wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2024 – 03:46 pm GMT

PDB ID	:	9FVE
Title	:	Crystal structure of VcSiaP W73A mutant in complex with sialic acid and a VHH antibody (VHH_VcP#2)
Authors	:	Schneberger, N.; Hagelueken, G.
Deposited on	:	2024-06-26
Resolution	:	2.81 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:

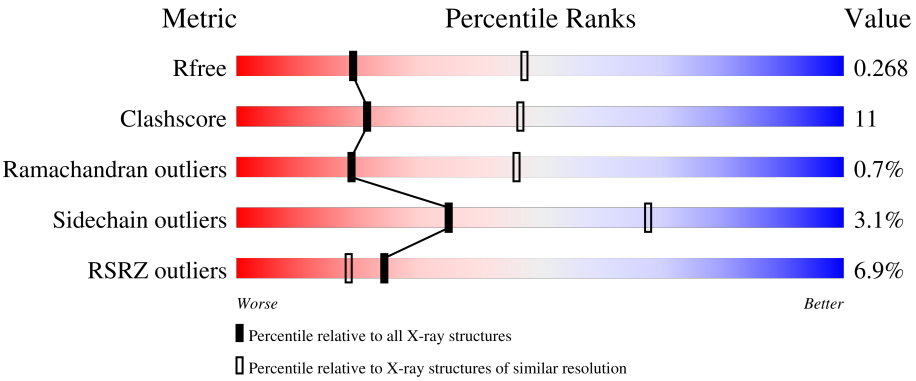
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.81 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}	164625	4293 (2.84-2.80)
Clashscore	180529	4801 (2.84-2.80)
Ramachandran outliers	177936	4739 (2.84-2.80)
Sidechain outliers	177891	4741 (2.84-2.80)
RSRZ outliers	164620	4295 (2.84-2.80)

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	A	303	44% 51% 38% 8%
1	C	303	6% 76% 19%
1	E	303	2% 83% 14%
1	G	303	4% 77% 17%
1	I	303	4% 78% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	303	9% 69% 24%
1	M	303	2% 82% 14%
1	O	303	5% 74% 21%
1	Q	303	% 86% 11%
1	U	303	2% 83% 12%
1	W	303	10% 77% 19%
1	Y	303	3% 82% 15%
2	B	123	6% 85% 10% 5%
2	D	123	6% 80% 13%
2	F	123	6% 89% 10%
2	H	123	6% 86% 10%
2	J	123	7% 83% 13%
2	L	123	4% 86% 9%
2	N	123	10% 82% 13%
2	P	123	2% 82% 15%
2	R	123	2% 88% 9%
2	T	123	5% 89% 8%
2	V	123	5% 89% 9%
2	X	123	2% 91% 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 79063 atoms, of which 39084 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 Sialic acid-binding periplasmic protein SiaP.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	299	Total	C	H	N	O	S	0	0	0
			4701	1489	2341	396	460	15			
1	C	299	Total	C	H	N	O	S	0	0	0
			4696	1489	2336	396	460	15			
1	E	301	Total	C	H	N	O	S	0	0	0
			4717	1494	2348	398	462	15			
1	G	299	Total	C	H	N	O	S	0	0	0
			4700	1489	2340	396	460	15			
1	I	299	Total	C	H	N	O	S	0	0	0
			4702	1489	2342	396	460	15			
1	K	299	Total	C	H	N	O	S	0	0	0
			4701	1489	2341	396	460	15			
1	M	299	Total	C	H	N	O	S	0	0	0
			4701	1489	2341	396	460	15			
1	O	299	Total	C	H	N	O	S	0	0	0
			4700	1489	2340	396	460	15			
1	Q	299	Total	C	H	N	O	S	0	0	0
			4701	1489	2341	396	460	15			
1	U	299	Total	C	H	N	O	S	0	0	0
			4701	1489	2341	396	460	15			
1	W	299	Total	C	H	N	O	S	0	0	0
			4700	1489	2340	396	460	15			
1	Y	299	Total	C	H	N	O	S	0	0	0
			4700	1489	2340	396	460	15			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q9KR64
A	-2	ALA	-	expression tag	UNP Q9KR64
A	-1	MET	-	expression tag	UNP Q9KR64
A	0	GLY	ALA	conflict	UNP Q9KR64
A	73	ALA	TRP	engineered mutation	UNP Q9KR64

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP Q9KR64
C	-2	ALA	-	expression tag	UNP Q9KR64
C	-1	MET	-	expression tag	UNP Q9KR64
C	0	GLY	ALA	conflict	UNP Q9KR64
C	73	ALA	TRP	engineered mutation	UNP Q9KR64
E	-3	GLY	-	expression tag	UNP Q9KR64
E	-2	ALA	-	expression tag	UNP Q9KR64
E	-1	MET	-	expression tag	UNP Q9KR64
E	0	GLY	ALA	conflict	UNP Q9KR64
E	73	ALA	TRP	engineered mutation	UNP Q9KR64
G	-3	GLY	-	expression tag	UNP Q9KR64
G	-2	ALA	-	expression tag	UNP Q9KR64
G	-1	MET	-	expression tag	UNP Q9KR64
G	0	GLY	ALA	conflict	UNP Q9KR64
G	73	ALA	TRP	engineered mutation	UNP Q9KR64
I	-3	GLY	-	expression tag	UNP Q9KR64
I	-2	ALA	-	expression tag	UNP Q9KR64
I	-1	MET	-	expression tag	UNP Q9KR64
I	0	GLY	ALA	conflict	UNP Q9KR64
I	73	ALA	TRP	engineered mutation	UNP Q9KR64
K	-3	GLY	-	expression tag	UNP Q9KR64
K	-2	ALA	-	expression tag	UNP Q9KR64
K	-1	MET	-	expression tag	UNP Q9KR64
K	0	GLY	ALA	conflict	UNP Q9KR64
K	73	ALA	TRP	engineered mutation	UNP Q9KR64
M	-3	GLY	-	expression tag	UNP Q9KR64
M	-2	ALA	-	expression tag	UNP Q9KR64
M	-1	MET	-	expression tag	UNP Q9KR64
M	0	GLY	ALA	conflict	UNP Q9KR64
M	73	ALA	TRP	engineered mutation	UNP Q9KR64
O	-3	GLY	-	expression tag	UNP Q9KR64
O	-2	ALA	-	expression tag	UNP Q9KR64
O	-1	MET	-	expression tag	UNP Q9KR64
O	0	GLY	ALA	conflict	UNP Q9KR64
O	73	ALA	TRP	engineered mutation	UNP Q9KR64
Q	-3	GLY	-	expression tag	UNP Q9KR64
Q	-2	ALA	-	expression tag	UNP Q9KR64
Q	-1	MET	-	expression tag	UNP Q9KR64
Q	0	GLY	ALA	conflict	UNP Q9KR64
Q	73	ALA	TRP	engineered mutation	UNP Q9KR64
U	-3	GLY	-	expression tag	UNP Q9KR64
U	-2	ALA	-	expression tag	UNP Q9KR64

Continued on next page...

Continued from previous page...

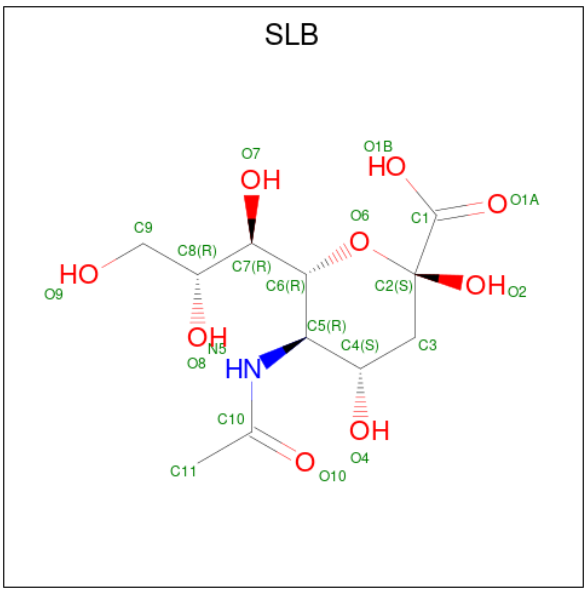
Chain	Residue	Modelled	Actual	Comment	Reference
U	-1	MET	-	expression tag	UNP Q9KR64
U	0	GLY	ALA	conflict	UNP Q9KR64
U	73	ALA	TRP	engineered mutation	UNP Q9KR64
W	-3	GLY	-	expression tag	UNP Q9KR64
W	-2	ALA	-	expression tag	UNP Q9KR64
W	-1	MET	-	expression tag	UNP Q9KR64
W	0	GLY	ALA	conflict	UNP Q9KR64
W	73	ALA	TRP	engineered mutation	UNP Q9KR64
Y	-3	GLY	-	expression tag	UNP Q9KR64
Y	-2	ALA	-	expression tag	UNP Q9KR64
Y	-1	MET	-	expression tag	UNP Q9KR64
Y	0	GLY	ALA	conflict	UNP Q9KR64
Y	73	ALA	TRP	engineered mutation	UNP Q9KR64

- Molecule 2 is a protein called VHH_VcP#2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	122	Total	C	H	N	O	S	0	0	0
			1836	590	891	168	183	4			
2	D	121	Total	C	H	N	O	S	0	0	0
			1826	587	887	167	181	4			
2	F	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	H	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	J	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	L	123	Total	C	H	N	O	S	0	0	0
			1847	592	898	169	184	4			
2	N	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	P	123	Total	C	H	N	O	S	0	0	0
			1846	592	897	169	184	4			
2	R	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	T	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	V	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			
2	X	123	Total	C	H	N	O	S	0	0	0
			1845	592	896	169	184	4			

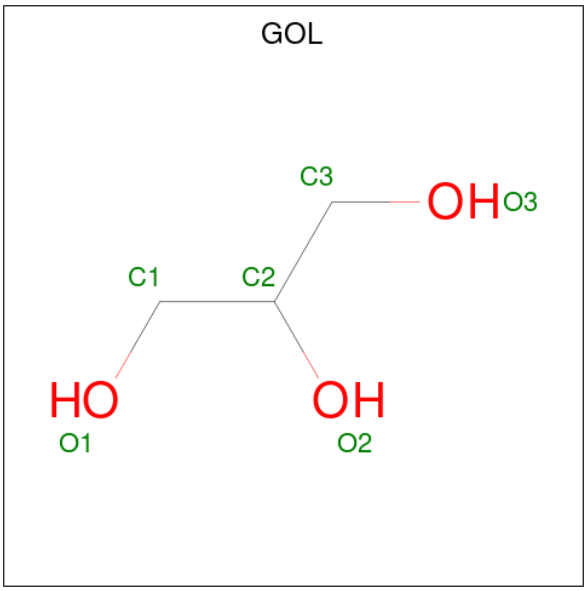
- Molecule 3 is N-acetyl-beta-neuraminic acid (three-letter code: SLB) (formula: $C_{11}H_{19}NO_9$)

(labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	C	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	E	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	G	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	I	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	K	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	M	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	O	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	Q	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	U	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	W	1	Total	C	H	N	O	0	0
			37	11	17	1	8		
3	Y	1	Total	C	H	N	O	0	0
			37	11	17	1	8		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



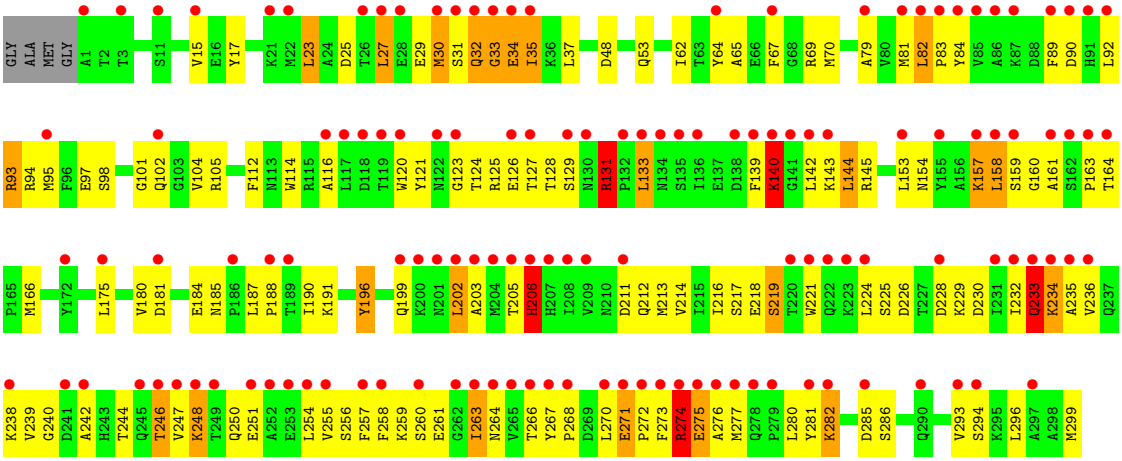
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	E	1	Total	C	H	O	0	0
			14	3	8	3		
4	G	1	Total	C	H	O	0	0
			14	3	8	3		
4	Q	1	Total	C	H	O	0	0
			14	3	8	3		
4	U	1	Total	C	H	O	0	0
			14	3	8	3		
4	U	1	Total	C	H	O	0	0
			14	3	8	3		
4	U	1	Total	C	H	O	0	0
			14	3	8	3		

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.

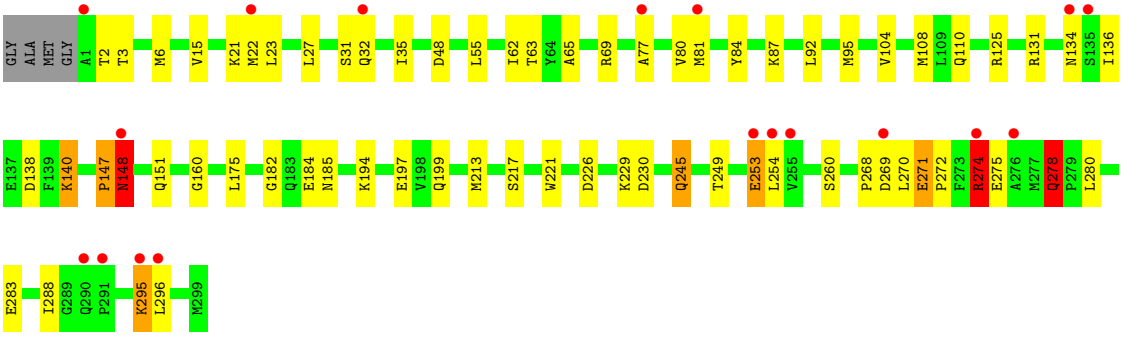
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain A:



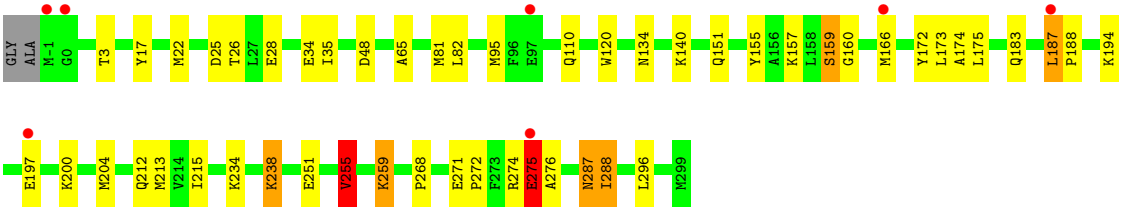
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain C:

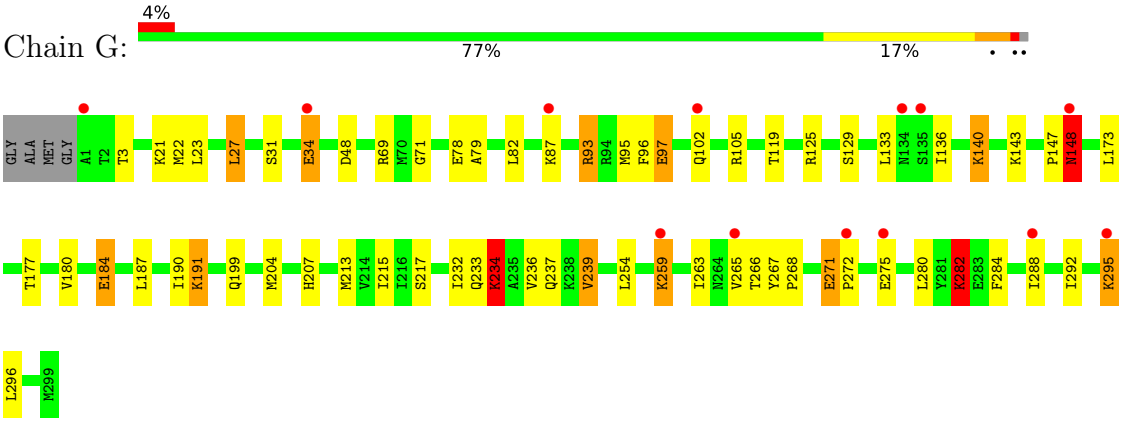


• Molecule 1: Sialic acid-binding periplasmic protein SiaP

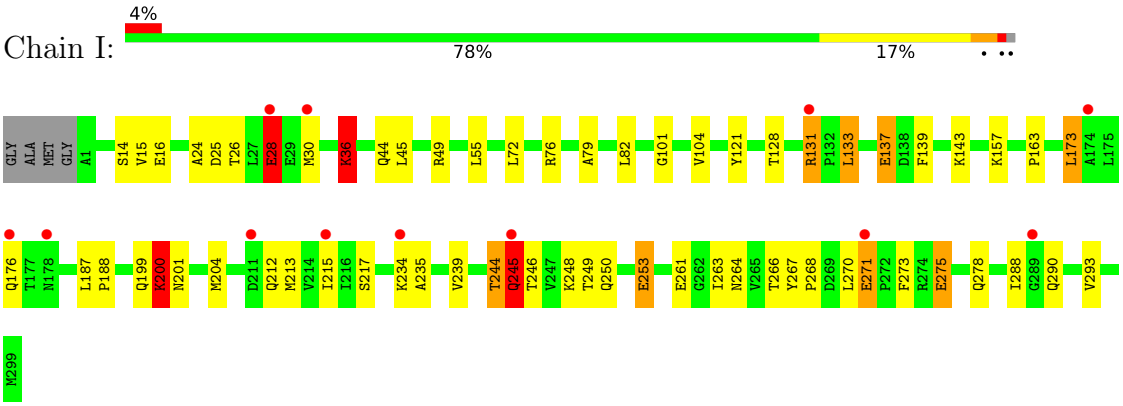
Chain E:



