wwPDB X-ray Structure Validation Summary Report (i)

Jul 6, 2022 – 04:13 am BST

PDB ID	:	7P2Z
Title	:	Crystal structure of human lysosomal acid-alpha-glucosidase, GAA, in complex with cyclosulfamidate 4
Authors	:	Roig-Zamboni, V.; Kok, K.; Overkleeft, H.; Artola, M.; Sulzenbacher, G.
Deposited on	:	2021-07-06
Resolution	:	1.85 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

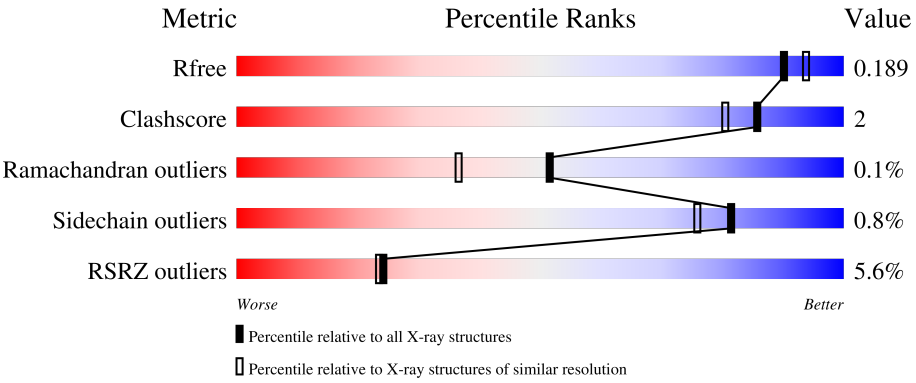
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

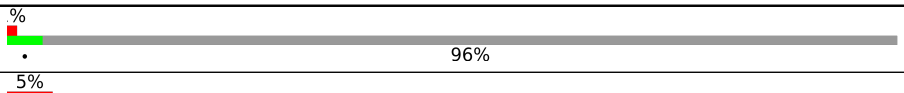
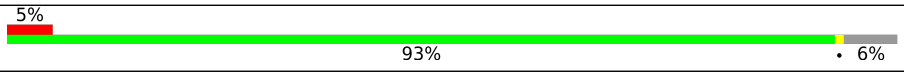
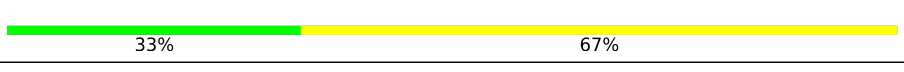
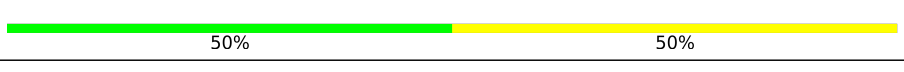
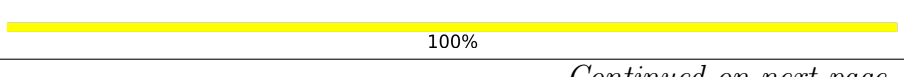
The reported resolution of this entry is 1.85 Å.

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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	AAA	872	
1	AaA	872	
2	BBB	3	
3	BcB	2	
3	BeB	2	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	BgB	3	33% 67%
5	BjB	2	50% 50%

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
10	YTW	AaA	1002	X	-	-	-
2	NAG	BBB	2	-	-	-	X
3	NAG	BeB	2	-	-	-	X
5	FUC	BjB	2	-	-	-	X

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 7791 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.

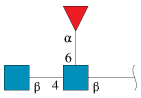
- Molecule 1 is a protein called Lysosomal alpha-glucosidase.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
1	AAA	35	Total C N O S 267 161 48 53 5	0	0	0
1	AaA	816	Total C N O S 6481 4168 1083 1201 29	0	10	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	199	ARG	HIS	conflict	UNP P10253
AAA	223	HIS	ARG	conflict	UNP P10253
AAA	780	ILE	VAL	conflict	UNP P10253
AaA	199	ARG	HIS	conflict	UNP P10253
AaA	223	HIS	ARG	conflict	UNP P10253
AaA	780	ILE	VAL	conflict	UNP P10253

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.

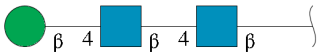


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	BBB	3	Total C N O 38 22 2 14	0	0	0

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

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

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

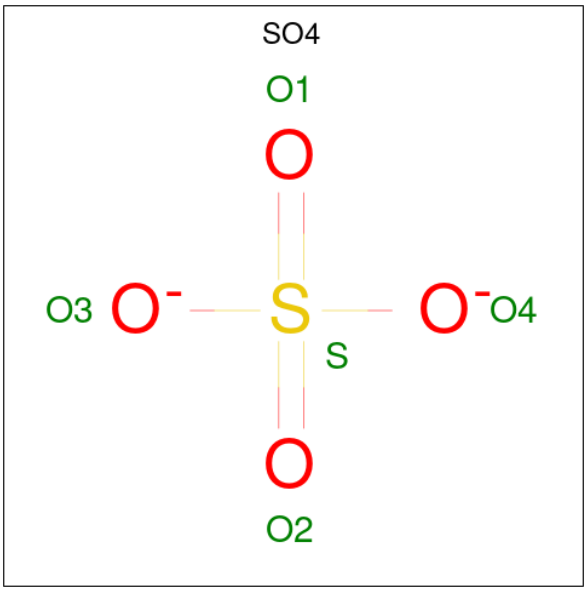


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
4	BgB	3	Total C N O 39 22 2 15	0	0	0

- Molecule 5 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	BjB	2	Total C N O 24 14 1 9	0	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).

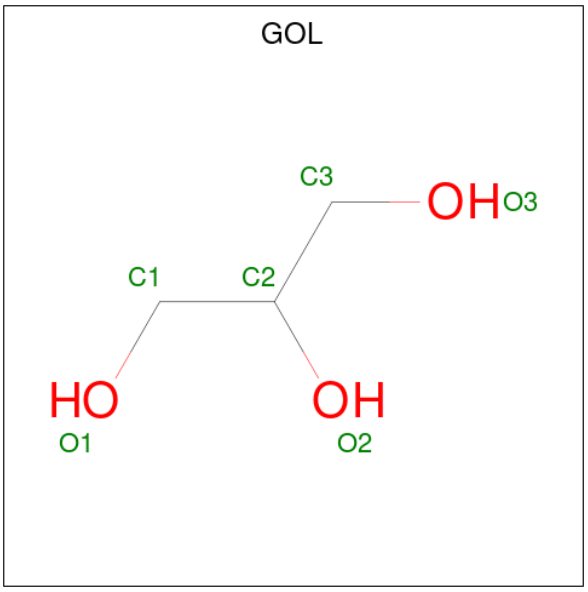


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	1	Total O S 5 4 1	0	0
6	AaA	1	Total O S 5 4 1	0	0
6	AaA	1	Total O S 5 4 1	0	0
6	AaA	1	Total O S 5 4 1	0	0
6	AaA	1	Total O S 5 4 1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

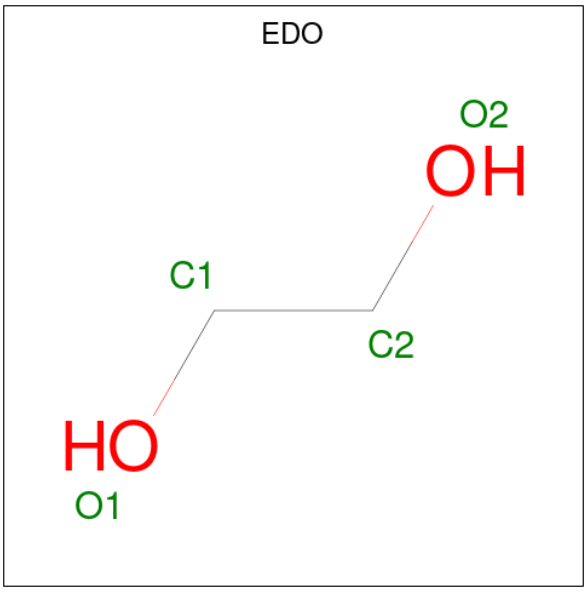
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	AAA	2	Total Cl 2 2	0	0
7	AaA	7	Total Cl 7 7	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$) (labeled as "Ligand of Interest" by depositor).



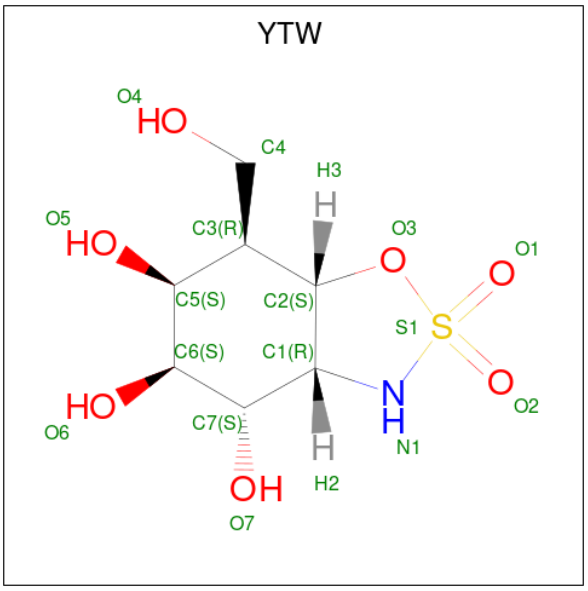
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	AAA	1	Total	C	O	0	0
			6	3	3		
8	AaA	1	Total	C	O	0	0
			6	3	3		
8	AaA	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$) (labeled as "Ligand of Interest" by depositor).



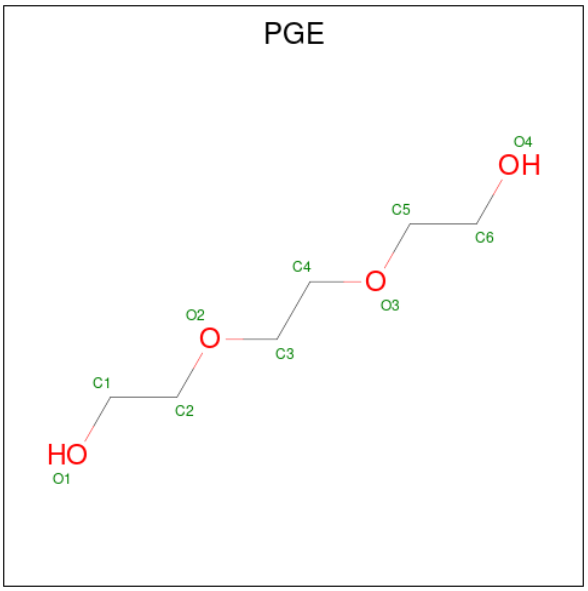
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	AAA	1	Total C O 4 2 2	0	0
9	AaA	1	Total C O 4 2 2	0	0
9	AaA	1	Total C O 4 2 2	0	0
9	AaA	1	Total C O 4 2 2	0	0
9	AaA	1	Total C O 4 2 2	0	0

- Molecule 10 is (3 {a} {R},4 {S},5 {S},6 {S},7 {R},7 {a} {S})-7-(hydroxymethyl)-2,2-bis(oxidanylidene)-3 {a},4,5,6,7,7 {a}-hexahydro-3 {H}-benzo[d][1,2,3]oxathiazole-4,5,6-triol (three-letter code: YTW) (formula: $C_7H_{13}NO_7S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	AaA	1	Total C N O S 16 7 1 7 1	0	0

- Molecule 11 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	AaA	1	Total C O 10 6 4	0	0
11	AaA	1	Total C O 10 6 4	0	0

- Molecule 12 is water.

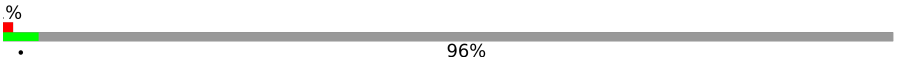
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	AAA	29	Total O 29 29	0	0
12	AaA	749	Total O 749 749	0	0

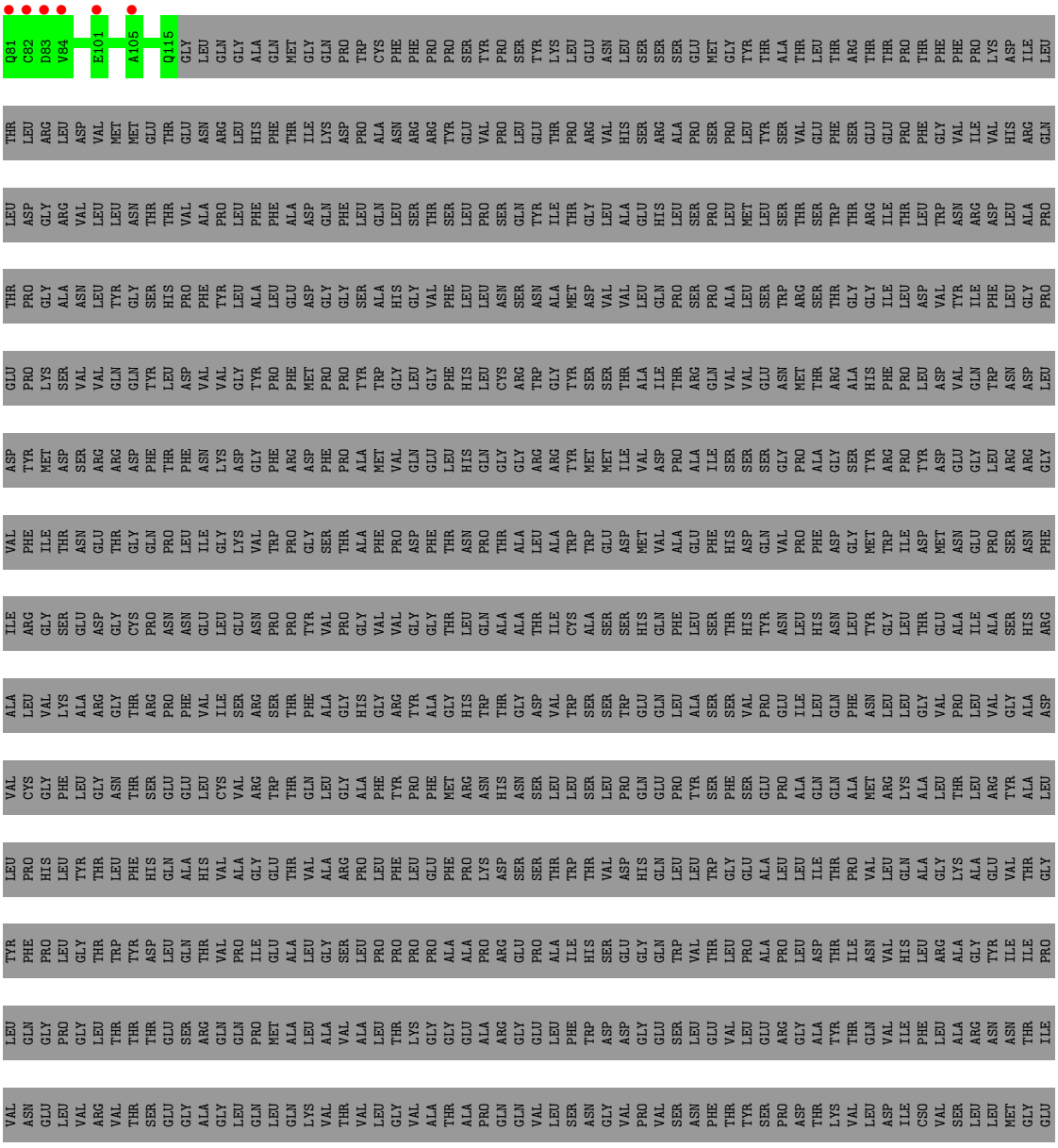
3 Residue-property plots (i)

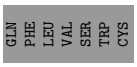
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Lysosomal alpha-glucosidase

Chain AAA:

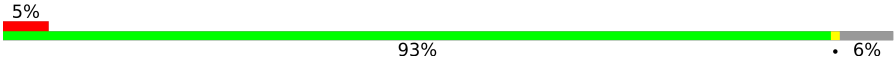


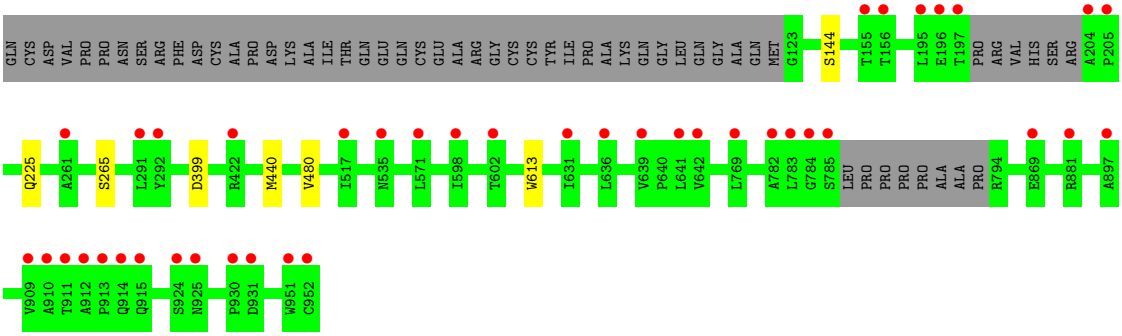




• Molecule 1: Lysosomal alpha-glucosidase

Chain AaA:





• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

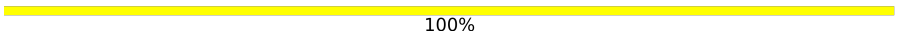
Chain BBB:

• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain BcB:

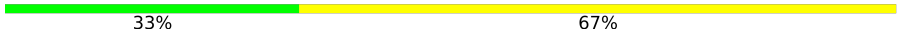
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain BeB:



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

Chain BgB:



• Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain BjB:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	97.16Å 102.54Å 129.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.00 – 1.85 47.65 – 1.85	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.00-1.85) 100.0 (47.65-1.85)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.19 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0266	Depositor
R, R_{free}	0.148 , 0.175 0.162 , 0.189	Depositor DCC
R_{free} test set	5503 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	27.7	Xtriage
Anisotropy	0.347	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	7791	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, FUC, SO4, PGE, CSO, YTW, CL, EDO, NAG, 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	AAA	0.64	0/272	0.78	0/368
1	AaA	0.64	0/6688	0.76	0/9140
All	All	0.64	0/6960	0.76	0/9508

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	AAA	267	0	245	0	0
1	AaA	6481	0	6284	0	0
2	BBB	38	0	34	0	0
3	BcB	28	0	25	0	0
3	BeB	28	0	25	0	0
4	BgB	39	0	34	0	0
5	BjB	24	0	22	0	0
6	AAA	5	0	0	0	0
6	AaA	20	0	0	0	0
7	AAA	2	0	0	0	0
7	AaA	7	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	AAA	6	0	8	0	0
8	AaA	12	0	16	0	0
9	AAA	4	0	6	0	0
9	AaA	16	0	24	0	0
10	AaA	16	0	0	0	0
11	AaA	20	0	28	0	0
12	AAA	29	0	0	0	0
12	AaA	749	0	0	0	0
All	All	7791	0	6751	0	0

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

There are no clashes within the asymmetric unit.

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	AAA	33/872 (4%)	33 (100%)	0	0	100 100
1	AaA	819/872 (94%)	795 (97%)	23 (3%)	1 (0%)	51 36
All	All	852/1744 (49%)	828 (97%)	23 (3%)	1 (0%)	51 36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AaA	480	VAL

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	AAA	30/741 (4%)	30 (100%)	0	100	100
1	AaA	705/741 (95%)	698 (99%)	7 (1%)	76	69
All	All	735/1482 (50%)	728 (99%)	7 (1%)	81	69

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AaA	265	SER
1	AaA	399	ASP
1	AaA	613	TRP
1	AaA	440	MET
1	AaA	225	GLN

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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	CSO	AaA	938	1	3,6,7	0.65	0	0,6,8	-	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSO	AaA	938	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

12 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	BBB	1	1,2	14,14,15	0.69	0	17,19,21	1.12	1 (5%)
2	NAG	BBB	2	2	14,14,15	0.55	0	17,19,21	0.75	0
2	FUC	BBB	3	2	10,10,11	0.57	0	14,14,16	1.10	1 (7%)
3	NAG	BcB	1	3,1	14,14,15	0.62	0	17,19,21	1.09	0
3	NAG	BcB	2	3	14,14,15	0.46	0	17,19,21	1.57	2 (11%)
3	NAG	BeB	1	3,1	14,14,15	0.57	0	17,19,21	1.45	2 (11%)
3	NAG	BeB	2	3	14,14,15	0.42	0	17,19,21	1.56	3 (17%)
4	NAG	BgB	1	4,1	14,14,15	0.66	0	17,19,21	1.23	1 (5%)
4	NAG	BgB	2	4	14,14,15	0.55	0	17,19,21	0.88	0
4	BMA	BgB	3	4	11,11,12	0.49	0	15,15,17	1.35	1 (6%)
5	NAG	BjB	1	1,5	14,14,15	0.72	0	17,19,21	1.06	0
5	FUC	BjB	2	5	10,10,11	0.49	0	14,14,16	1.51	3 (21%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	BBB	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	BBB	2	2	-	2/6/23/26	0/1/1/1
2	FUC	BBB	3	2	-	-	0/1/1/1
3	NAG	BcB	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	BcB	2	3	-	0/6/23/26	0/1/1/1
3	NAG	BeB	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	BeB	2	3	-	0/6/23/26	0/1/1/1
4	NAG	BgB	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	BgB	2	4	-	0/6/23/26	0/1/1/1
4	BMA	BgB	3	4	-	2/2/19/22	0/1/1/1
5	NAG	BjB	1	1,5	-	0/6/23/26	0/1/1/1
5	FUC	BjB	2	5	-	-	0/1/1/1

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	BeB	2	NAG	C1-O5-C5	4.17	117.84	112.19
4	BgB	3	BMA	O5-C5-C6	3.79	113.15	107.20
3	BeB	1	NAG	O5-C1-C2	-3.72	105.41	111.29
2	BBB	1	NAG	O5-C1-C2	-3.54	105.70	111.29
3	BcB	2	NAG	C1-C2-N2	-3.25	104.93	110.49

There are no chirality outliers.

