



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 24, 2022 – 03:59 pm GMT

PDB ID : 6S0W  
Title : The crystal structure of kanamycin B dioxygenase (KanJ) from *Streptomyces kanamyceticus* in complex with nickel and kanamycin B sulfate  
Authors : Mrugala, B.; Niedzialkowska, E.; Minor, W.; Borowski, T.  
Deposited on : 2019-06-18  
Resolution : 2.36 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

The 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 : 2.27  
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) : 2.27

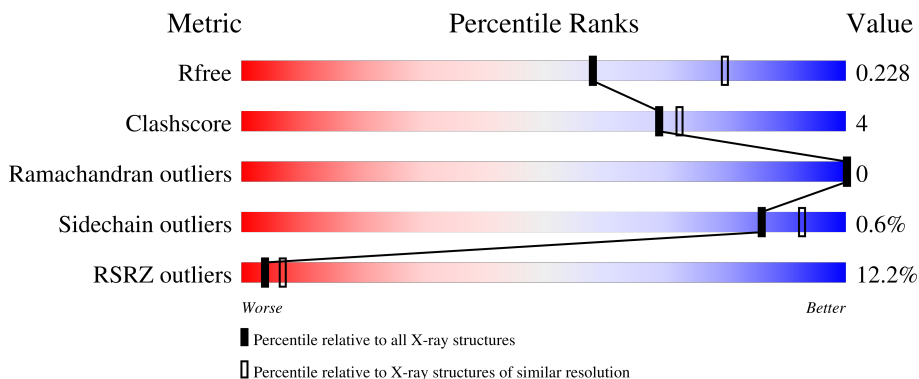
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.36 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	288	 11% 85% 10% 5%
1	B	288	 3% 88% 10% ..
1	C	288	 11% 89% 8% .
1	D	288	 3% 87% 10% .
1	E	288	 28% 86% 9% 5%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	288	

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
2	9CS	A	301	-	-	-	X
4	SO4	C	305	-	-	-	X
4	SO4	F	305	-	-	-	X

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 14359 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 Kanamycin B dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	275	Total	C	N	O	S	0	1	0
			2156	1370	379	398	9			
1	B	282	Total	C	N	O	S	0	0	0
			2197	1397	386	405	9			
1	C	277	Total	C	N	O	S	0	1	0
			2169	1378	381	400	10			
1	D	279	Total	C	N	O	S	0	0	0
			2173	1380	382	401	10			
1	E	274	Total	C	N	O	S	0	1	0
			2149	1366	378	396	9			
1	F	283	Total	C	N	O	S	0	0	0
			2202	1402	387	404	9			

There are 18 discrepancies between the modelled and reference sequences:

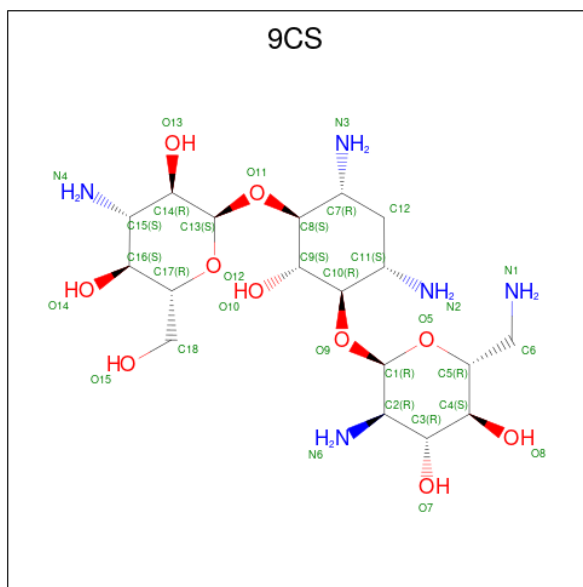
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q6L732
A	-1	ASN	-	expression tag	UNP Q6L732
A	0	ALA	-	expression tag	UNP Q6L732
B	-2	SER	-	expression tag	UNP Q6L732
B	-1	ASN	-	expression tag	UNP Q6L732
B	0	ALA	-	expression tag	UNP Q6L732
C	-2	SER	-	expression tag	UNP Q6L732
C	-1	ASN	-	expression tag	UNP Q6L732
C	0	ALA	-	expression tag	UNP Q6L732
D	-2	SER	-	expression tag	UNP Q6L732
D	-1	ASN	-	expression tag	UNP Q6L732
D	0	ALA	-	expression tag	UNP Q6L732
E	-2	SER	-	expression tag	UNP Q6L732
E	-1	ASN	-	expression tag	UNP Q6L732
E	0	ALA	-	expression tag	UNP Q6L732
F	-2	SER	-	expression tag	UNP Q6L732
F	-1	ASN	-	expression tag	UNP Q6L732

