	0.78	2/13058 (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	A	0	1
2	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	325	GLU	CB-CG	-6.18	1.40	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	353	MET	CB-CA-C	-6.72	96.96	110.40
2	B	396	ASP	CB-CG-OD1	5.36	123.13	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	346	ARG	Sidechain
2	B	346	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	4617	4552	4444	20	0
2	B	4587	4539	4436	19	2
3	A	40	32	34	0	0
3	B	20	16	17	0	0
4	A	5	5	5	1	0
5	A	12	18	18	3	0
5	B	12	18	18	0	0
6	A	2	0	0	0	0
6	B	2	0	0	3	1
7	A	1	0	0	0	0
7	B	1	0	0	0	0
8	A	5	0	0	0	0
9	B	20	16	16	0	0
10	A	701	0	0	4	0
10	B	715	0	0	6	0
All	All	10740	9196	8988	40	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (40) 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)
2:B:193:ARG:NH1	10:B:701:HOH:O	1.61	1.26
2:B:100:LEU:O	6:B:607:CL:CL	2.31	0.85
2:B:344:GLU:OE1	10:B:702:HOH:O	2.02	0.75
1:A:137[A]:ASP:OD2	10:A:801:HOH:O	2.13	0.67
2:B:102[B]:ARG:HG3	6:B:607:CL:CL	2.32	0.65
2:B:508[B]:LEU:HD21	2:B:565:SER:HB2	1.81	0.62
4:A:703:IMD:N3	10:A:802:HOH:O	2.29	0.60

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:228[B]:THR:HG21	1:A:246:LYS:NZ	2.19	0.57
1:A:116:LEU:HD22	5:A:704:EDO:H22	1.88	0.55
2:B:102[B]:ARG:CG	6:B:607:CL:CL	2.92	0.54
2:B:191:VAL:O	10:B:701:HOH:O	2.19	0.53
2:B:508[B]:LEU:CD2	2:B:565:SER:HB2	2.41	0.51
2:B:358:LYS:HE3	10:B:1288:HOH:O	2.11	0.51
1:A:116:LEU:HD22	5:A:704:EDO:C2	2.41	0.50
2:B:26:PRO:HG2	10:B:773:HOH:O	2.12	0.50
1:A:585:ILE:C	1:A:585:ILE:HD12	2.33	0.49
2:B:59:GLU:H	2:B:59:GLU:CD	2.18	0.47
2:B:157:PRO:HA	2:B:167[A]:ILE:O	2.16	0.46
1:A:228[B]:THR:HG22	1:A:246:LYS:HD3	1.99	0.45
1:A:470:VAL:HG23	1:A:542:TRP:HZ3	1.82	0.45
1:A:228[B]:THR:HG21	1:A:246:LYS:HZ2	1.80	0.45
1:A:226:ASP:OD1	1:A:228[B]:THR:HG23	2.17	0.45
1:A:470:VAL:HG23	1:A:542:TRP:CZ3	2.51	0.45
2:B:367:GLN:O	2:B:371:HIS:HB2	2.17	0.44
2:B:176:GLY:HA2	2:B:215:ASN:O	2.18	0.44
1:A:351:SER:HA	1:A:364:SER:O	2.18	0.43
1:A:176:GLY:HA2	1:A:215:ASN:O	2.18	0.43
1:A:342[B]:GLY:O	1:A:346:ARG:NH2	2.30	0.43
2:B:342[A]:GLY:O	2:B:346:ARG:NH2	2.33	0.43
2:B:424:GLN:NE2	2:B:429:GLY:O	2.44	0.43
2:B:60:PRO:HG2	10:B:855:HOH:O	2.19	0.42
1:A:254:LYS:HD2	5:A:705:EDO:O1	2.20	0.42
1:A:70:LYS:HE3	10:A:826:HOH:O	2.19	0.42
1:A:540:LEU:HD23	1:A:549:VAL:HG22	2.02	0.41
1:A:157:PRO:HA	1:A:167[B]:ILE:O	2.20	0.41
1:A:167[B]:ILE:HG23	1:A:179:PHE:HB3	2.01	0.41
1:A:367:GLN:O	1:A:371:HIS:HB2	2.21	0.41
1:A:60:PRO:HG2	10:A:1001:HOH:O	2.20	0.41
2:B:351:SER:HA	2:B:364:SER:O	2.21	0.41
2:B:585:ILE:HD12	2:B:585:ILE:C	2.42	0.40

All (2) 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)
2:B:102[A]:ARG:HE	2:B:193:ARG:HH22[2_557]	1.09	0.51
2:B:193:ARG:NH2	6:B:607:CL:CL[2_547]	2.05	0.15

## 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	596/578 (103%)	576 (97%)	20 (3%)	0	100	100
2	B	593/576 (103%)	573 (97%)	18 (3%)	2 (0%)	41	17
All	All	1189/1154 (103%)	1149 (97%)	38 (3%)	2 (0%)	47	19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	368	HIS
2	B	343	PRO

### 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	502/481 (104%)	495 (99%)	7 (1%)	67	34
2	B	499/482 (104%)	496 (99%)	3 (1%)	86	65
All	All	1001/963 (104%)	991 (99%)	10 (1%)	78	48