All (4) torsion outliers are listed below:

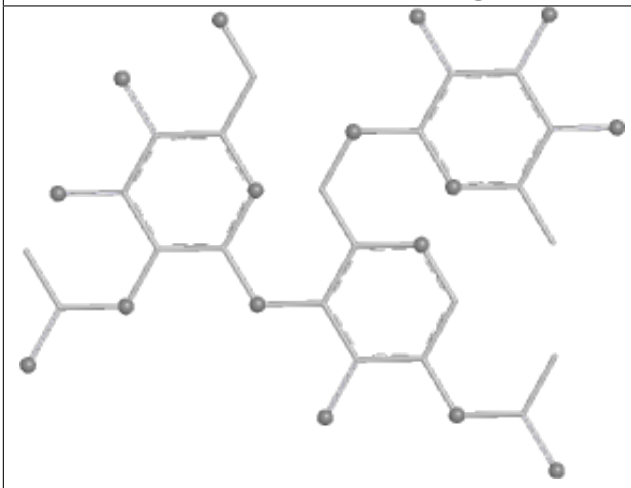
Mol	Chain	Res	Type	Atoms
4	BgB	3	BMA	O5-C5-C6-O6
4	BgB	3	BMA	C4-C5-C6-O6
2	BBB	2	NAG	C4-C5-C6-O6
2	BBB	2	NAG	O5-C5-C6-O6

There are no ring outliers.

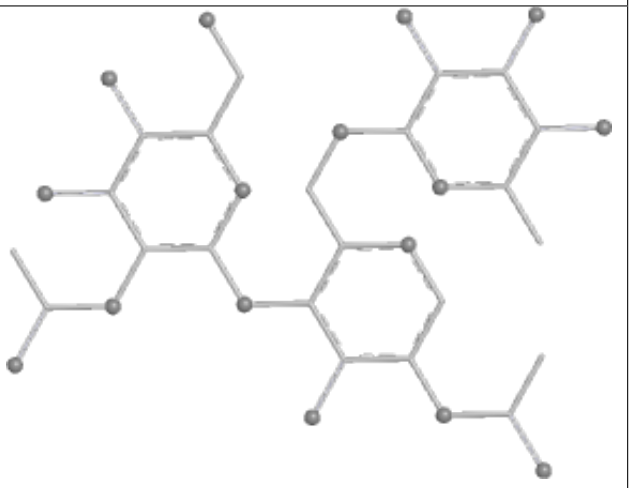
No monomer is involved in short contacts.

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

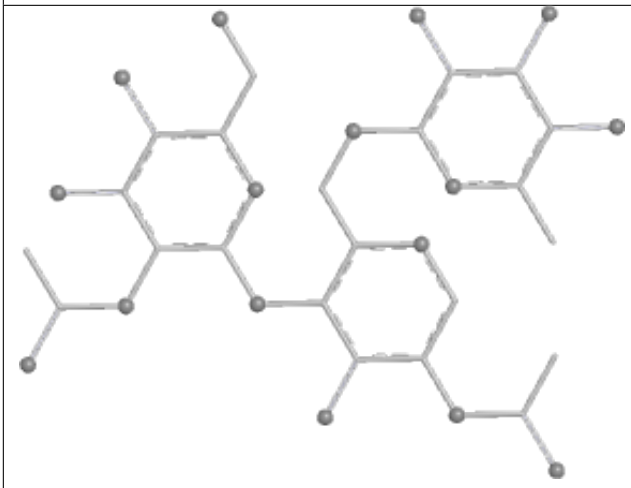
Oligosaccharide Chain BBB



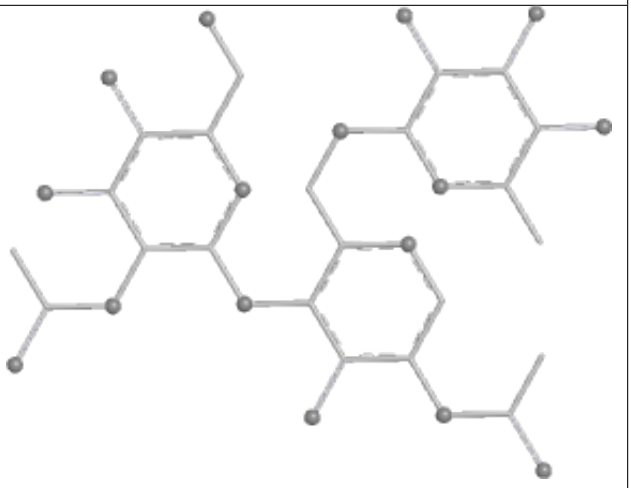
Bond lengths



Bond angles

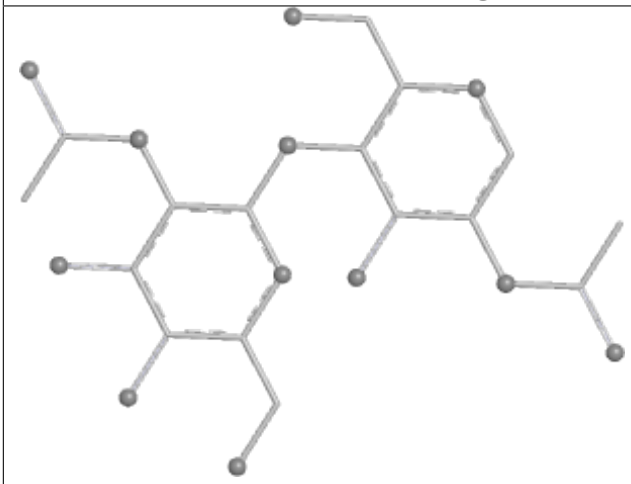


Torsions

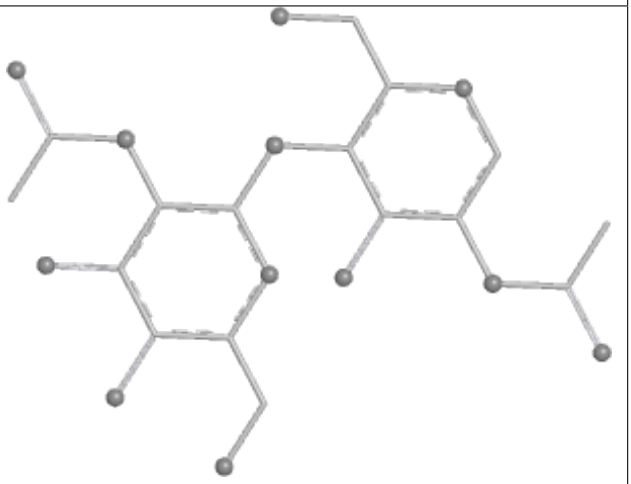


Rings

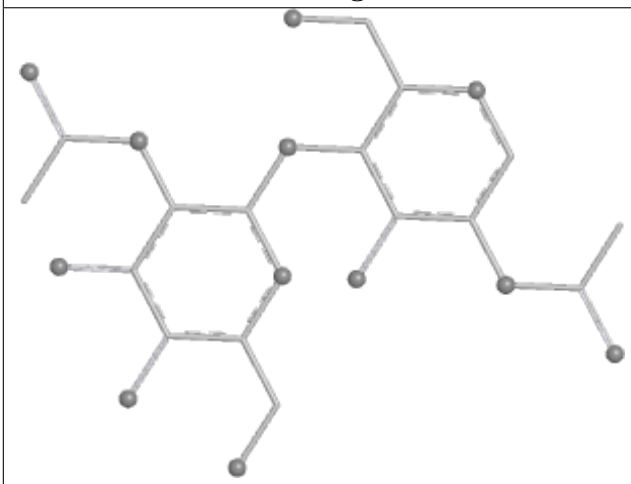
Oligosaccharide Chain BcB



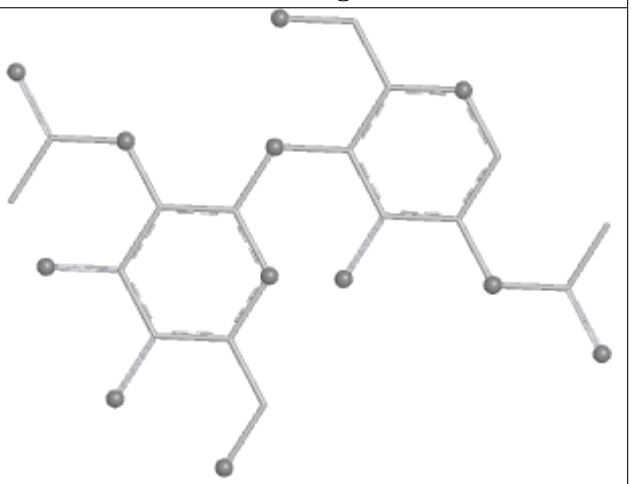
Bond lengths



Bond angles

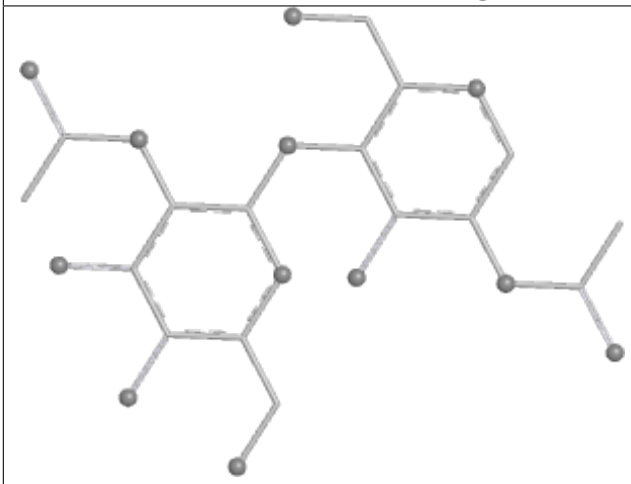


Torsions

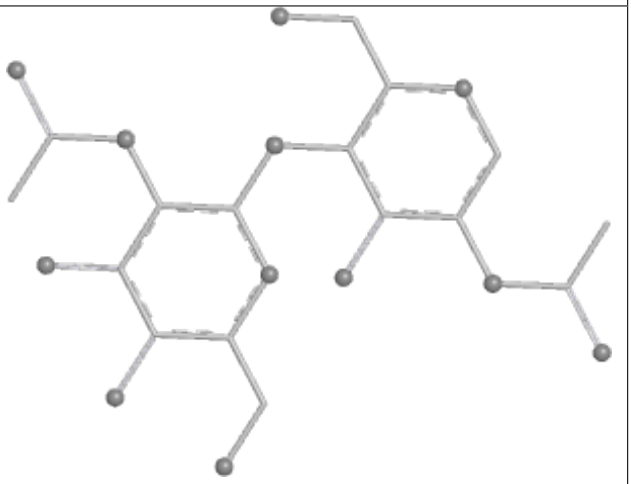


Rings

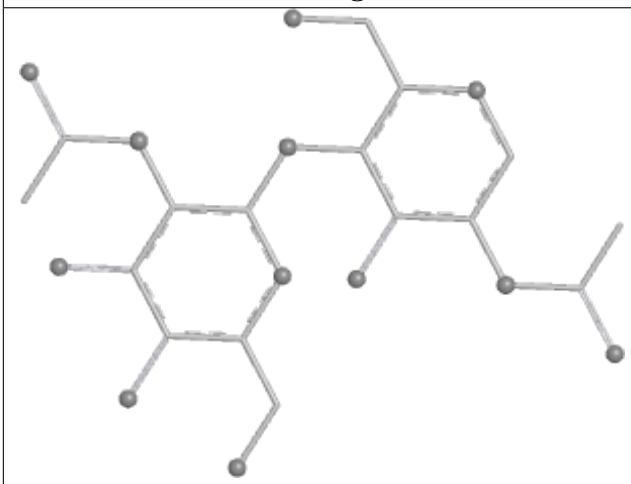
Oligosaccharide Chain BeB



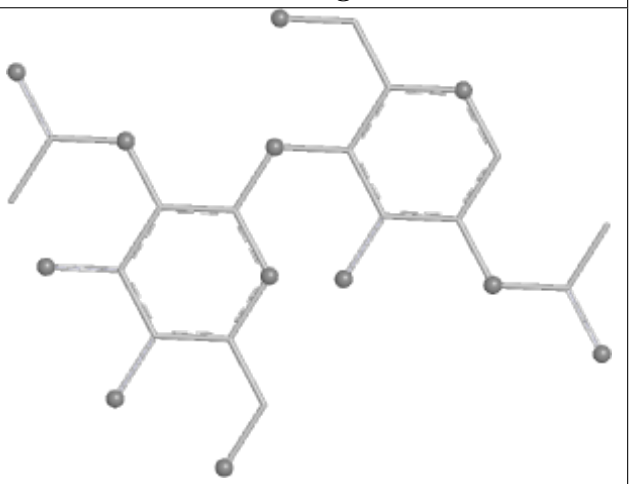
Bond lengths



Bond angles

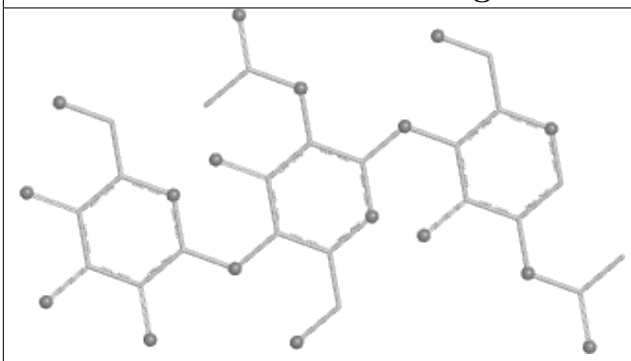


Torsions

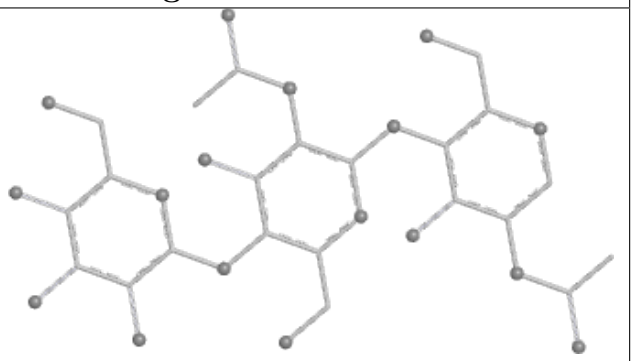


Rings

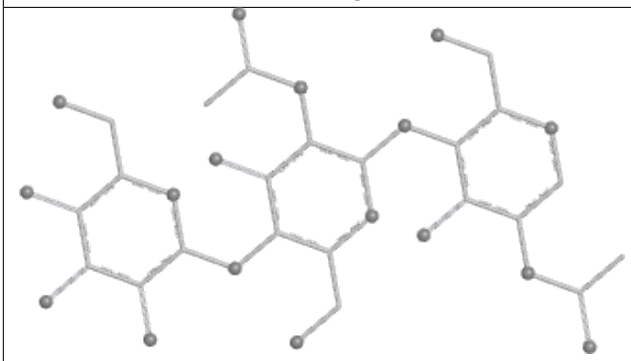
Oligosaccharide Chain BgB



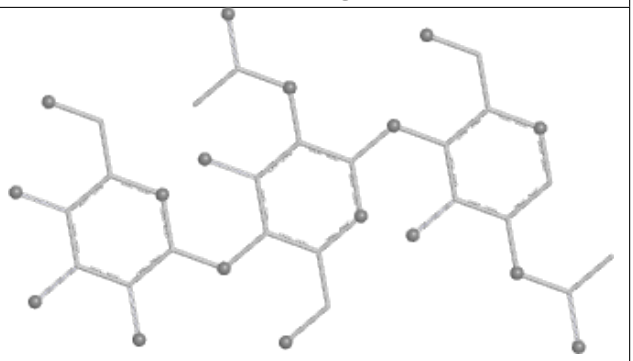
Bond lengths



Bond angles

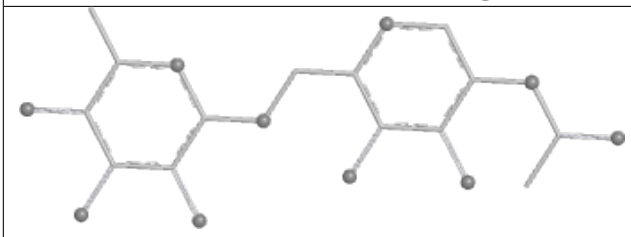


Torsions

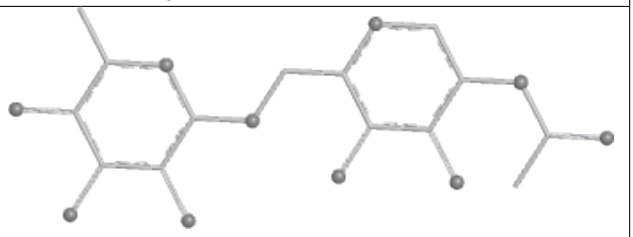


Rings

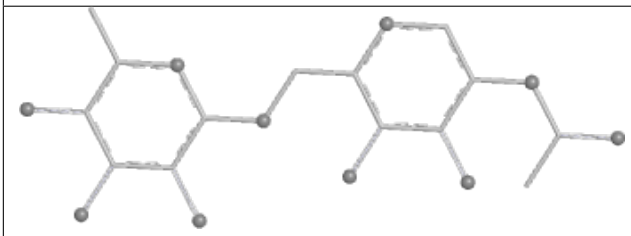
Oligosaccharide Chain BjB



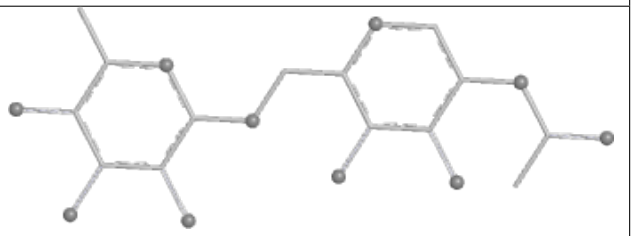
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 25 ligands modelled in this entry, 9 are monoatomic - leaving 16 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond

length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	SO4	AAA	1001	-	4,4,4	0.43	0	6,6,6	0.11	0
9	EDO	AaA	1001	-	3,3,3	0.19	0	2,2,2	0.61	0
11	PGE	AaA	1020	-	9,9,9	0.44	0	8,8,8	0.29	0
6	SO4	AaA	1006	-	4,4,4	0.32	0	6,6,6	0.33	0
8	GOL	AAA	1004	-	5,5,5	0.22	0	5,5,5	0.57	0
6	SO4	AaA	1003	-	4,4,4	0.32	0	6,6,6	0.17	0
11	PGE	AaA	1019	-	9,9,9	0.25	0	8,8,8	0.10	0
9	EDO	AaA	1016	-	3,3,3	0.19	0	2,2,2	0.37	0
8	GOL	AaA	1015	-	5,5,5	0.20	0	5,5,5	0.34	0
9	EDO	AaA	1018	-	3,3,3	0.15	0	2,2,2	0.36	0
9	EDO	AAA	1005	-	3,3,3	0.22	0	2,2,2	0.23	0
6	SO4	AaA	1004	-	4,4,4	0.27	0	6,6,6	0.12	0
10	YTW	AaA	1002	-	14,17,17	1.05	2 (14%)	18,27,27	1.42	2 (11%)
9	EDO	AaA	1017	-	3,3,3	0.11	0	2,2,2	0.04	0
8	GOL	AaA	1014	-	5,5,5	0.08	0	5,5,5	0.17	0
6	SO4	AaA	1005	-	4,4,4	0.44	0	6,6,6	0.21	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
9	EDO	AaA	1001	-	-	0/1/1/1	-
11	PGE	AaA	1020	-	-	4/7/7/7	-
8	GOL	AAA	1004	-	-	2/4/4/4	-
11	PGE	AaA	1019	-	-	1/7/7/7	-
9	EDO	AaA	1016	-	-	0/1/1/1	-
8	GOL	AaA	1015	-	-	2/4/4/4	-
9	EDO	AaA	1018	-	-	0/1/1/1	-
9	EDO	AAA	1005	-	-	0/1/1/1	-
10	YTW	AaA	1002	-	1/1/7/8	0/2/36/36	0/2/2/2
9	EDO	AaA	1017	-	-	0/1/1/1	-
8	GOL	AaA	1014	-	-	0/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	AaA	1002	YTW	O2-S1	2.36	1.44	1.42
10	AaA	1002	YTW	O1-S1	2.32	1.44	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	AaA	1002	YTW	C2-C1-N1	2.94	104.35	100.15
10	AaA	1002	YTW	C7-C6-C5	-2.50	106.45	110.82

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
10	AaA	1002	YTW	C5

5 of 9 torsion outliers are listed below:

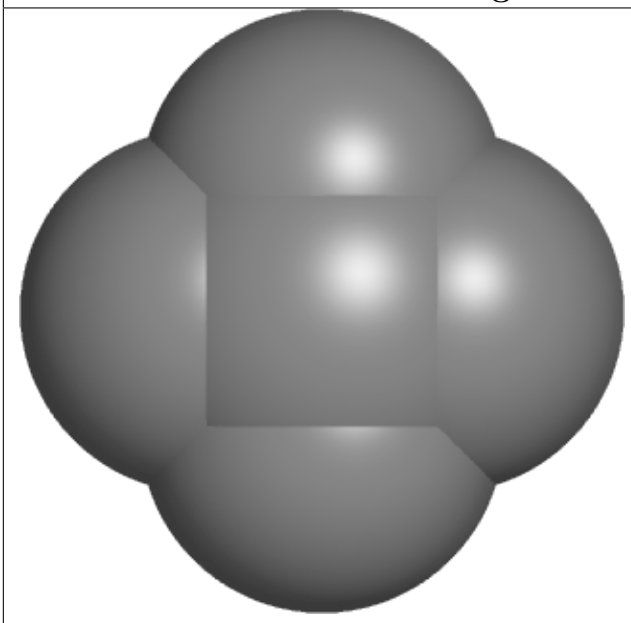
Mol	Chain	Res	Type	Atoms
8	AAA	1004	GOL	O1-C1-C2-C3
11	AaA	1020	PGE	O2-C3-C4-O3
8	AAA	1004	GOL	O1-C1-C2-O2
11	AaA	1020	PGE	O3-C5-C6-O4
11	AaA	1020	PGE	C3-C4-O3-C5

There are no ring outliers.