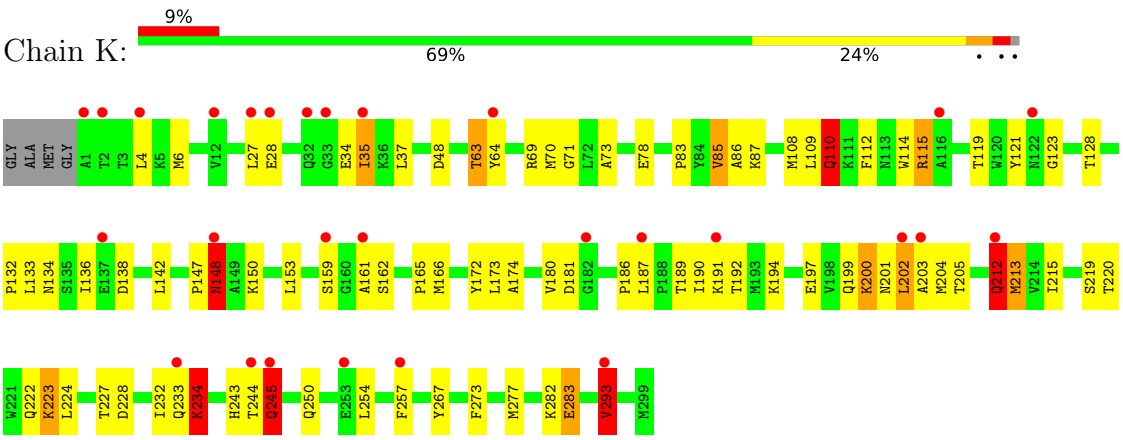
• Molecule 1: Sialic acid-binding periplasmic protein SiaP



• Molecule 1: Sialic acid-binding periplasmic protein SiaP

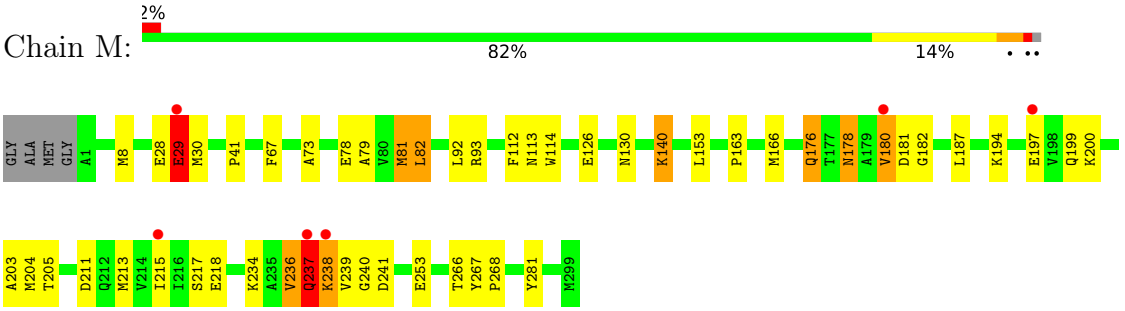


• Molecule 1: Sialic acid-binding periplasmic protein SiaP



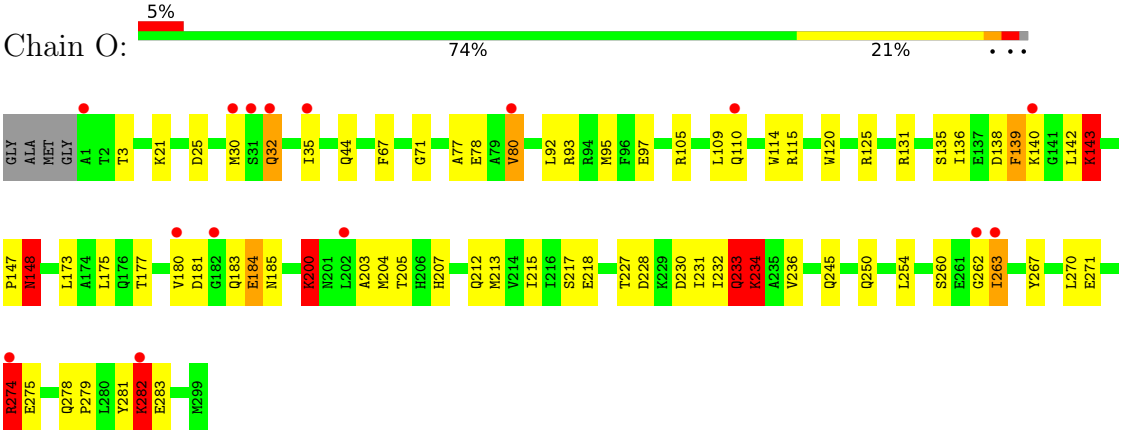
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain M:



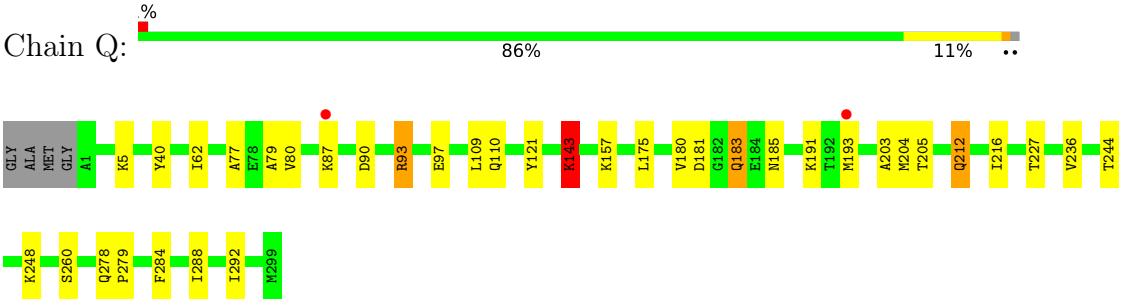
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain O:



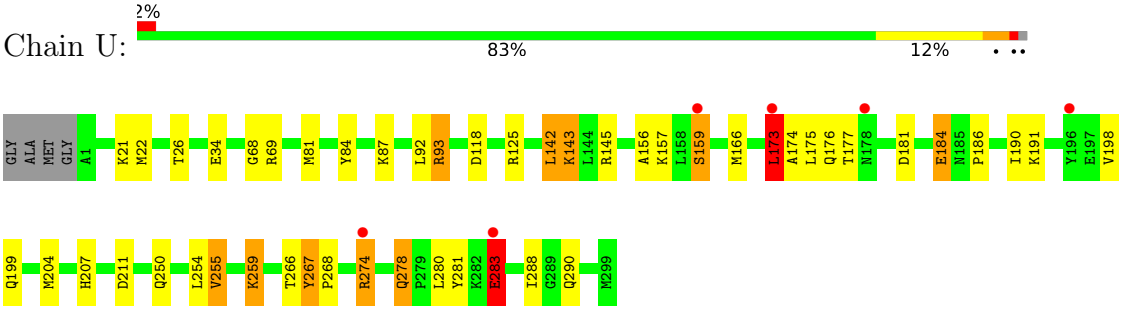
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain Q:



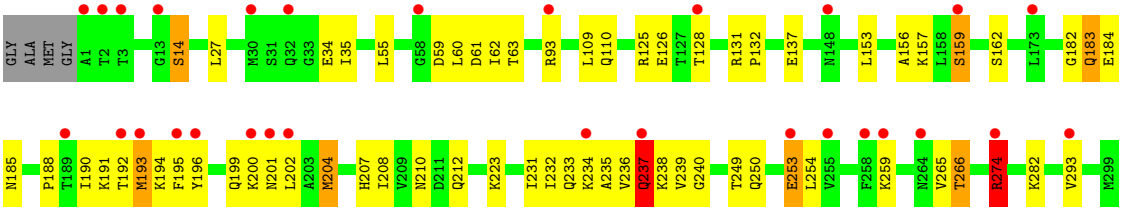
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain U:

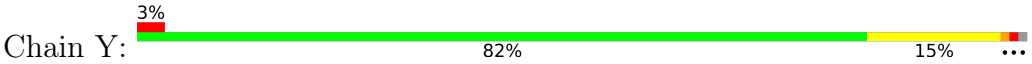


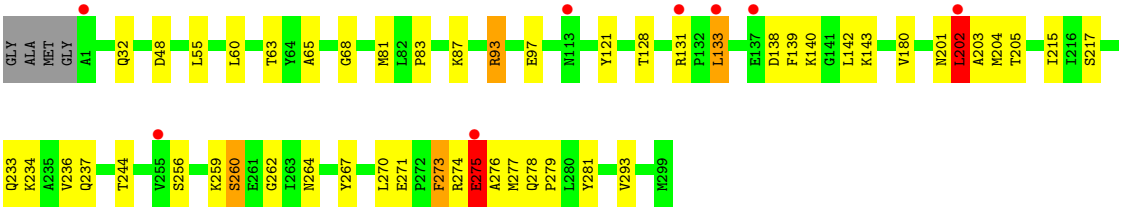
• Molecule 1: Sialic acid-binding periplasmic protein SiaP

Chain W:

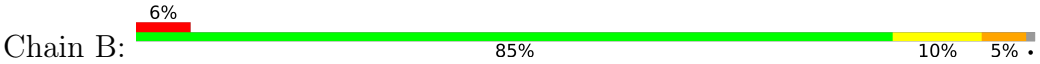


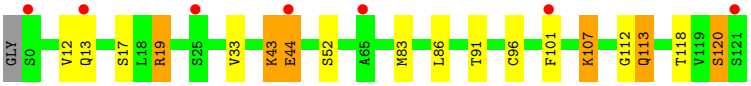
• Molecule 1: Sialic acid-binding periplasmic protein SiaP



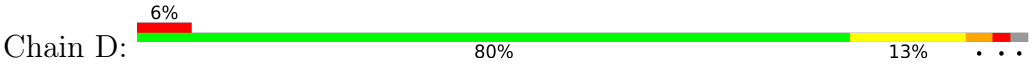


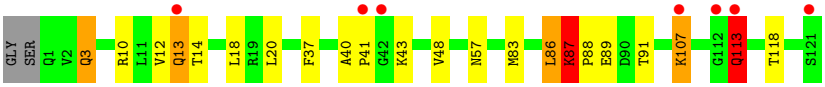
• Molecule 2: VHH_VcP#2



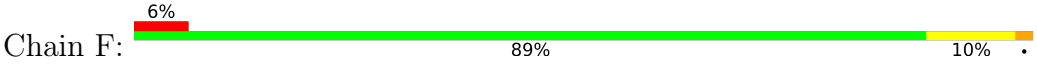


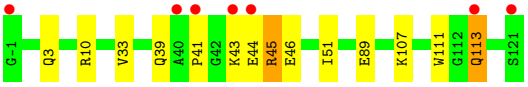
• Molecule 2: VHH_VcP#2



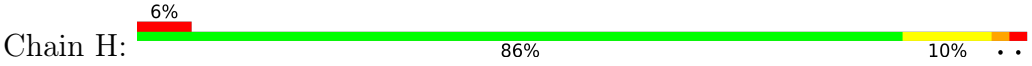


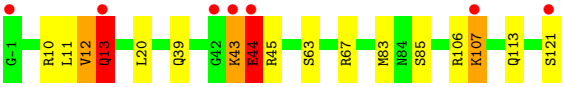
• Molecule 2: VHH_VcP#2





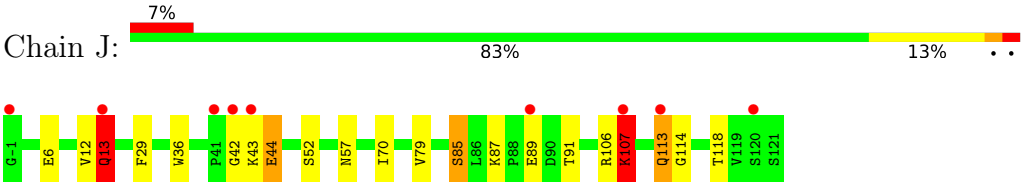
• Molecule 2: VHH_VcP#2





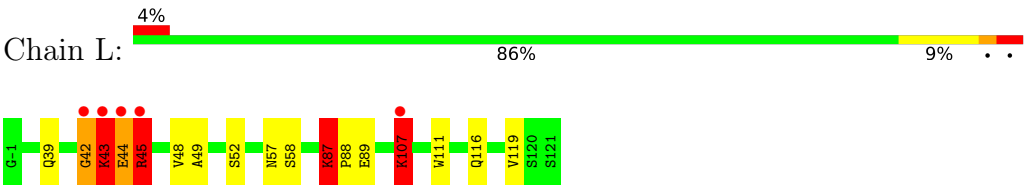
• Molecule 2: VHH_VcP#2

Chain J:



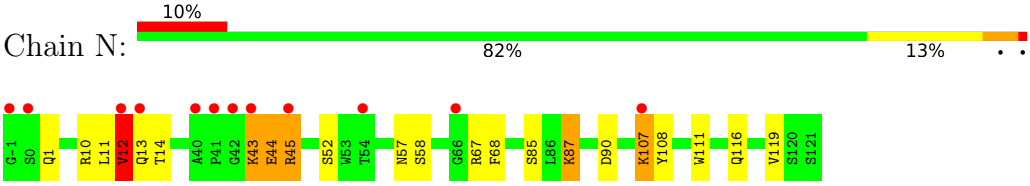
• Molecule 2: VHH_VcP#2

Chain L:



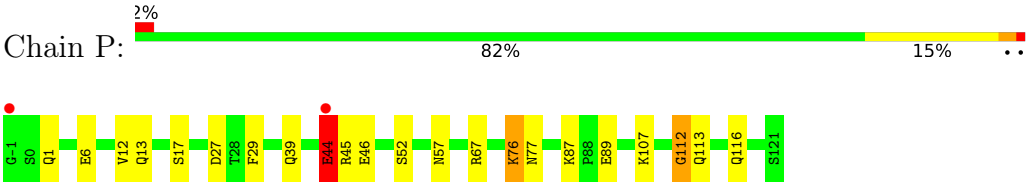
• Molecule 2: VHH_VcP#2

Chain N:



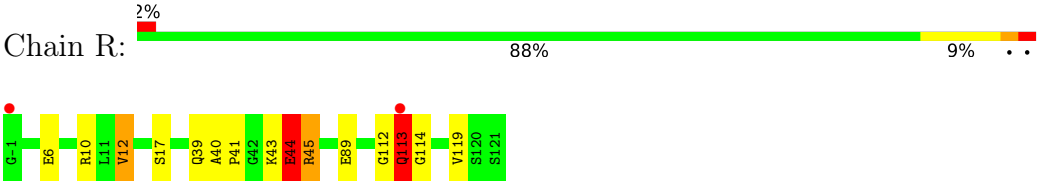
• Molecule 2: VHH_VcP#2

Chain P:



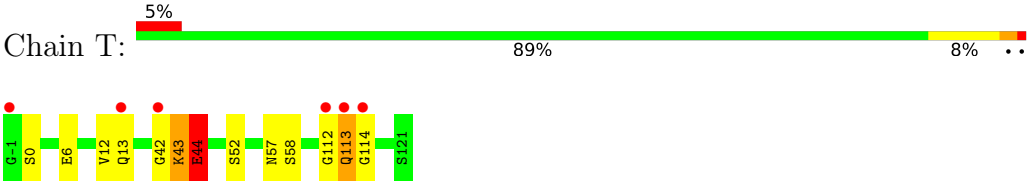
• Molecule 2: VHH_VcP#2

Chain R:



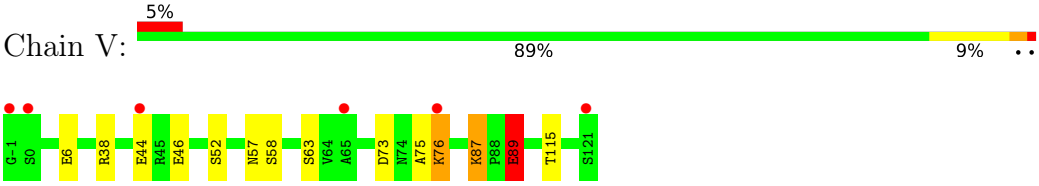
• Molecule 2: VHH_VcP#2

Chain T:



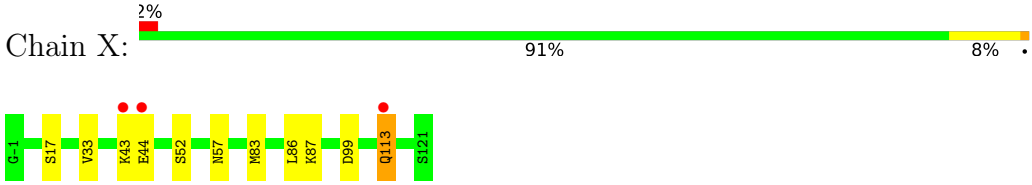
• Molecule 2: VHH_VcP#2

Chain V:



• Molecule 2: VHH_VcP#2

Chain X:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	223.57Å 153.11Å 210.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.24 – 2.81 34.24 – 2.81	Depositor EDS
% Data completeness (in resolution range)	99.1 (34.24-2.81) 99.0 (34.24-2.81)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.25 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.234 , 0.267 0.233 , 0.268	Depositor DCC
R_{free} test set	170340 reflections (1.17%)	wwPDB-VP
Wilson B-factor (Å^2)	44.9	Xtriage
Anisotropy	0.149	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 38.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.008 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	79063	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.07 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3500e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SLB, GOL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.80	9/2404 (0.4%)	1.18	24/3248 (0.7%)
1	C	0.61	7/2404 (0.3%)	0.86	12/3248 (0.4%)
1	E	0.70	13/2413 (0.5%)	0.85	12/3260 (0.4%)
1	G	0.92	16/2404 (0.7%)	0.97	19/3248 (0.6%)
1	I	0.94	14/2404 (0.6%)	1.11	27/3248 (0.8%)
1	K	0.86	14/2404 (0.6%)	1.15	34/3248 (1.0%)
1	M	0.67	7/2404 (0.3%)	1.02	13/3248 (0.4%)
1	O	0.83	13/2404 (0.5%)	1.02	25/3248 (0.8%)
1	Q	0.38	1/2404 (0.0%)	0.70	8/3248 (0.2%)
1	U	0.66	7/2404 (0.3%)	0.83	11/3248 (0.3%)
1	W	0.58	6/2404 (0.2%)	0.92	17/3248 (0.5%)
1	Y	0.51	4/2404 (0.2%)	0.79	12/3248 (0.4%)
2	B	1.23	7/967 (0.7%)	1.03	5/1311 (0.4%)
2	D	0.65	4/961 (0.4%)	1.02	8/1303 (0.6%)
2	F	0.71	4/971 (0.4%)	1.09	14/1316 (1.1%)
2	H	0.85	4/971 (0.4%)	0.90	8/1316 (0.6%)
2	J	0.62	3/971 (0.3%)	0.85	4/1316 (0.3%)
2	L	0.70	4/971 (0.4%)	0.95	9/1316 (0.7%)
2	N	0.95	5/971 (0.5%)	1.00	8/1316 (0.6%)
2	P	0.74	5/971 (0.5%)	1.05	10/1316 (0.8%)
2	R	0.81	5/971 (0.5%)	0.91	7/1316 (0.5%)
2	T	0.38	0/971	0.74	2/1316 (0.2%)
2	V	0.68	5/971 (0.5%)	0.75	3/1316 (0.2%)
2	X	0.53	3/971 (0.3%)	0.68	2/1316 (0.2%)
All	All	0.74	160/40495 (0.4%)	0.95	294/54762 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
1	C	0	3
1	E	0	2
1	G	0	2
1	I	0	2
1	K	0	3
1	M	0	5
1	O	0	6
1	U	0	3
1	W	0	5
1	Y	0	3
2	B	0	1
2	D	0	3
2	H	0	1
2	J	0	2
2	L	0	3
2	N	0	1
2	P	0	3
2	R	0	3
2	T	0	1
All	All	0	57

The worst 5 of 160 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	107	LYS	CE-NZ	29.34	2.22	1.49
1	I	271	GLU	CD-OE2	22.21	1.50	1.25
1	I	271	GLU	CD-OE1	18.72	1.46	1.25
2	N	107	LYS	CD-CE	16.91	1.93	1.51
1	K	234	LYS	CD-CE	15.20	1.89	1.51

The worst 5 of 294 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	M	29	GLU	CG-CD-OE1	22.50	163.29	118.30
1	M	29	GLU	CG-CD-OE2	-20.57	77.16	118.30
1	M	29	GLU	OE1-CD-OE2	-19.80	99.54	123.30
1	U	173	LEU	CB-CG-CD2	18.87	143.08	111.00
1	I	173	LEU	CB-CG-CD1	18.14	141.83	111.00

There are no chirality outliers.

