



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 20, 2022 – 01:21 pm BST

PDB ID : 7OZE
Title : Sulfated host glycan recognition by carbohydrate sulfatases of the human gut microbiota (BT1624-S1_15)
Authors : Cartmell, A.
Deposited on : 2021-06-27
Resolution : 2.70 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.28.1
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) : 2.28.1

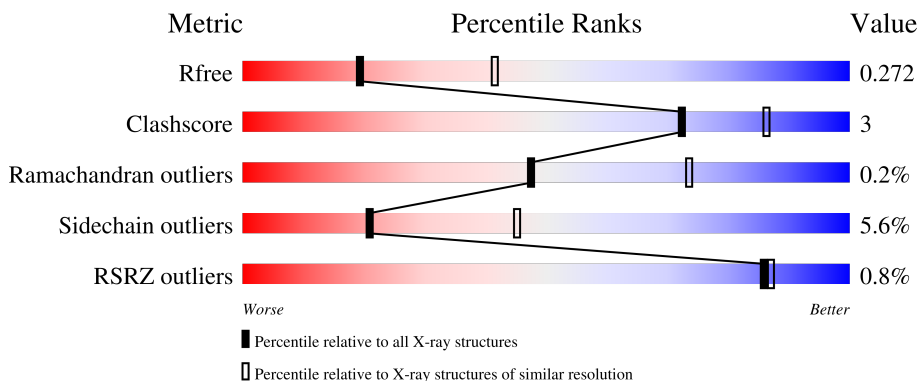
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 $\leq 5\%$. 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	AAA	513	
1	EEE	513	
1	FFF	513	
1	GGG	513	

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 15253 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 Putative secreted sulfatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	487	3736	2365	646	710	15	0	1	0
1	EEE	487	3733	2364	648	706	15	0	1	0
1	FFF	487	3715	2354	645	701	15	0	0	0
1	GGG	487	3734	2366	648	705	15	0	1	0

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	3	MET	-	initiating methionine	UNP Q8A7A1
AAA	4	GLY	-	expression tag	UNP Q8A7A1
AAA	5	SER	-	expression tag	UNP Q8A7A1
AAA	6	SER	-	expression tag	UNP Q8A7A1
AAA	7	HIS	-	expression tag	UNP Q8A7A1
AAA	8	HIS	-	expression tag	UNP Q8A7A1
AAA	9	HIS	-	expression tag	UNP Q8A7A1
AAA	10	HIS	-	expression tag	UNP Q8A7A1
AAA	11	HIS	-	expression tag	UNP Q8A7A1
AAA	12	HIS	-	expression tag	UNP Q8A7A1
AAA	13	SER	-	expression tag	UNP Q8A7A1
AAA	14	SER	-	expression tag	UNP Q8A7A1
AAA	15	GLY	-	expression tag	UNP Q8A7A1
AAA	16	LEU	-	expression tag	UNP Q8A7A1
AAA	17	VAL	-	expression tag	UNP Q8A7A1
AAA	18	PRO	-	expression tag	UNP Q8A7A1
AAA	19	ARG	-	expression tag	UNP Q8A7A1
AAA	20	GLY	-	expression tag	UNP Q8A7A1
AAA	21	SER	-	expression tag	UNP Q8A7A1
AAA	22	HIS	-	expression tag	UNP Q8A7A1
AAA	23	MET	-	expression tag	UNP Q8A7A1

Continued on next page...

Continued from previous page...

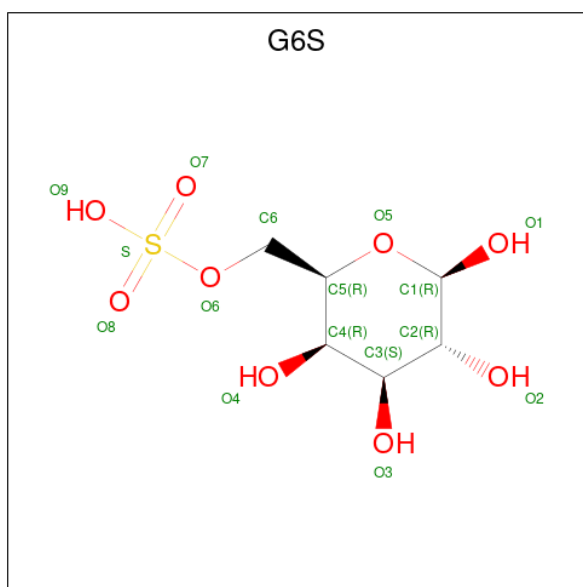
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	24	ALA	-	expression tag	UNP Q8A7A1
AAA	25	SER	-	expression tag	UNP Q8A7A1
EEE	3	MET	-	initiating methionine	UNP Q8A7A1
EEE	4	GLY	-	expression tag	UNP Q8A7A1
EEE	5	SER	-	expression tag	UNP Q8A7A1
EEE	6	SER	-	expression tag	UNP Q8A7A1
EEE	7	HIS	-	expression tag	UNP Q8A7A1
EEE	8	HIS	-	expression tag	UNP Q8A7A1
EEE	9	HIS	-	expression tag	UNP Q8A7A1
EEE	10	HIS	-	expression tag	UNP Q8A7A1
EEE	11	HIS	-	expression tag	UNP Q8A7A1
EEE	12	HIS	-	expression tag	UNP Q8A7A1
EEE	13	SER	-	expression tag	UNP Q8A7A1
EEE	14	SER	-	expression tag	UNP Q8A7A1
EEE	15	GLY	-	expression tag	UNP Q8A7A1
EEE	16	LEU	-	expression tag	UNP Q8A7A1
EEE	17	VAL	-	expression tag	UNP Q8A7A1
EEE	18	PRO	-	expression tag	UNP Q8A7A1
EEE	19	ARG	-	expression tag	UNP Q8A7A1
EEE	20	GLY	-	expression tag	UNP Q8A7A1
EEE	21	SER	-	expression tag	UNP Q8A7A1
EEE	22	HIS	-	expression tag	UNP Q8A7A1
EEE	23	MET	-	expression tag	UNP Q8A7A1
EEE	24	ALA	-	expression tag	UNP Q8A7A1
EEE	25	SER	-	expression tag	UNP Q8A7A1
FFF	3	MET	-	initiating methionine	UNP Q8A7A1
FFF	4	GLY	-	expression tag	UNP Q8A7A1
FFF	5	SER	-	expression tag	UNP Q8A7A1
FFF	6	SER	-	expression tag	UNP Q8A7A1
FFF	7	HIS	-	expression tag	UNP Q8A7A1
FFF	8	HIS	-	expression tag	UNP Q8A7A1
FFF	9	HIS	-	expression tag	UNP Q8A7A1
FFF	10	HIS	-	expression tag	UNP Q8A7A1
FFF	11	HIS	-	expression tag	UNP Q8A7A1
FFF	12	HIS	-	expression tag	UNP Q8A7A1
FFF	13	SER	-	expression tag	UNP Q8A7A1
FFF	14	SER	-	expression tag	UNP Q8A7A1
FFF	15	GLY	-	expression tag	UNP Q8A7A1
FFF	16	LEU	-	expression tag	UNP Q8A7A1
FFF	17	VAL	-	expression tag	UNP Q8A7A1
FFF	18	PRO	-	expression tag	UNP Q8A7A1
FFF	19	ARG	-	expression tag	UNP Q8A7A1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
FFF	20	GLY	-	expression tag	UNP Q8A7A1
FFF	21	SER	-	expression tag	UNP Q8A7A1
FFF	22	HIS	-	expression tag	UNP Q8A7A1
FFF	23	MET	-	expression tag	UNP Q8A7A1
FFF	24	ALA	-	expression tag	UNP Q8A7A1
FFF	25	SER	-	expression tag	UNP Q8A7A1
GGG	3	MET	-	initiating methionine	UNP Q8A7A1
GGG	4	GLY	-	expression tag	UNP Q8A7A1
GGG	5	SER	-	expression tag	UNP Q8A7A1
GGG	6	SER	-	expression tag	UNP Q8A7A1
GGG	7	HIS	-	expression tag	UNP Q8A7A1
GGG	8	HIS	-	expression tag	UNP Q8A7A1
GGG	9	HIS	-	expression tag	UNP Q8A7A1
GGG	10	HIS	-	expression tag	UNP Q8A7A1
GGG	11	HIS	-	expression tag	UNP Q8A7A1
GGG	12	HIS	-	expression tag	UNP Q8A7A1
GGG	13	SER	-	expression tag	UNP Q8A7A1
GGG	14	SER	-	expression tag	UNP Q8A7A1
GGG	15	GLY	-	expression tag	UNP Q8A7A1
GGG	16	LEU	-	expression tag	UNP Q8A7A1
GGG	17	VAL	-	expression tag	UNP Q8A7A1
GGG	18	PRO	-	expression tag	UNP Q8A7A1
GGG	19	ARG	-	expression tag	UNP Q8A7A1
GGG	20	GLY	-	expression tag	UNP Q8A7A1
GGG	21	SER	-	expression tag	UNP Q8A7A1
GGG	22	HIS	-	expression tag	UNP Q8A7A1
GGG	23	MET	-	expression tag	UNP Q8A7A1
GGG	24	ALA	-	expression tag	UNP Q8A7A1
GGG	25	SER	-	expression tag	UNP Q8A7A1