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ALA	-	expression tag	UNP Q6L732

- Molecule 2 is (1R,2S,3S,4R,6S)-4,6-DIAMINO-3-[(3-AMINO-3-DEOXY-ALPHA-D-GLUCOPYRANOSYL)OXY]-2-HYDROXYCYCLOHEXYL 2,6-DIAMINO-2,6-DIDEOXY-ALPHA-D-GLUCOPYRANOSIDE (three-letter code: 9CS) (formula: C<sub>18</sub>H<sub>37</sub>N<sub>5</sub>O<sub>10</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
2	A	1	Total	C	N	O	0	0
			33	18	5	10		
2	B	1	Total	C	N	O	0	0
			33	18	5	10		
2	C	1	Total	C	N	O	0	0
			33	18	5	10		
2	D	1	Total	C	N	O	0	0
			33	18	5	10		
2	E	1	Total	C	N	O	0	0
			33	18	5	10		
2	F	1	Total	C	N	O	0	0
			33	18	5	10		

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni) (labeled as "Ligand of Interest" by depositor).

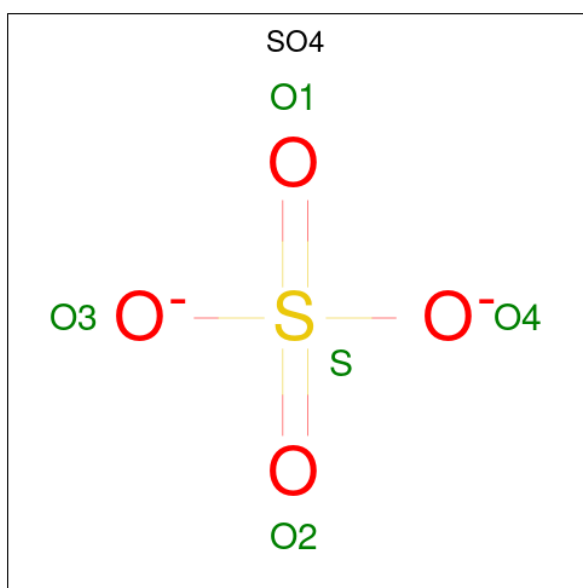
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ni	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Ni 1 1	0	0
3	C	1	Total Ni 1 1	0	0
3	D	1	Total Ni 1 1	0	0
3	E	1	Total Ni 1 1	0	0
3	F	1	Total Ni 1 1	0	0

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



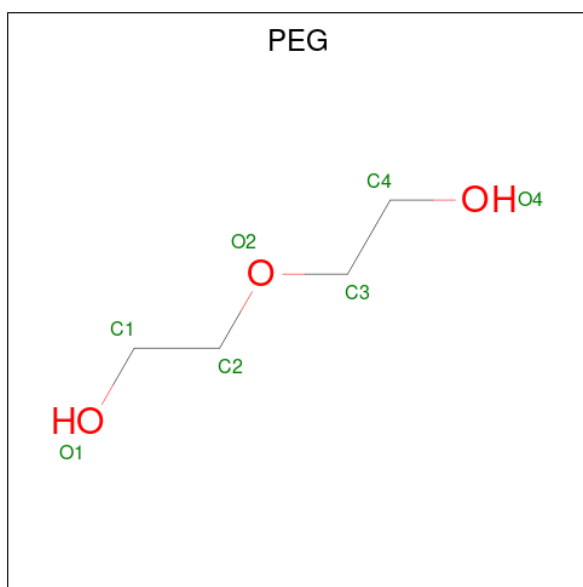
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	E	1	Total 5	O 4	S 1	0	0
4	E	1	Total 5	O 4	S 1	0	0
4	E	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	1	Total C O 7 4 3	0	0

- Molecule 6 is water.