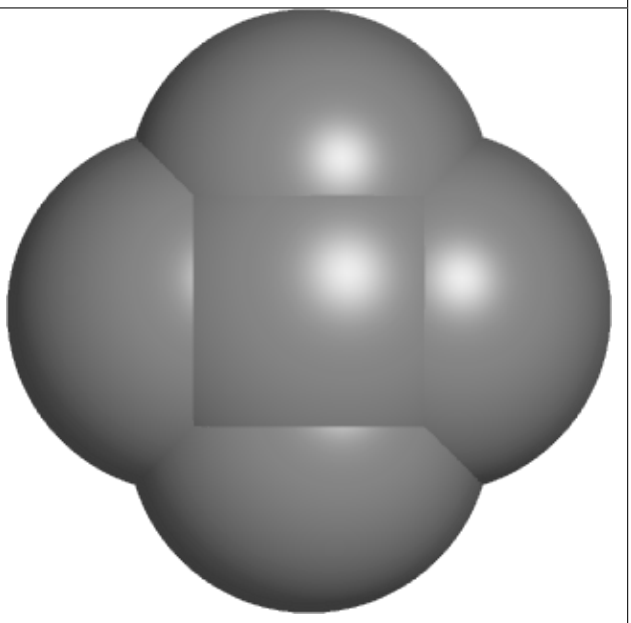
No monomer is involved in short contacts.

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 then 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.

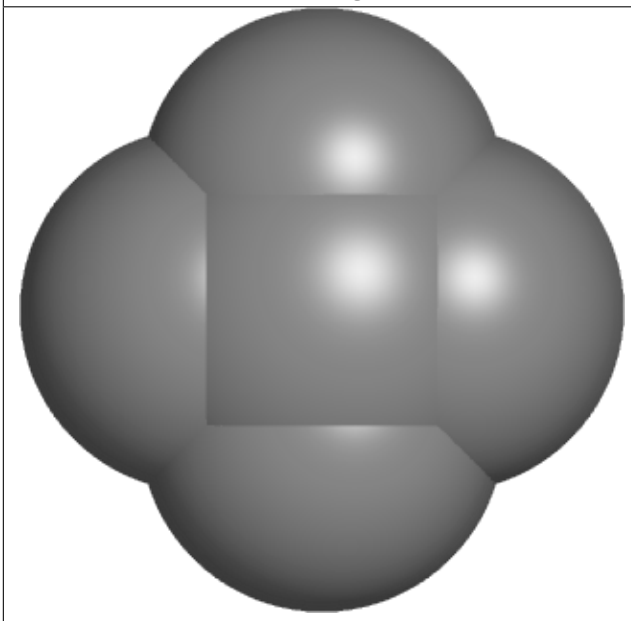
Ligand SO4 AAA 1001



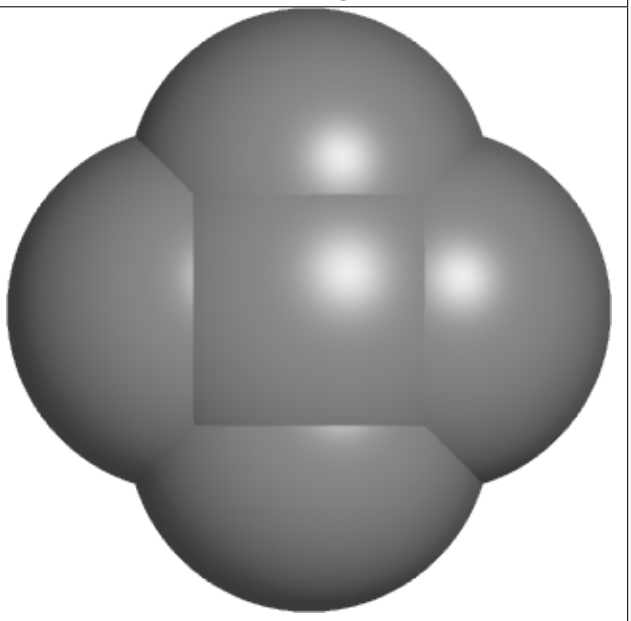
Bond lengths



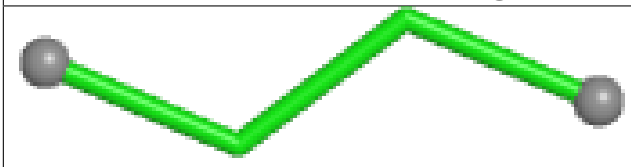
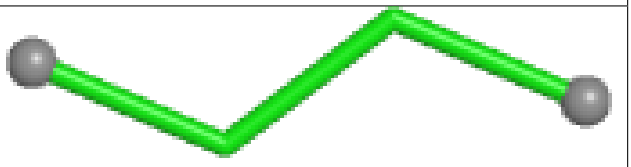
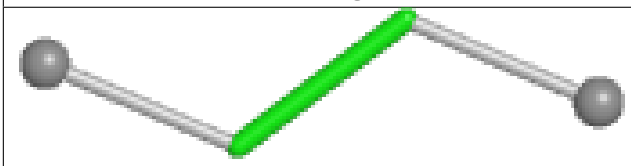
Bond angles

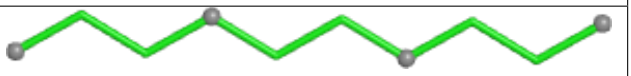
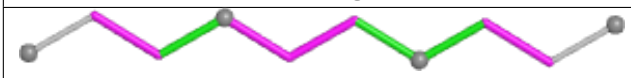
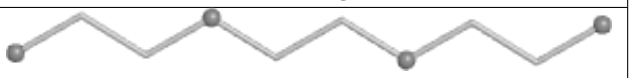


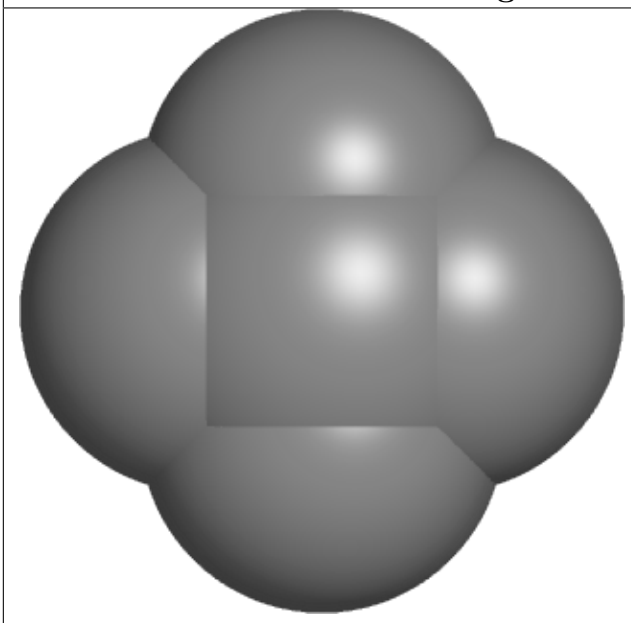
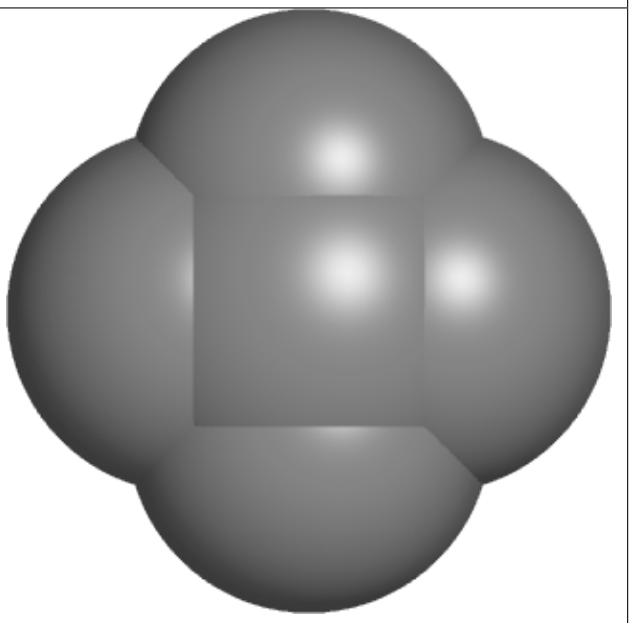
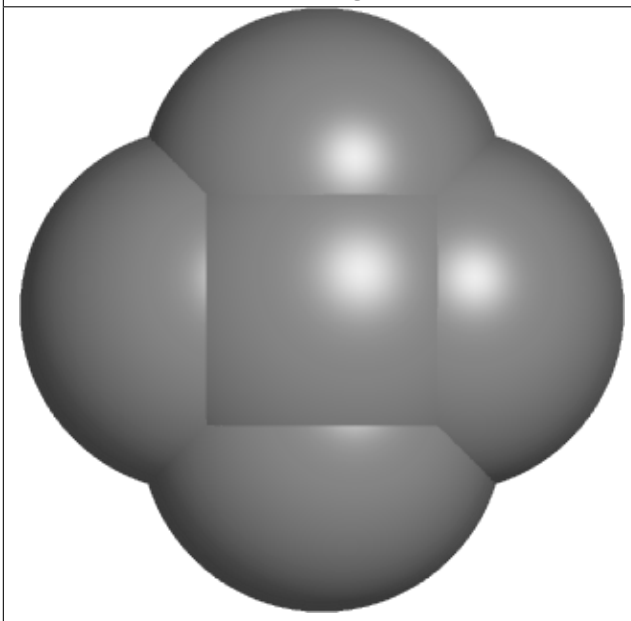
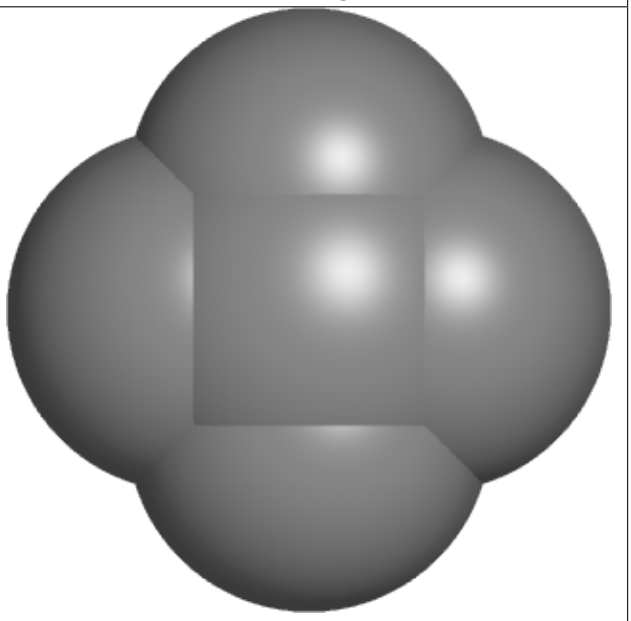
Torsions



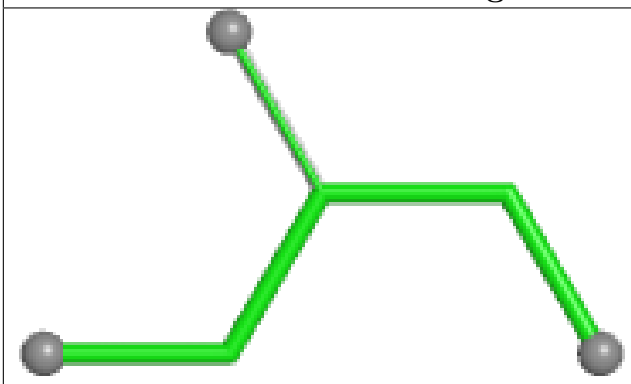
Rings

Ligand EDO AaA 1001	
 Bond lengths	 Bond angles
 Torsions	 Rings

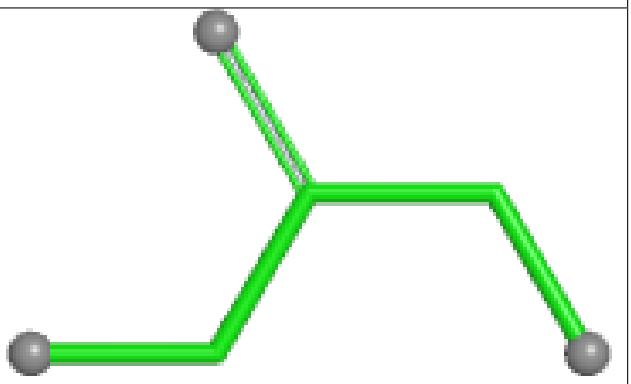
Ligand PGE AaA 1020	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand SO4 AaA 1006	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