- Molecule 2 is 6-O-sulfo-beta-D-galactopyranose (three-letter code: G6S) (formula: C₆H₁₂O₉S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total	C	O	S	0	0
			16	6	9	1		
2	EEE	1	Total	C	O	S	0	0
			16	6	9	1		
2	FFF	1	Total	C	O	S	0	0
			16	6	9	1		
2	GGG	1	Total	C	O	S	0	0
			16	6	9	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	1	Total	Ca	0	0
			1	1		
3	EEE	1	Total	Ca	0	0
			1	1		
3	FFF	1	Total	Ca	0	0
			1	1		
3	GGG	1	Total	Ca	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	88	Total	O	0	0
			88	88		

Continued on next page...

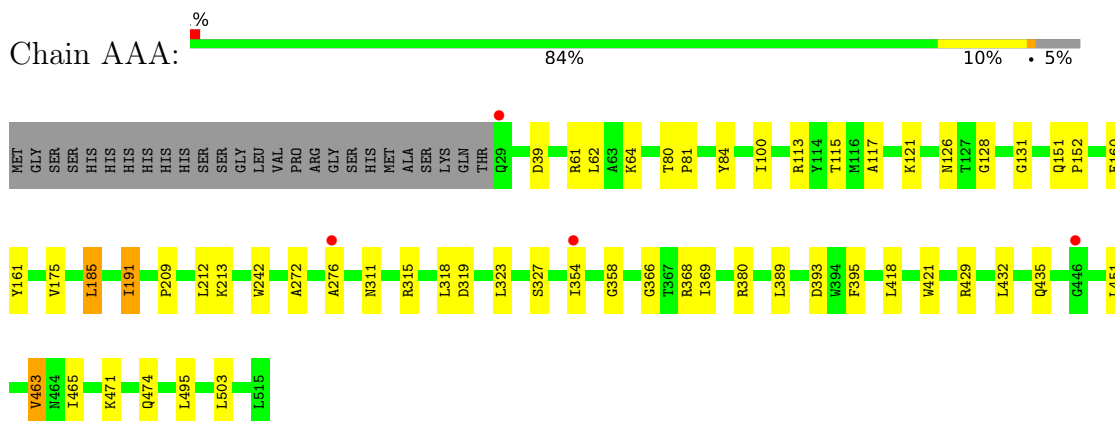
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	EEE	55	Total 55	O 55	0	0
4	FFF	75	Total 75	O 75	0	0
4	GGG	49	Total 49	O 49	0	0

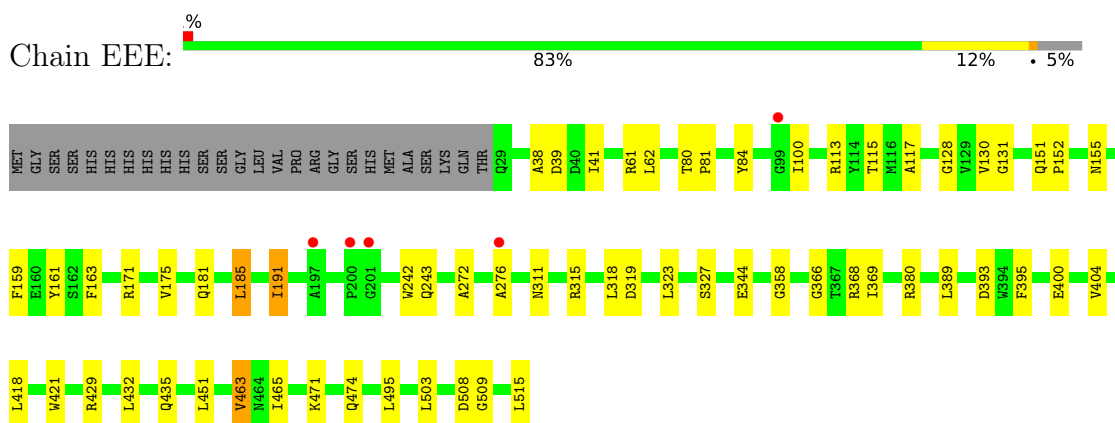
3 Residue-property plots [i](#)

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.