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

Mol	Chain	Res	Type
1	A	102[A]	ARG
1	A	102[B]	ARG
1	A	239	GLN
1	A	349	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	353	MET
1	A	494	ARG
1	A	555	LYS
2	B	349	HIS
2	B	494	ARG
2	B	593	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	239	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](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 6 are monoatomic - leaving 12 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	SLB	A	702	-	20,20,21	5.31	10 (50%)	24,28,31	2.38	6 (25%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	EDO	B	603	-	3,3,3	0.79	0	2,2,2	0.32	0
9	DAN	B	602	-	20,20,20	6.28	10 (50%)	23,28,28	1.48	4 (17%)
5	EDO	B	604	-	3,3,3	0.56	0	2,2,2	0.39	0
5	EDO	B	605	-	3,3,3	0.37	0	2,2,2	0.66	0
8	SO4	A	710	-	4,4,4	0.15	0	6,6,6	0.23	0
5	EDO	A	706	-	3,3,3	0.45	0	2,2,2	0.49	0
5	EDO	A	705	-	3,3,3	0.50	0	2,2,2	0.21	0
3	SLB	A	701	-	20,20,21	4.40	7 (35%)	24,28,31	2.07	6 (25%)
5	EDO	A	704	-	3,3,3	0.61	0	2,2,2	0.25	0
3	SLB	B	601	-	20,20,21	4.19	9 (45%)	24,28,31	2.05	4 (16%)
4	IMD	A	703	-	3,5,5	0.68	0	4,5,5	0.27	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	SLB	A	702	-	-	0/18/34/38	0/1/1/1
5	EDO	B	603	-	-	0/1/1/1	-
9	DAN	B	602	-	-	2/18/34/34	0/1/1/1
5	EDO	B	604	-	-	0/1/1/1	-
5	EDO	B	605	-	-	1/1/1/1	-
5	EDO	A	706	-	-	0/1/1/1	-
5	EDO	A	705	-	-	1/1/1/1	-
3	SLB	A	701	-	-	0/18/34/38	0/1/1/1
5	EDO	A	704	-	-	0/1/1/1	-
3	SLB	B	601	-	-	0/18/34/38	0/1/1/1
4	IMD	A	703	-	-	-	0/1/1/1

All (36) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	602	DAN	C3-C2	19.20	1.63	1.33
3	A	702	SLB	C4-C5	14.53	1.65	1.53
3	A	701	SLB	C4-C5	10.98	1.62	1.53
9	B	602	DAN	C4-C5	10.63	1.65	1.53
3	A	702	SLB	O6-C6	10.30	1.60	1.44
3	B	601	SLB	C4-C5	9.88	1.61	1.53
9	B	602	DAN	C6-C5	-9.71	1.37	1.53
3	A	701	SLB	C6-C5	-9.45	1.37	1.53

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	702	SLB	C6-C5	-9.38	1.37	1.53
3	B	601	SLB	C6-C5	-8.97	1.38	1.53
9	B	602	DAN	O6-C6	8.79	1.60	1.46
3	B	601	SLB	O6-C6	8.24	1.56	1.44
9	B	602	DAN	C4-C3	-7.85	1.40	1.50
3	A	701	SLB	O6-C6	7.24	1.55	1.44
3	A	701	SLB	C3-C4	-7.14	1.40	1.52
3	A	702	SLB	C3-C4	-6.92	1.40	1.52
3	B	601	SLB	C3-C4	-6.34	1.41	1.52
3	A	701	SLB	O6-C2	-5.62	1.36	1.43
3	A	702	SLB	C3-C2	5.41	1.61	1.52
9	B	602	DAN	C10-N5	5.39	1.52	1.34
3	A	702	SLB	C10-N5	5.08	1.51	1.34
3	B	601	SLB	O6-C2	-5.01	1.37	1.43
3	A	702	SLB	O6-C2	-4.89	1.37	1.43
9	B	602	DAN	C5-N5	4.16	1.52	1.45
3	B	601	SLB	C10-N5	3.79	1.47	1.34
3	A	702	SLB	C5-N5	3.55	1.51	1.45
3	A	701	SLB	C10-N5	3.47	1.46	1.34
3	A	701	SLB	C2-C1	-2.80	1.50	1.52
9	B	602	DAN	C2-C1	2.67	1.55	1.48
3	B	601	SLB	C3-C2	2.66	1.56	1.52
9	B	602	DAN	C11-C10	2.60	1.55	1.50
3	B	601	SLB	C5-N5	2.49	1.49	1.45
3	A	702	SLB	C11-C10	2.36	1.55	1.50
3	A	702	SLB	O8-C8	-2.11	1.38	1.43
9	B	602	DAN	O8-C8	-2.09	1.38	1.43
3	B	601	SLB	O1B-C1	-2.04	1.23	1.30