5 of 57 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	131	ARG	Sidechain
1	A	159	SER	Mainchain
1	A	219	SER	Mainchain
1	A	274	ARG	Sidechain
1	A	33	GLY	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2360	2341	2342	146	4
1	C	2360	2336	2342	48	0
1	E	2369	2348	2347	50	1
1	G	2360	2340	2341	64	0
1	I	2360	2342	2342	52	1
1	K	2360	2341	2342	68	0
1	M	2360	2341	2342	44	0
1	O	2360	2340	2342	81	3
1	Q	2360	2341	2342	20	0
1	U	2360	2341	2342	40	2
1	W	2360	2340	2342	46	1
1	Y	2360	2340	2342	36	1
2	B	945	891	893	19	0
2	D	939	887	888	22	2
2	F	949	896	896	10	0
2	H	949	896	895	10	0
2	J	949	896	896	11	3
2	L	949	898	896	19	1
2	N	949	896	896	34	2
2	P	949	897	896	9	0
2	R	949	896	896	13	0
2	T	949	896	896	10	3
2	V	949	896	896	10	0
2	X	949	896	896	4	0
3	A	20	17	17	0	0
3	C	20	17	17	2	0
3	E	20	17	17	0	0
3	G	20	17	17	0	0
3	I	20	17	17	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	K	20	17	17	0	0
3	M	20	17	17	0	0
3	O	20	17	17	0	0
3	Q	20	17	17	1	0
3	U	20	17	17	0	0
3	W	20	17	17	1	0
3	Y	20	17	17	0	0
4	E	6	8	8	0	0
4	G	6	8	8	0	0
4	Q	6	8	7	1	0
4	U	18	24	24	0	0
All	All	39979	39084	39099	845	12

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 11.

The worst 5 of 845 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:I:131:ARG:CG	1:I:131:ARG:CB	1.74	1.64
1:G:191:LYS:CE	1:G:191:LYS:CD	1.75	1.63
2:L:43:LYS:CB	2:L:43:LYS:CG	1.75	1.62
1:K:234:LYS:CD	1:K:234:LYS:CG	1.76	1.61
1:O:263:ILE:CD1	1:O:263:ILE:CG1	1.75	1.59

The worst 5 of 12 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:282:LYS:NZ	2:J:44:GLU:OE1[4_446]	1.62	0.58
1:O:275:GLU:HG3	2:T:43:LYS:HZ1[4_456]	1.18	0.42
1:O:274:ARG:HH11	2:T:42:GLY:O[4_456]	1.22	0.38
1:A:17:TYR:OH	1:I:290:GLN:OE1[3_445]	1.90	0.30
1:O:274:ARG:NH1	2:T:42:GLY:O[4_456]	1.90	0.30

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	297/303 (98%)	286 (96%)	9 (3%)	2 (1%)	19	46
1	C	297/303 (98%)	290 (98%)	6 (2%)	1 (0%)	37	65
1	E	299/303 (99%)	289 (97%)	9 (3%)	1 (0%)	37	65
1	G	297/303 (98%)	286 (96%)	9 (3%)	2 (1%)	19	46
1	I	297/303 (98%)	290 (98%)	5 (2%)	2 (1%)	19	46
1	K	297/303 (98%)	286 (96%)	8 (3%)	3 (1%)	13	37
1	M	297/303 (98%)	288 (97%)	4 (1%)	5 (2%)	7	24
1	O	297/303 (98%)	285 (96%)	8 (3%)	4 (1%)	10	30
1	Q	297/303 (98%)	292 (98%)	4 (1%)	1 (0%)	37	65
1	U	297/303 (98%)	289 (97%)	5 (2%)	3 (1%)	13	37
1	W	297/303 (98%)	290 (98%)	5 (2%)	2 (1%)	19	46
1	Y	297/303 (98%)	292 (98%)	3 (1%)	2 (1%)	19	46
2	B	120/123 (98%)	120 (100%)	0	0	100	100
2	D	119/123 (97%)	117 (98%)	2 (2%)	0	100	100
2	F	121/123 (98%)	121 (100%)	0	0	100	100
2	H	121/123 (98%)	119 (98%)	2 (2%)	0	100	100
2	J	121/123 (98%)	119 (98%)	1 (1%)	1 (1%)	16	42
2	L	121/123 (98%)	117 (97%)	3 (2%)	1 (1%)	16	42
2	N	121/123 (98%)	121 (100%)	0	0	100	100
2	P	121/123 (98%)	115 (95%)	5 (4%)	1 (1%)	16	42
2	R	121/123 (98%)	119 (98%)	1 (1%)	1 (1%)	16	42
2	T	121/123 (98%)	120 (99%)	1 (1%)	0	100	100
2	V	121/123 (98%)	119 (98%)	1 (1%)	1 (1%)	16	42
2	X	121/123 (98%)	118 (98%)	3 (2%)	0	100	100
All	All	5015/5112 (98%)	4888 (98%)	94 (2%)	33 (1%)	19	46

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	204	MET
1	G	267	TYR
1	I	204	MET
1	I	267	TYR
1	K	204	MET

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	258/259 (100%)	236 (92%)	22 (8%)	8 26
1	C	258/259 (100%)	249 (96%)	9 (4%)	31 63
1	E	258/259 (100%)	254 (98%)	4 (2%)	58 84
1	G	258/259 (100%)	251 (97%)	7 (3%)	40 72
1	I	258/259 (100%)	253 (98%)	5 (2%)	52 81
1	K	258/259 (100%)	250 (97%)	8 (3%)	35 68
1	M	258/259 (100%)	253 (98%)	5 (2%)	52 81
1	O	258/259 (100%)	250 (97%)	8 (3%)	35 68
1	Q	258/259 (100%)	249 (96%)	9 (4%)	31 63
1	U	258/259 (100%)	254 (98%)	4 (2%)	58 84
1	W	258/259 (100%)	251 (97%)	7 (3%)	40 72
1	Y	258/259 (100%)	252 (98%)	6 (2%)	45 77
2	B	98/98 (100%)	93 (95%)	5 (5%)	20 49
2	D	97/98 (99%)	96 (99%)	1 (1%)	73 91
2	F	98/98 (100%)	96 (98%)	2 (2%)	50 80
2	H	98/98 (100%)	93 (95%)	5 (5%)	20 49
2	J	98/98 (100%)	94 (96%)	4 (4%)	26 57
2	L	98/98 (100%)	94 (96%)	4 (4%)	26 57
2	N	98/98 (100%)	93 (95%)	5 (5%)	20 49

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	P	98/98 (100%)	95 (97%)	3 (3%)	35 68
2	R	98/98 (100%)	96 (98%)	2 (2%)	50 80
2	T	98/98 (100%)	95 (97%)	3 (3%)	35 68
2	V	98/98 (100%)	96 (98%)	2 (2%)	50 80
2	X	98/98 (100%)	96 (98%)	2 (2%)	50 80
All	All	4271/4284 (100%)	4139 (97%)	132 (3%)	35 68

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	V	63	SER
1	W	162	SER
1	Y	256	SER
1	G	275	GLU
1	G	234	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	W	102	GLN
1	W	233	GLN
1	Y	201	ASN
1	W	237	GLN
1	W	212	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

18 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	SLB	G	301	-	20,20,21	1.36	3 (15%)	24,28,31	1.40	3 (12%)
3	SLB	U	301	-	20,20,21	1.43	3 (15%)	24,28,31	1.22	3 (12%)
3	SLB	E	301	-	20,20,21	1.41	3 (15%)	24,28,31	1.25	2 (8%)
3	SLB	Y	301	-	20,20,21	1.35	3 (15%)	24,28,31	1.25	3 (12%)
3	SLB	Q	301	-	20,20,21	1.36	3 (15%)	24,28,31	1.17	3 (12%)
3	SLB	W	301	-	20,20,21	1.43	3 (15%)	24,28,31	1.19	4 (16%)
4	GOL	E	302	-	5,5,5	0.92	0	5,5,5	0.99	0
3	SLB	O	301	-	20,20,21	1.37	3 (15%)	24,28,31	1.31	5 (20%)
3	SLB	I	301	-	20,20,21	1.33	3 (15%)	24,28,31	1.23	5 (20%)
4	GOL	U	302	-	5,5,5	0.86	0	5,5,5	0.98	0
3	SLB	A	301	-	20,20,21	1.45	3 (15%)	24,28,31	1.43	4 (16%)
3	SLB	M	301	-	20,20,21	1.37	3 (15%)	24,28,31	1.36	3 (12%)
4	GOL	G	302	-	5,5,5	1.19	1 (20%)	5,5,5	1.42	1 (20%)
4	GOL	U	303	-	5,5,5	0.88	0	5,5,5	1.08	0
4	GOL	U	304	-	5,5,5	0.83	0	5,5,5	1.06	0
4	GOL	Q	302	-	5,5,5	1.30	1 (20%)	5,5,5	1.35	1 (20%)
3	SLB	K	301	-	20,20,21	1.37	2 (10%)	24,28,31	1.31	4 (16%)
3	SLB	C	301	-	20,20,21	1.37	3 (15%)	24,28,31	1.37	4 (16%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SLB	G	301	-	-	3/18/34/38	0/1/1/1
3	SLB	U	301	-	-	7/18/34/38	0/1/1/1
3	SLB	E	301	-	-	2/18/34/38	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SLB	Y	301	-	-	1/18/34/38	0/1/1/1
3	SLB	Q	301	-	-	1/18/34/38	0/1/1/1
3	SLB	W	301	-	-	2/18/34/38	0/1/1/1
4	GOL	E	302	-	-	2/4/4/4	-
3	SLB	O	301	-	-	1/18/34/38	0/1/1/1
3	SLB	I	301	-	-	3/18/34/38	0/1/1/1
4	GOL	U	302	-	-	0/4/4/4	-
3	SLB	A	301	-	-	6/18/34/38	0/1/1/1
3	SLB	M	301	-	-	1/18/34/38	0/1/1/1
4	GOL	G	302	-	-	0/4/4/4	-
4	GOL	U	303	-	-	0/4/4/4	-
4	GOL	U	304	-	-	0/4/4/4	-
4	GOL	Q	302	-	-	2/4/4/4	-
3	SLB	K	301	-	-	1/18/34/38	0/1/1/1
3	SLB	C	301	-	-	3/18/34/38	0/1/1/1

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	U	301	SLB	C10-N5	3.47	1.46	1.34
3	A	301	SLB	C10-N5	3.41	1.46	1.34
3	E	301	SLB	C10-N5	3.31	1.45	1.34
3	W	301	SLB	C10-N5	3.31	1.45	1.34
3	M	301	SLB	C10-N5	3.30	1.45	1.34

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	M	301	SLB	C6-O6-C2	3.74	119.33	111.34
3	G	301	SLB	C6-O6-C2	3.72	119.29	111.34
3	A	301	SLB	C9-C8-C7	-3.56	104.69	112.41
3	K	301	SLB	C6-O6-C2	3.23	118.26	111.34
3	C	301	SLB	C6-O6-C2	3.21	118.21	111.34

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	301	SLB	O1A-C1-C2-O6
3	A	301	SLB	C6-C7-C8-C9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	301	SLB	O7-C7-C8-C9
3	E	301	SLB	O1A-C1-C2-O6
3	K	301	SLB	O1A-C1-C2-O6

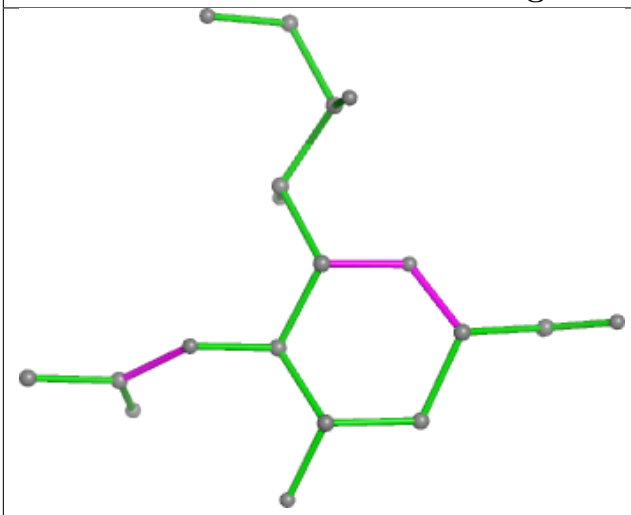
There are no ring outliers.

4 monomers are involved in 5 short contacts:

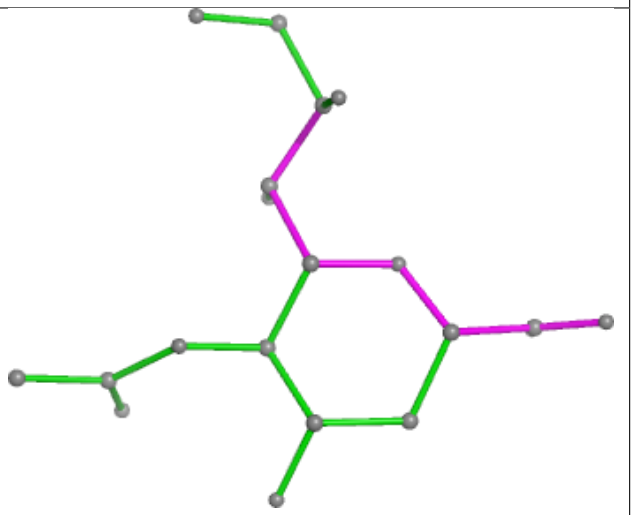
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	Q	301	SLB	1	0
3	W	301	SLB	1	0
4	Q	302	GOL	1	0
3	C	301	SLB	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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.