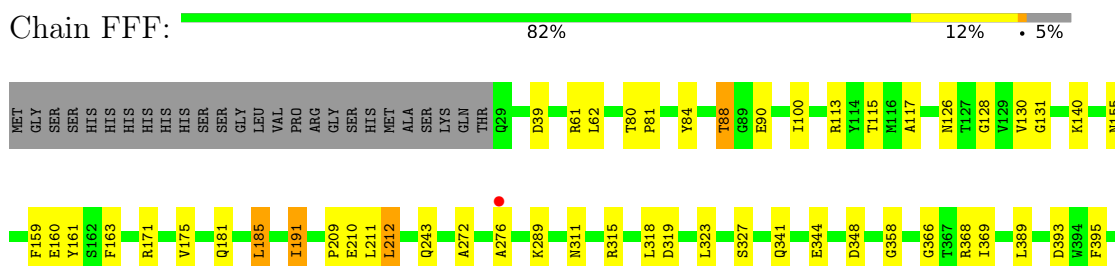
- Molecule 1: Putative secreted sulfatase

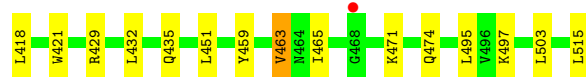


- Molecule 1: Putative secreted sulfatase

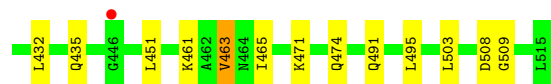
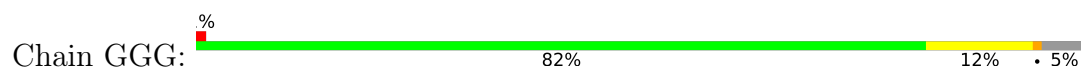


- Molecule 1: Putative secreted sulfatase





- Molecule 1: Putative secreted sulfatase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	86.40Å 124.72Å 97.14Å 90.00° 95.32° 90.00°	Depositor
Resolution (Å)	50.50 – 2.70 50.49 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.50-2.70) 99.5 (50.49-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.66 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.229 , 0.264 0.237 , 0.272	Depositor DCC
R_{free} test set	2736 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å ²)	18.6	Xtriage
Anisotropy	0.842	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	15253	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: CA, G6S

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	0.65	0/3830	0.77	0/5205
1	EEE	0.65	0/3827	0.77	0/5201
1	FFF	0.65	0/3809	0.77	0/5179
1	GGG	0.65	0/3828	0.77	0/5200
All	All	0.65	0/15294	0.77	0/20785

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 [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	AAA	3736	0	3599	24	0
1	EEE	3733	0	3610	25	0
1	FFF	3715	0	3584	30	0
1	GGG	3734	0	3614	26	0
2	AAA	16	0	11	0	0
2	EEE	16	0	11	0	0
2	FFF	16	0	11	0	0
2	GGG	16	0	11	0	0
3	AAA	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	EEE	1	0	0	0	0
3	FFF	1	0	0	0	0
3	GGG	1	0	0	0	0
4	AAA	88	0	0	1	0
4	EEE	55	0	0	0	0
4	FFF	75	0	0	1	0
4	GGG	49	0	0	0	0
All	All	15253	0	14451	100	0