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	702	SLB	O6-C2-C3	8.26	121.83	110.46
3	B	601	SLB	O6-C2-C3	7.02	120.11	110.46
3	A	701	SLB	O6-C2-C3	5.83	118.47	110.46
3	A	702	SLB	C4-C3-C2	4.28	117.48	109.81
3	B	601	SLB	O6-C2-C1	3.85	115.26	107.70
3	A	702	SLB	O6-C2-C1	3.64	114.84	107.70
3	A	701	SLB	C4-C5-N5	-3.62	103.21	110.38
3	A	701	SLB	C4-C3-C2	3.54	116.15	109.81
3	A	701	SLB	O6-C2-C1	3.39	114.35	107.70
3	B	601	SLB	C4-C3-C2	3.38	115.86	109.81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	601	SLB	C4-C5-N5	-3.23	103.98	110.38
9	B	602	DAN	O6-C2-C1	3.16	118.38	112.06
3	A	701	SLB	C6-C5-N5	2.91	115.75	110.91
3	A	701	SLB	C5-N5-C10	-2.45	117.21	123.18
3	A	702	SLB	C5-N5-C10	-2.43	117.27	123.18
9	B	602	DAN	C11-C10-N5	2.40	120.17	116.10
3	A	702	SLB	C11-C10-N5	2.21	119.84	116.10
9	B	602	DAN	C4-C3-C2	-2.12	118.01	121.60
3	A	702	SLB	O4-C4-C3	-2.06	104.84	109.94
9	B	602	DAN	C3-C2-C1	-2.04	119.25	123.65

There are no chirality outliers.

All (4) torsion outliers are listed below:

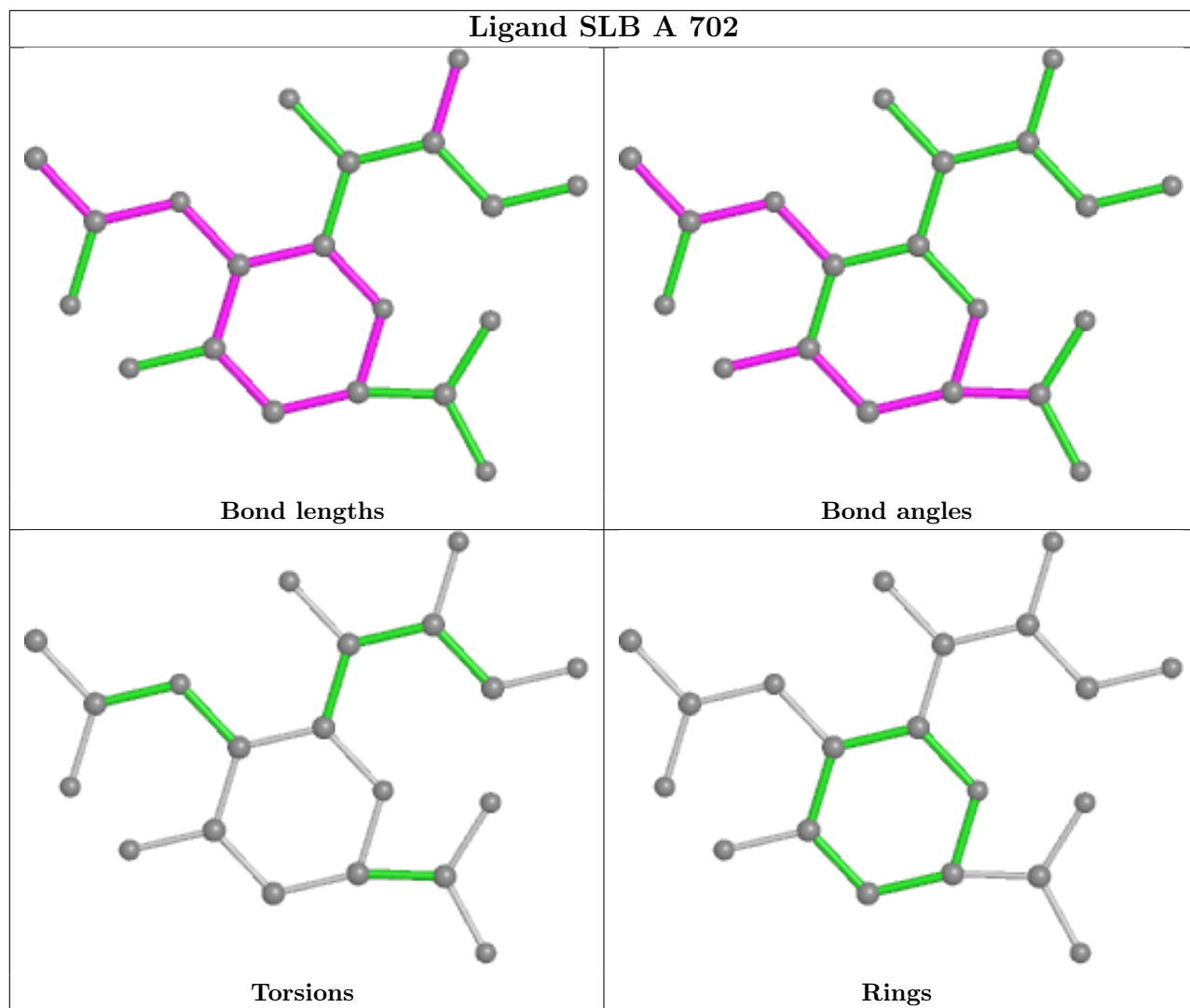
Mol	Chain	Res	Type	Atoms
9	B	602	DAN	O8-C8-C9-O9
9	B	602	DAN	C7-C8-C9-O9
5	B	605	EDO	O1-C1-C2-O2
5	A	705	EDO	O1-C1-C2-O2