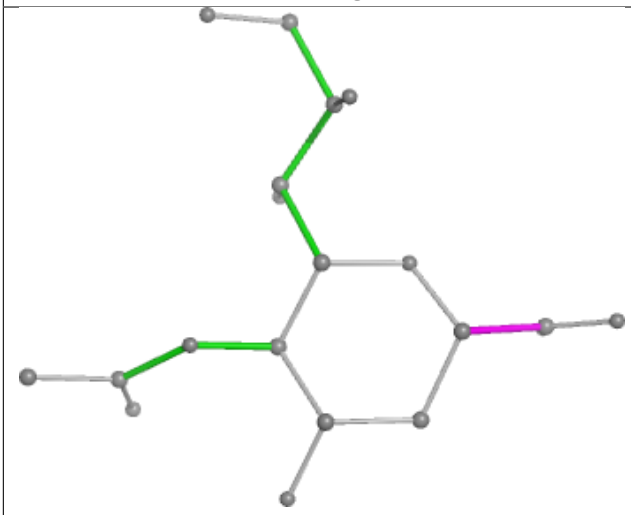
Ligand SLB G 301



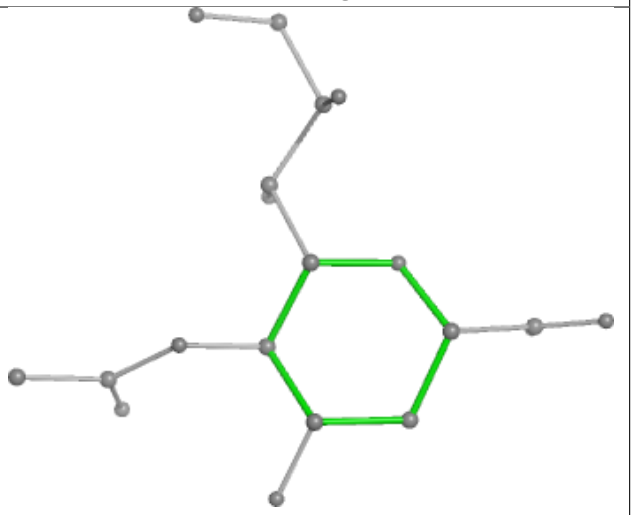
Bond lengths



Bond angles

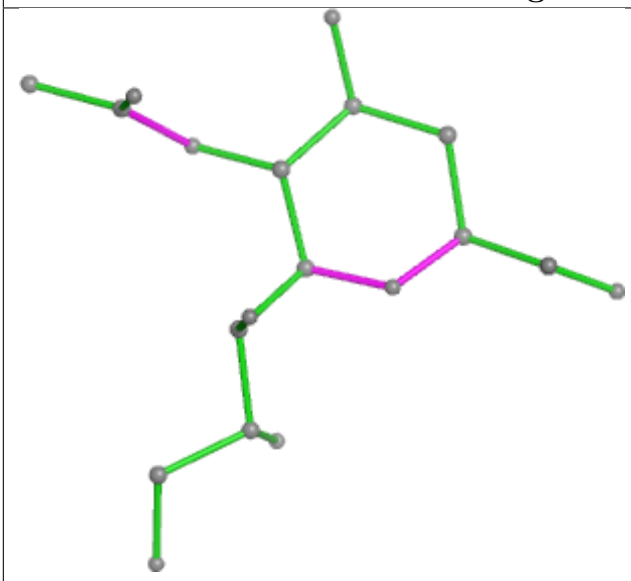


Torsions

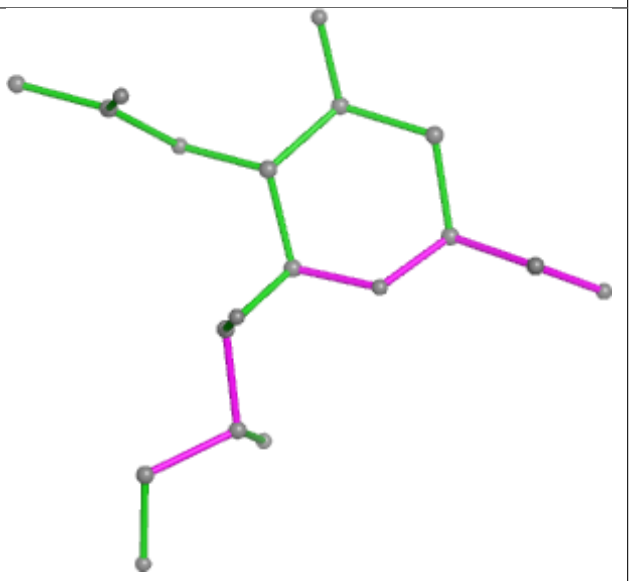


Rings

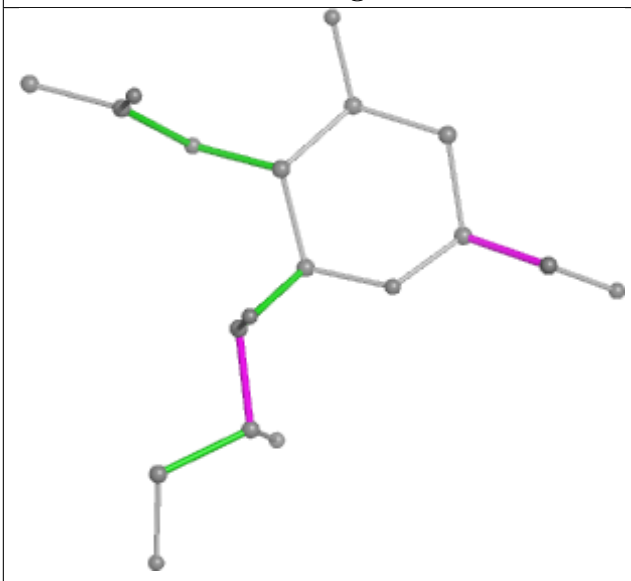
Ligand SLB U 301



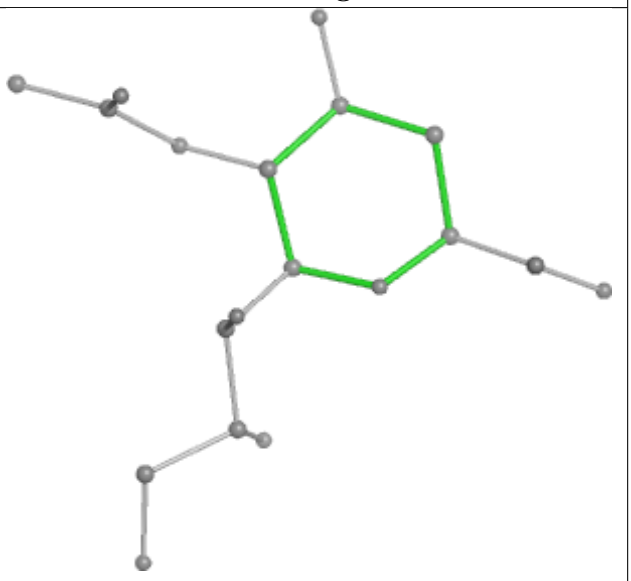
Bond lengths



Bond angles

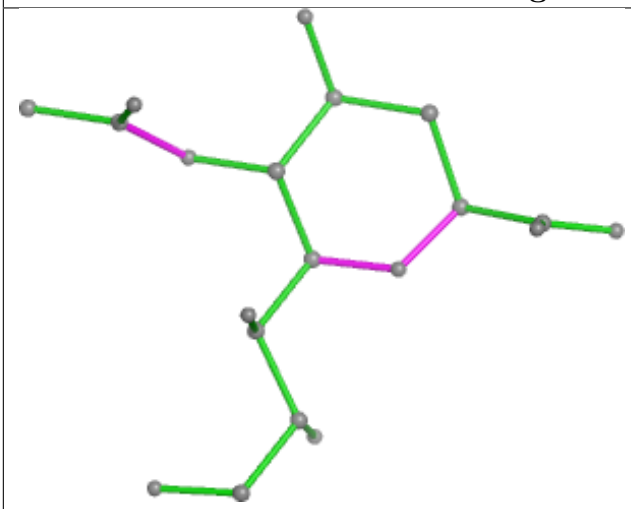


Torsions

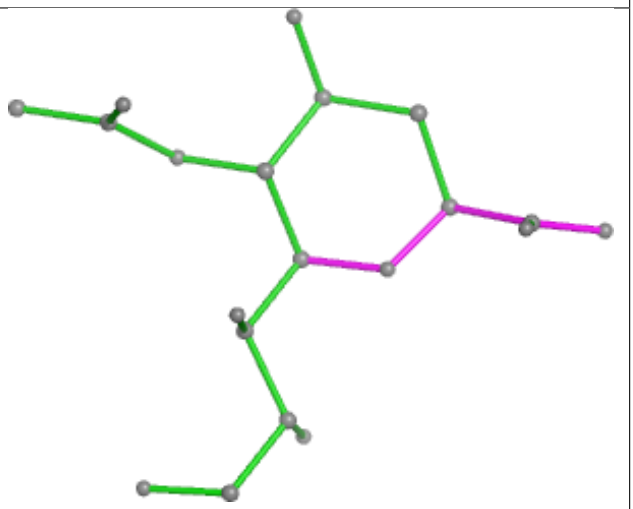


Rings

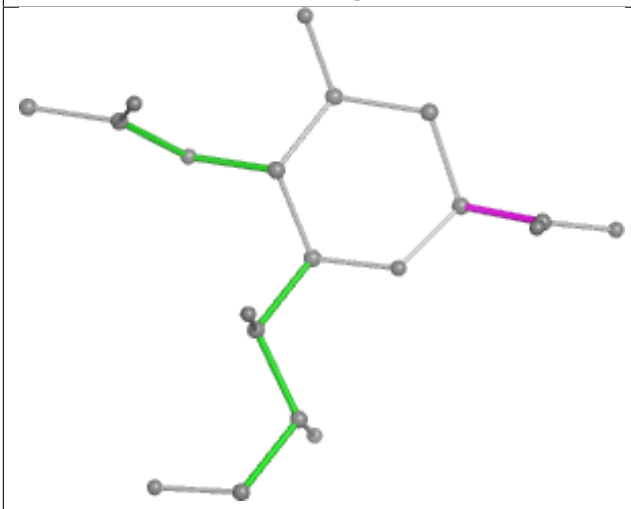
Ligand SLB E 301



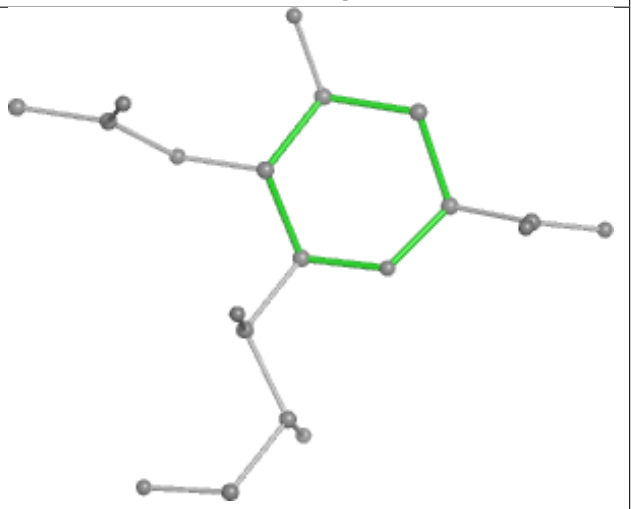
Bond lengths



Bond angles

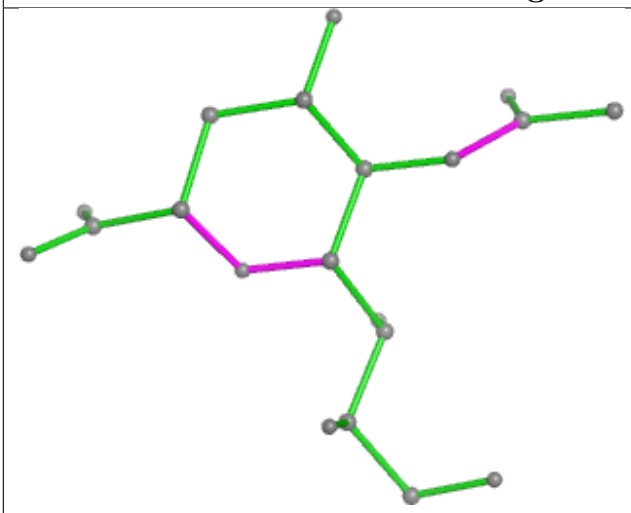


Torsions

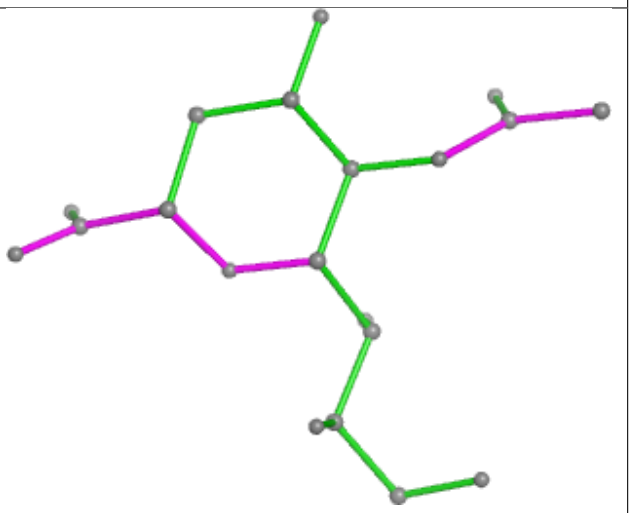


Rings

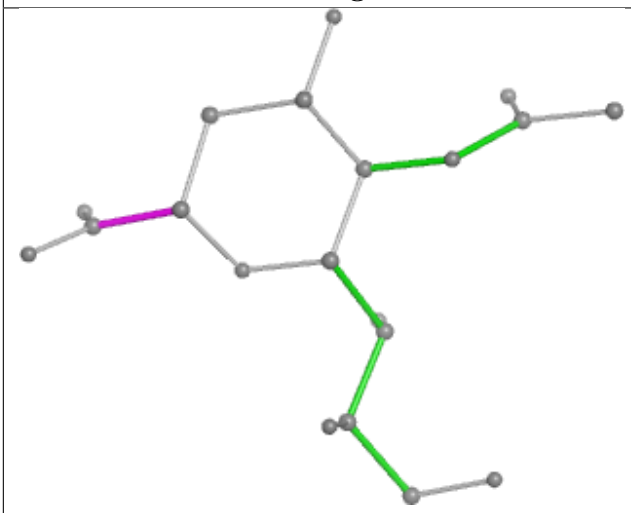
Ligand SLB Y 301



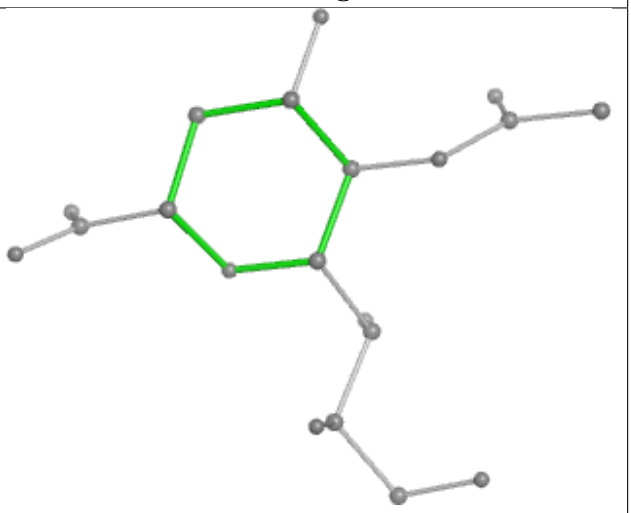
Bond lengths



Bond angles

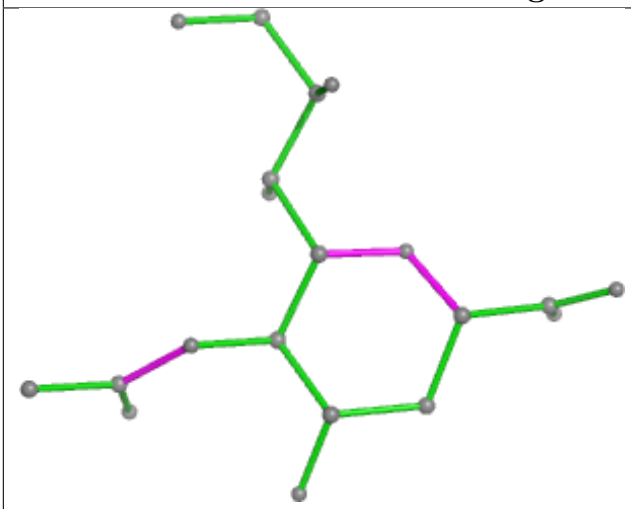


Torsions

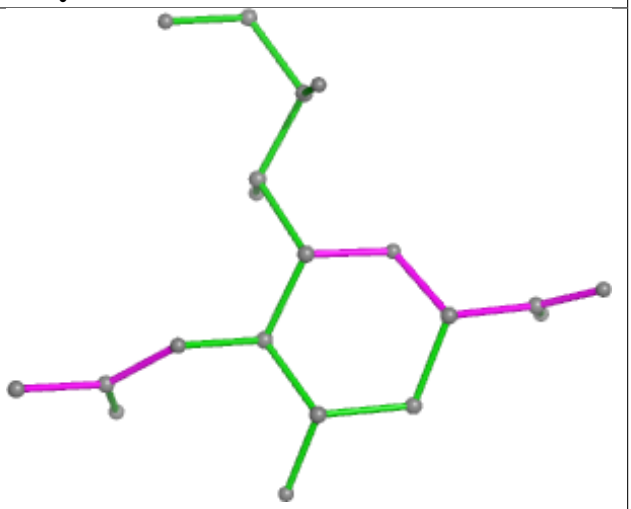


Rings

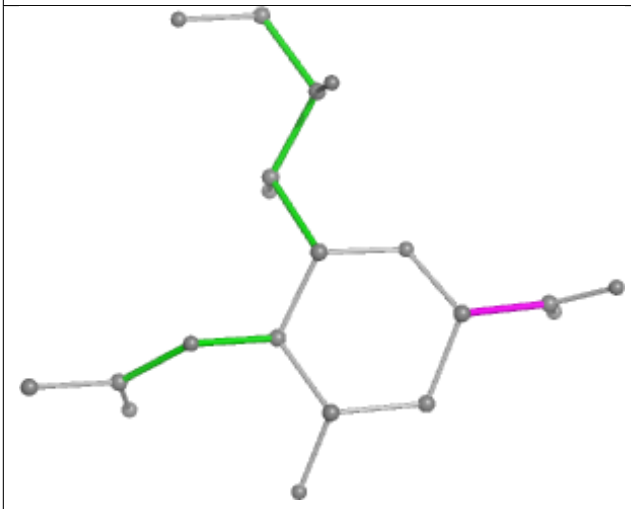
Ligand SLB Q 301



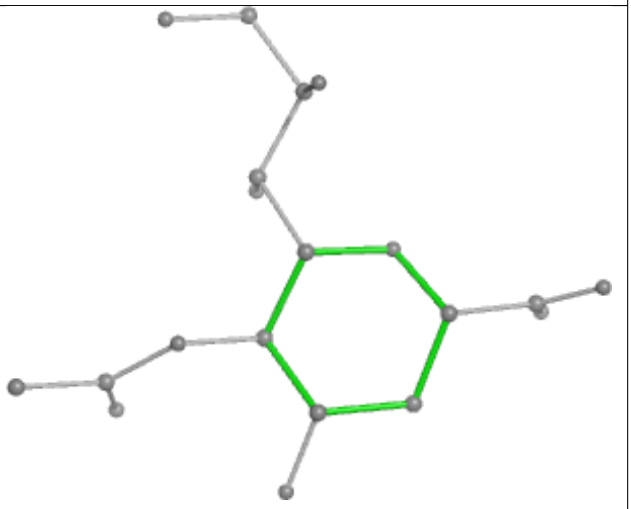
Bond lengths



Bond angles

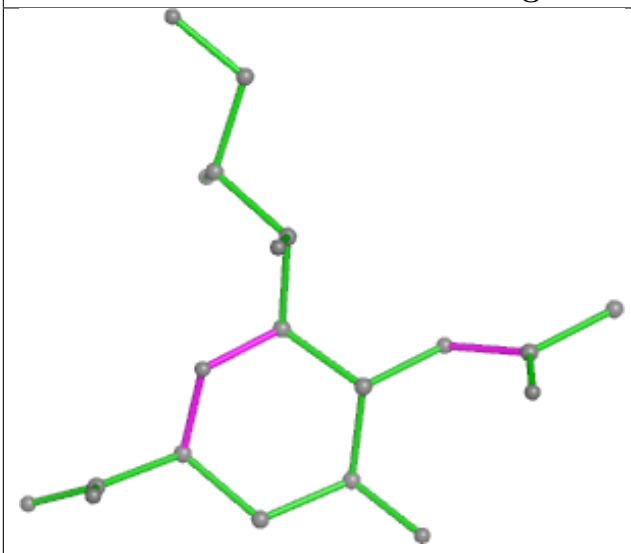


Torsions

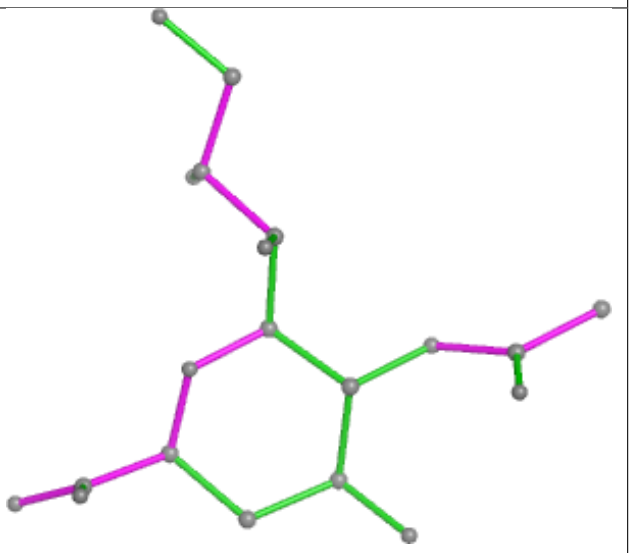


Rings

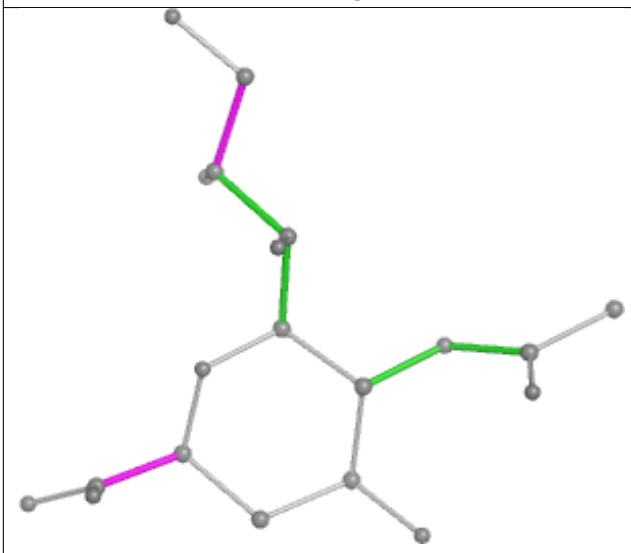
Ligand SLB W 301



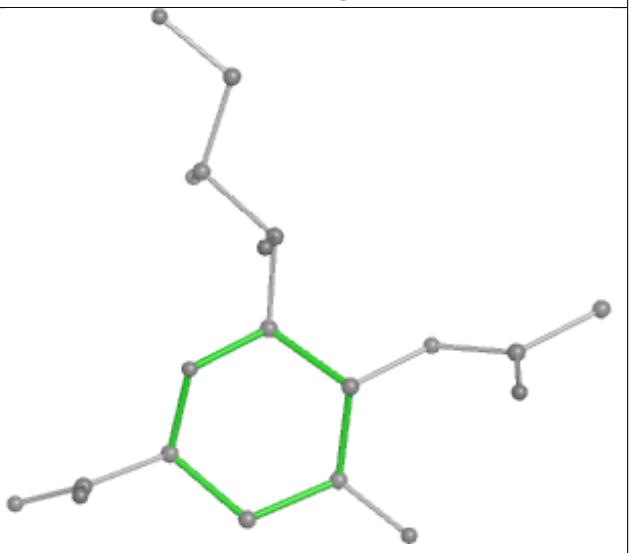
Bond lengths



Bond angles

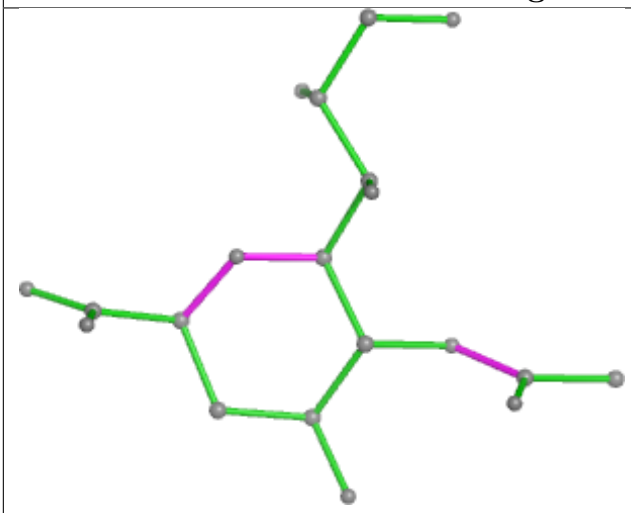


Torsions

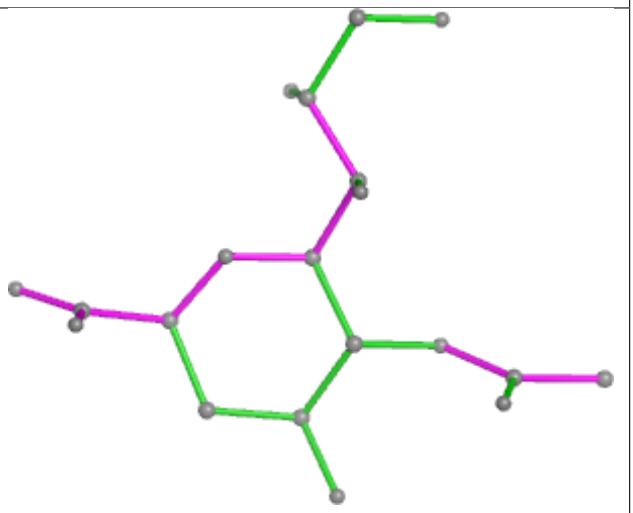


Rings

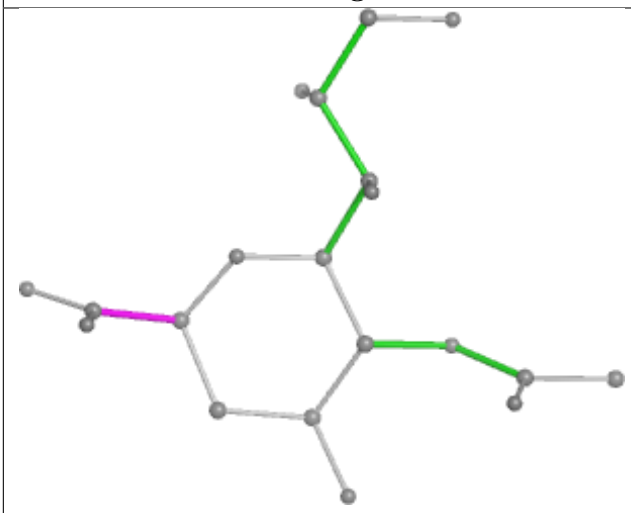
Ligand SLB O 301



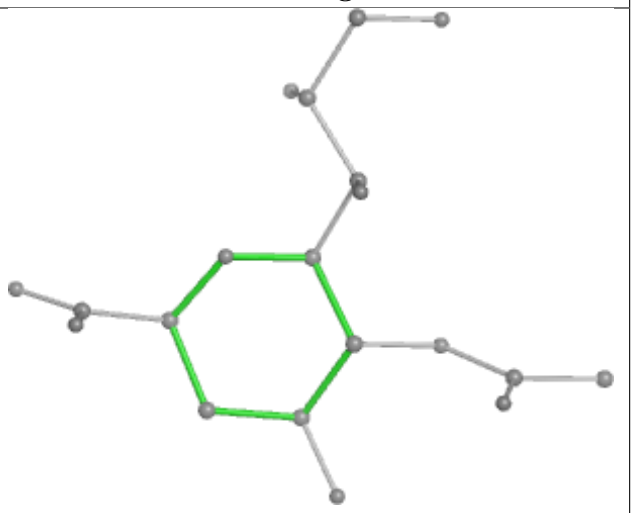
Bond lengths



Bond angles

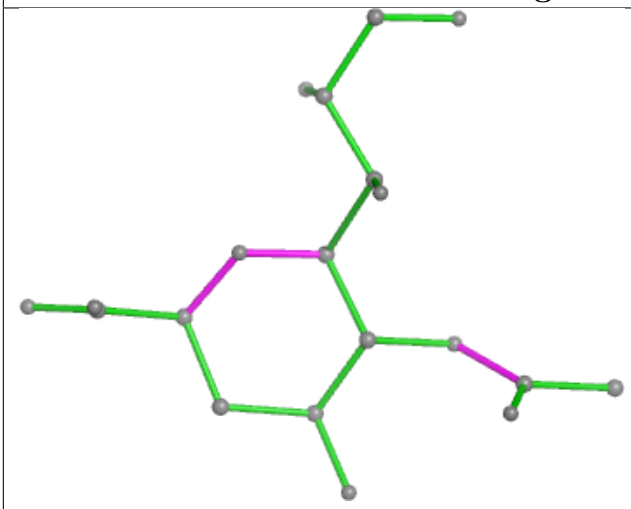


Torsions

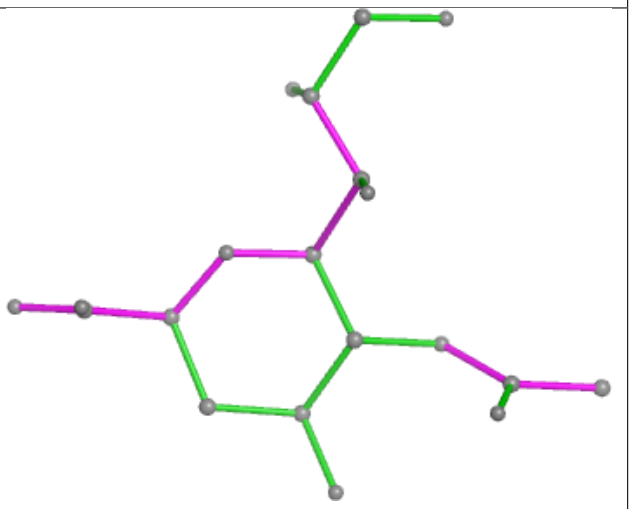


Rings