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

All (100) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:GGG:175:VAL:HG11	1:GGG:191:ILE:HD11	1.77	0.66
1:AAA:315:ARG:NH2	4:AAA:701:HOH:O	2.29	0.64
1:FFF:315:ARG:NH2	4:FFF:702:HOH:O	2.33	0.61
1:FFF:88:THR:HG22	1:FFF:90:GLU:H	1.67	0.59
1:FFF:210:GLU:HG2	1:FFF:211:LEU:HD23	1.84	0.59
1:EEE:393:ASP:OD1	1:EEE:429:ARG:NH1	2.36	0.59
1:FFF:61:ARG:NH2	1:FFF:311:ASN:OD1	2.37	0.58
1:GGG:185:LEU:HD11	1:GGG:191:ILE:HG13	1.86	0.57
1:EEE:61:ARG:NH2	1:EEE:311:ASN:OD1	2.38	0.56
1:FFF:88:THR:CG2	1:FFF:90:GLU:HB3	2.36	0.56
1:AAA:61:ARG:NH2	1:AAA:311:ASN:OD1	2.39	0.56
1:AAA:393:ASP:OD1	1:AAA:429:ARG:NH1	2.39	0.56
1:GGG:393:ASP:OD1	1:GGG:429:ARG:NH1	2.39	0.56
1:EEE:389:LEU:HB3	1:EEE:421:TRP:CZ2	2.41	0.55
1:GGG:61:ARG:NH2	1:GGG:311:ASN:OD1	2.39	0.55
1:FFF:319:ASP:OD2	1:GGG:319:ASP:OD2	2.24	0.55
1:AAA:389:LEU:HB3	1:AAA:421:TRP:CZ2	2.42	0.55
1:FFF:389:LEU:HB3	1:FFF:421:TRP:CZ2	2.42	0.54
1:GGG:115:THR:HG22	1:GGG:117:ALA:H	1.72	0.54
1:AAA:115:THR:HG22	1:AAA:117:ALA:H	1.73	0.54
1:FFF:115:THR:HG22	1:FFF:117:ALA:H	1.73	0.54
1:GGG:389:LEU:HB3	1:GGG:421:TRP:CZ2	2.43	0.54
1:EEE:115:THR:HG22	1:EEE:117:ALA:H	1.73	0.53
1:FFF:393:ASP:OD1	1:FFF:429:ARG:NH2	2.42	0.52
1:AAA:175:VAL:HG11	1:AAA:191:ILE:HD11	1.92	0.52
1:GGG:39:ASP:O	1:GGG:276:ALA:HB3	2.11	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:EEE:39:ASP:O	1:EEE:276:ALA:HB3	2.10	0.51
1:AAA:39:ASP:O	1:AAA:276:ALA:HB3	2.10	0.51
1:EEE:451:LEU:HB2	1:EEE:474:GLN:HB2	1.93	0.51
1:FFF:175:VAL:HG11	1:FFF:191:ILE:HD11	1.93	0.50
1:FFF:128:GLY:HA2	1:FFF:161:TYR:O	2.12	0.49
1:FFF:39:ASP:O	1:FFF:276:ALA:HB3	2.12	0.49
1:AAA:128:GLY:HA2	1:AAA:161:TYR:O	2.13	0.49
1:GGG:128:GLY:HA2	1:GGG:161:TYR:O	2.13	0.49
1:EEE:128:GLY:HA2	1:EEE:161:TYR:O	2.13	0.49
1:EEE:175:VAL:HG11	1:EEE:191:ILE:HD11	1.94	0.49
1:AAA:185:LEU:HD11	1:AAA:191:ILE:HG13	1.96	0.48
1:FFF:451:LEU:HB2	1:FFF:474:GLN:HB2	1.95	0.47
1:FFF:348:ASP:OD1	1:FFF:348:ASP:N	2.44	0.47
1:GGG:451:LEU:HB2	1:GGG:474:GLN:HB2	1.96	0.47
1:AAA:451:LEU:HB2	1:AAA:474:GLN:HB2	1.96	0.46
1:AAA:463:VAL:CG1	1:AAA:465:ILE:HD12	2.45	0.46
1:EEE:115:THR:HG22	1:EEE:117:ALA:N	2.31	0.46
1:FFF:131:GLY:HA3	1:FFF:272:ALA:O	2.16	0.46
1:GGG:185:LEU:HD12	1:GGG:185:LEU:HA	1.87	0.45
1:EEE:185:LEU:HD11	1:EEE:191:ILE:HG13	1.98	0.45
1:EEE:131:GLY:HA3	1:EEE:272:ALA:O	2.16	0.45
1:AAA:115:THR:HG22	1:AAA:117:ALA:N	2.30	0.45
1:FFF:115:THR:HG22	1:FFF:117:ALA:N	2.31	0.45
1:FFF:185:LEU:HD12	1:FFF:185:LEU:HA	1.84	0.45
1:AAA:185:LEU:HD12	1:AAA:185:LEU:HA	1.88	0.45
1:GGG:115:THR:HG22	1:GGG:117:ALA:N	2.30	0.45
1:FFF:463:VAL:CG1	1:FFF:465:ILE:HD12	2.47	0.44
1:EEE:155:ASN:HA	1:EEE:159:PHE:O	2.18	0.44
1:AAA:131:GLY:HA3	1:AAA:272:ALA:O	2.17	0.44
1:EEE:327:SER:OG	1:EEE:369:ILE:HD13	2.18	0.44
1:GGG:131:GLY:HA3	1:GGG:272:ALA:O	2.17	0.44
1:AAA:80:THR:N	1:AAA:81:PRO:HD2	2.33	0.43
1:AAA:327:SER:OG	1:AAA:369:ILE:HD13	2.18	0.43
1:FFF:80:THR:N	1:FFF:81:PRO:HD2	2.33	0.43
1:GGG:80:THR:N	1:GGG:81:PRO:HD2	2.34	0.43
1:FFF:185:LEU:HD11	1:FFF:191:ILE:HG13	1.99	0.43
1:EEE:151:GLN:HA	1:EEE:152:PRO:HA	1.88	0.43
1:EEE:463:VAL:CG1	1:EEE:465:ILE:HD12	2.48	0.43
1:GGG:84:TYR:OH	1:GGG:113:ARG:NH2	2.50	0.43
1:FFF:327:SER:OG	1:FFF:369:ILE:HD13	2.18	0.43
1:GGG:327:SER:OG	1:GGG:369:ILE:HD13	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:EEE:185:LEU:HD12	1:EEE:185:LEU:HA	1.86	0.42
1:EEE:80:THR:N	1:EEE:81:PRO:HD2	2.34	0.42
1:EEE:508:ASP:HB3	1:EEE:509:GLY:H	1.78	0.42
1:FFF:126:ASN:HA	1:FFF:160:GLU:OE1	2.19	0.42
1:EEE:400:GLU:HA	1:GGG:210:GLU:O	2.20	0.42
1:FFF:88:THR:CG2	1:FFF:90:GLU:H	2.31	0.42
1:GGG:151:GLN:HA	1:GGG:152:PRO:HA	1.87	0.42
1:EEE:358:GLY:O	1:EEE:366:GLY:HA2	2.20	0.41
1:FFF:84:TYR:OH	1:FFF:113:ARG:NH2	2.51	0.41
1:GGG:358:GLY:O	1:GGG:366:GLY:HA2	2.21	0.41
1:GGG:126:ASN:HA	1:GGG:160:GLU:OE1	2.21	0.41
1:AAA:64:LYS:O	1:EEE:315:ARG:HD3	2.21	0.41
1:AAA:121:LYS:HE3	1:FFF:459:TYR:CE1	2.55	0.41
1:AAA:126:ASN:HA	1:AAA:160:GLU:OE1	2.21	0.41
1:FFF:155:ASN:HA	1:FFF:159:PHE:O	2.20	0.41
1:GGG:209:PRO:HA	1:GGG:212:LEU:HD22	2.03	0.41
1:FFF:130:VAL:HA	1:FFF:163:PHE:O	2.21	0.41
1:GGG:463:VAL:CG1	1:GGG:465:ILE:HD12	2.50	0.41
1:AAA:358:GLY:O	1:AAA:366:GLY:HA2	2.21	0.41
1:EEE:130:VAL:HA	1:EEE:163:PHE:O	2.21	0.41
1:GGG:155:ASN:HA	1:GGG:159:PHE:O	2.20	0.41
1:FFF:209:PRO:HA	1:FFF:212:LEU:HD22	2.02	0.41
1:FFF:341:GLN:HA	1:FFF:344:GLU:OE2	2.21	0.41
1:GGG:62:LEU:HD23	1:GGG:65:MET:HE3	2.03	0.41
1:AAA:151:GLN:HA	1:AAA:152:PRO:HA	1.88	0.40
1:EEE:84:TYR:OH	1:EEE:113:ARG:NH2	2.50	0.40
1:AAA:209:PRO:HA	1:AAA:212:LEU:HD22	2.03	0.40
1:GGG:508:ASP:HB3	1:GGG:509:GLY:H	1.76	0.40
1:FFF:358:GLY:O	1:FFF:366:GLY:HA2	2.21	0.40
1:AAA:84:TYR:OH	1:AAA:113:ARG:NH2	2.53	0.40
1:GGG:30:ARG:N	1:GGG:31:PRO:CD	2.83	0.40
1:AAA:319:ASP:OD2	1:EEE:319:ASP:OD2	2.39	0.40
1:EEE:38:ALA:HB1	1:EEE:41:ILE:HD12	2.02	0.40

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	AAA	486/513 (95%)	471 (97%)	14 (3%)	1 (0%)	47	73
1	EEE	486/513 (95%)	471 (97%)	14 (3%)	1 (0%)	47	73
1	FFF	485/513 (94%)	470 (97%)	14 (3%)	1 (0%)	47	73
1	GGG	486/513 (95%)	470 (97%)	15 (3%)	1 (0%)	47	73
All	All	1943/2052 (95%)	1882 (97%)	57 (3%)	4 (0%)	47	73

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	100	ILE
1	EEE	100	ILE
1	FFF	100	ILE
1	GGG	100	ILE

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	383/414 (92%)	365 (95%)	18 (5%)	26	54
1	EEE	384/414 (93%)	362 (94%)	22 (6%)	20	44
1	FFF	380/414 (92%)	357 (94%)	23 (6%)	18	41
1	GGG	383/414 (92%)	361 (94%)	22 (6%)	20	44
All	All	1530/1656 (92%)	1445 (94%)	85 (6%)	21	45