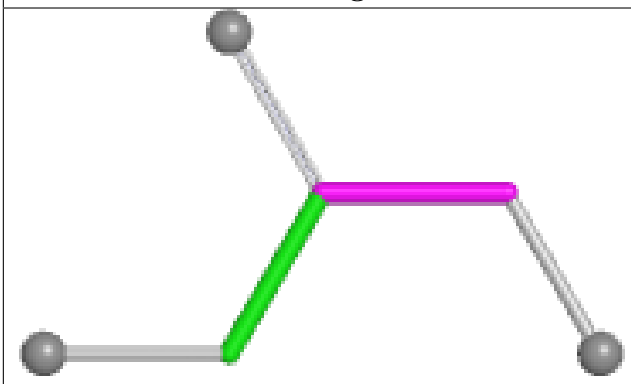
Ligand GOL AAA 1004



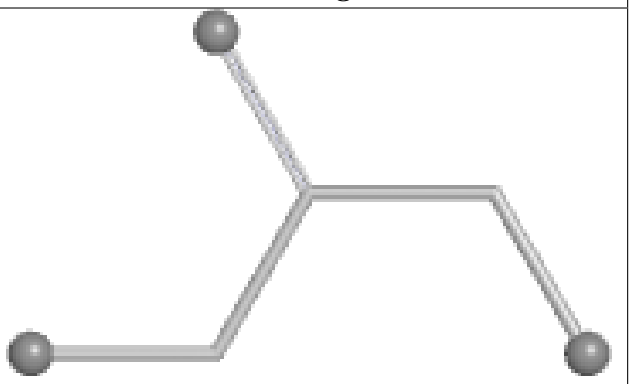
Bond lengths



Bond angles

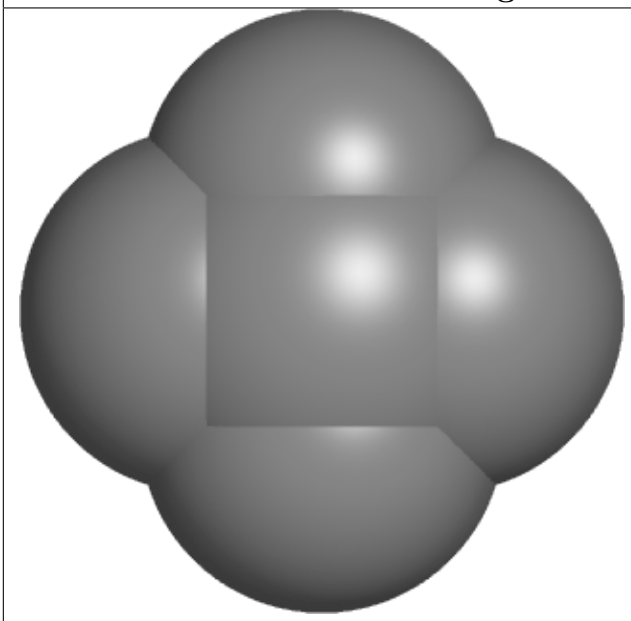


Torsions

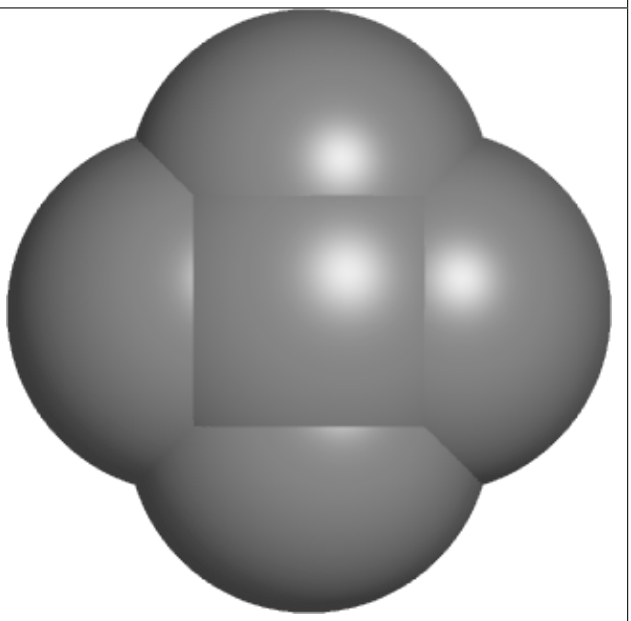


Rings

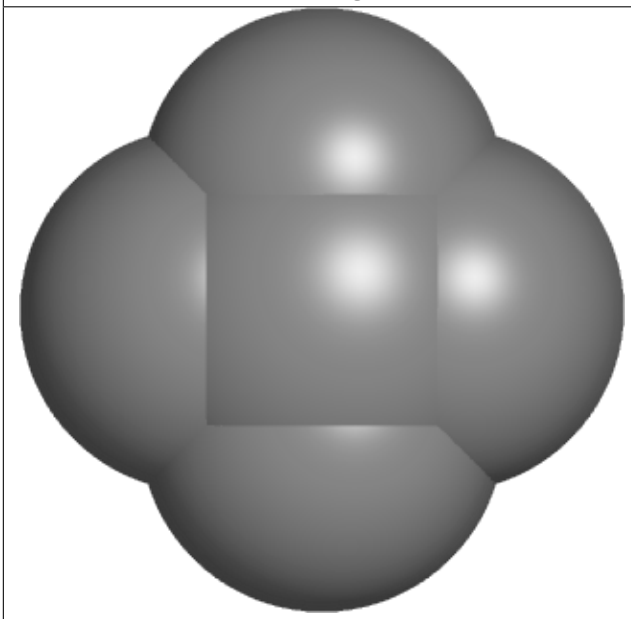
Ligand SO4 AaA 1003



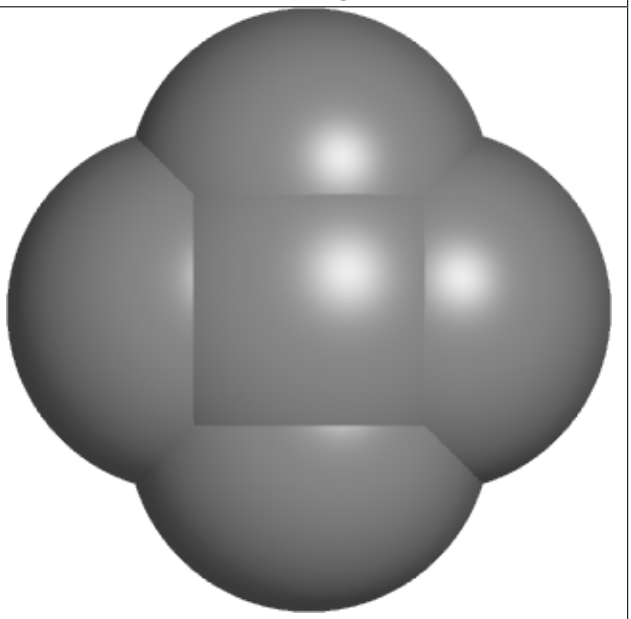
Bond lengths



Bond angles



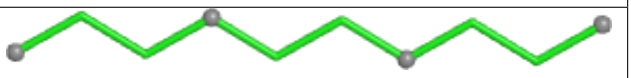
Torsions



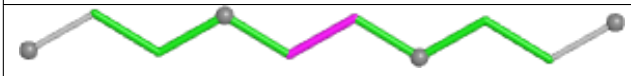
Rings

Ligand PGE AaA 1019

Bond lengths



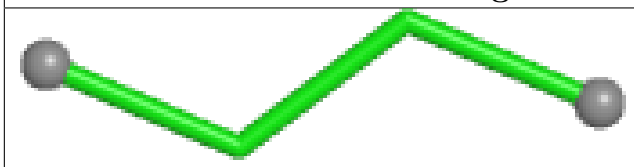
Bond angles



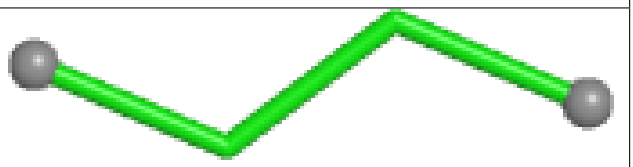
Torsions

Rings

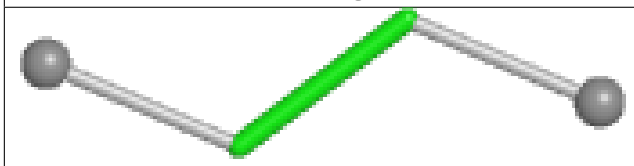
Ligand EDO AaA 1016



Bond lengths



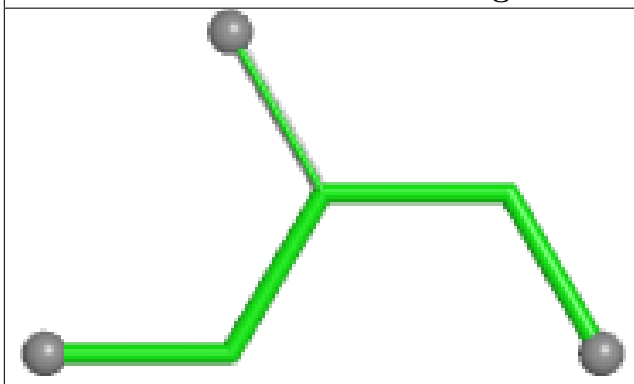
Bond angles



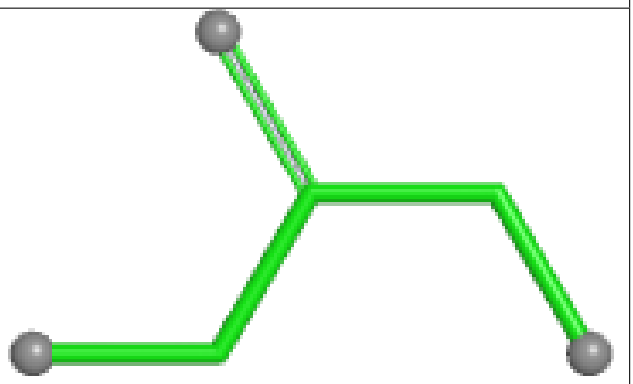
Torsions

Rings

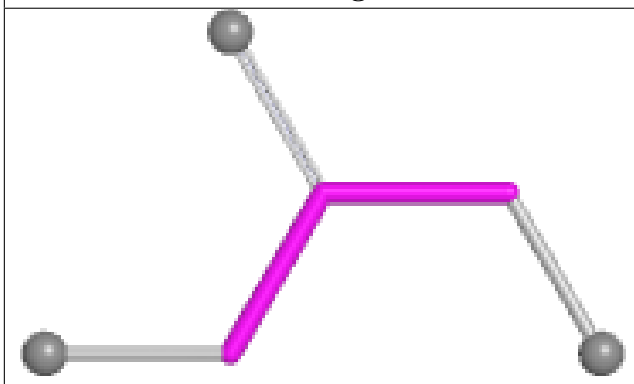
Ligand GOL AaA 1015



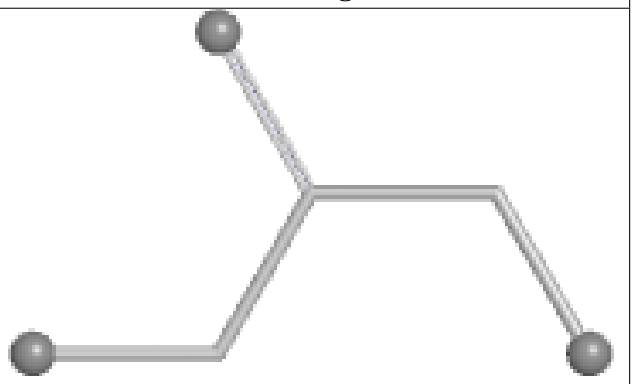
Bond lengths



Bond angles

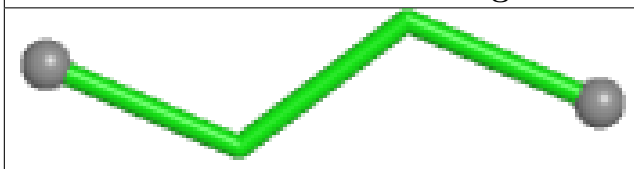


Torsions

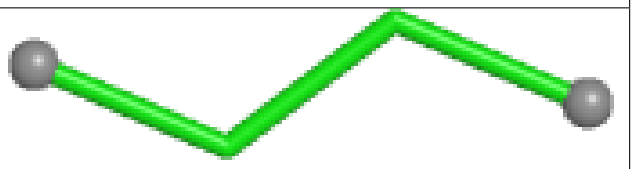


Rings

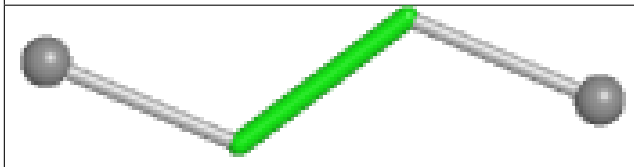
Ligand EDO AaA 1018



Bond lengths

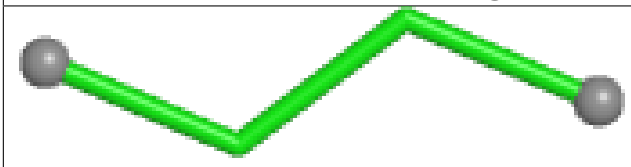
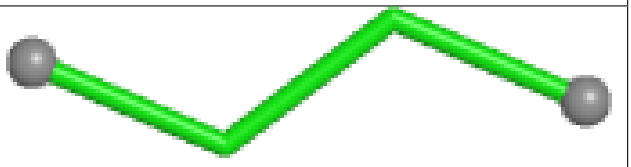
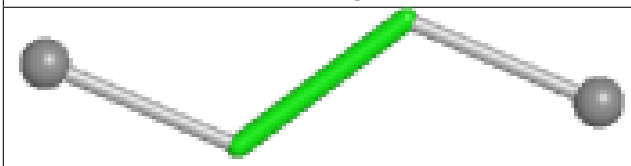


Bond angles

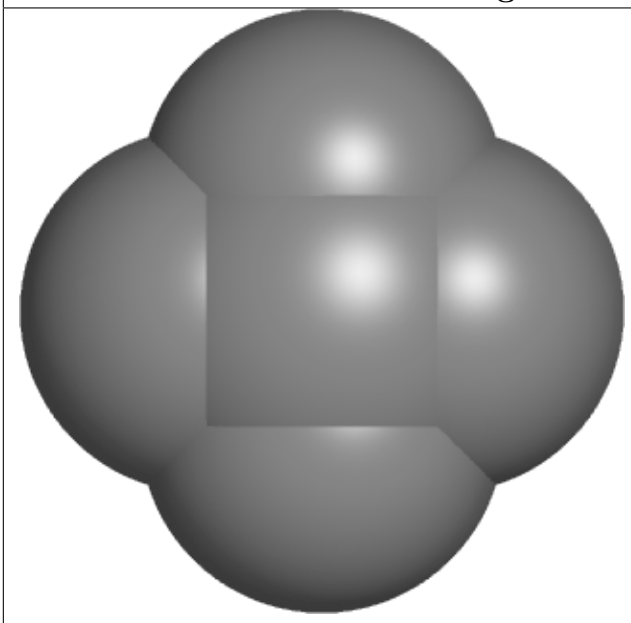


Torsions

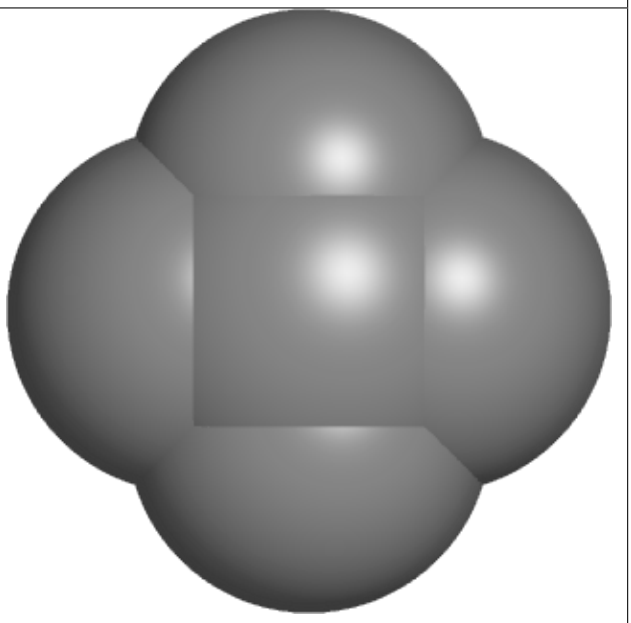
Rings

Ligand EDO AAA 1005	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

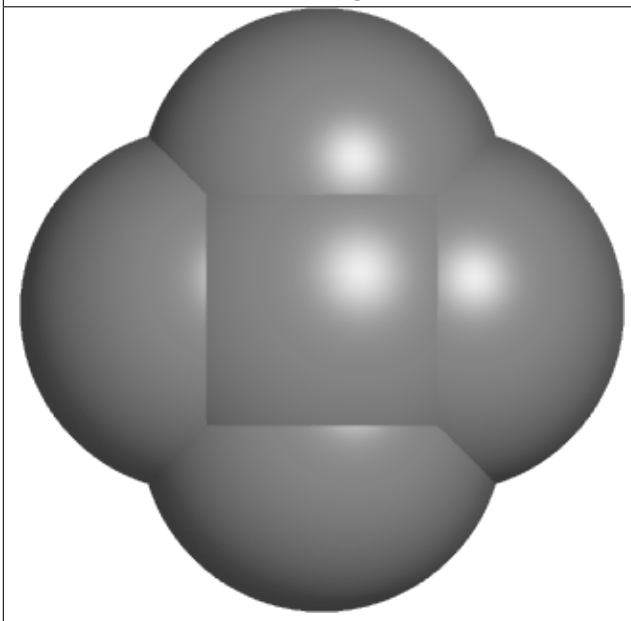
Ligand SO4 AaA 1004



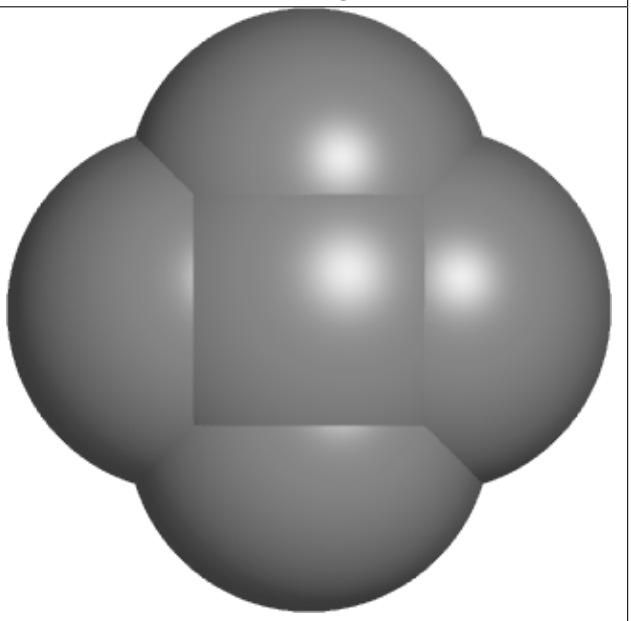
Bond lengths



Bond angles

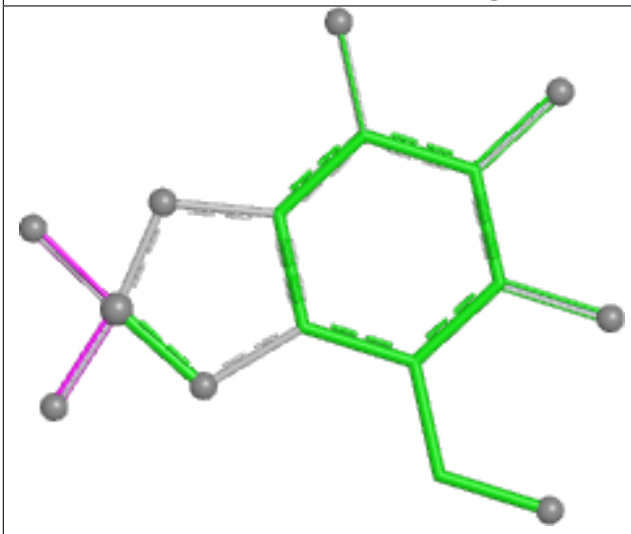


Torsions

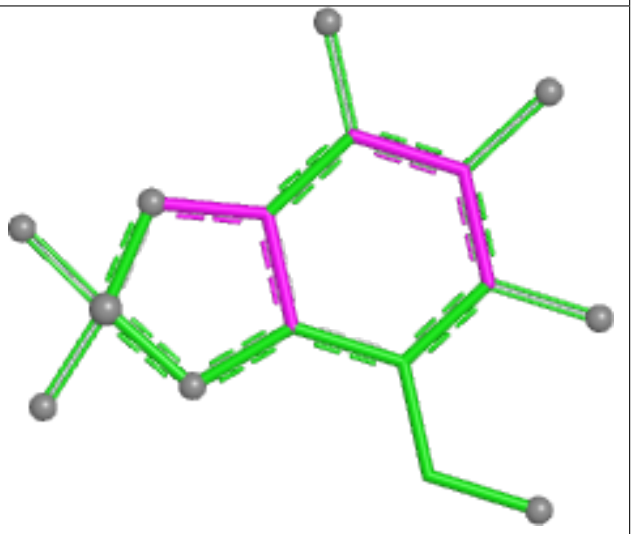


Rings

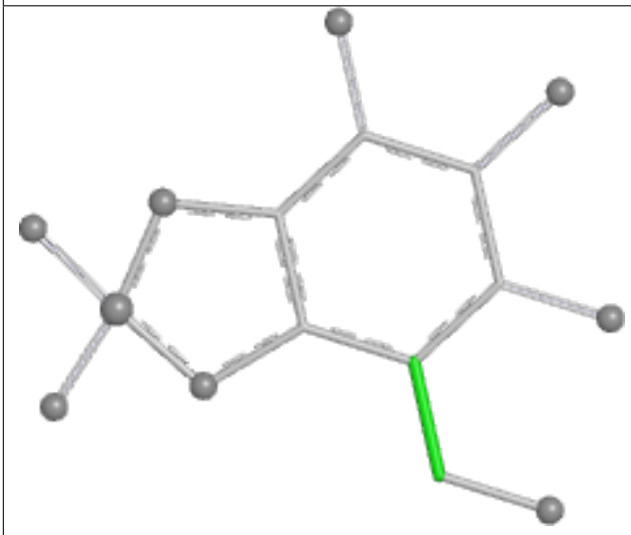
Ligand YTW AaA 1002



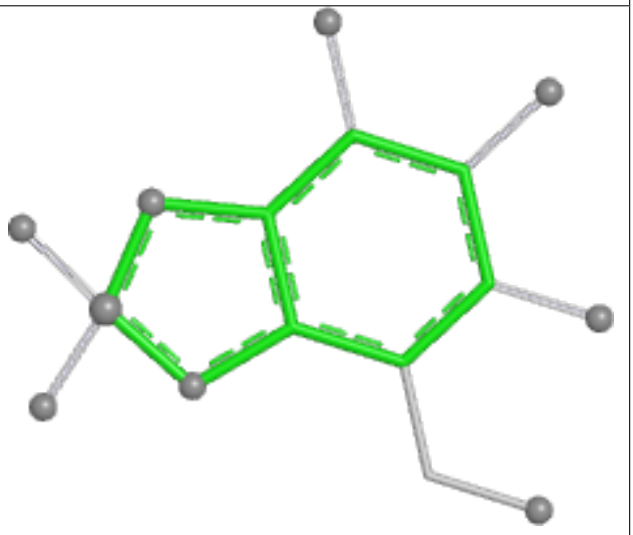
Bond lengths



Bond angles

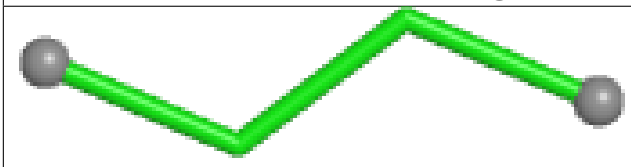


Torsions

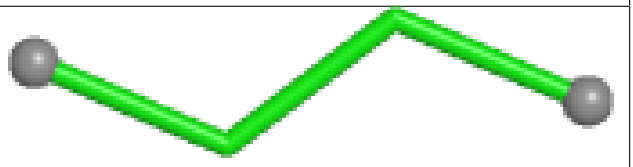


Rings

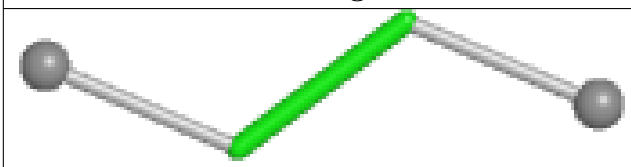
Ligand EDO AaA 1017



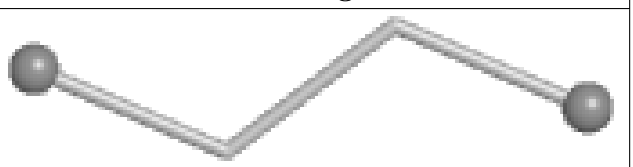
Bond lengths



Bond angles

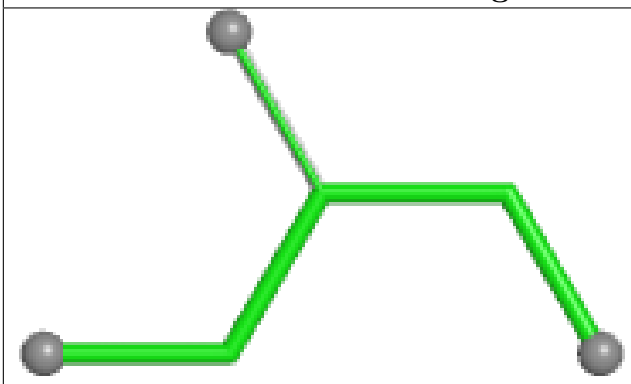


Torsions

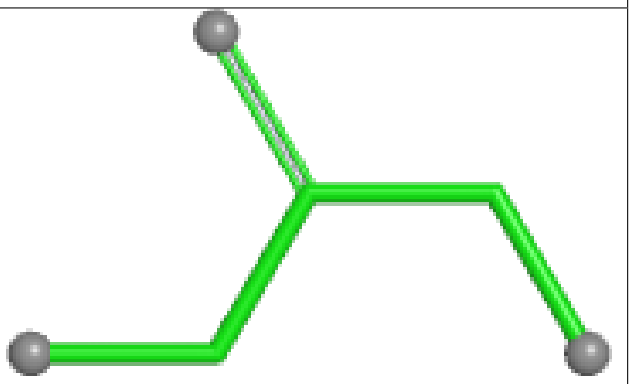


Rings

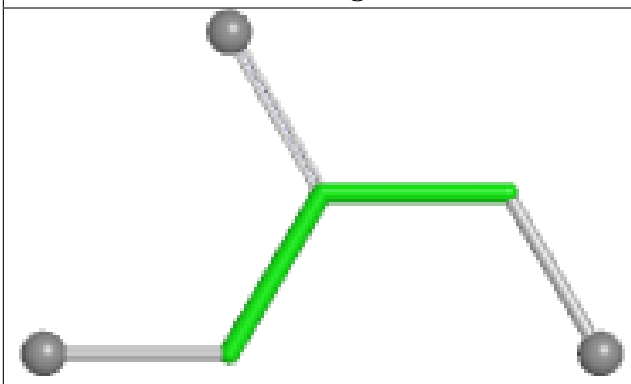
Ligand GOL AaA 1014



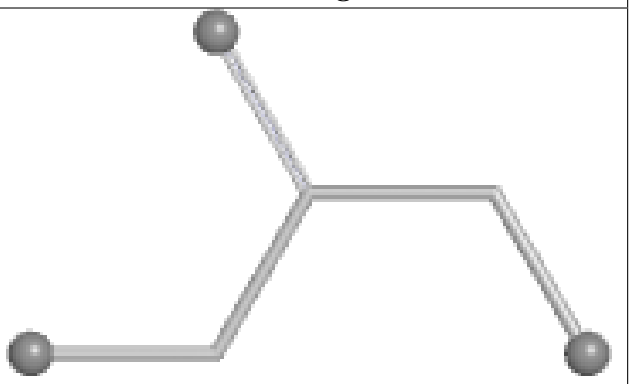
Bond lengths



Bond angles

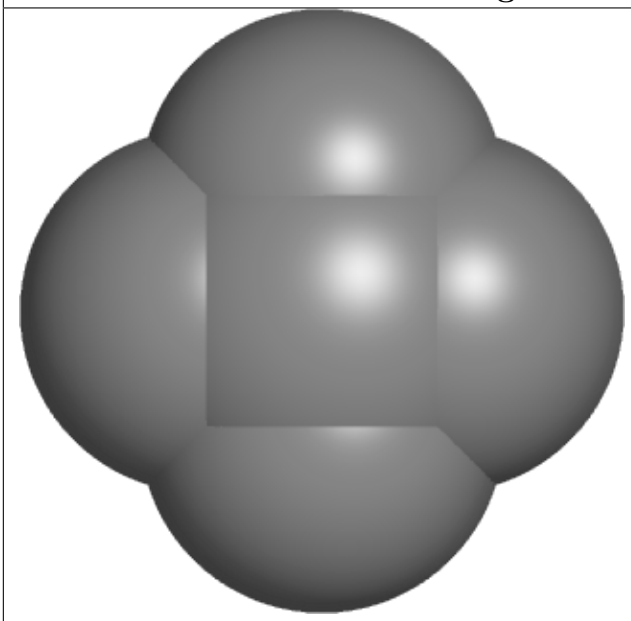


Torsions

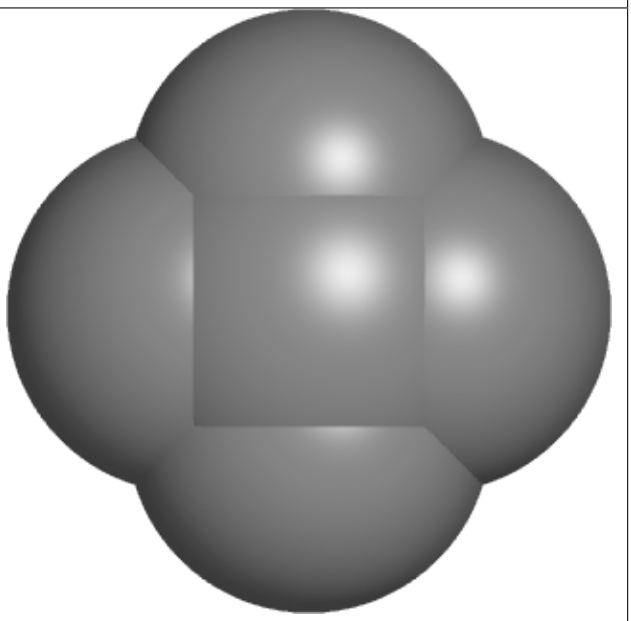


Rings

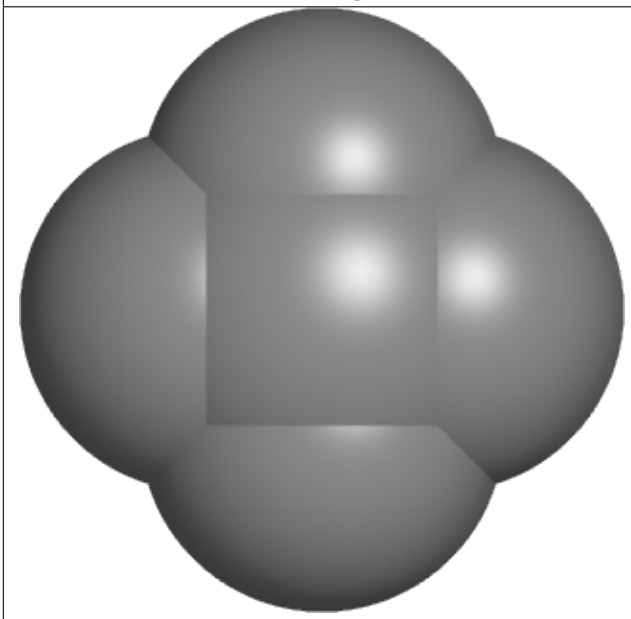
Ligand SO4 AaA 1005



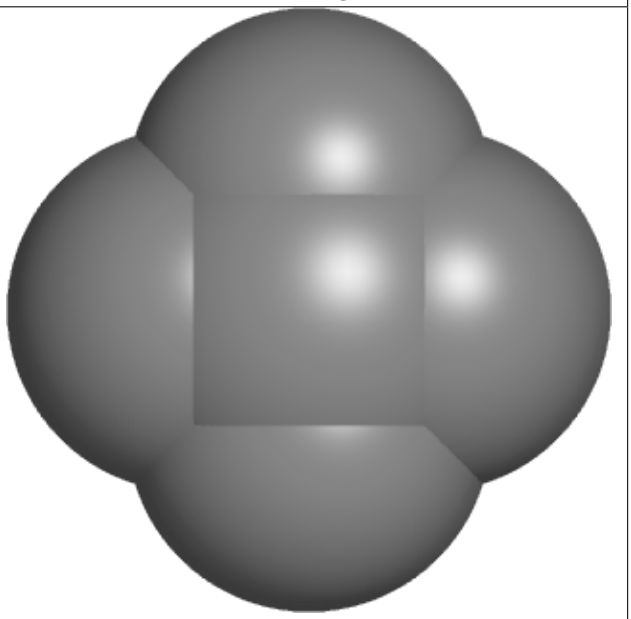
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	AAA	35/872 (4%)	0.97	6 (17%)	1	1	24, 34, 63, 79	0
1	AaA	815/872 (93%)	0.40	42 (5%)	27	26	18, 25, 48, 102	0
All	All	850/1744 (48%)	0.42	48 (5%)	24	23	18, 26, 51, 102	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AaA	783	LEU	9.5
1	AaA	784	GLY	7.4
1	AaA	205	PRO	5.8
1	AaA	952	CYS	5.5
1	AaA	782	ALA	4.9