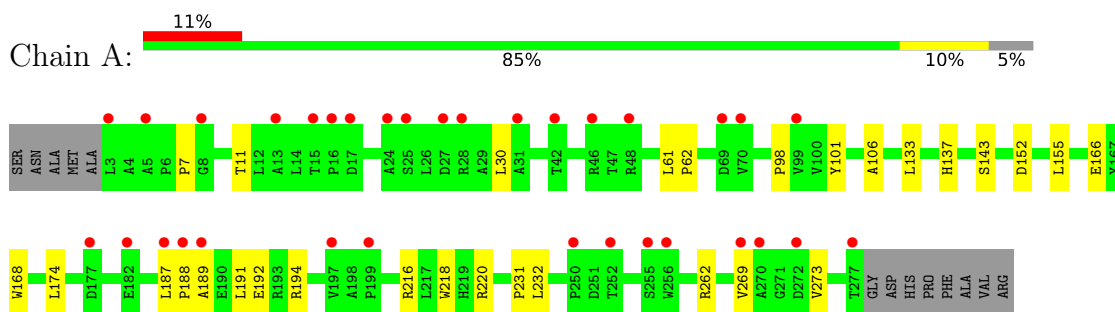
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	142	Total O 142 142	0	0
6	B	269	Total O 269 269	0	0
6	C	161	Total O 161 161	0	0
6	D	216	Total O 216 216	0	0
6	E	82	Total O 82 82	0	0
6	F	137	Total O 137 137	0	0



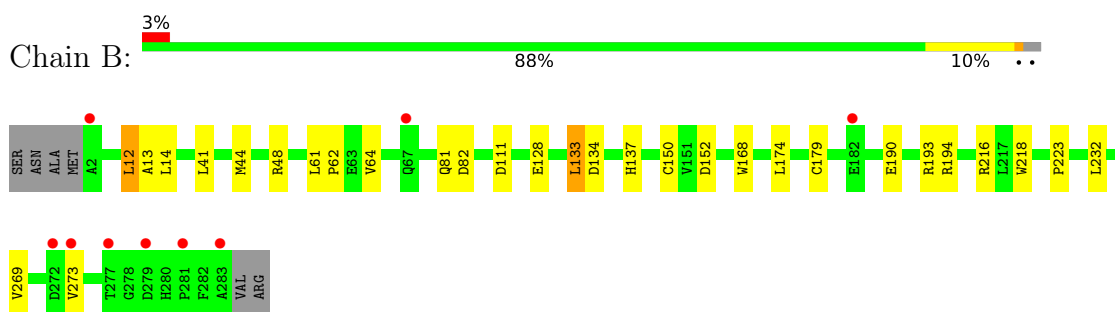
### 3 Residue-property plots

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.

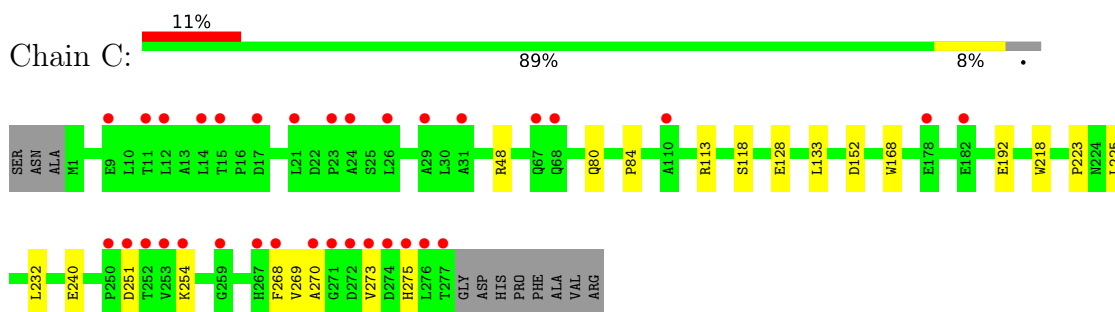
- Molecule 1: Kanamycin B dioxygenase



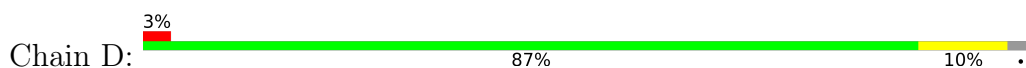
- Molecule 1: Kanamycin B dioxygenase



- Molecule 1: Kanamycin B dioxygenase



- Molecule 1: Kanamycin B dioxygenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.44Å 186.06Å 110.17Å 90.00° 95.20° 90.00°	Depositor
Resolution (Å)	44.22 – 2.36 44.19 – 2.36	Depositor EDS
% Data completeness (in resolution range)	95.1 (44.22-2.36) 95.1 (44.19-2.36)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.60 (at 2.37Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.185 , 0.225 0.191 , 0.228	Depositor DCC
$R_{free}$ test set	3946 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.5	Xtriage
Anisotropy	0.194	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	14359	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: NI, PEG, SO4, 9CS