All (85) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	62	LEU
1	AAA	185	LEU
1	AAA	191	ILE
1	AAA	213	LYS
1	AAA	242	TRP
1	AAA	318	LEU
1	AAA	323	LEU
1	AAA	354	ILE
1	AAA	368	ARG
1	AAA	380	ARG
1	AAA	395	PHE
1	AAA	418	LEU
1	AAA	432	LEU
1	AAA	435	GLN
1	AAA	463	VAL
1	AAA	471	LYS
1	AAA	495	LEU
1	AAA	503	LEU
1	EEE	62	LEU
1	EEE	171	ARG
1	EEE	181	GLN
1	EEE	185	LEU
1	EEE	191	ILE
1	EEE	242	TRP
1	EEE	243	GLN
1	EEE	318	LEU
1	EEE	323	LEU
1	EEE	344	GLU
1	EEE	368	ARG
1	EEE	380	ARG
1	EEE	395	PHE
1	EEE	404	VAL
1	EEE	418	LEU
1	EEE	432	LEU
1	EEE	435	GLN
1	EEE	463	VAL
1	EEE	471	LYS
1	EEE	495	LEU
1	EEE	503	LEU
1	EEE	515	LEU
1	FFF	62	LEU
1	FFF	88	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	FFF	140	LYS
1	FFF	171	ARG
1	FFF	181	GLN
1	FFF	185	LEU
1	FFF	191	ILE
1	FFF	212	LEU
1	FFF	243	GLN
1	FFF	289	LYS
1	FFF	318	LEU
1	FFF	323	LEU
1	FFF	368	ARG
1	FFF	395	PHE
1	FFF	418	LEU
1	FFF	432	LEU
1	FFF	435	GLN
1	FFF	463	VAL
1	FFF	471	LYS
1	FFF	495	LEU
1	FFF	497	LYS
1	FFF	503	LEU
1	FFF	515	LEU
1	GGG	30	ARG
1	GGG	61	ARG
1	GGG	62	LEU
1	GGG	140	LYS
1	GGG	171	ARG
1	GGG	213	LYS
1	GGG	318	LEU
1	GGG	323	LEU
1	GGG	330	ASN
1	GGG	368	ARG
1	GGG	380	ARG
1	GGG	395	PHE
1	GGG	405	ARG
1	GGG	418	LEU
1	GGG	432	LEU
1	GGG	435	GLN
1	GGG	461	LYS
1	GGG	463	VAL
1	GGG	471	LYS
1	GGG	491	GLN
1	GGG	495	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	GGG	503	LEU

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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
2	G6S	FFF	601	3	16,16,16	1.21	1 (6%)	22,24,24	1.03	2 (9%)
2	G6S	GGG	601	3	16,16,16	1.31	1 (6%)	22,24,24	0.90	1 (4%)
2	G6S	EEE	601	3	16,16,16	1.19	1 (6%)	22,24,24	1.20	1 (4%)
2	G6S	AAA	601	3	16,16,16	1.25	1 (6%)	22,24,24	1.32	3 (13%)

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	G6S	FFF	601	3	-	0/6/26/26	0/1/1/1
2	G6S	GGG	601	3	-	0/6/26/26	0/1/1/1
2	G6S	EEE	601	3	-	5/6/26/26	0/1/1/1
2	G6S	AAA	601	3	-	0/6/26/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	GGG	601	G6S	O7-S	4.45	1.64	1.45
2	FFF	601	G6S	O8-S	4.28	1.63	1.45
2	EEE	601	G6S	O7-S	4.11	1.62	1.45
2	AAA	601	G6S	O7-S	4.00	1.62	1.45

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AAA	601	G6S	O5-C5-C6	3.56	113.86	106.67
2	FFF	601	G6S	O5-C5-C6	2.89	112.50	106.67
2	AAA	601	G6S	O5-C1-C2	2.65	115.01	110.28
2	EEE	601	G6S	O5-C5-C6	2.39	111.50	106.67
2	FFF	601	G6S	C3-C4-C5	-2.31	106.12	110.24
2	GGG	601	G6S	O5-C5-C6	2.11	110.92	106.67
2	AAA	601	G6S	C3-C4-C5	-2.04	106.59	110.24

There are no chirality outliers.

All (5) torsion outliers are listed below:

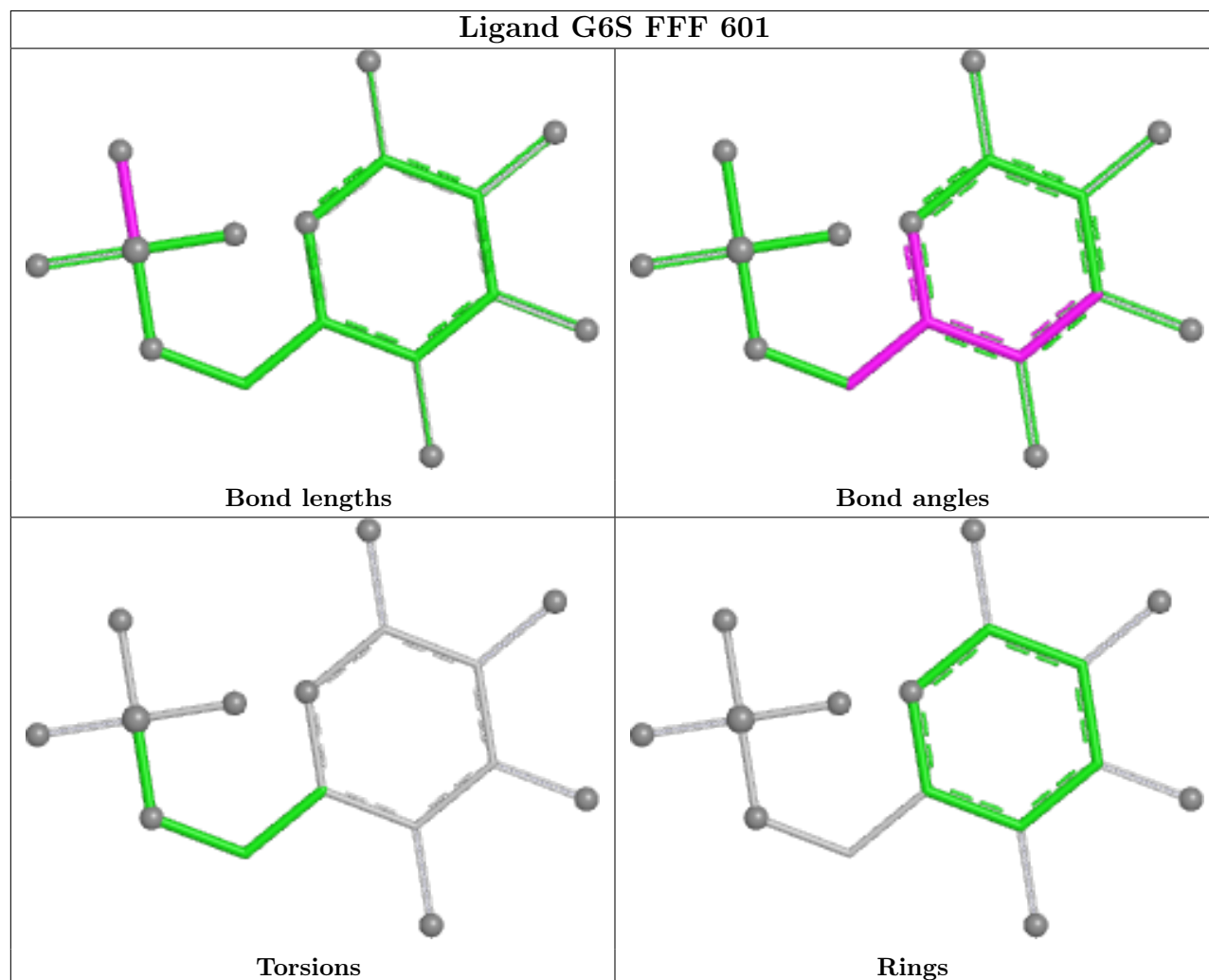
Mol	Chain	Res	Type	Atoms
2	EEE	601	G6S	C4-C5-C6-O6
2	EEE	601	G6S	O5-C5-C6-O6
2	EEE	601	G6S	C6-O6-S-O7
2	EEE	601	G6S	C6-O6-S-O8
2	EEE	601	G6S	C6-O6-S-O9