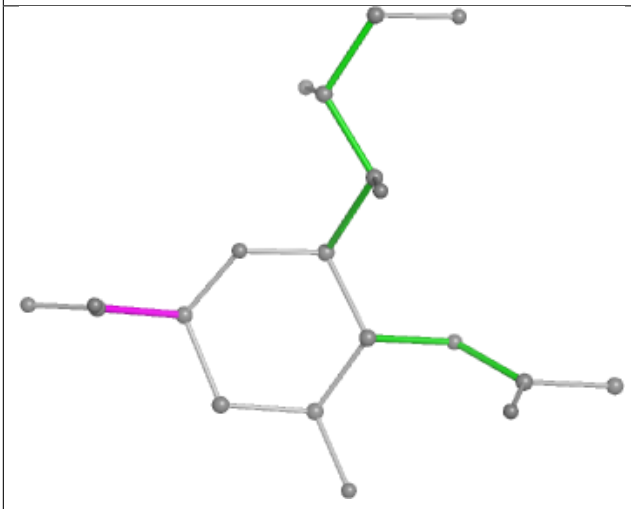
Ligand SLB I 301



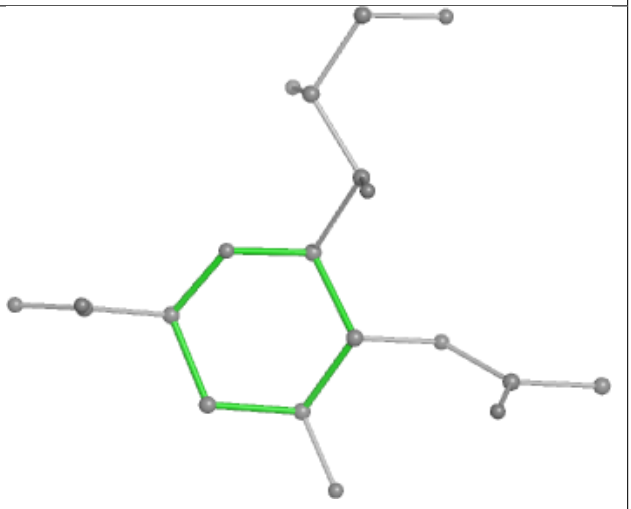
Bond lengths



Bond angles

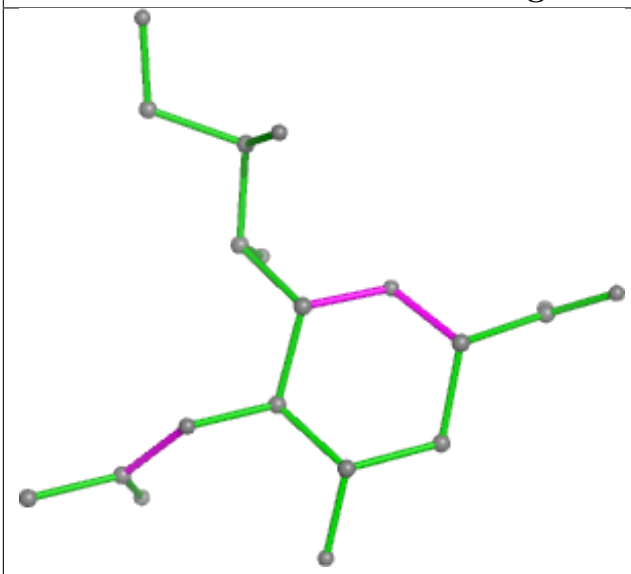


Torsions

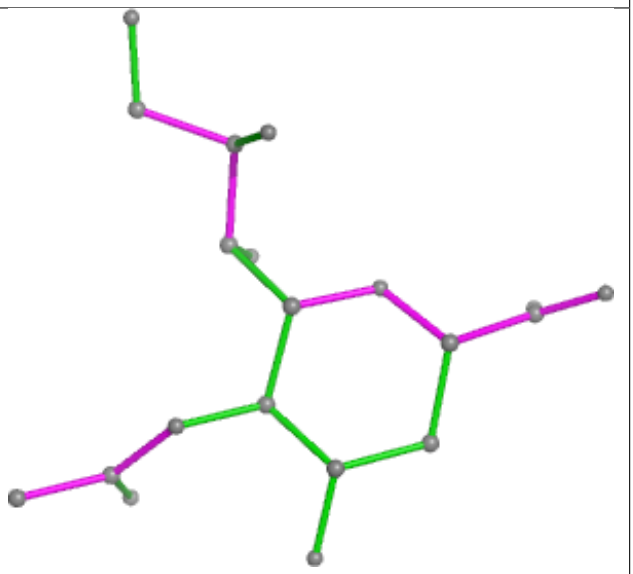


Rings

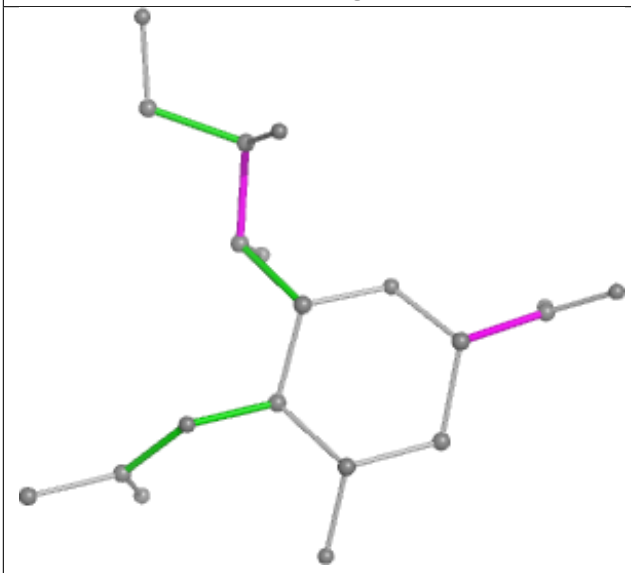
Ligand SLB A 301



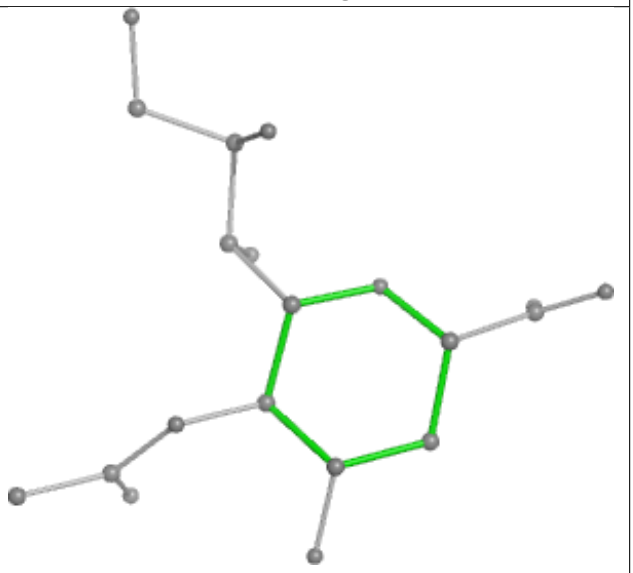
Bond lengths



Bond angles

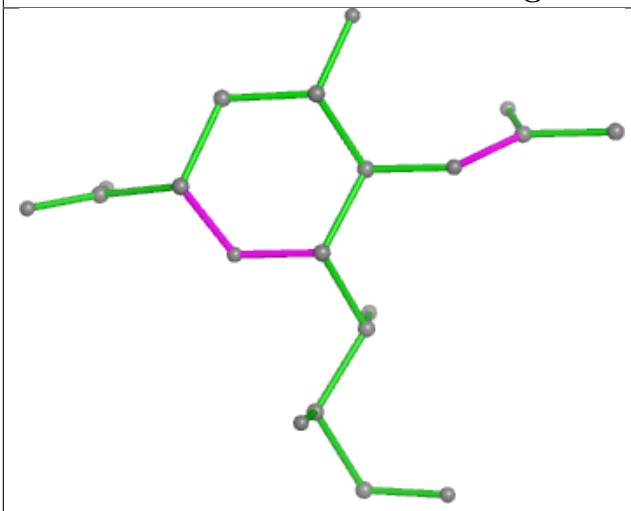


Torsions

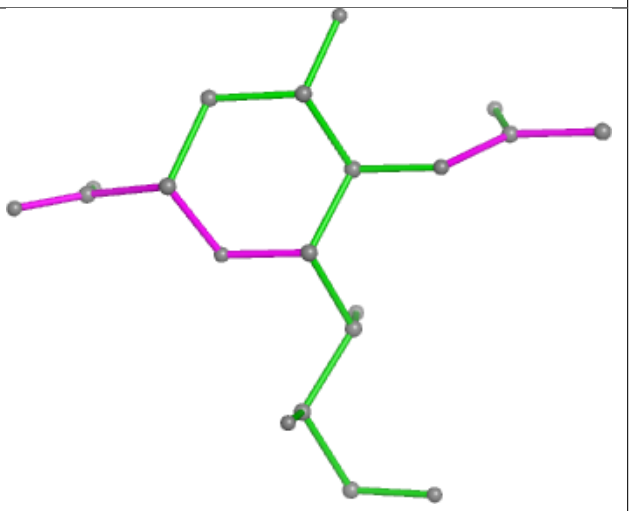


Rings

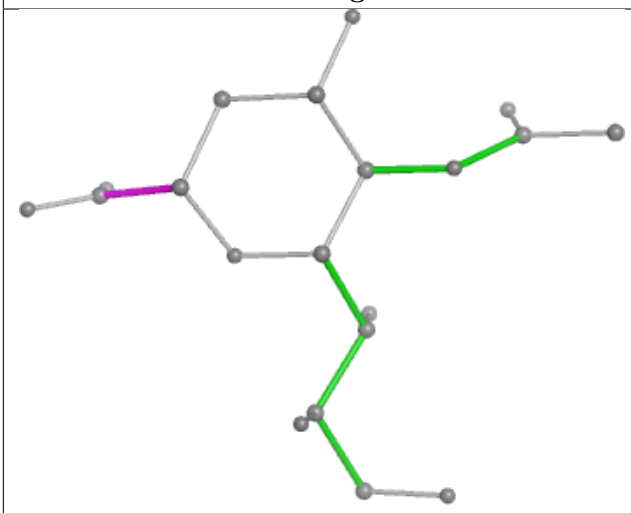
Ligand SLB M 301



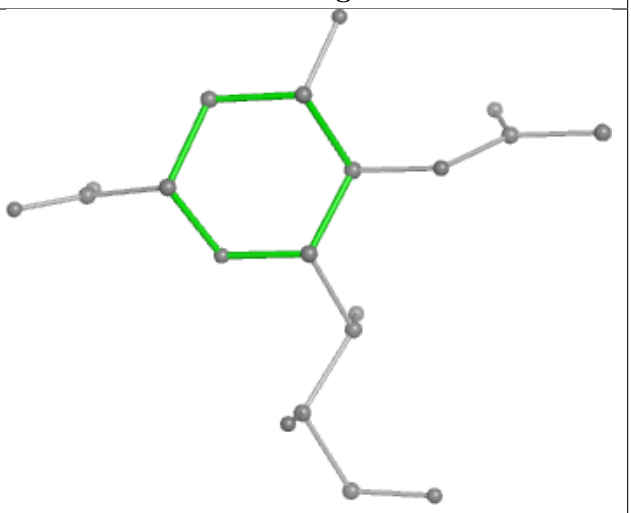
Bond lengths



Bond angles

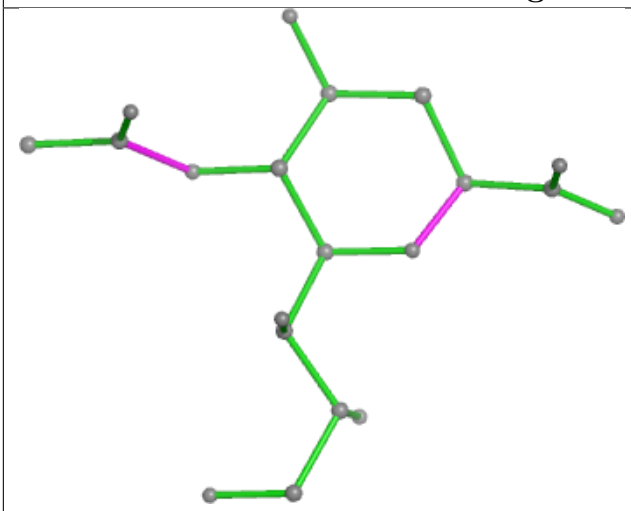


Torsions

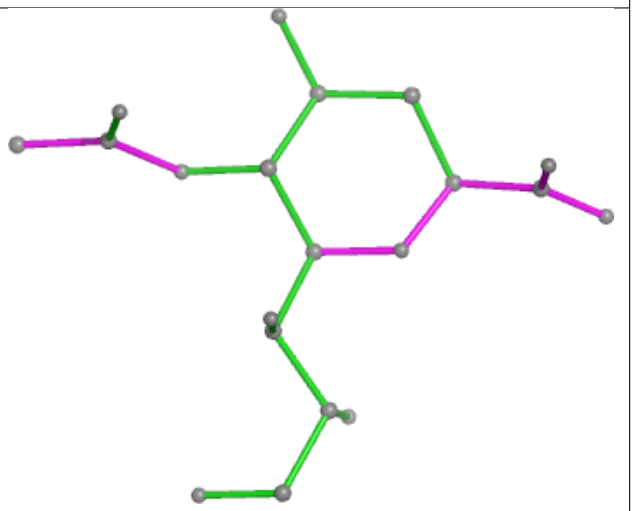


Rings

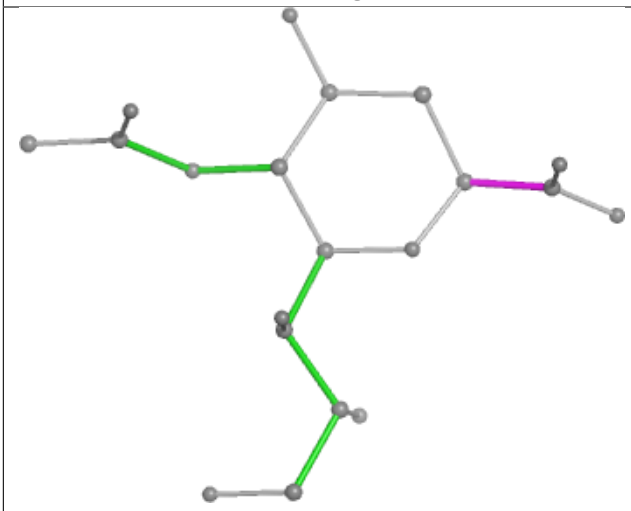
Ligand SLB K 301



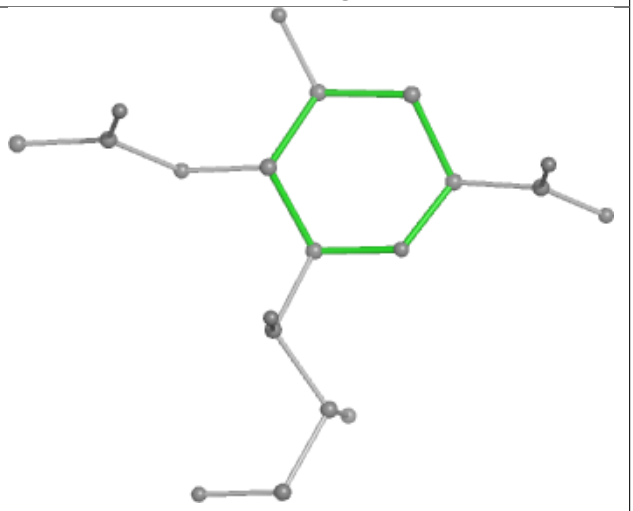
Bond lengths



Bond angles

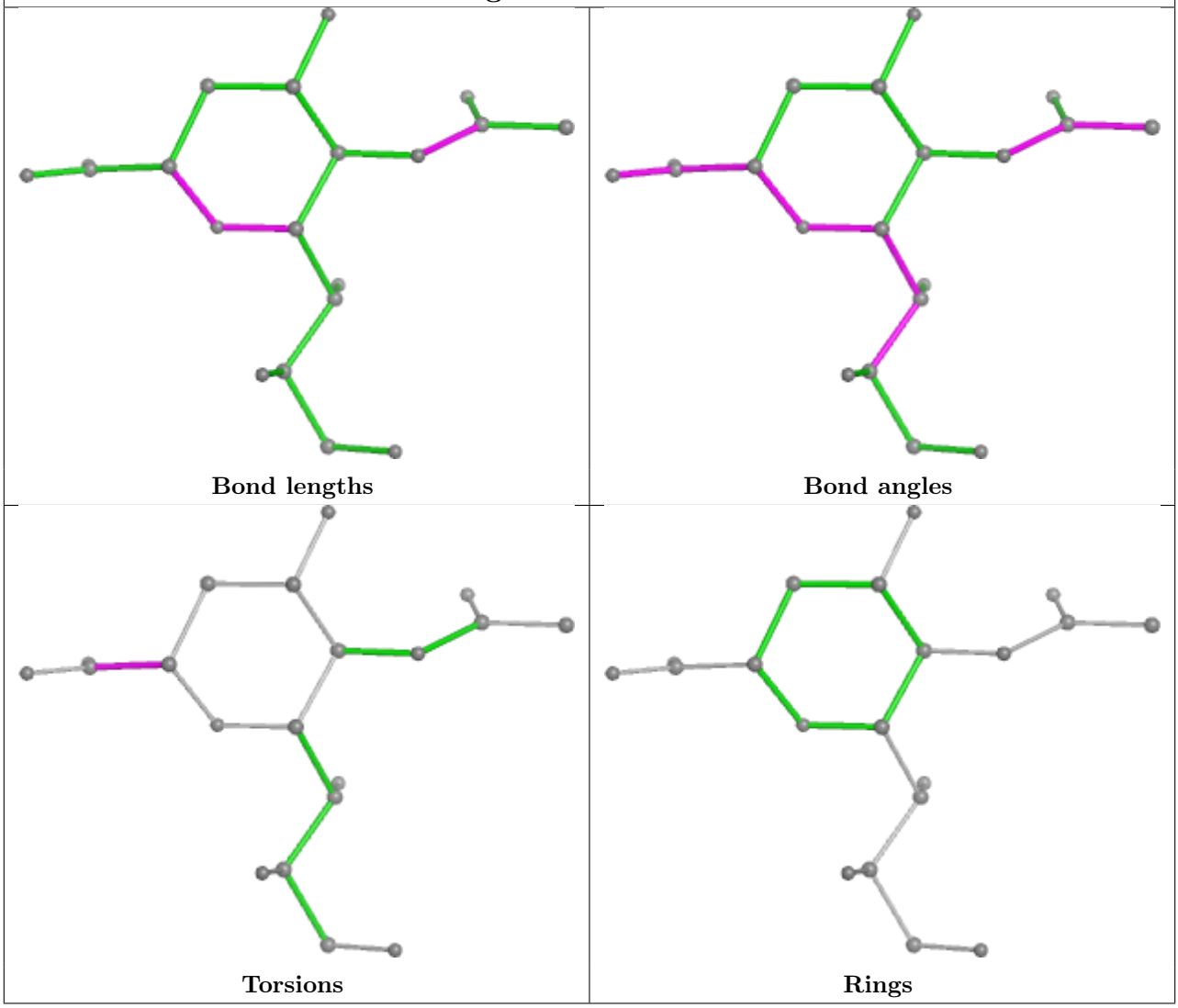


Torsions



Rings

Ligand SLB C 301



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	A	299/303 (98%)	1.85	132 (44%) 1 1	37, 82, 117, 133	0
1	C	299/303 (98%)	0.69	18 (6%) 29 22	33, 60, 81, 93	0
1	E	301/303 (99%)	0.04	7 (2%) 61 53	24, 42, 62, 98	0
1	G	299/303 (98%)	0.45	13 (4%) 40 33	27, 56, 81, 93	0
1	I	299/303 (98%)	0.49	12 (4%) 43 35	32, 55, 82, 100	0
1	K	299/303 (98%)	0.86	28 (9%) 15 12	30, 63, 85, 96	0
1	M	299/303 (98%)	-0.03	6 (2%) 64 57	26, 42, 58, 77	0
1	O	299/303 (98%)	0.43	15 (5%) 35 28	29, 54, 79, 92	0
1	Q	299/303 (98%)	-0.08	2 (0%) 84 79	26, 42, 63, 78	0
1	U	299/303 (98%)	0.11	6 (2%) 64 57	23, 44, 68, 81	0
1	W	299/303 (98%)	0.66	29 (9%) 15 11	32, 61, 88, 103	0
1	Y	299/303 (98%)	0.13	8 (2%) 56 48	23, 45, 70, 91	0