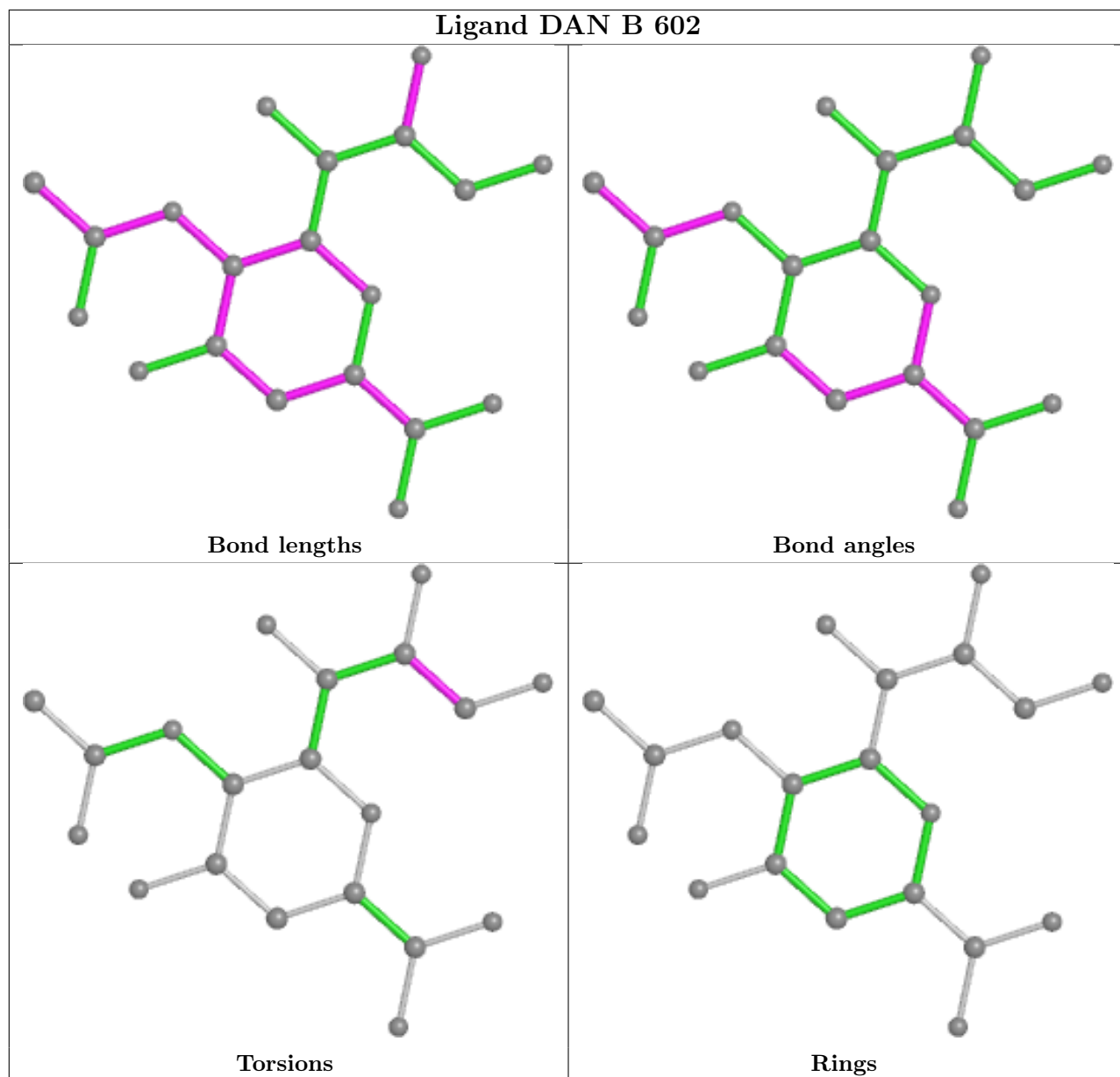
There are no ring outliers.