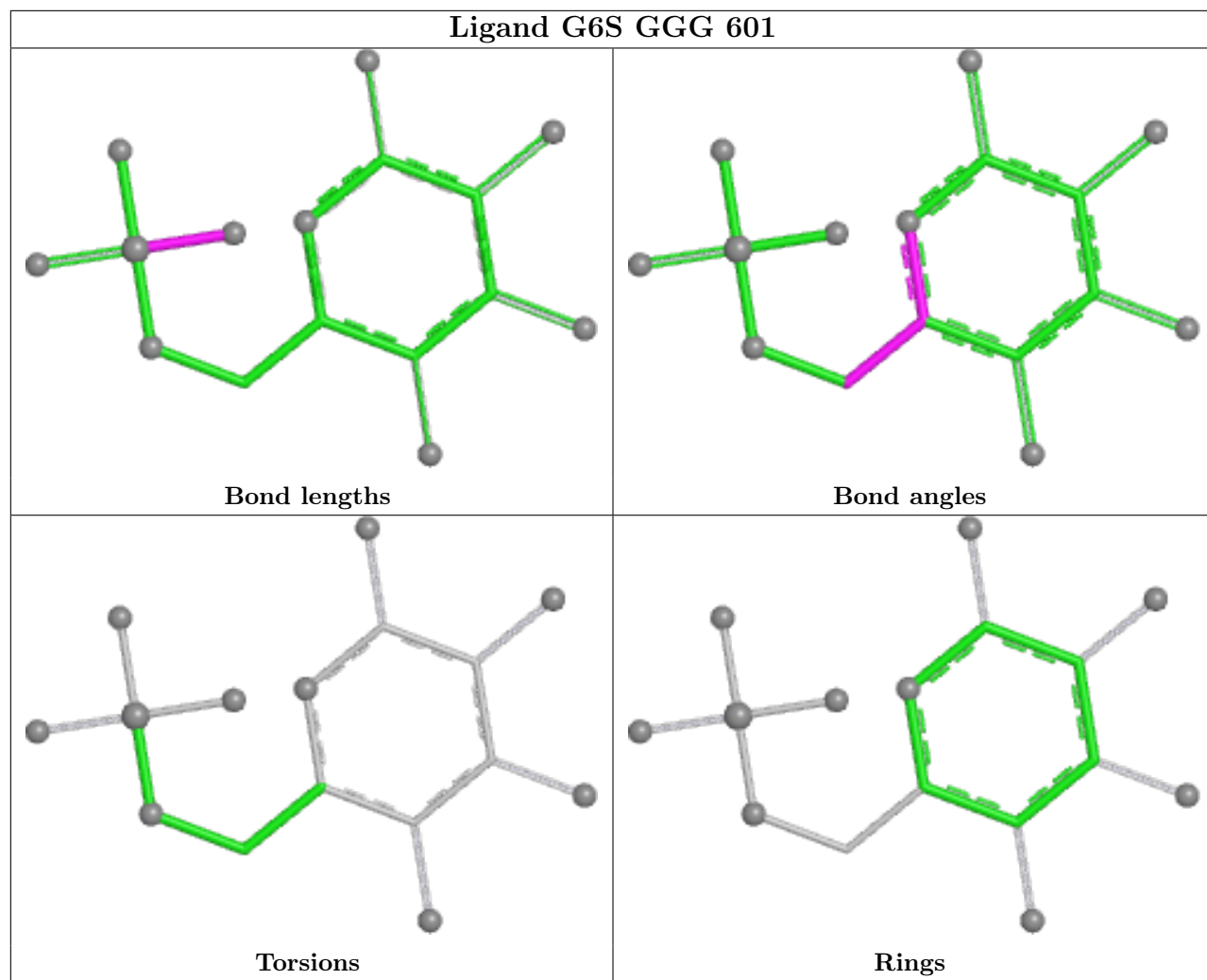
There are no ring outliers.

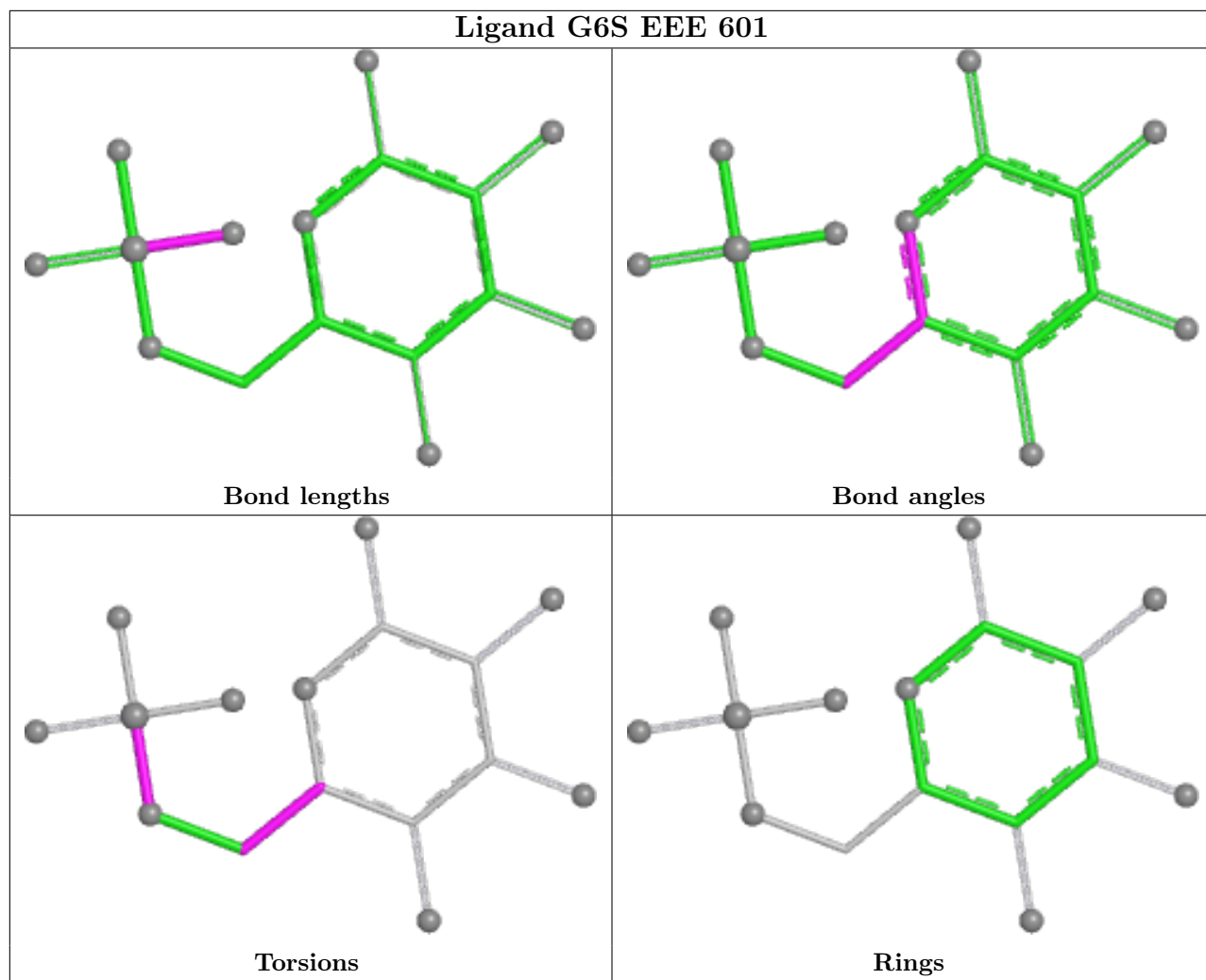
No monomer is involved in short contacts.