2	B	122/123 (99%)	0.41	7 (5%) 30 24	30, 45, 72, 94	0
2	D	121/123 (98%)	0.38	7 (5%) 30 23	34, 48, 76, 97	0
2	F	123/123 (100%)	0.13	7 (5%) 30 24	25, 40, 74, 92	0
2	H	123/123 (100%)	0.02	7 (5%) 30 24	25, 35, 70, 86	0
2	J	123/123 (100%)	0.34	9 (7%) 22 17	27, 42, 76, 93	0
2	L	123/123 (100%)	0.07	5 (4%) 42 34	28, 38, 67, 83	0
2	N	123/123 (100%)	0.29	12 (9%) 14 11	26, 39, 70, 92	0
2	P	123/123 (100%)	-0.02	2 (1%) 70 63	25, 34, 55, 78	0
2	R	123/123 (100%)	-0.19	2 (1%) 70 63	22, 32, 59, 90	0
2	T	123/123 (100%)	-0.11	6 (4%) 36 28	23, 34, 59, 92	0
2	V	123/123 (100%)	-0.09	6 (4%) 36 28	23, 32, 56, 70	0
2	X	123/123 (100%)	-0.09	3 (2%) 59 52	28, 38, 57, 72	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	5063/5112 (99%)	0.36	349 (6%)	24	18	22, 48, 84, 133	0

The worst 5 of 349 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	V	0	SER	8.2
2	L	43	LYS	5.9
2	H	42	GLY	5.8
1	E	187	LEU	5.7
2	B	0	SER	5.4

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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
4	GOL	U	303	6/6	0.14	0.24	100,121,134,136	0
4	GOL	U	302	6/6	0.39	0.27	81,102,117,123	0
4	GOL	Q	302	6/6	0.78	0.18	34,43,52,59	0
4	GOL	U	304	6/6	0.78	0.17	50,60,68,80	0
4	GOL	E	302	6/6	0.79	0.20	27,44,53,63	0
4	GOL	G	302	6/6	0.79	0.25	75,90,98,103	0
3	SLB	K	301	20/21	0.86	0.14	39,53,62,66	0
3	SLB	W	301	20/21	0.88	0.15	41,51,63,74	0
3	SLB	A	301	20/21	0.89	0.13	42,53,69,75	0
3	SLB	U	301	20/21	0.89	0.11	21,31,39,44	0
3	SLB	E	301	20/21	0.89	0.10	23,35,46,53	0
3	SLB	M	301	20/21	0.90	0.11	24,32,42,46	0
3	SLB	I	301	20/21	0.90	0.12	29,36,44,47	0

Continued on next page...

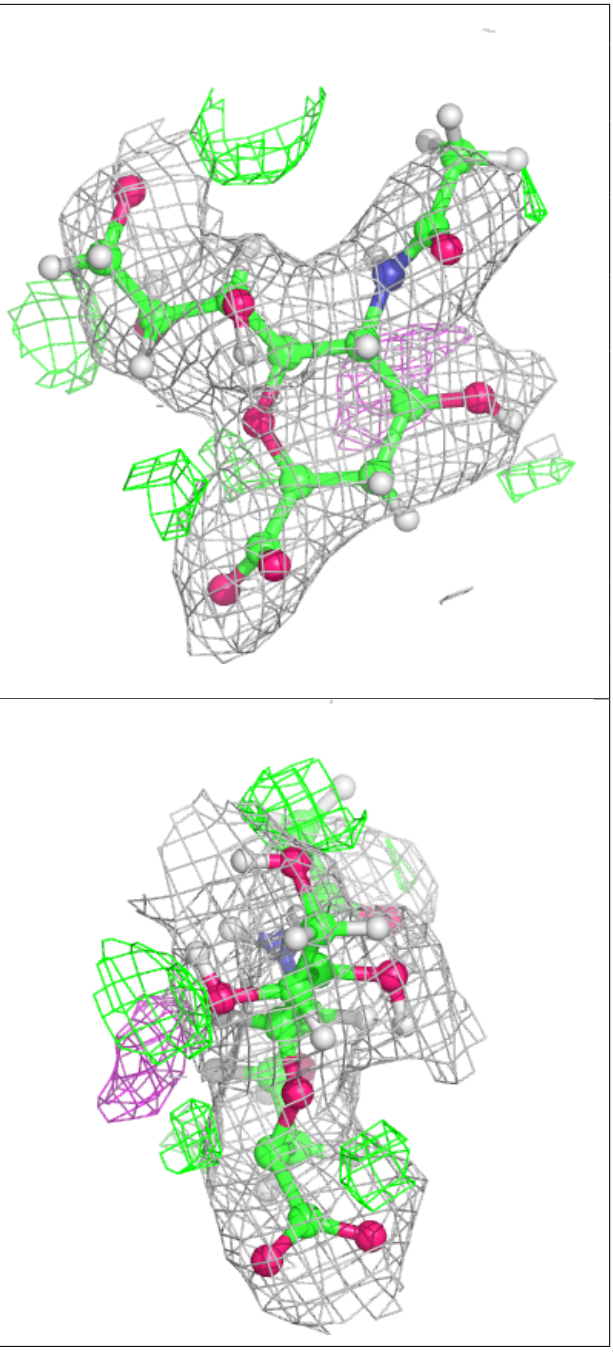
Continued from previous page...