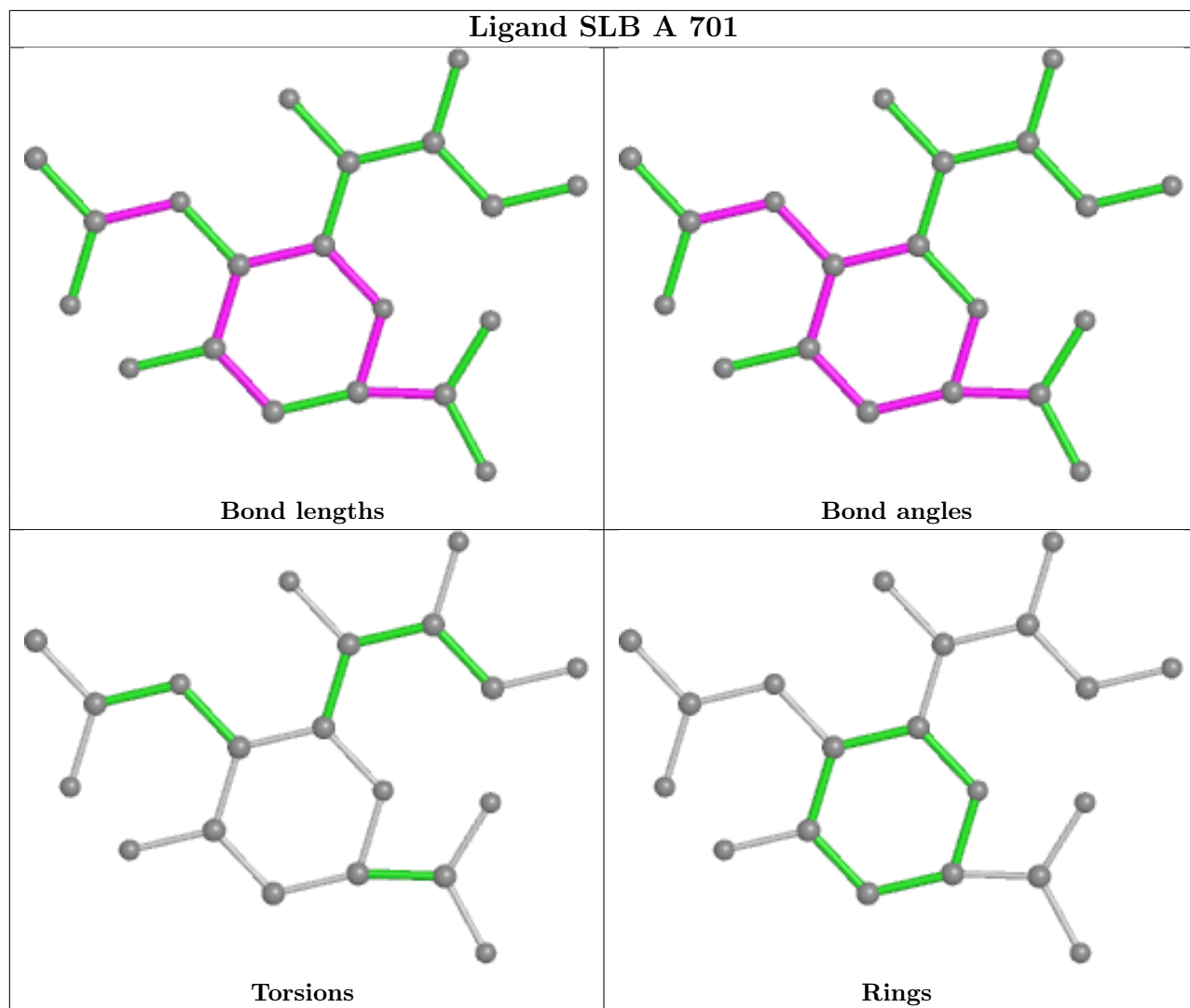
3 monomers are involved in 4 short contacts:

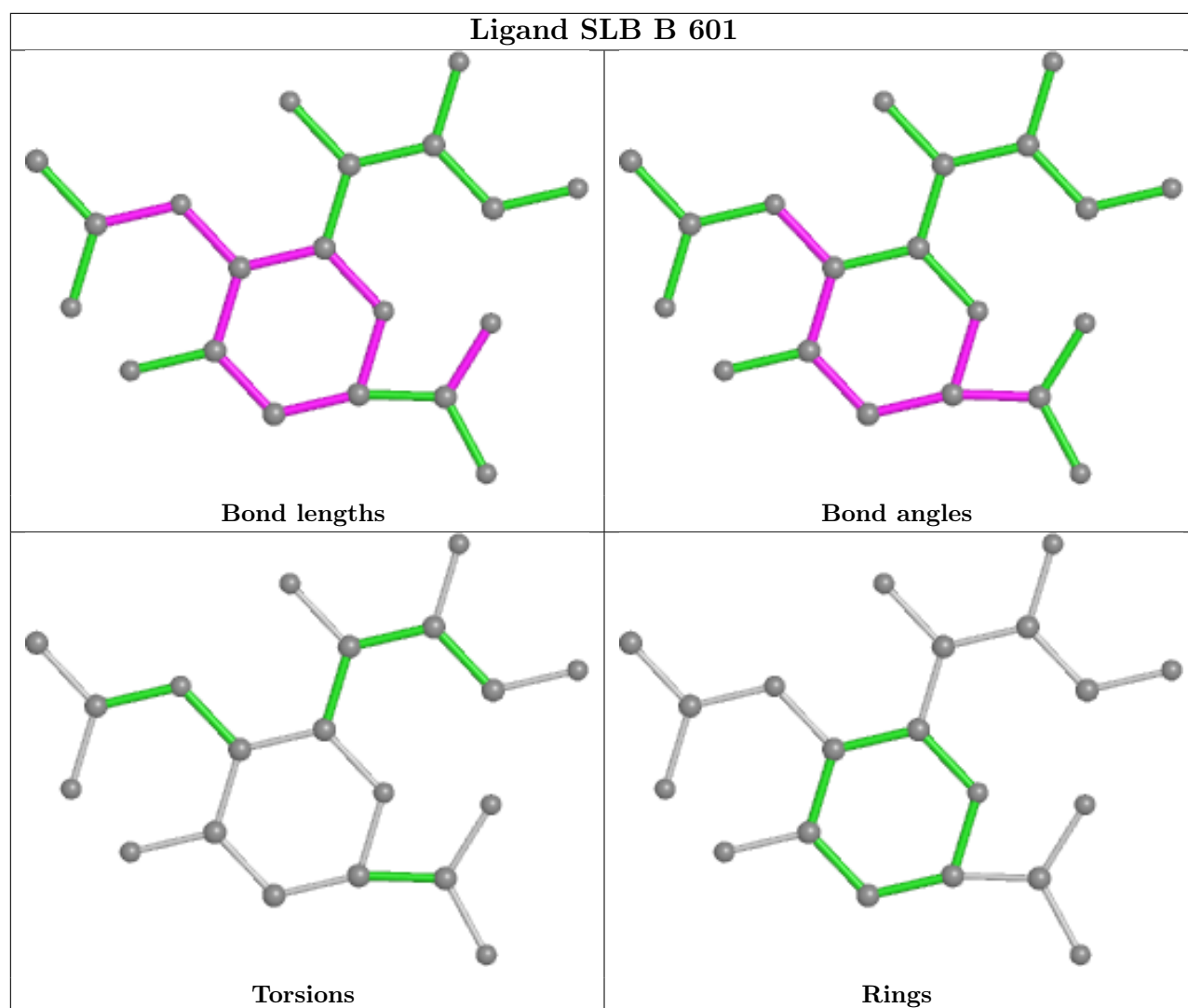
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	705	EDO	1	0
5	A	704	EDO	2	0
4	A	703	IMD	1	0