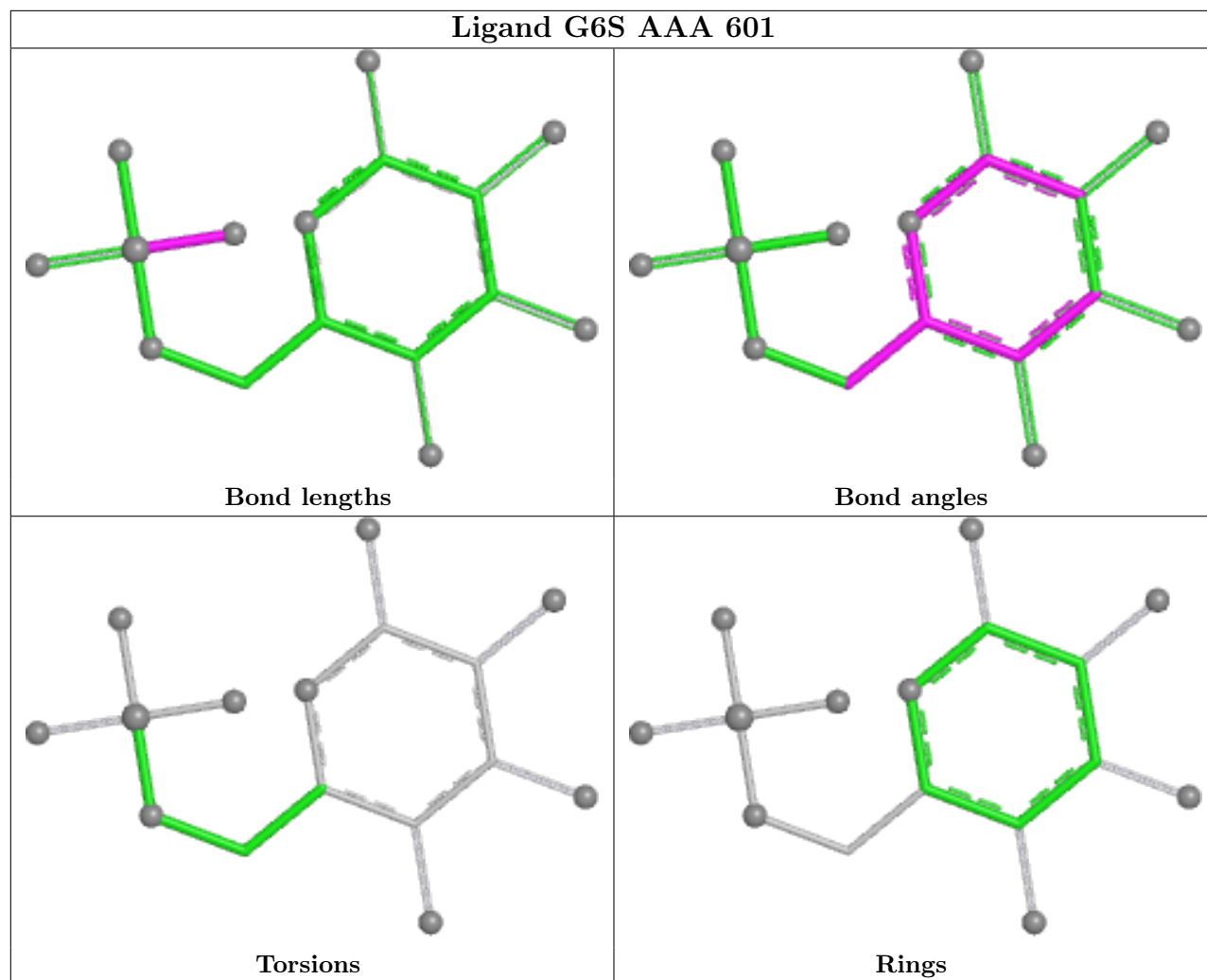
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.









5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [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, 95th 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	487/513 (94%)	0.18	4 (0%) 86 87	10, 17, 28, 43	0
1	EEE	487/513 (94%)	0.18	5 (1%) 82 83	10, 17, 28, 47	0
1	FFF	487/513 (94%)	0.20	2 (0%) 92 93	10, 17, 28, 40	0
1	GGG	487/513 (94%)	0.16	4 (0%) 86 87	10, 17, 26, 44	0
All	All	1948/2052 (94%)	0.18	15 (0%) 86 87	10, 17, 28, 47	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	EEE	276	ALA	3.1
1	FFF	276	ALA	2.9
1	EEE	201	GLY	2.7
1	AAA	446	GLY	2.7
1	EEE	197	ALA	2.5
1	GGG	276	ALA	2.5
1	EEE	200	PRO	2.4
1	AAA	276	ALA	2.4
1	EEE	99	GLY	2.4
1	FFF	468	GLY	2.2
1	GGG	355	TYR	2.1
1	AAA	29	GLN	2.1
1	GGG	446	GLY	2.1
1	AAA	354	ILE	2.0
1	GGG	29	GLN	2.0

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

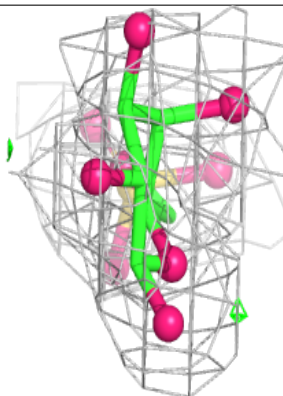
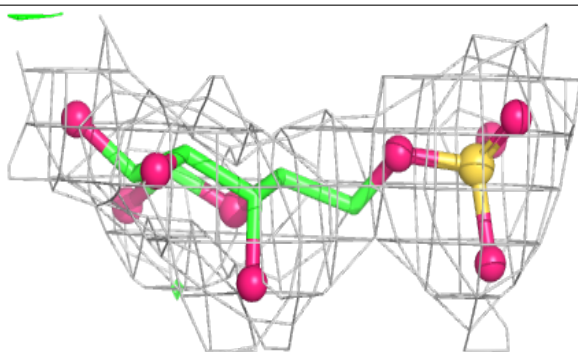
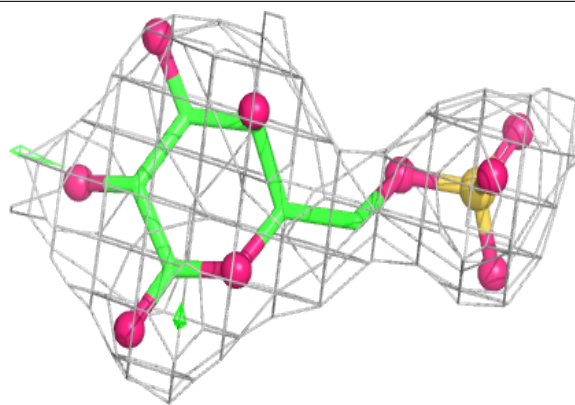
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	CA	AAA	602	1/1	0.74	0.15	46,46,46,46	0
3	CA	FFF	602	1/1	0.78	0.18	38,38,38,38	0
3	CA	EEE	602	1/1	0.92	0.16	45,45,45,45	0
2	G6S	FFF	601	16/16	0.94	0.17	17,18,18,19	0
2	G6S	AAA	601	16/16	0.94	0.19	18,23,25,26	0
3	CA	GGG	602	1/1	0.94	0.14	35,35,35,35	0
2	G6S	EEE	601	16/16	0.96	0.17	12,13,13,13	0
2	G6S	GGG	601	16/16	0.96	0.18	18,19,21,21	0