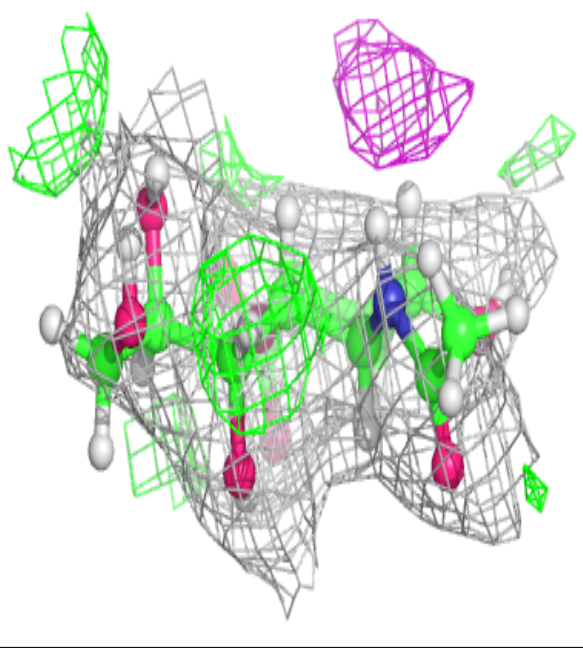
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	SLB	C	301	20/21	0.90	0.12	31,43,56,58	0
3	SLB	G	301	20/21	0.92	0.12	25,33,44,52	0
3	SLB	Q	301	20/21	0.92	0.11	28,35,42,46	0
3	SLB	O	301	20/21	0.93	0.12	30,40,48,53	0
3	SLB	Y	301	20/21	0.93	0.09	26,34,47,47	0

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 SLB K 301:

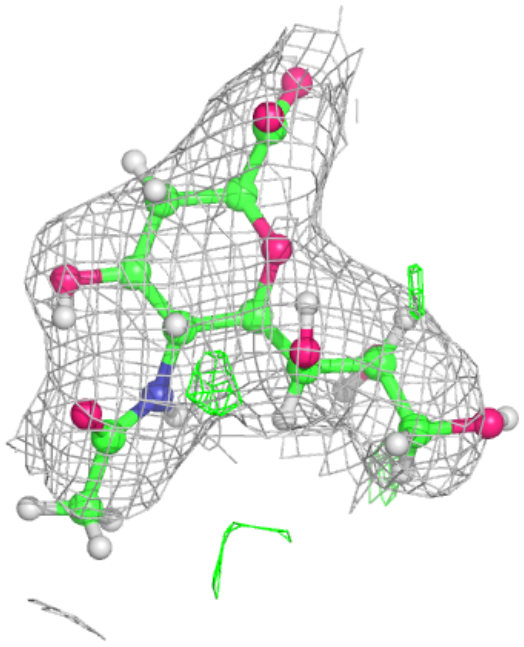
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

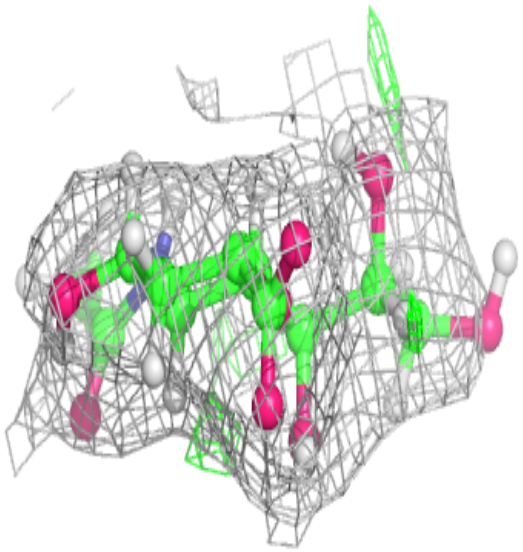




Electron density around SLB W 301:

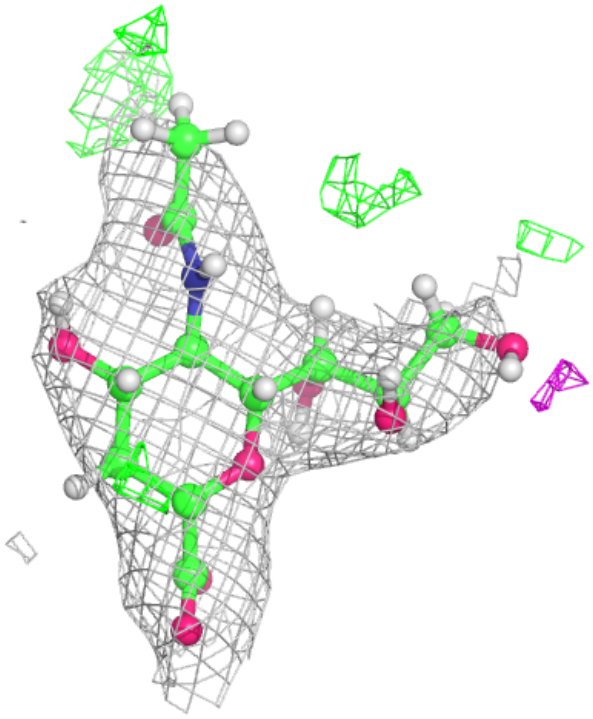
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

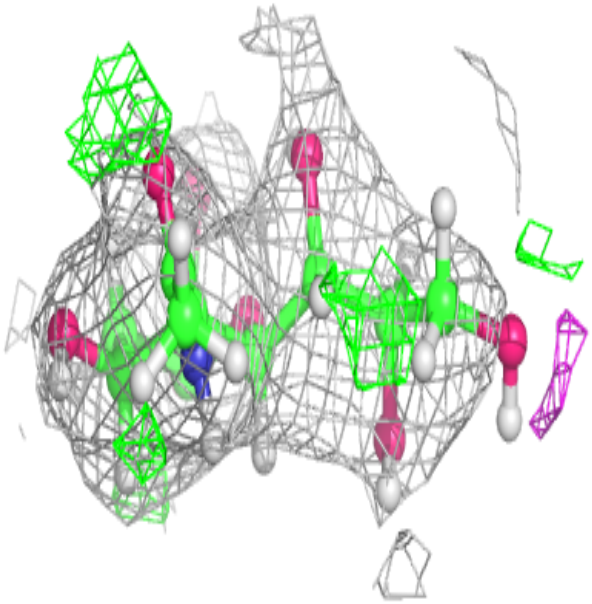


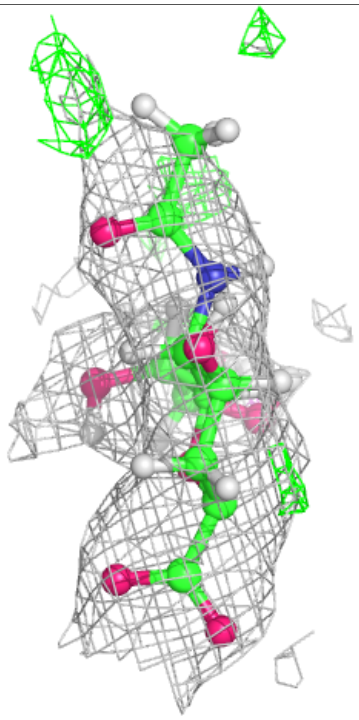


Electron density around SLB A 301:

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

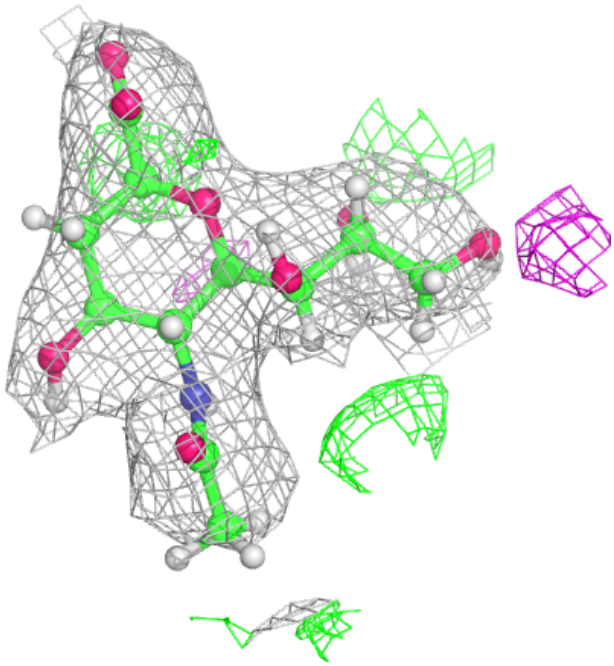


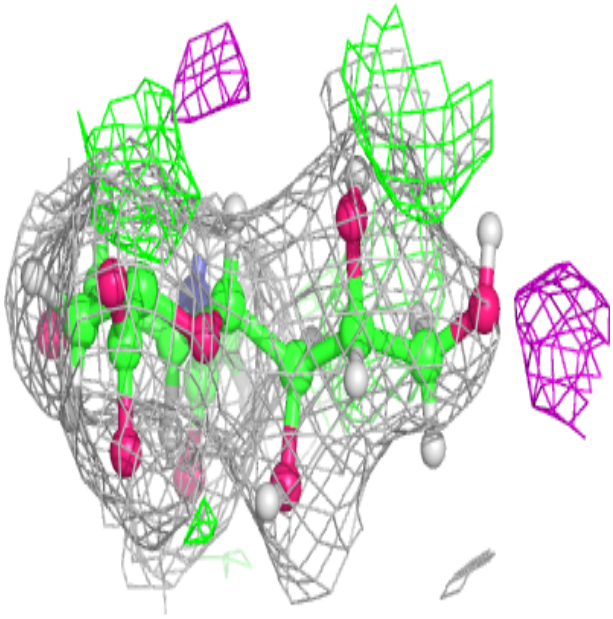




Electron density around SLB U 301:

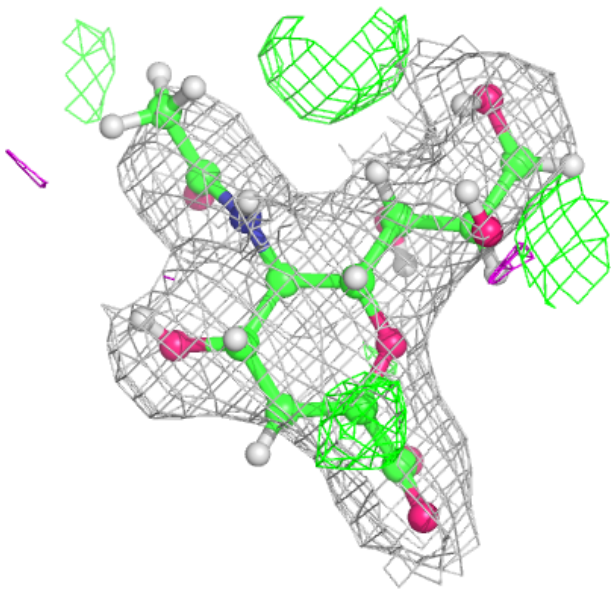
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

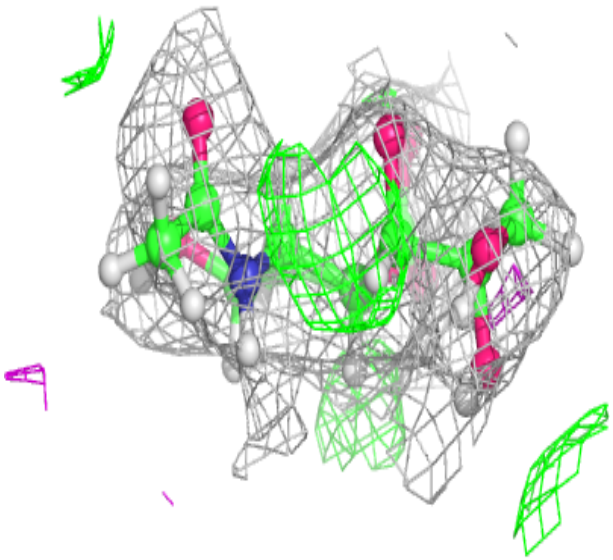


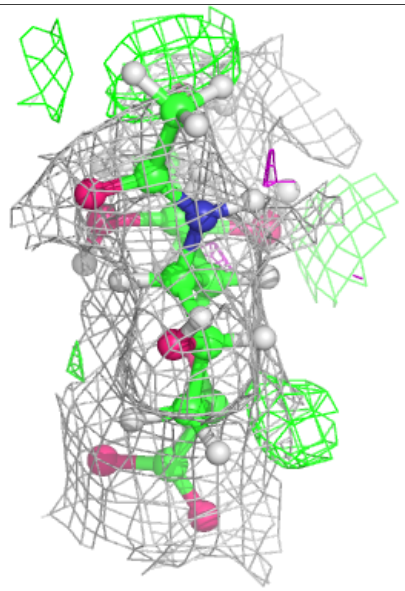


Electron density around SLB E 301:

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

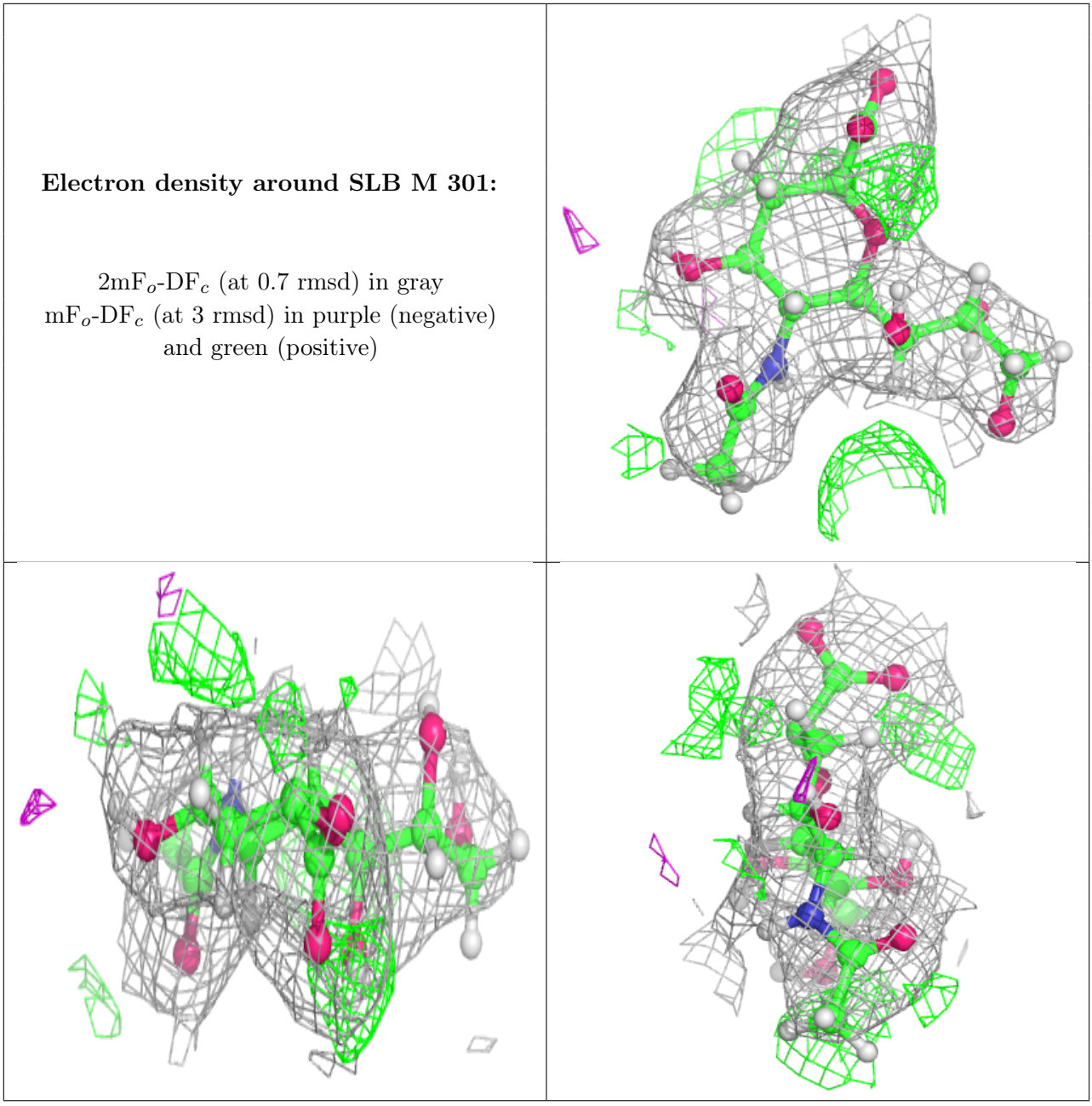






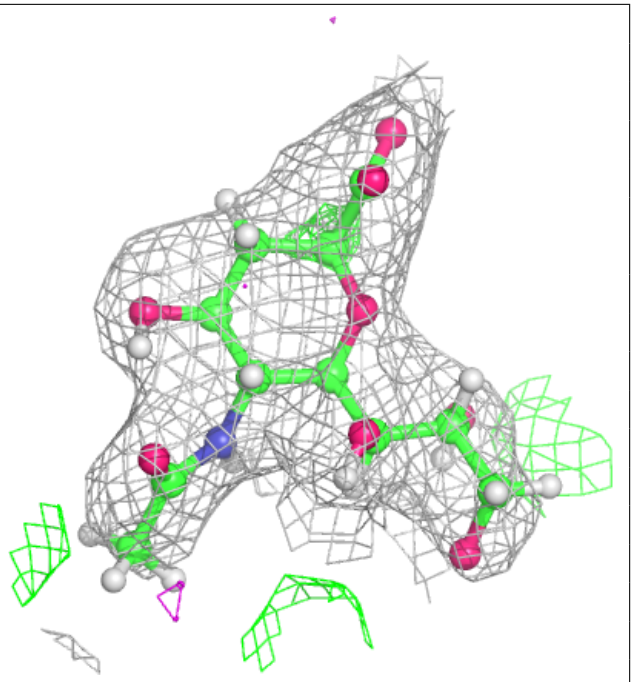
Electron density around SLB M 301:

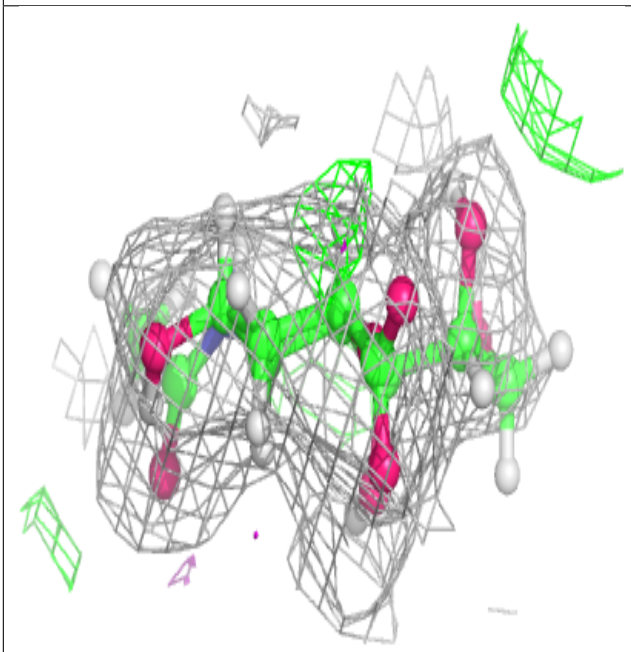
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