6.2 Non-standard residues in protein, DNA, RNA chains

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	CSO	AaA	938	7/8	0.93	0.13	28,33,46,57	0

6.3 Carbohydrates

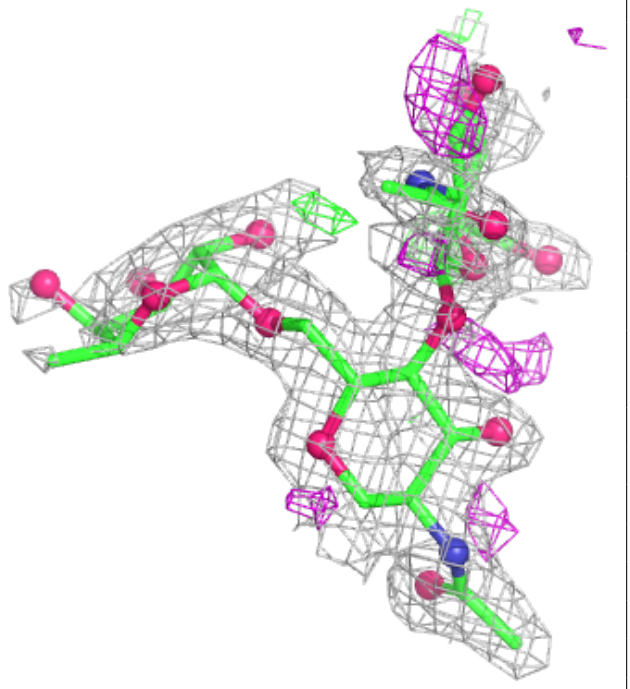
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

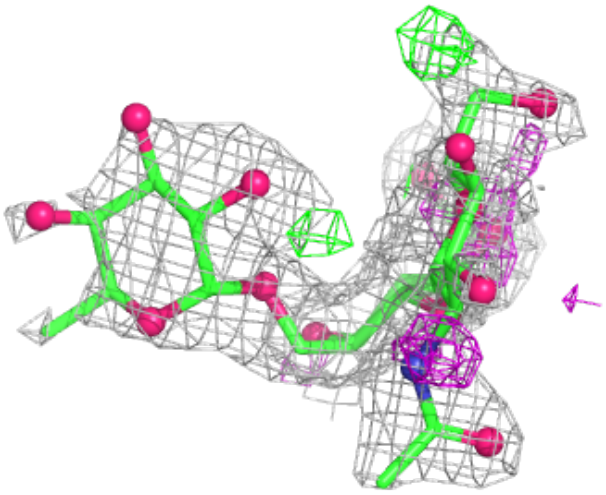
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	BBB	2	14/15	0.68	0.47	53,76,137,156	0
3	NAG	BeB	2	14/15	0.68	0.51	67,90,124,138	0
3	NAG	BcB	2	14/15	0.76	0.31	49,68,94,97	0
5	NAG	BjB	1	14/15	0.78	0.28	35,51,66,71	0
5	FUC	BjB	2	10/11	0.79	0.49	67,85,124,144	0
2	FUC	BBB	3	10/11	0.83	0.38	61,65,97,105	0
4	BMA	BgB	3	11/12	0.84	0.27	52,75,103,116	0
3	NAG	BeB	1	14/15	0.88	0.21	29,37,62,62	0
2	NAG	BBB	1	14/15	0.89	0.31	31,41,51,55	0
3	NAG	BcB	1	14/15	0.91	0.13	22,32,40,41	0
4	NAG	BgB	1	14/15	0.92	0.13	24,28,34,40	0
4	NAG	BgB	2	14/15	0.93	0.15	28,35,52,54	0

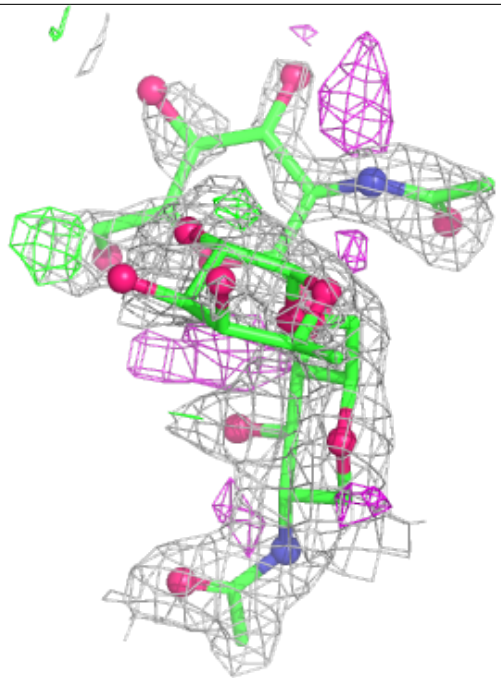
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain BBB:

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

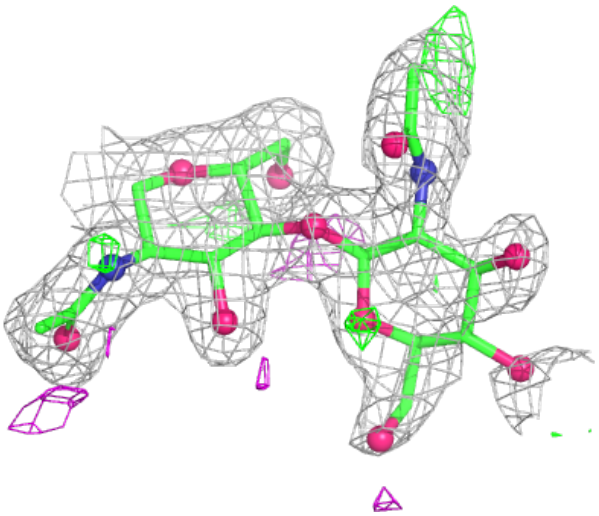


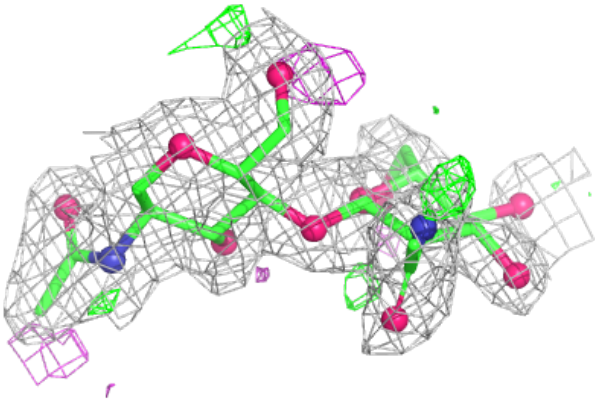


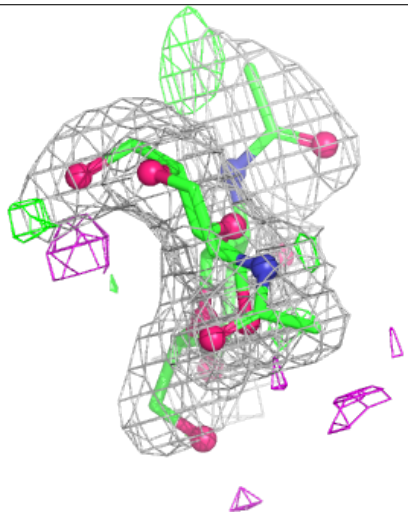


Electron density around Chain BcB:

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

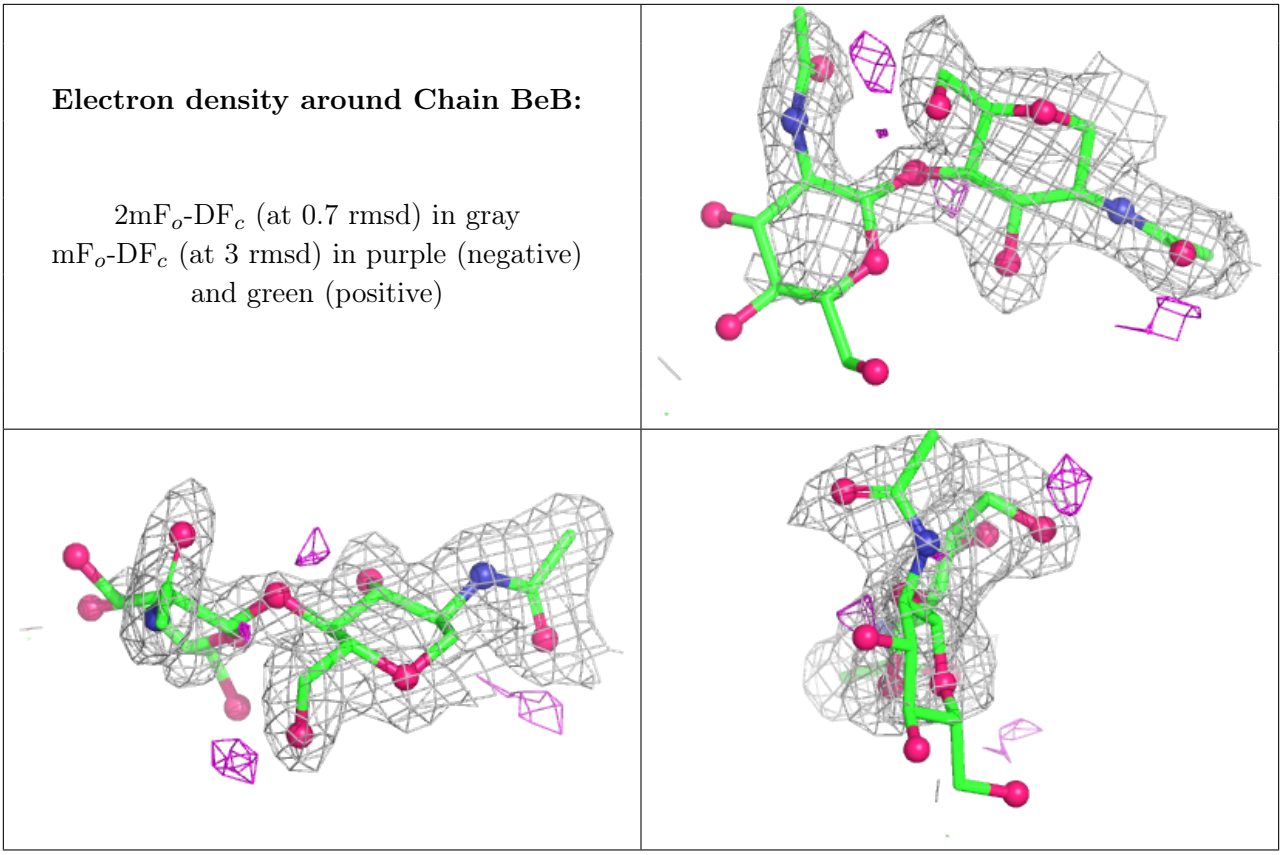






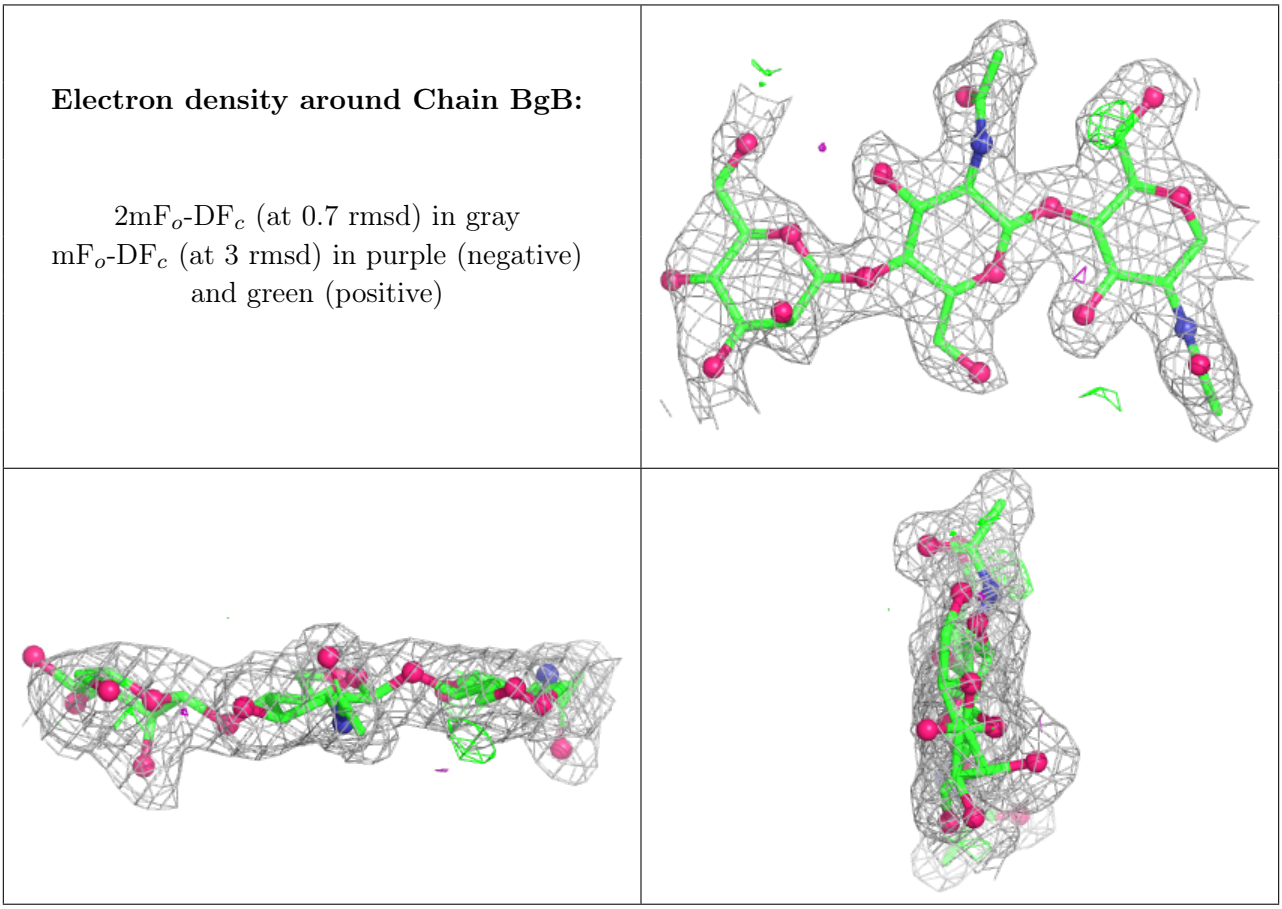
Electron density around Chain BeB:

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



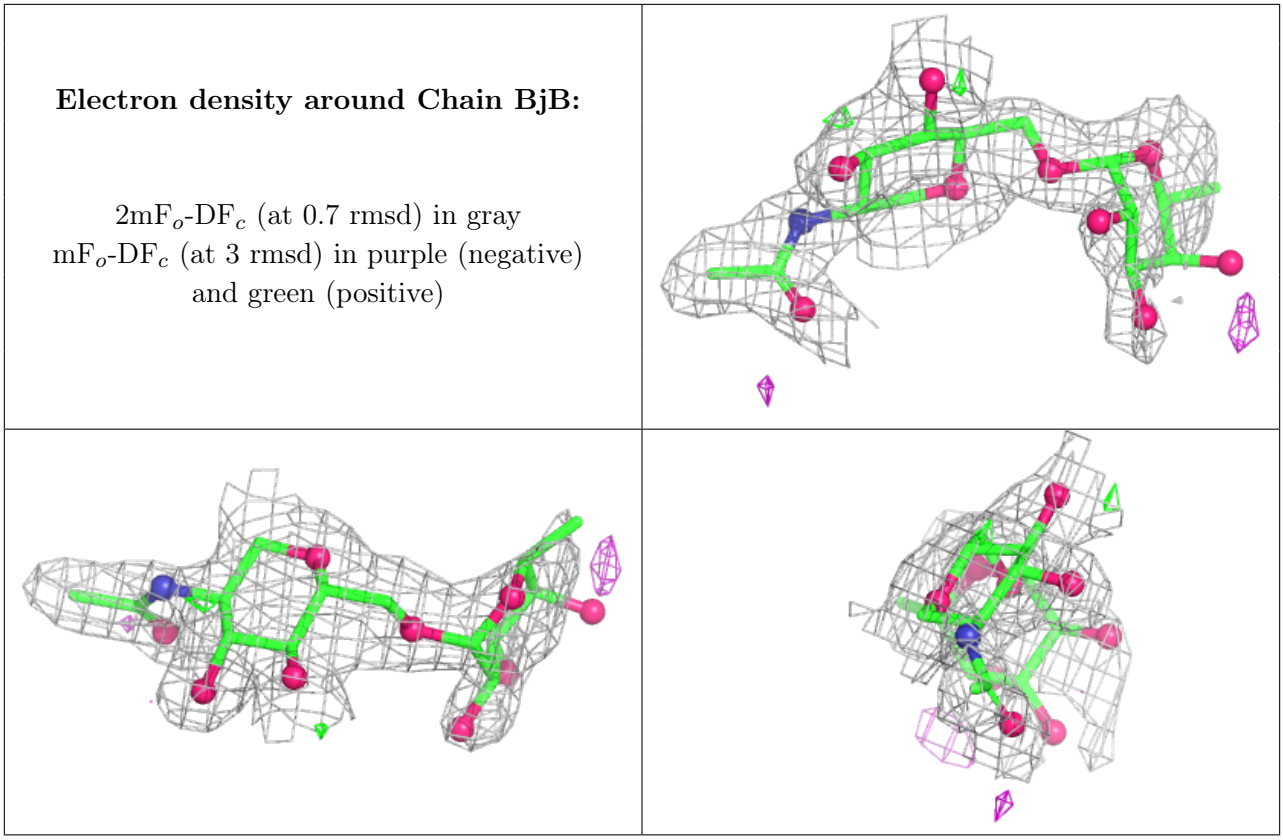
Electron density around Chain BgB:

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



Electron density around Chain BjB:

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



6.4 Ligands

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
11	PGE	AaA	1019	10/10	0.57	0.30	66,71,84,84	0
11	PGE	AaA	1020	10/10	0.63	0.22	60,65,80,91	0
7	CL	AaA	1012	1/1	0.71	0.33	80,80,80,80	0
7	CL	AaA	1013	1/1	0.72	0.12	82,82,82,82	0
7	CL	AaA	1010	1/1	0.77	0.31	82,82,82,82	0
6	SO4	AaA	1004	5/5	0.84	0.27	51,64,73,96	0
8	GOL	AaA	1015	6/6	0.84	0.21	28,30,35,47	6
8	GOL	AAA	1004	6/6	0.86	0.18	31,45,50,56	0
7	CL	AAA	1003	1/1	0.87	0.13	58,58,58,58	0
7	CL	AaA	1011	1/1	0.88	0.12	73,73,73,73	0
9	EDO	AaA	1001	4/4	0.88	0.10	43,46,48,50	0
9	EDO	AaA	1018	4/4	0.88	0.31	54,59,61,62	0
7	CL	AAA	1002	1/1	0.88	0.17	64,64,64,64	0
8	GOL	AaA	1014	6/6	0.88	0.11	32,37,40,54	0

Continued on next page...

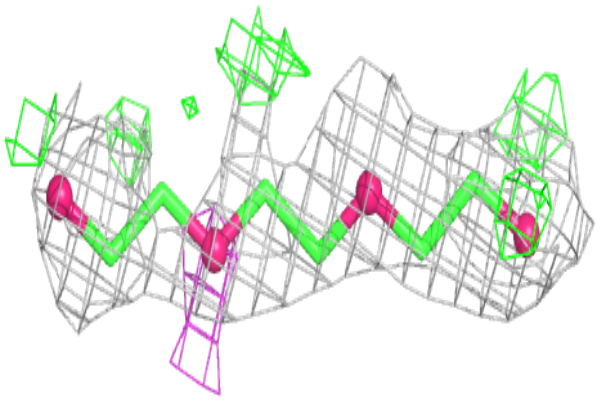
Continued from previous page...