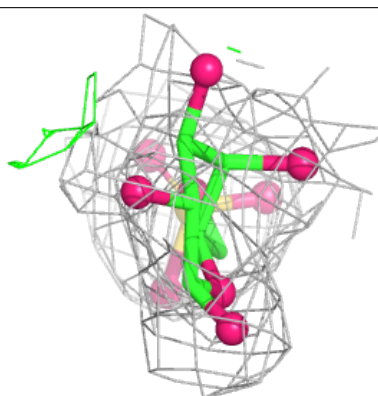
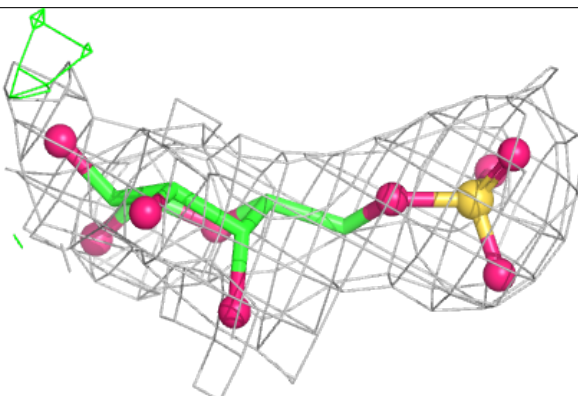
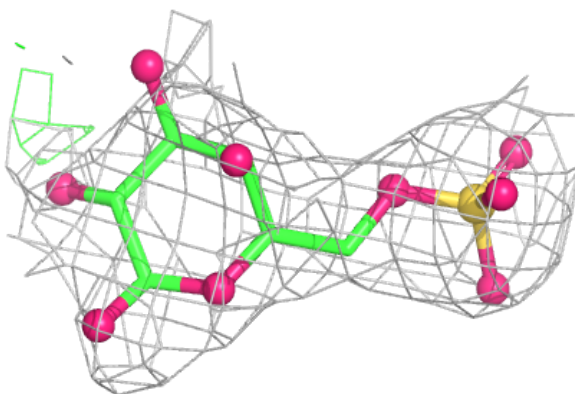
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around G6S FFF 601:

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

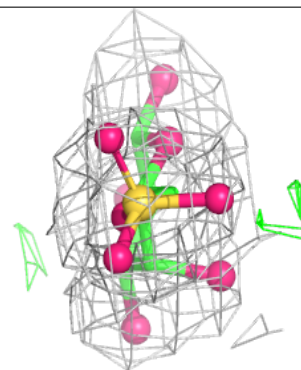
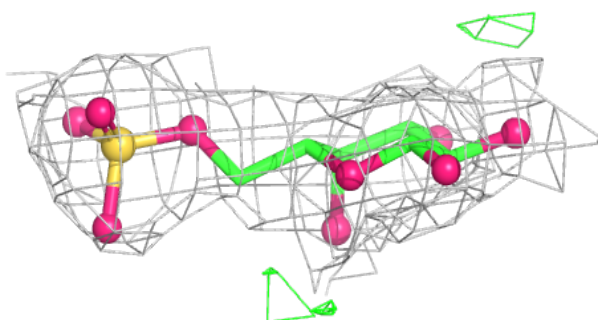
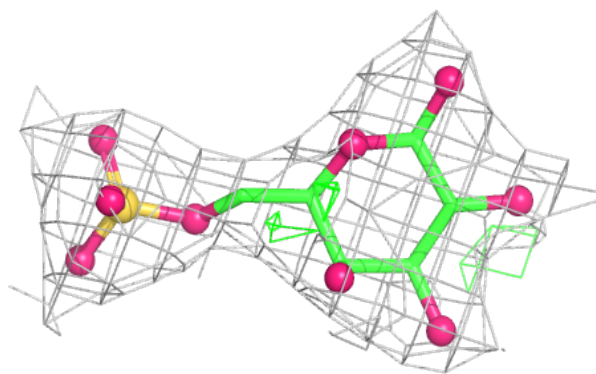
**Electron density around G6S AAA 601:**

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

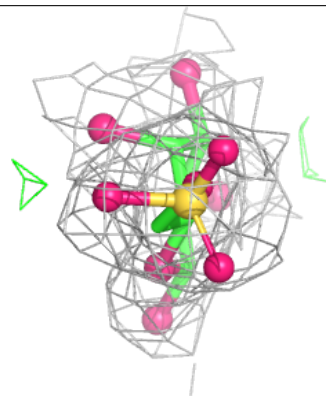
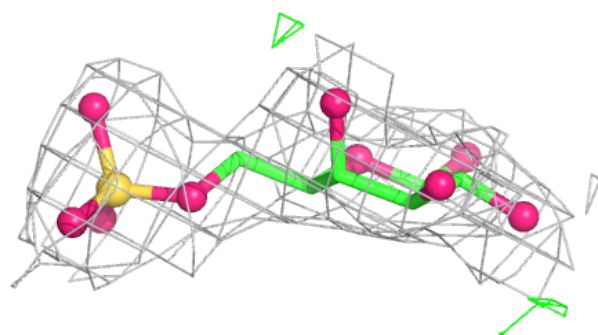
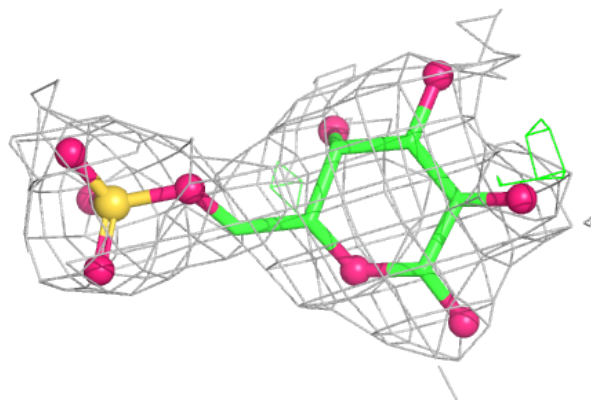


Electron density around G6S EEE 601:

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

**Electron density around G6S GGG 601:**

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



6.5 Other polymers [i](#)

There are no such residues in this entry.