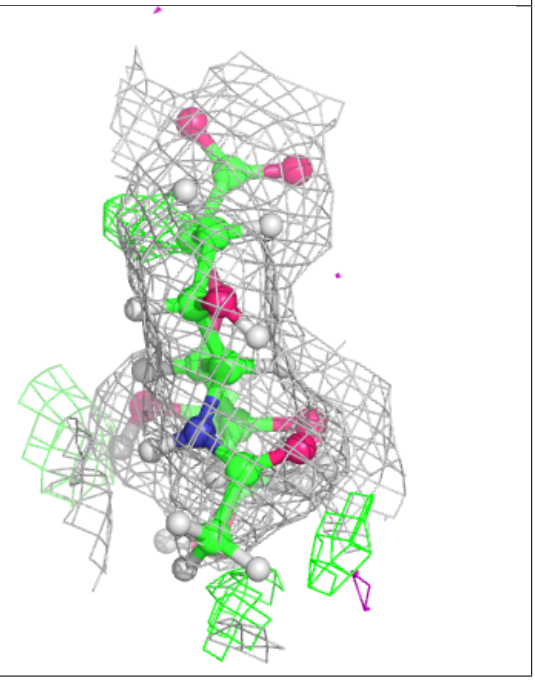


Electron density around SLB I 301:

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

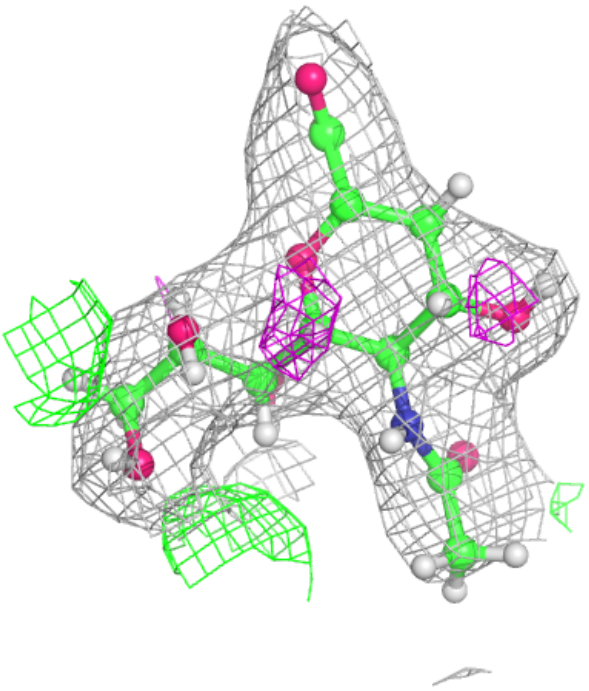


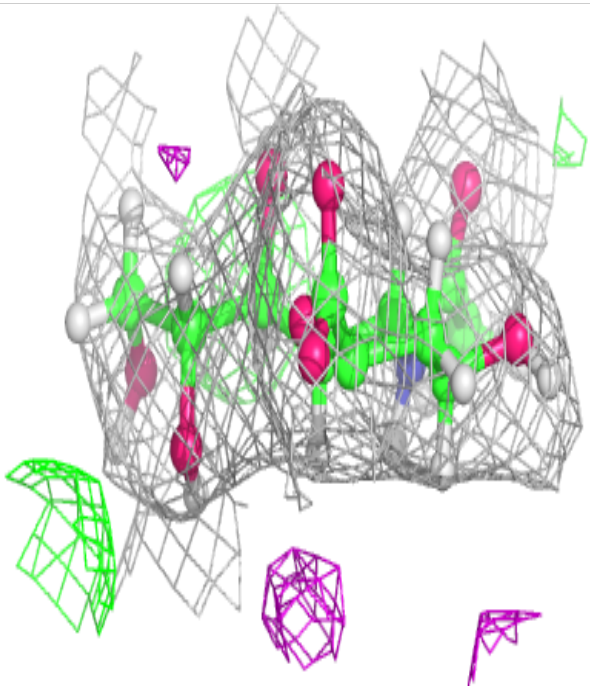


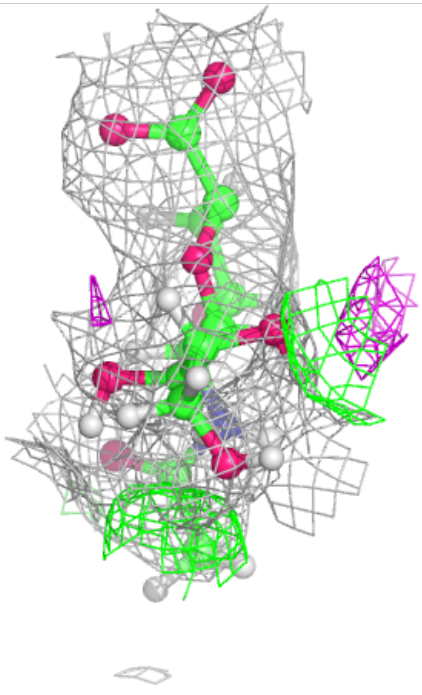


Electron density around SLB C 301:

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

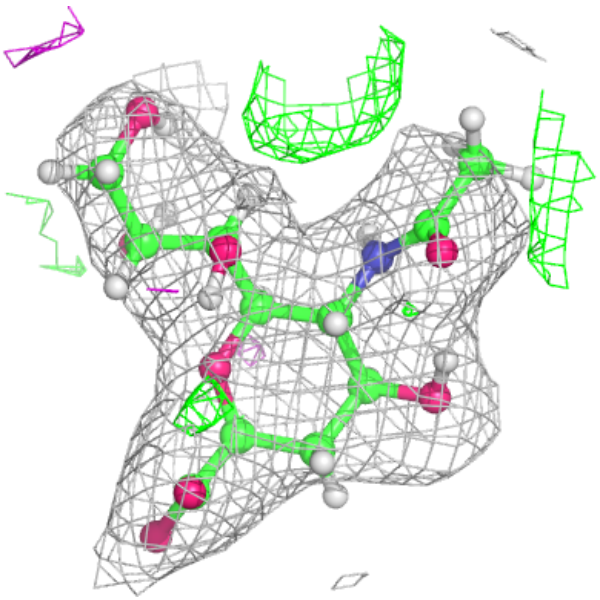


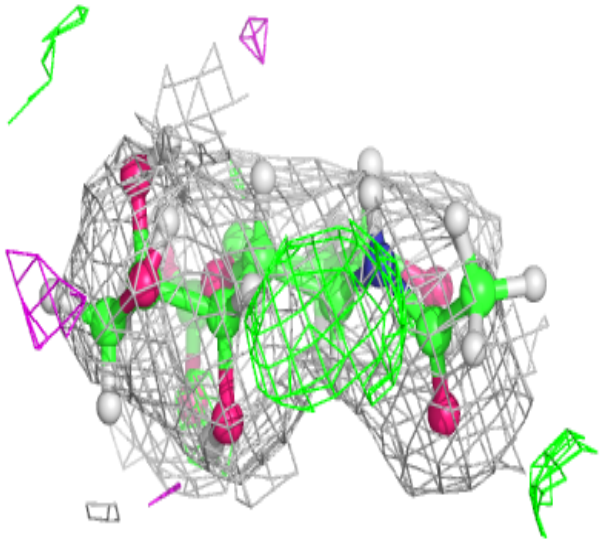


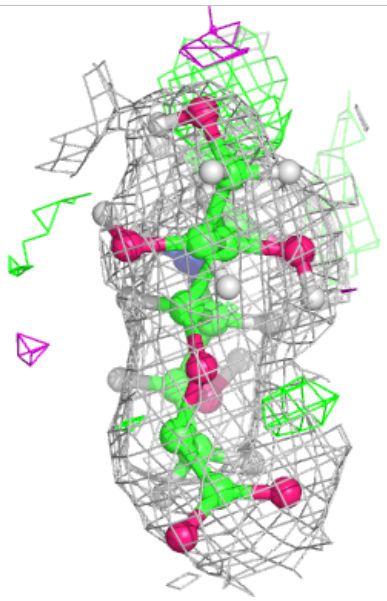


Electron density around SLB G 301:

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

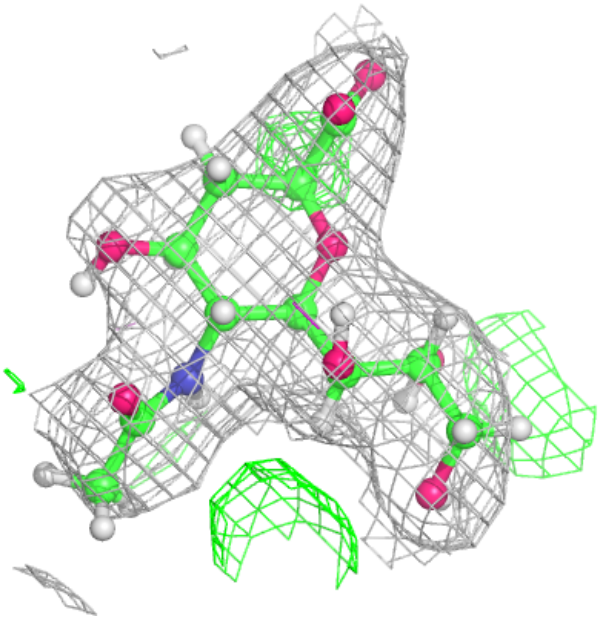


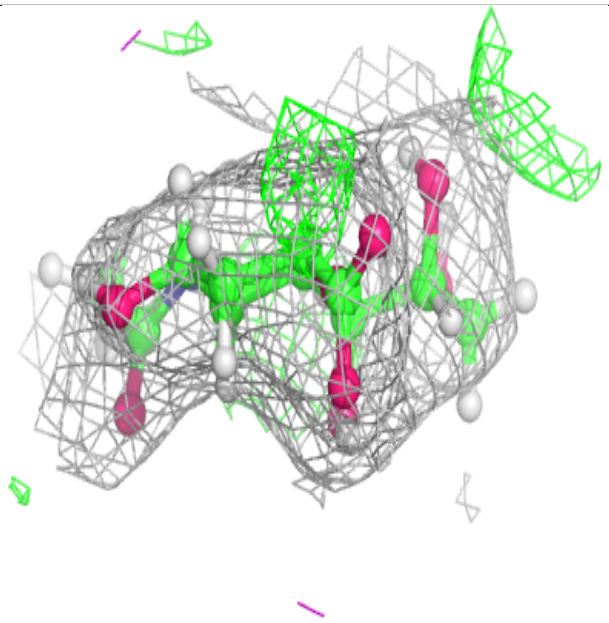


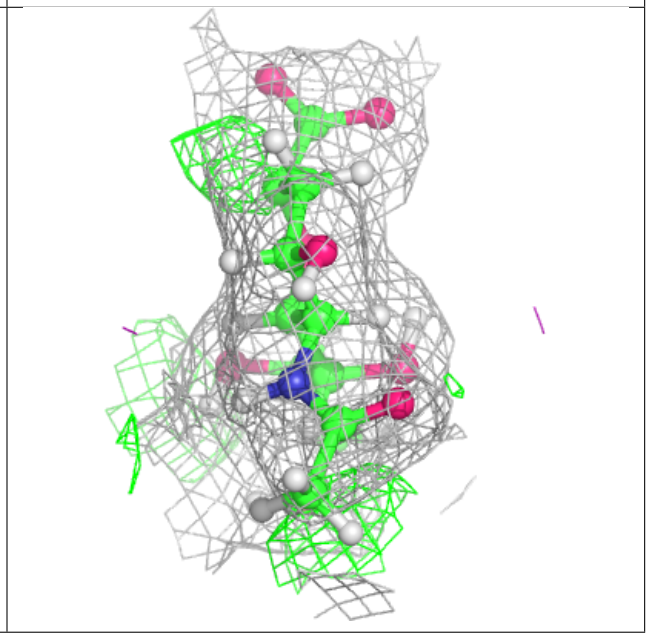


Electron density around SLB Q 301:

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

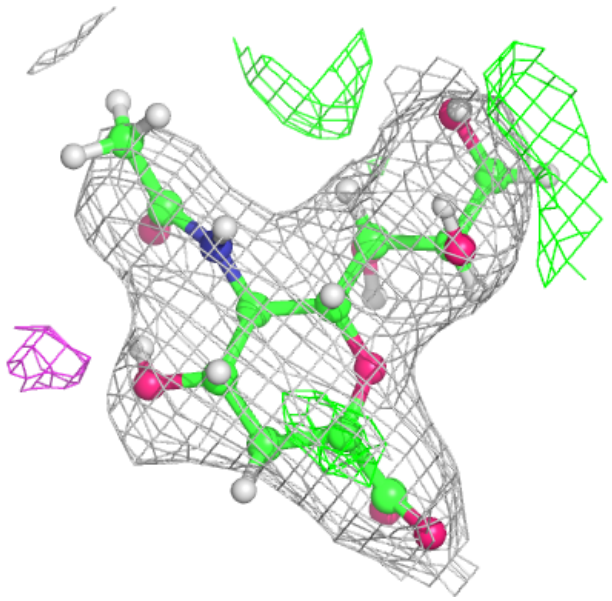


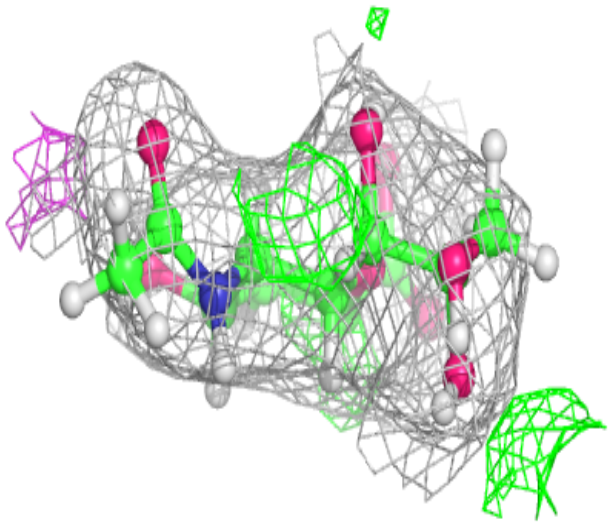


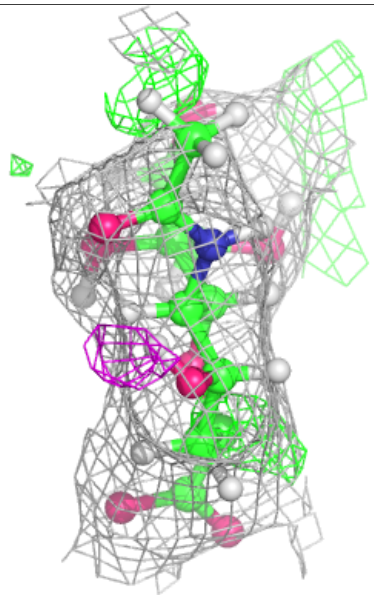


Electron density around SLB O 301:

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

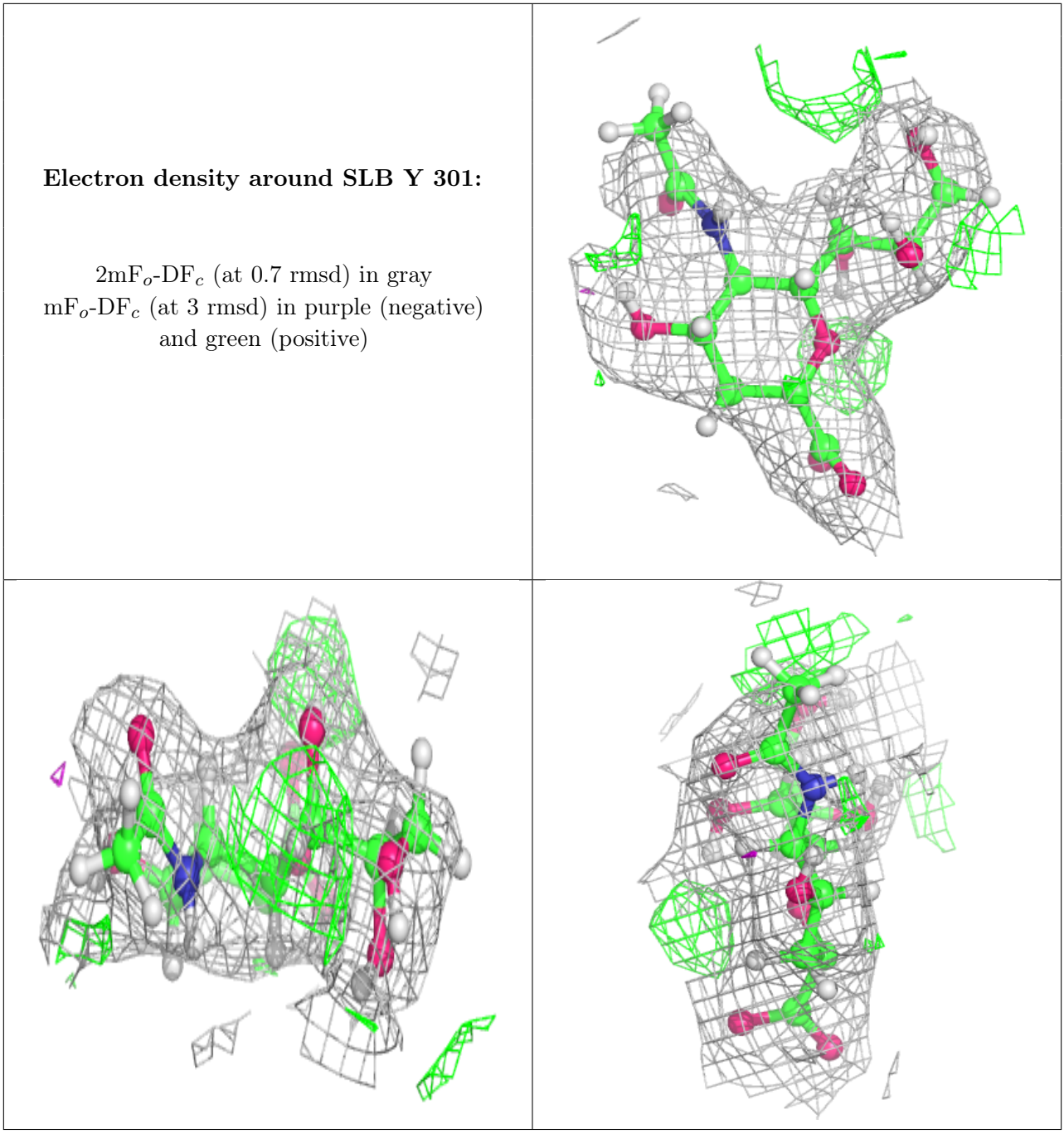






Electron density around SLB Y 301:

$2mF_o$-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.