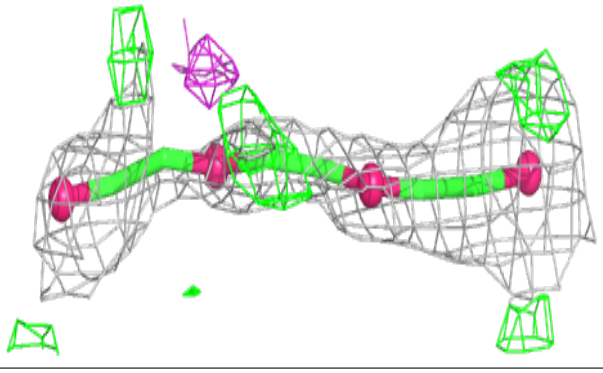
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
9	EDO	AaA	1016	4/4	0.89	0.12	44,52,63,64	0
7	CL	AaA	1007	1/1	0.89	0.25	65,65,65,65	0
6	SO4	AaA	1005	5/5	0.91	0.23	41,41,62,73	5
9	EDO	AaA	1017	4/4	0.92	0.14	31,58,59,77	0
6	SO4	AAA	1001	5/5	0.93	0.14	39,41,42,46	5
10	YTW	AaA	1002	16/16	0.95	0.10	19,22,26,27	0
7	CL	AaA	1008	1/1	0.95	0.13	71,71,71,71	0
9	EDO	AAA	1005	4/4	0.95	0.16	41,53,60,61	0
7	CL	AaA	1009	1/1	0.96	0.06	60,60,60,60	0
6	SO4	AaA	1006	5/5	0.96	0.11	23,25,33,43	5
6	SO4	AaA	1003	5/5	0.99	0.21	39,39,41,41	0

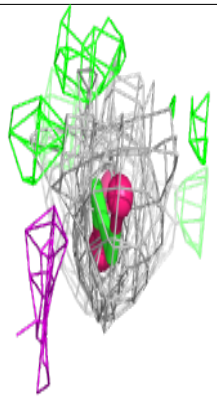
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around PGE AaA 1019:

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

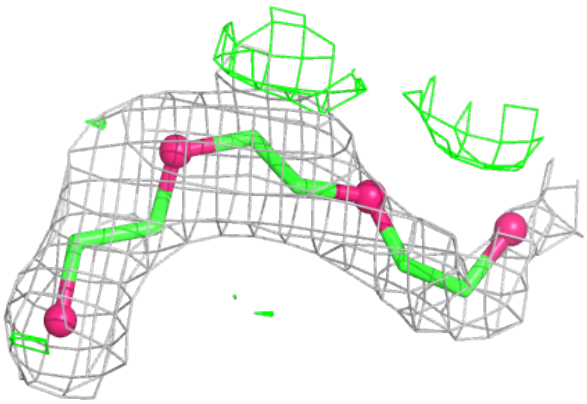


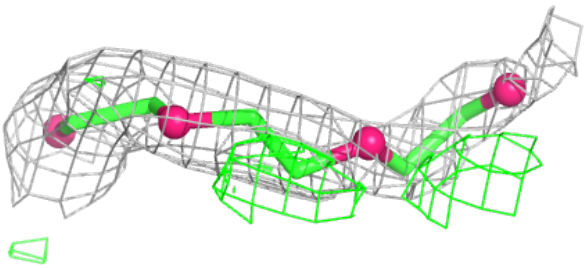




Electron density around PGE AaA 1020:

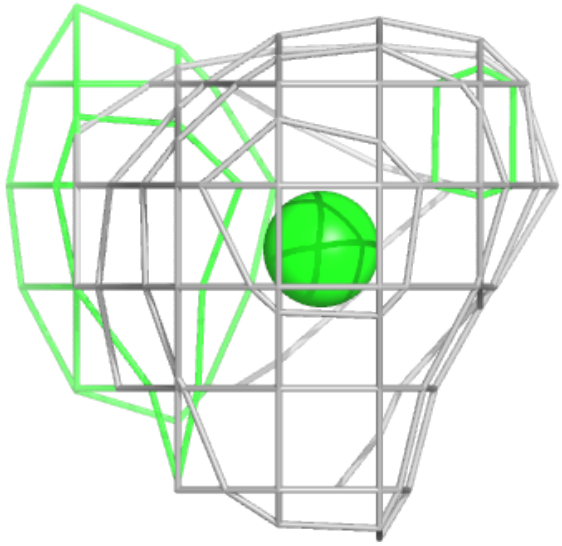
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

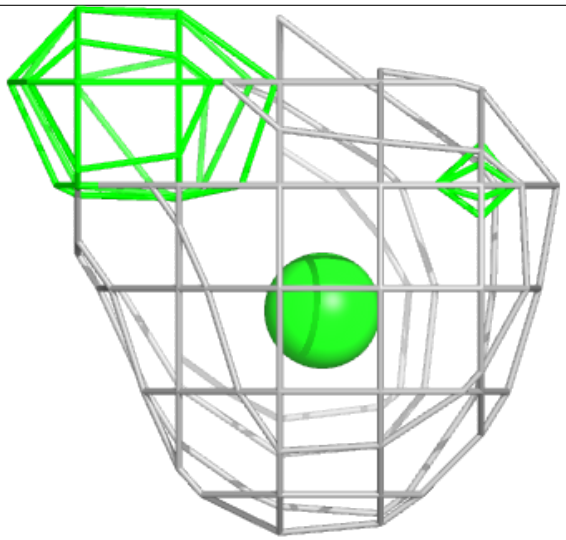


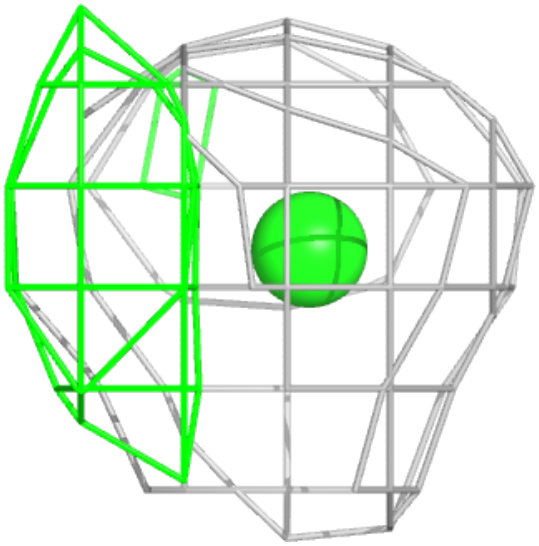


Electron density around CL AaA 1012:

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

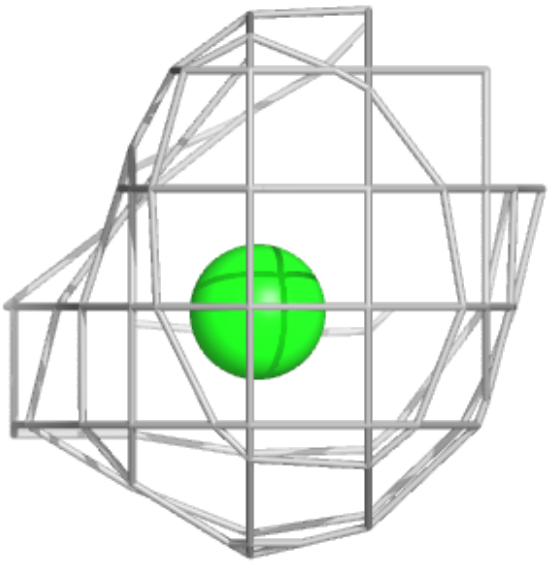


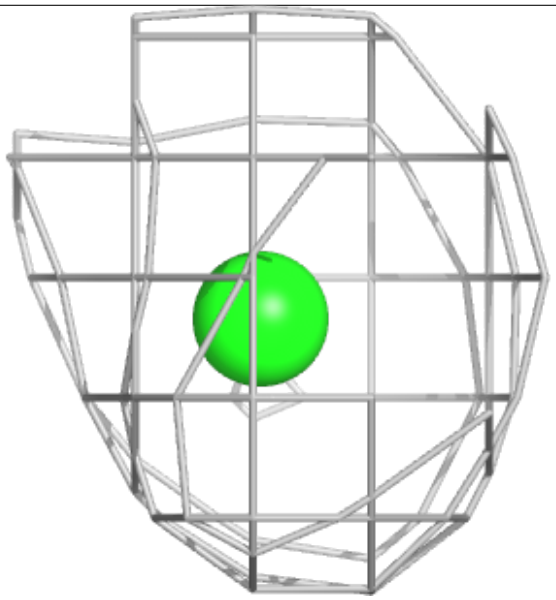


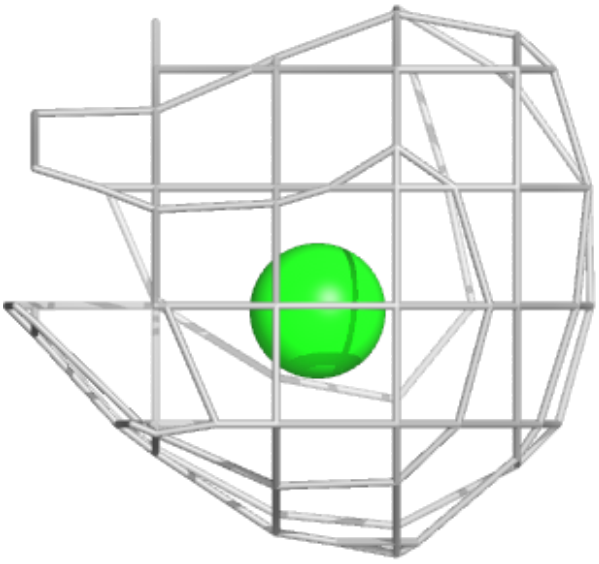


Electron density around CL AaA 1013:

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

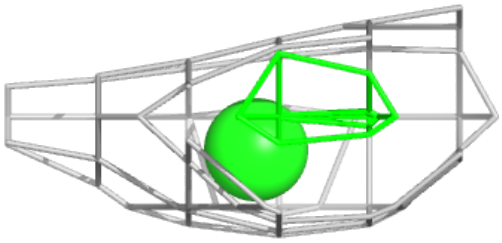


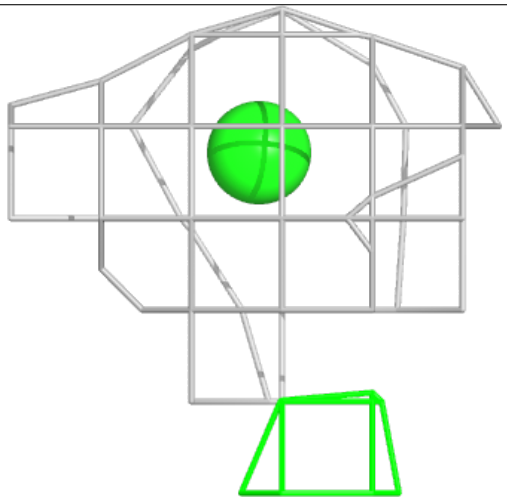


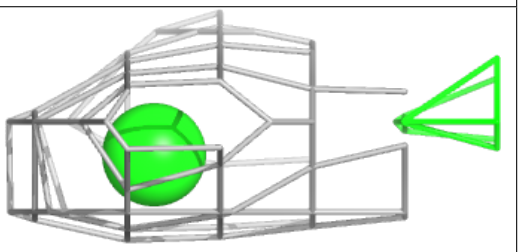


Electron density around CL AaA 1010:

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

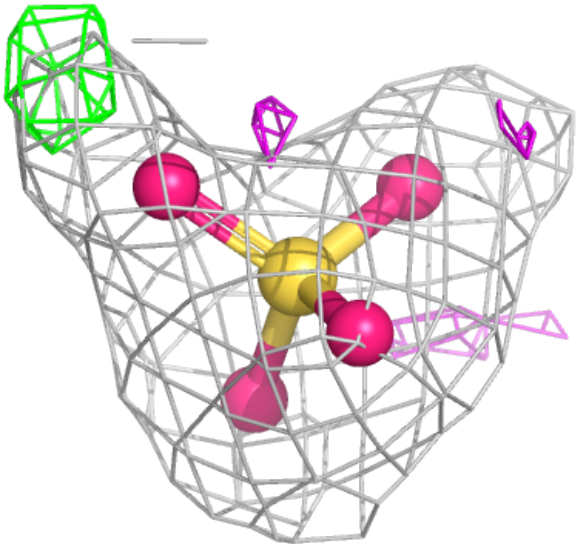


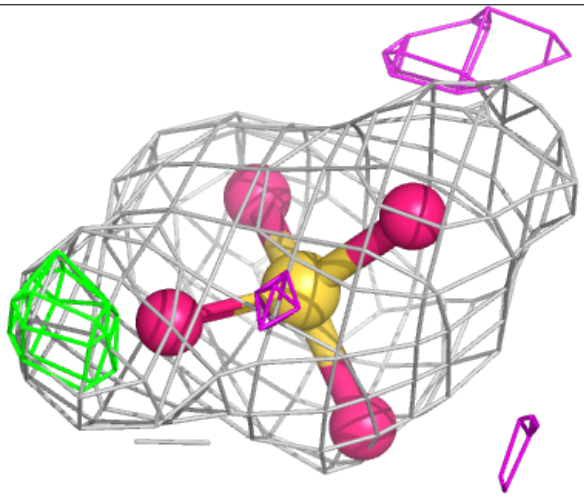


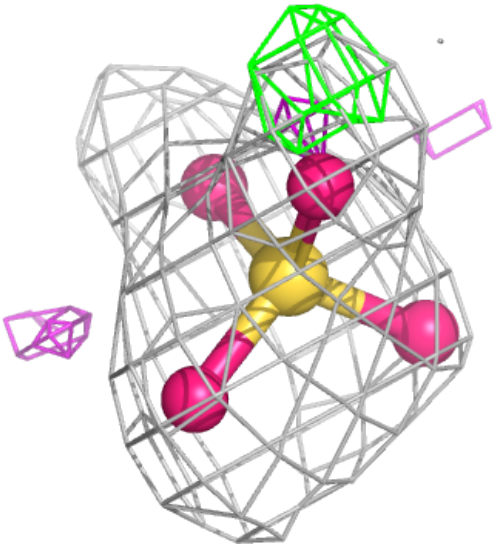


Electron density around SO4 AaA 1004:

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

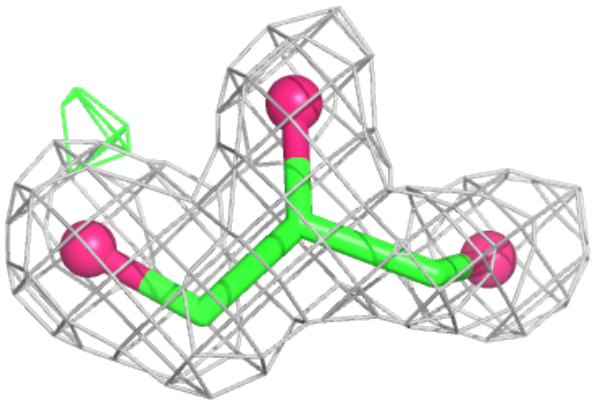


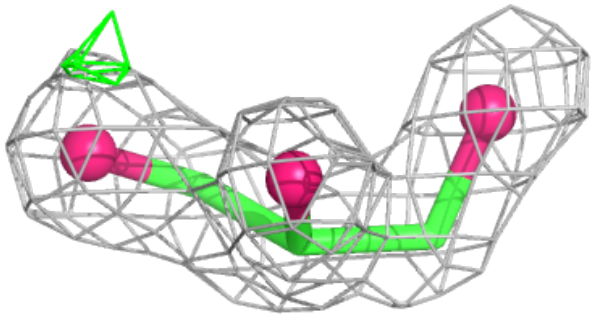


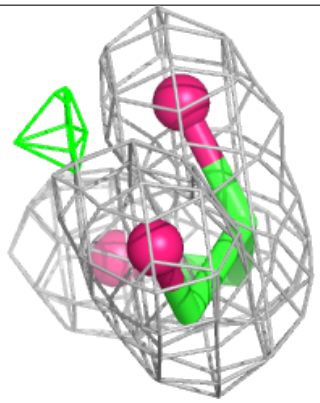


Electron density around GOL AaA 1015:

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

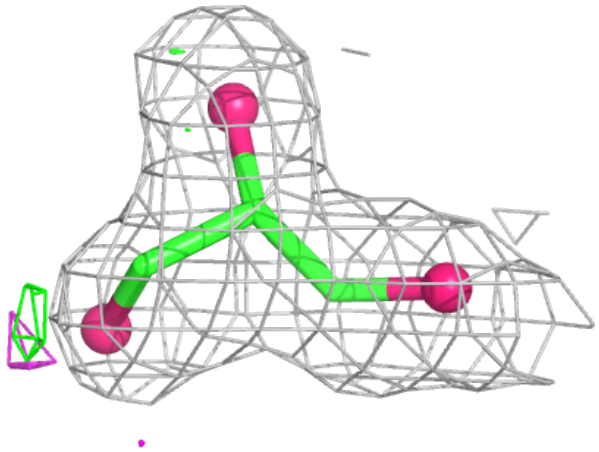


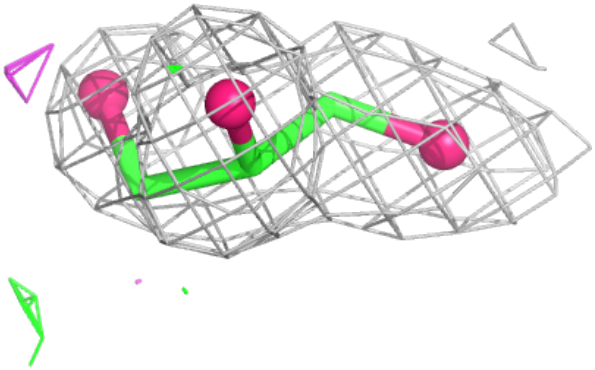


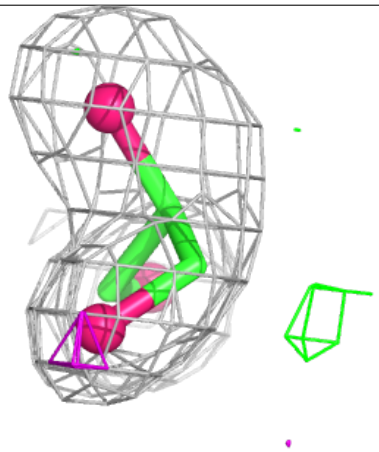


Electron density around GOL AAA 1004:

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

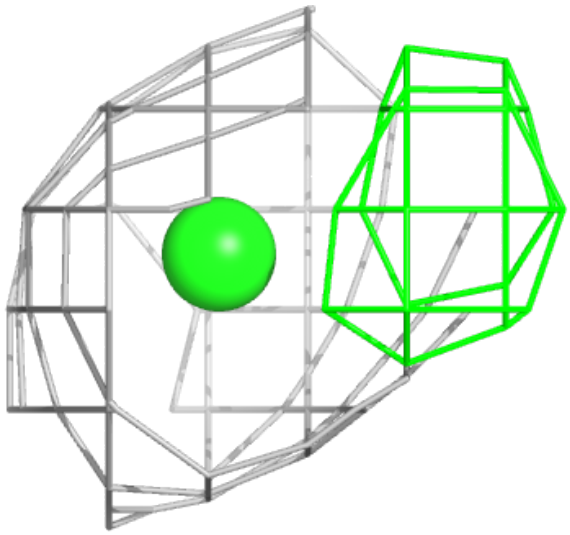


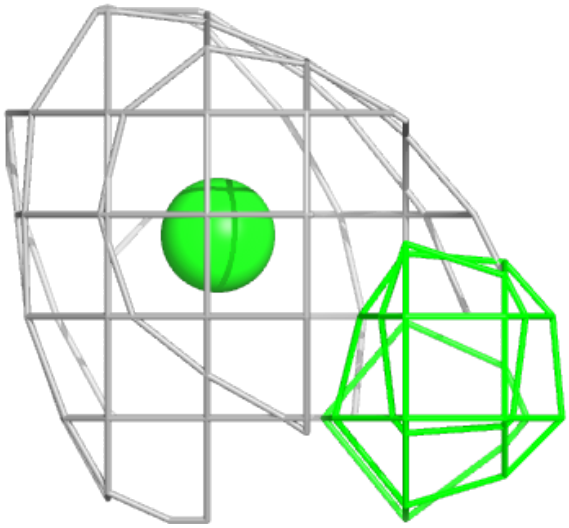


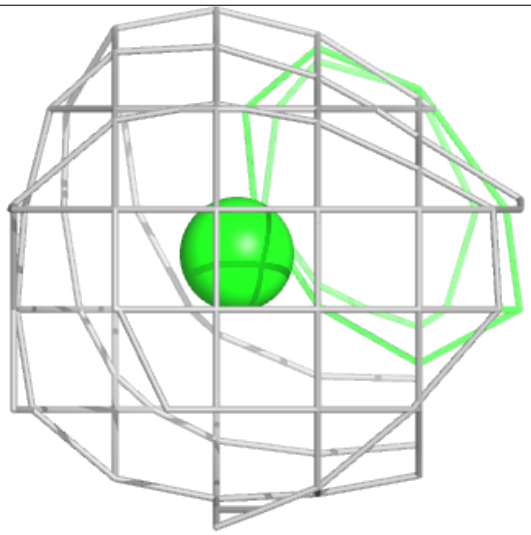


Electron density around CL AAA 1003:

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

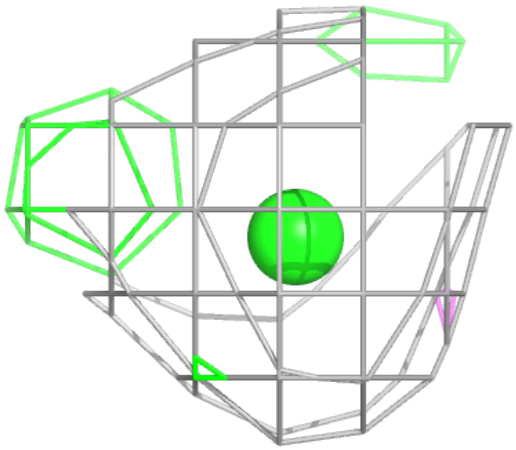


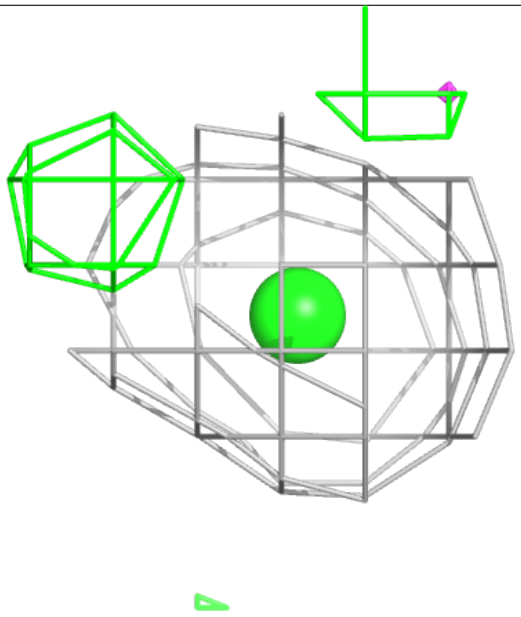


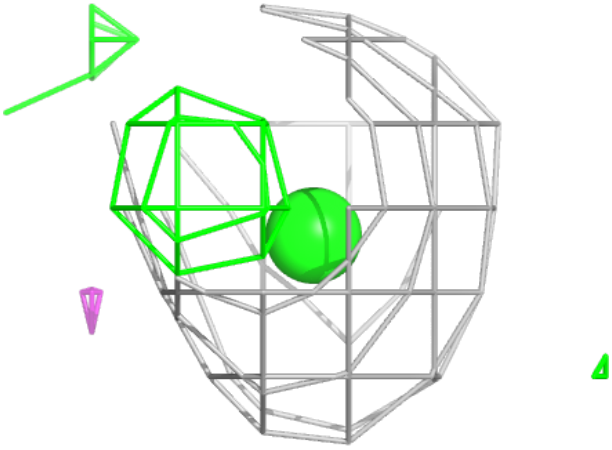


Electron density around CL AaA 1011:

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

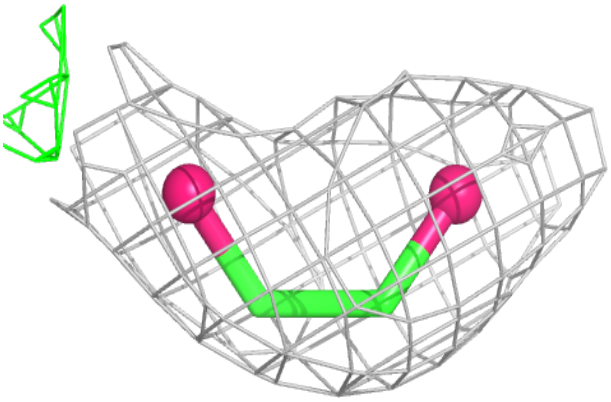


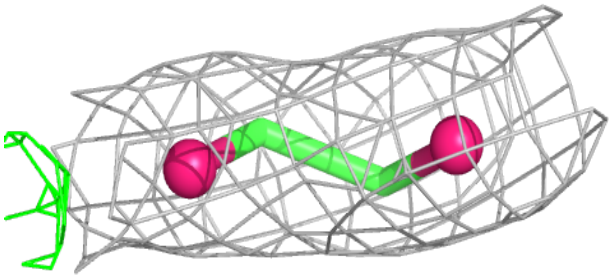


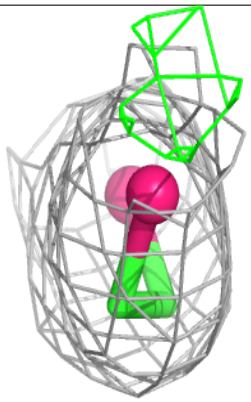


Electron density around EDO AaA 1001:

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

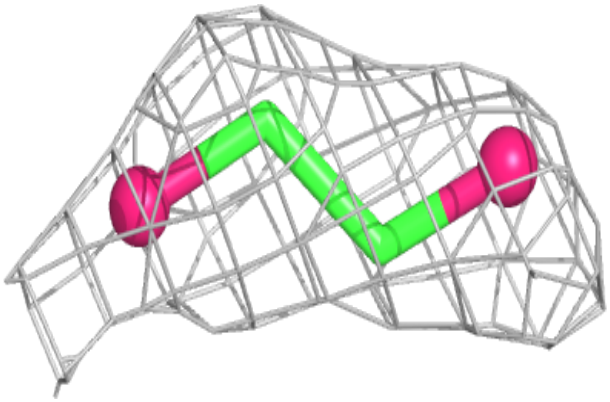


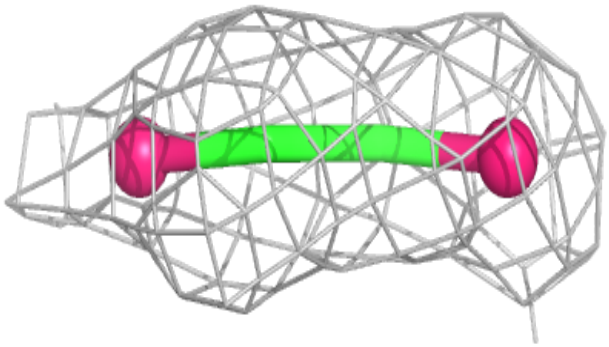


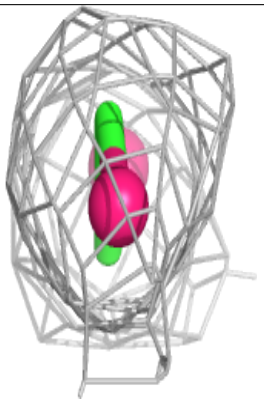


Electron density around EDO AaA 1018:

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

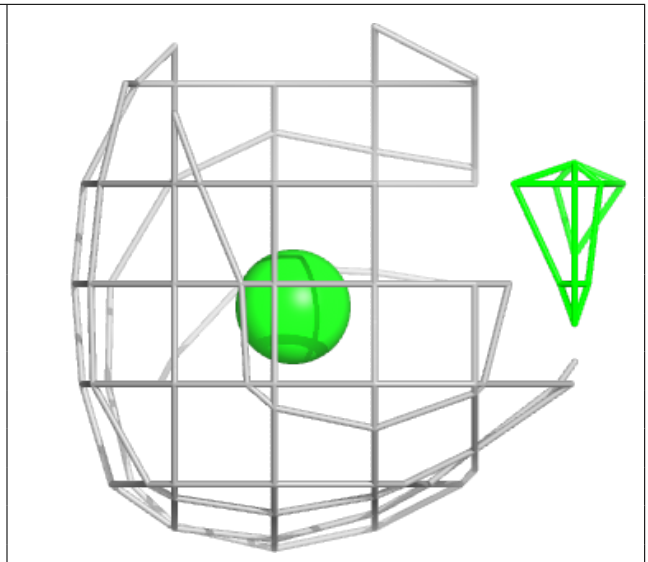


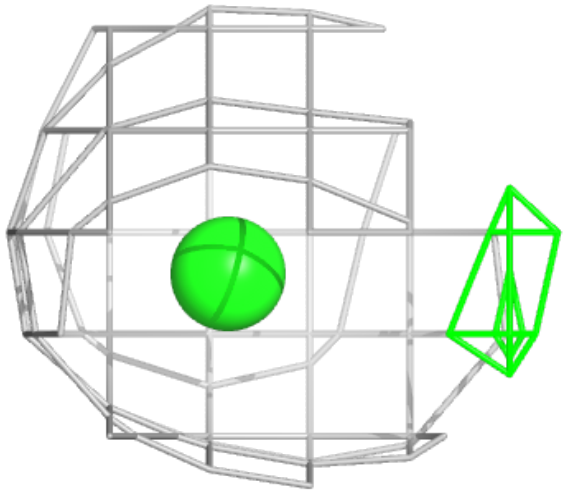


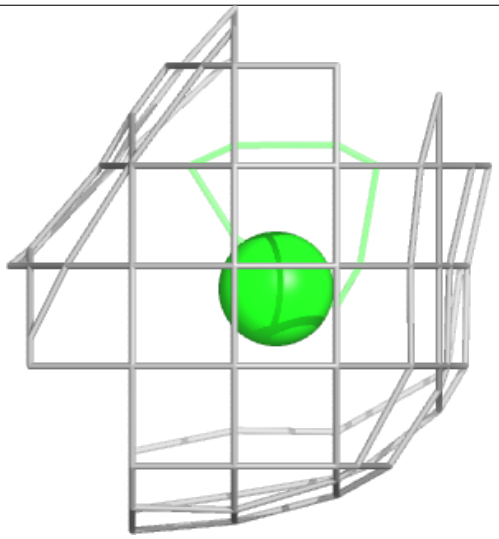


Electron density around CL AAA 1002:

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

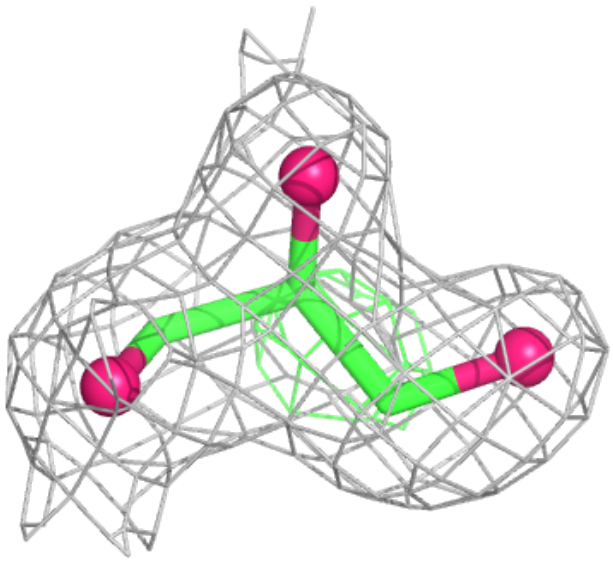


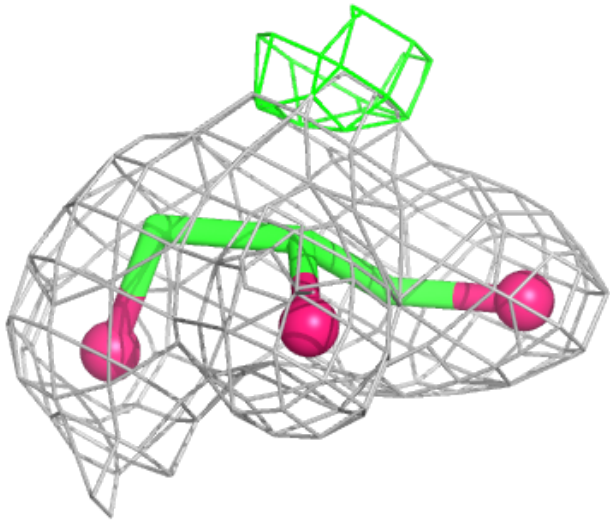


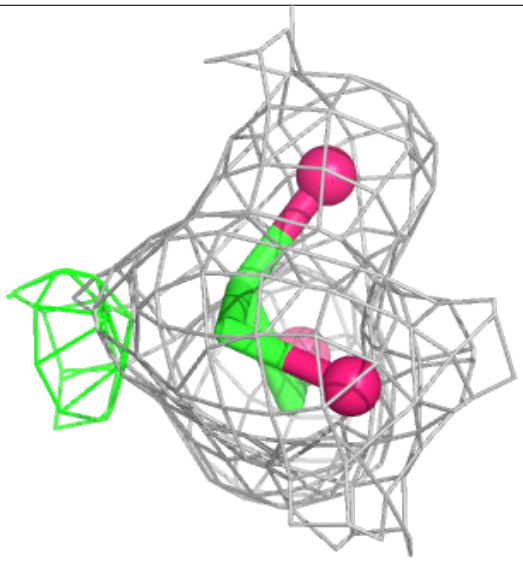


Electron density around GOL AaA 1014:

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

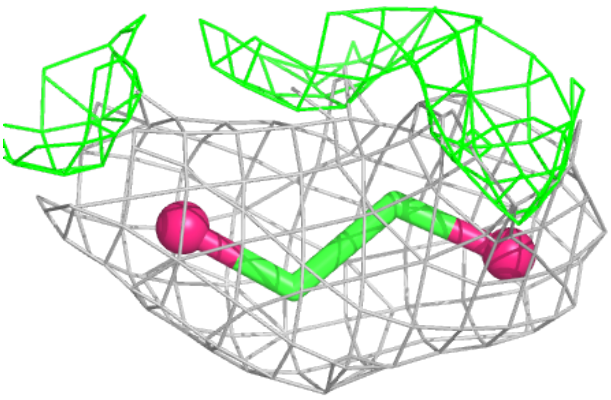


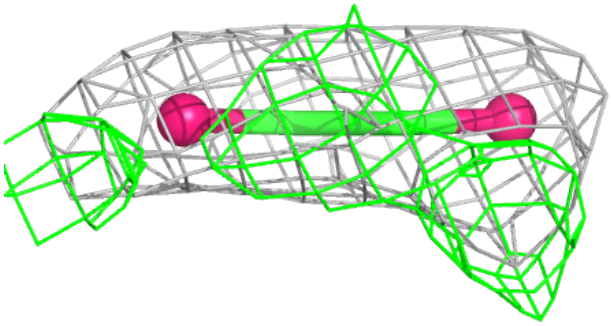


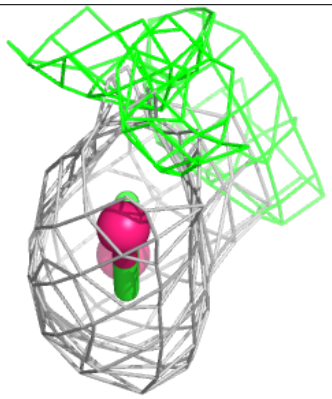


Electron density around EDO AaA 1016:

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

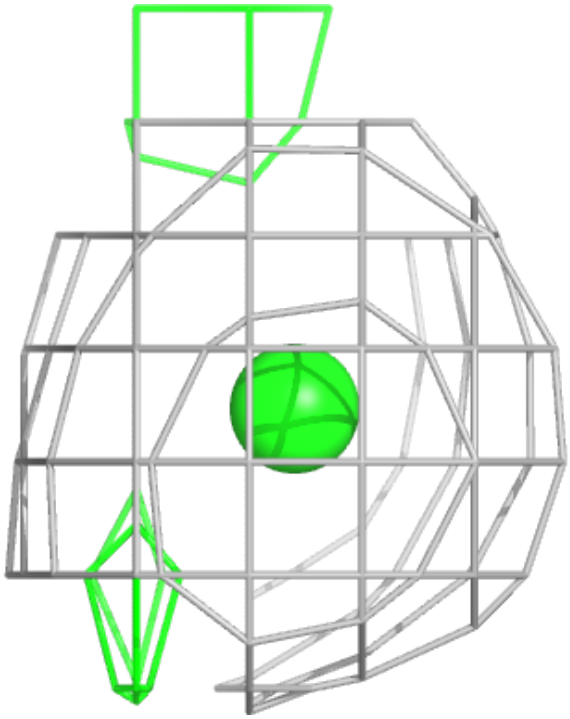


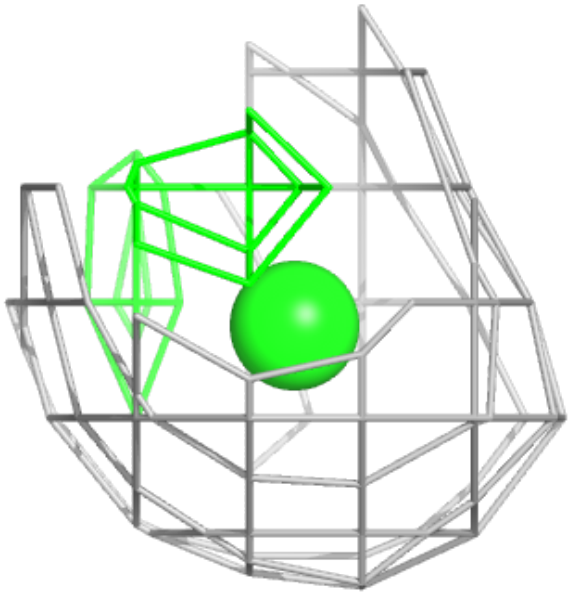


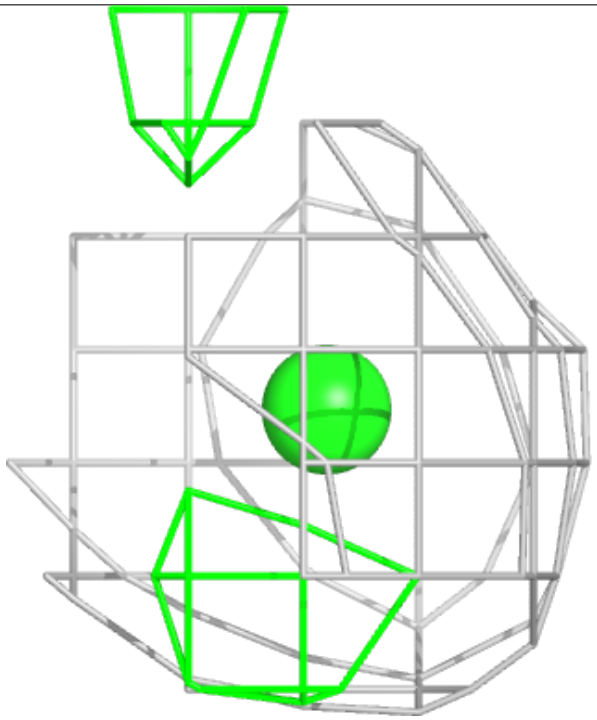


Electron density around CL AaA 1007:

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

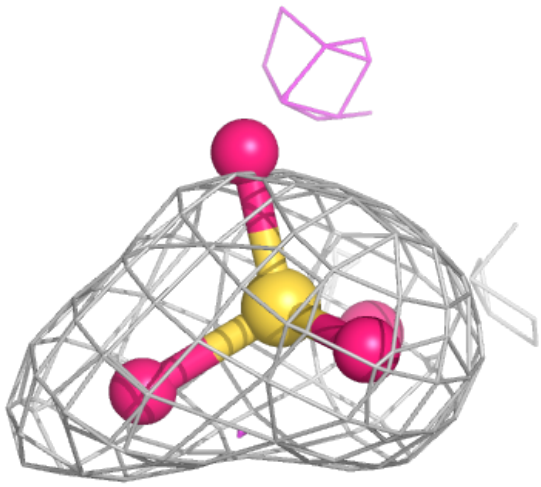


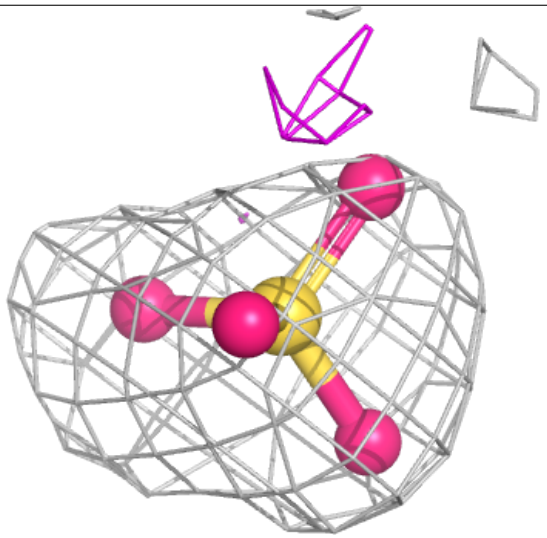


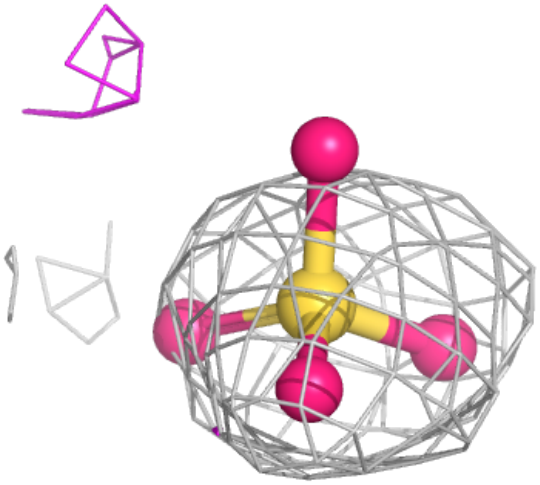


Electron density around SO4 AaA 1005:

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

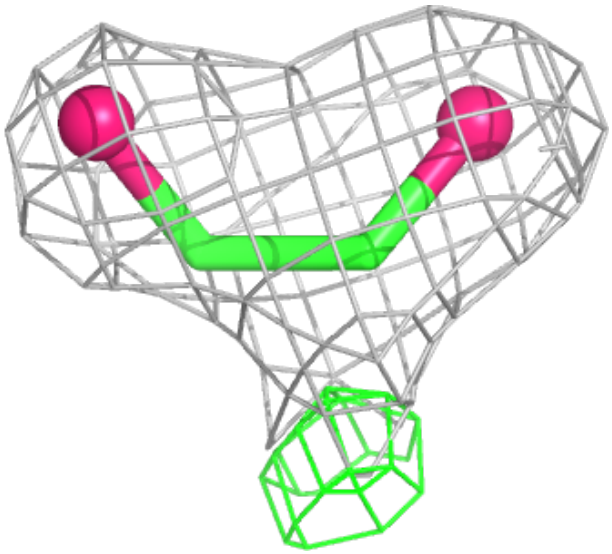


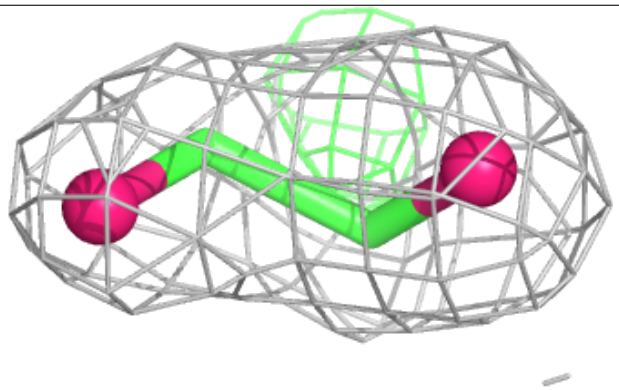


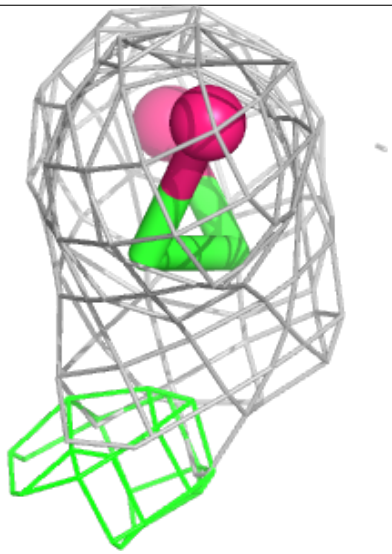


Electron density around EDO AaA 1017:

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

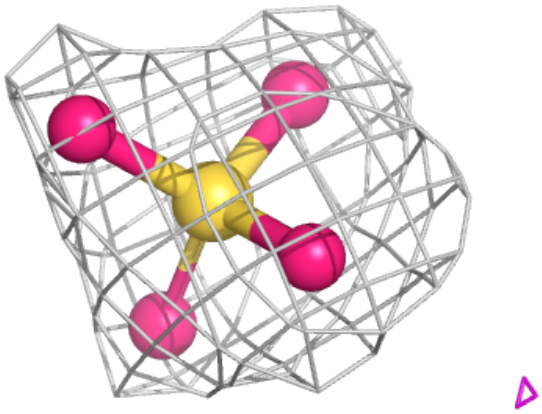


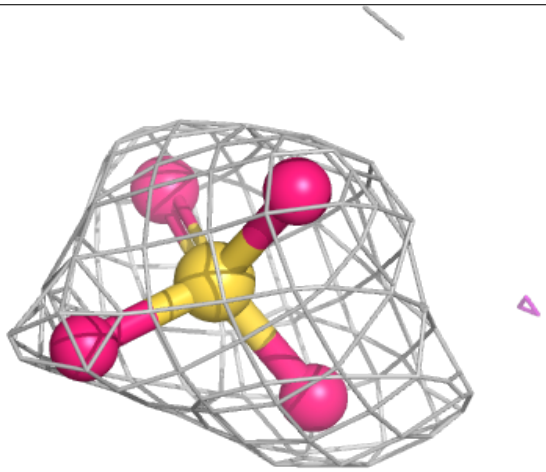


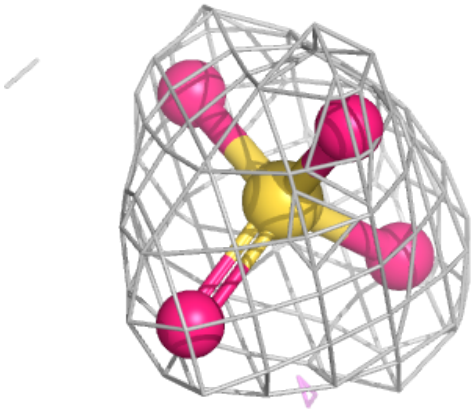


Electron density around SO4 AAA 1001:

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

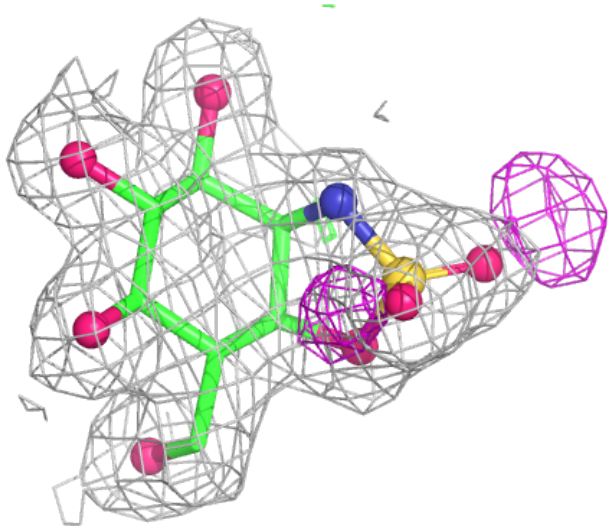


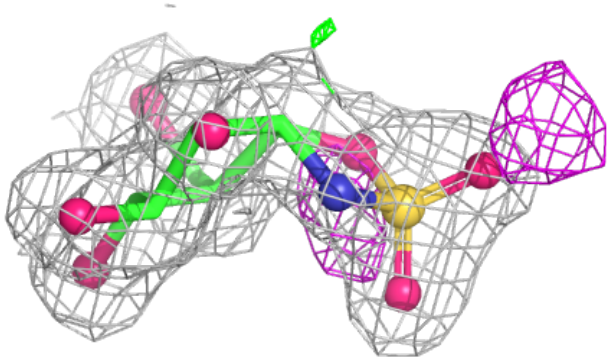


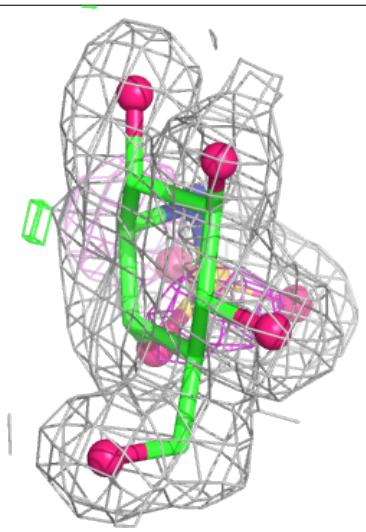


Electron density around YTW AaA 1002:

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

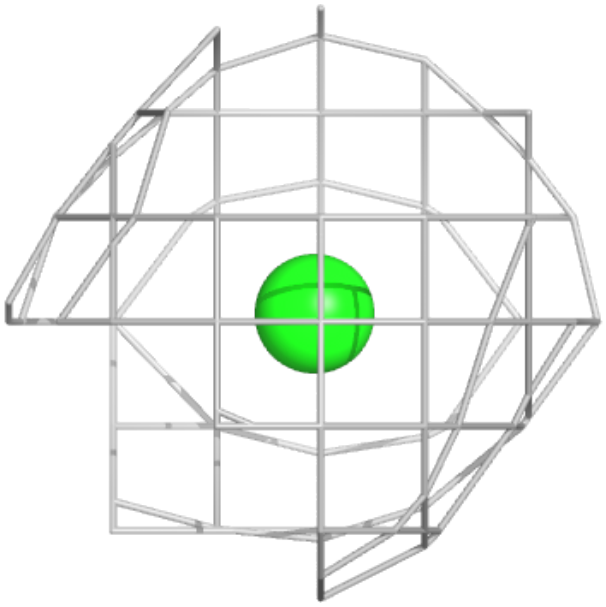


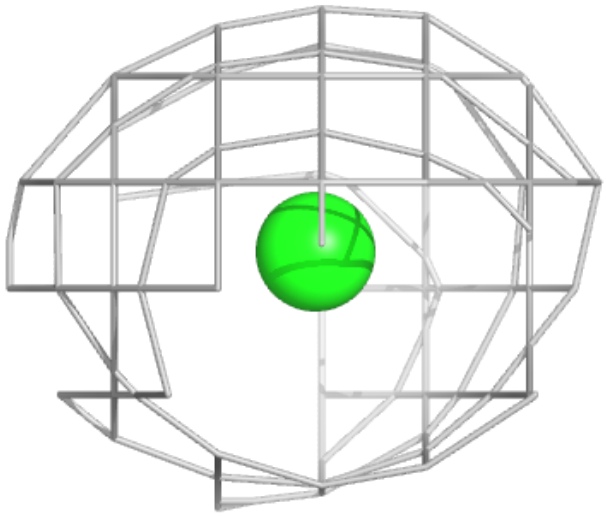


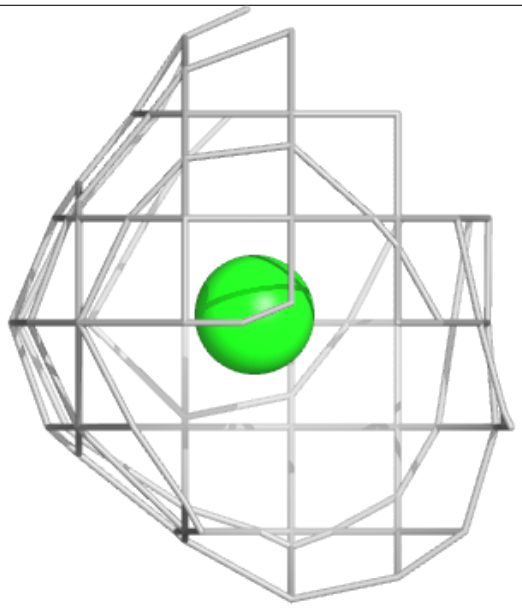


Electron density around CL AaA 1008:

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

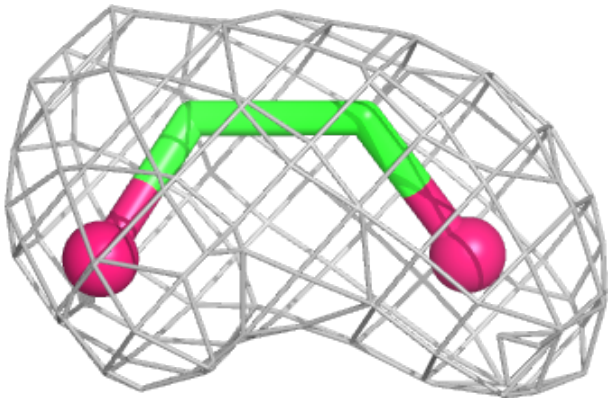


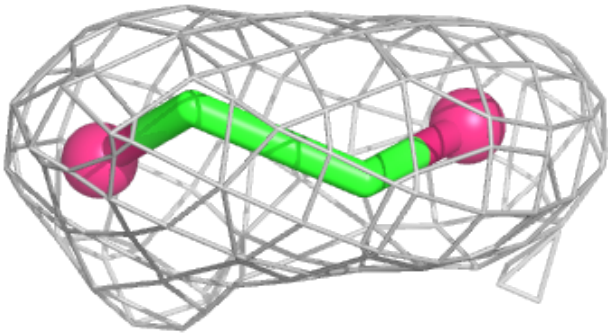


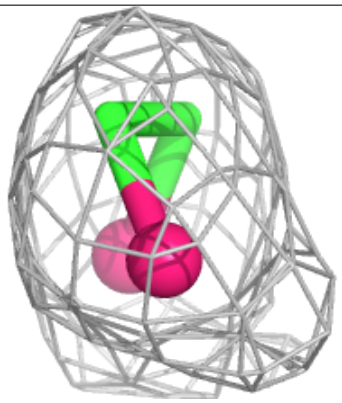


Electron density around EDO AAA 1005:

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

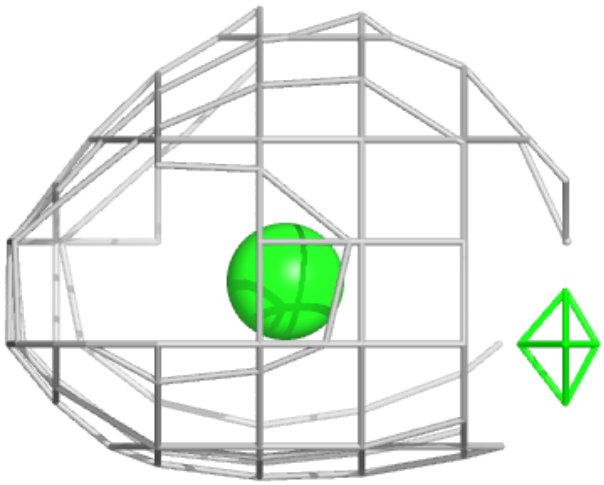


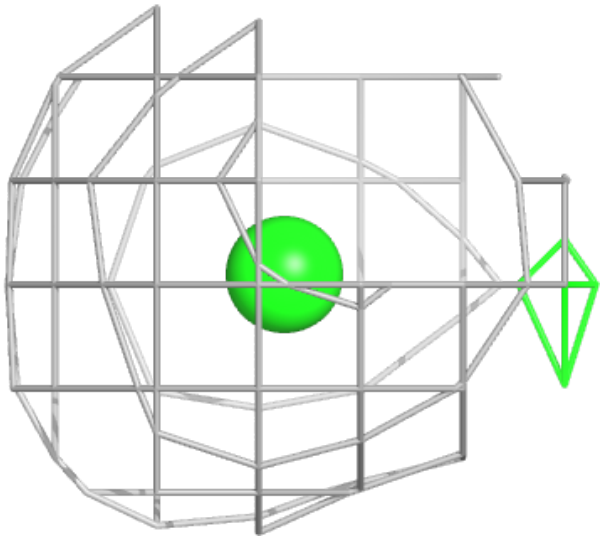


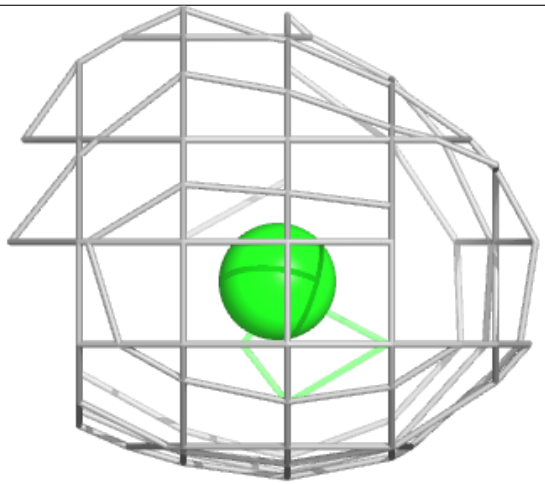


Electron density around CL AaA 1009:

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

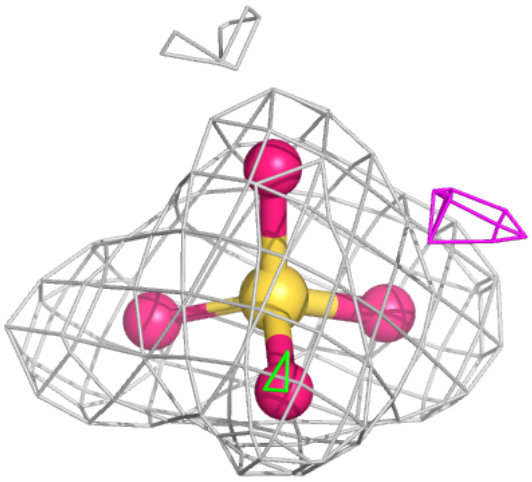


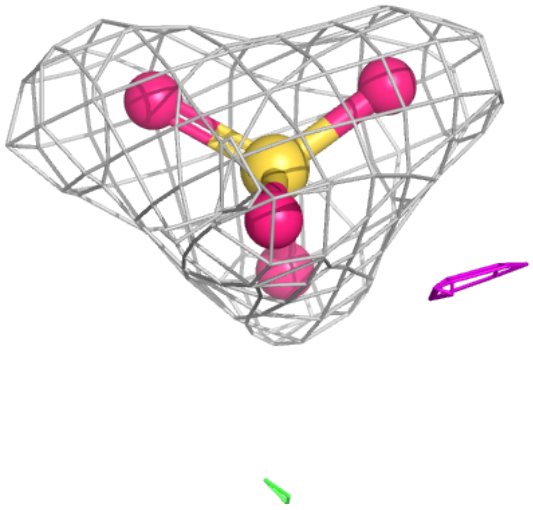


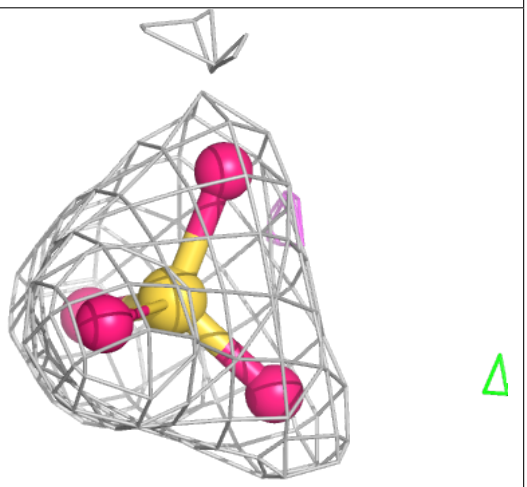


Electron density around SO4 AaA 1006:

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

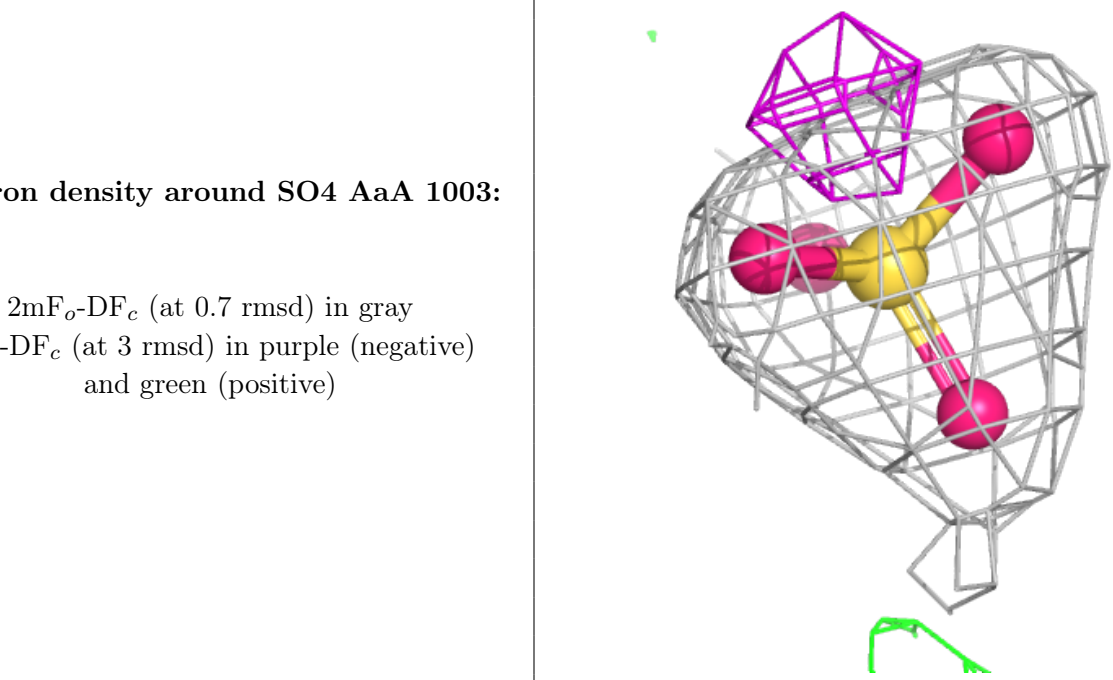


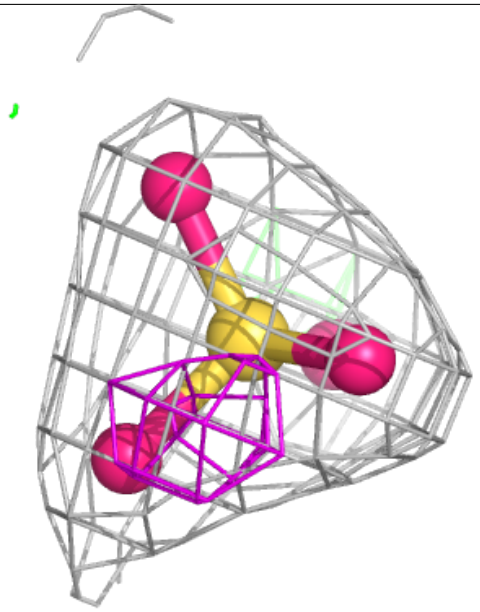


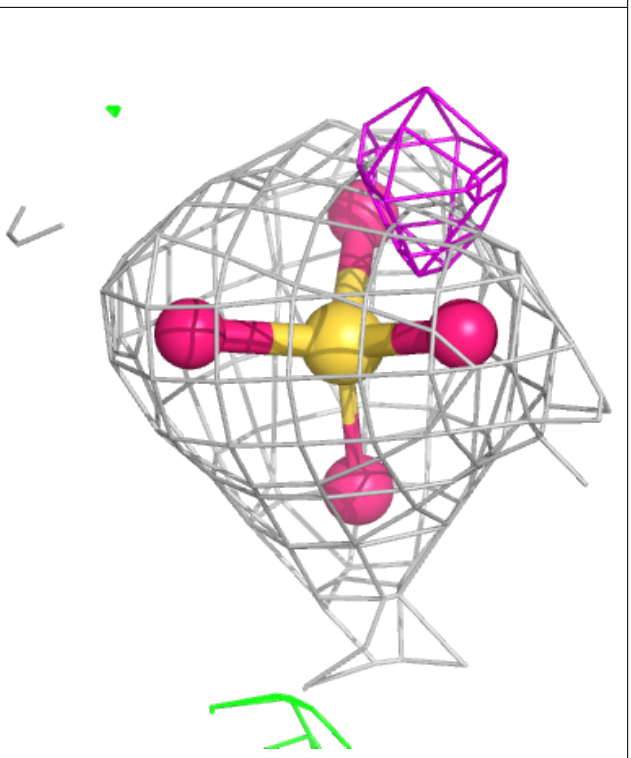


Electron density around SO4 AaA 1003:

$2mF_o$-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

There are no such residues in this entry.