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	595:LYS	C	598:ALA	N	3.08

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

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

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	573/578 (99%)	-0.15	4 (0%) 87 87	14, 20, 38, 64	1 (0%)
2	B	572/576 (99%)	-0.19	2 (0%) 94 93	13, 20, 36, 56	0
All	All	1145/1154 (99%)	-0.17	6 (0%) 91 91	13, 20, 37, 64	1 (0%)

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	596	ALA	6.3
1	A	598	ALA	3.7
2	B	25	VAL	3.5
1	A	25	VAL	3.0
1	A	203	ALA	2.8
1	A	204	ARG	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](#)

There are no monosaccharides in this entry.

### 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.

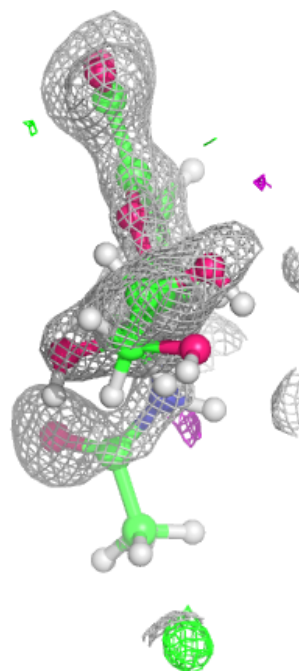
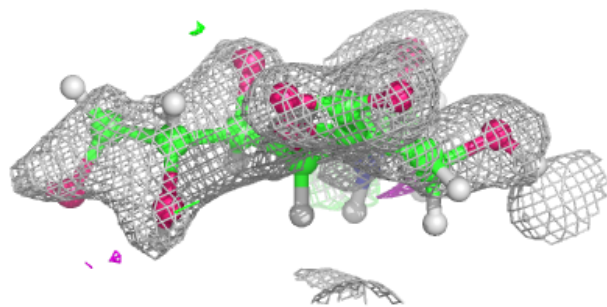
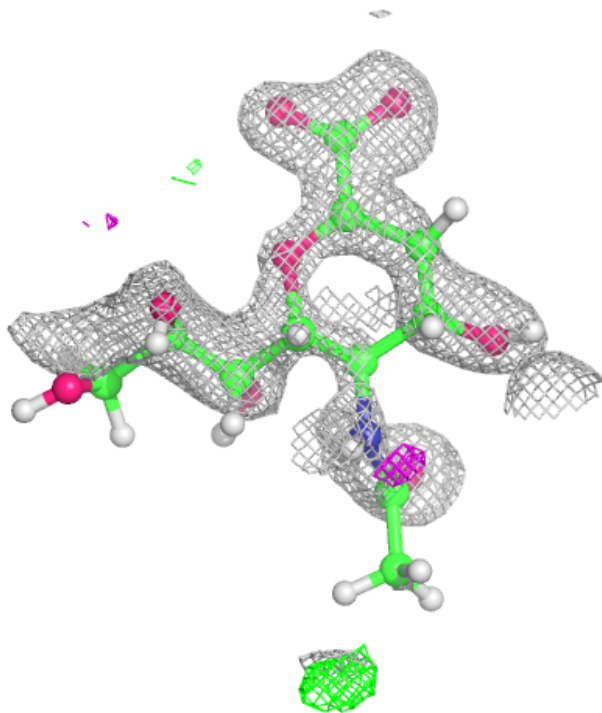


Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	EDO	A	705	4/4	0.67	0.35	50,60,63,63	0
9	DAN	B	602	20/20	0.72	0.24	36,59,77,95	0
5	EDO	A	706	4/4	0.74	0.10	34,43,55,55	0
6	CL	B	607	1/1	0.75	2.48	86,86,86,86	0
4	IMD	A	703	5/5	0.76	0.24	22,28,49,59	0
8	SO4	A	710	5/5	0.79	0.27	51,51,68,90	0
5	EDO	B	605	4/4	0.79	0.32	28,42,51,54	0
3	SLB	A	702	20/21	0.80	0.18	27,49,72,81	0
5	EDO	B	604	4/4	0.86	0.10	19,24,33,37	0
5	EDO	B	603	4/4	0.91	0.11	23,27,34,34	0
5	EDO	A	704	4/4	0.92	0.13	19,44,54,54	0
3	SLB	A	701	20/21	0.97	0.07	13,16,21,21	0
3	SLB	B	601	20/21	0.97	0.07	14,18,26,26	0
6	CL	B	606	1/1	0.99	0.10	20,20,20,20	0
6	CL	A	708	1/1	0.99	0.05	29,29,29,29	0
7	CA	B	608	1/1	1.00	0.10	15,15,15,15	0
6	CL	A	707	1/1	1.00	0.14	16,16,16,16	0
7	CA	A	709	1/1	1.00	0.08	16,16,16,16	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.

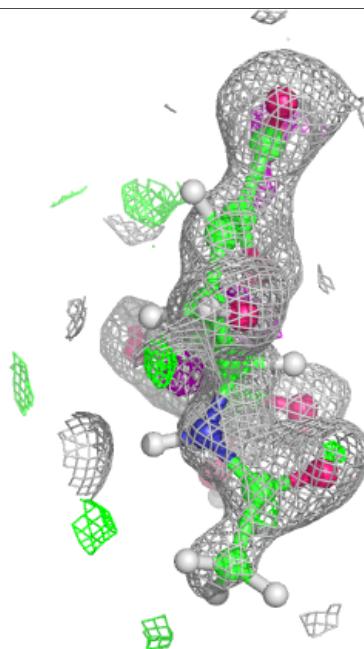
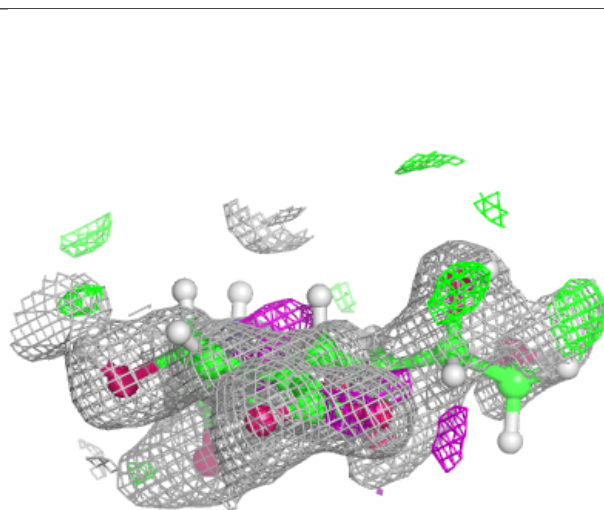
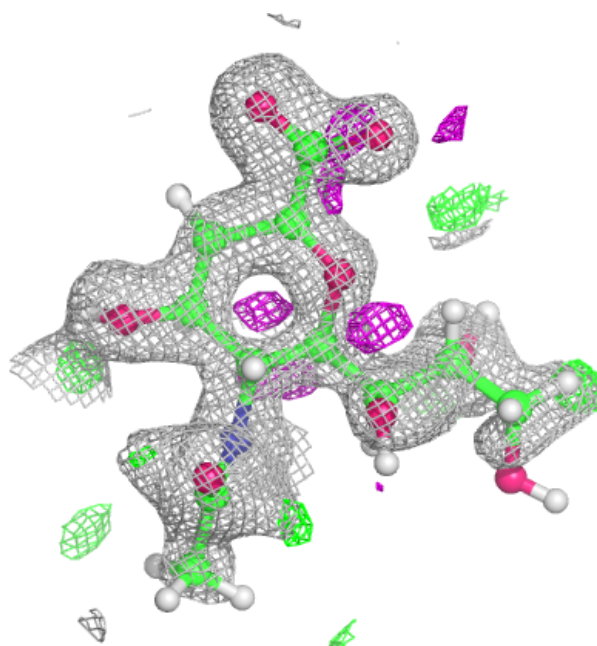
**Electron density around DAN B 602:**

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



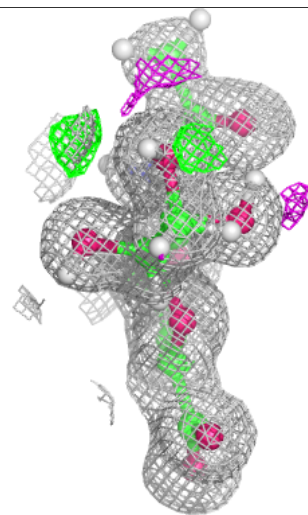
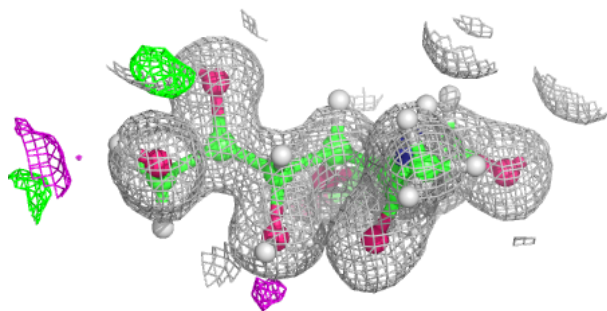
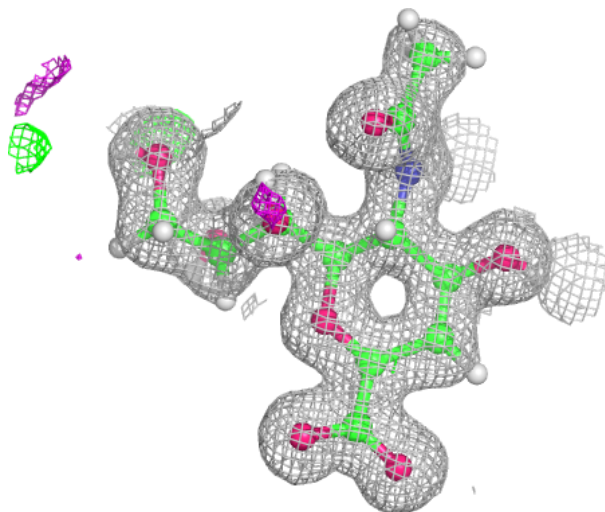
**Electron density around SLB A 702:**

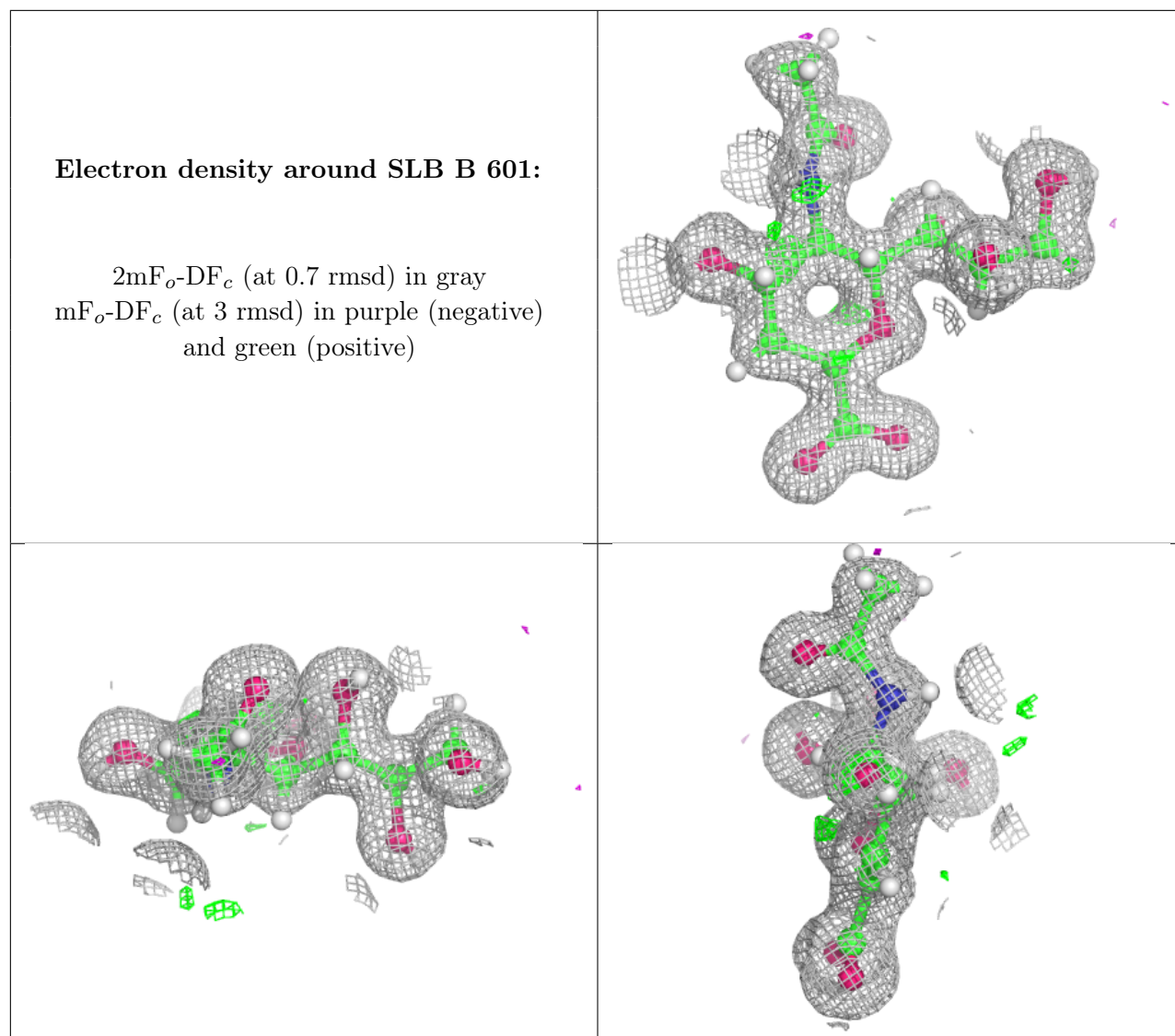
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around SLB A 701:**

$2mF_o-DF_c$  (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.