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	0.62	0/2222	0.77	0/3057
1	B	0.65	0/2266	0.76	0/3118
1	C	0.64	0/2235	0.76	0/3074
1	D	0.63	0/2239	0.79	0/3080
1	E	0.62	0/2215	0.77	0/3047
1	F	0.62	0/2270	0.81	0/3123
All	All	0.63	0/13447	0.78	0/18499

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	A	2156	0	2106	24	0
1	B	2197	0	2139	20	0
1	C	2169	0	2123	13	0
1	D	2173	0	2127	19	0
1	E	2149	0	2099	18	0
1	F	2202	0	2147	9	0
2	A	33	0	37	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	33	0	37	3	0
2	C	33	0	37	0	0
2	D	33	0	37	1	0
2	E	33	0	37	0	0
2	F	33	0	37	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	10	0	0	0	0
4	B	15	0	0	0	0
4	C	15	0	0	0	0
4	D	15	0	0	0	0
4	E	15	0	0	0	0
4	F	25	0	0	0	0
5	D	7	0	10	1	0
6	A	142	0	0	0	0
6	B	269	0	0	2	0
6	C	161	0	0	0	0
6	D	216	0	0	0	0
6	E	82	0	0	0	0
6	F	137	0	0	1	0
All	All	14359	0	12973	98	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 4.

All (98) 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:C:251:ASP:OD2	1:C:270:ALA:HA	1.78	0.83
1:D:187:LEU:HD12	1:D:188:PRO:HD2	1.67	0.76
1:A:187:LEU:HD23	1:A:189:ALA:H	1.51	0.75
1:E:73:ASN:HD21	1:E:80[B]:GLN:HG2	1.57	0.69
1:F:166:GLU:HB3	1:F:220:ARG:HG3	1.79	0.65
1:B:269:VAL:CG2	1:B:273:VAL:HG22	2.31	0.60
1:A:152:ASP:HB3	1:A:232:LEU:HD11	1.83	0.59
1:B:133:LEU:HD12	1:B:133:LEU:O	2.04	0.58
1:F:192:GLU:HA	1:F:192:GLU:OE1	2.03	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:188:PRO:HG2	1:E:15:THR:HG22	1.86	0.57
1:B:152:ASP:HB3	1:B:232:LEU:HD11	1.86	0.57
1:C:269:VAL:CG2	1:C:273:VAL:HG22	2.34	0.57
1:E:192:GLU:HA	1:E:192:GLU:OE1	2.05	0.57
1:B:190:GLU:OE1	1:B:193:ARG:NH1	2.27	0.56
1:A:192:GLU:HA	1:A:192:GLU:OE1	2.05	0.55
1:A:101:TYR:CD2	1:A:262:ARG:HB3	2.42	0.54
1:E:131:VAL:HB	1:E:191:LEU:HD11	1.90	0.53
1:B:269:VAL:HG21	1:B:273:VAL:HG22	1.89	0.53
1:D:168:TRP:CH2	1:D:191:LEU:HD22	2.44	0.53
1:D:192:GLU:OE1	1:D:192:GLU:HA	2.09	0.53
2:B:301:9CS:O10	2:B:301:9CS:H1	2.09	0.53
1:E:73:ASN:ND2	1:E:80[B]:GLN:HG2	2.24	0.52
1:A:269:VAL:CG2	1:A:273:VAL:HG22	2.41	0.51
1:B:168:TRP:HB2	1:B:218:TRP:HB2	1.93	0.51
1:E:181:ASP:OD2	1:E:185:CYS:HB2	2.11	0.51
1:D:166:GLU:HB3	1:D:220:ARG:HG3	1.92	0.51
1:D:64:VAL:HG21	1:D:81:GLN:HG3	1.92	0.50
1:A:7:PRO:O	1:A:11:THR:OG1	2.20	0.50
1:D:3:LEU:CD1	1:D:7:PRO:HG3	2.42	0.50
1:E:269:VAL:CG2	1:E:273:VAL:HG22	2.42	0.50
1:A:269:VAL:HG21	1:A:273:VAL:HG22	1.94	0.49
1:E:17:ASP:OD1	1:E:46:ARG:NH2	2.46	0.49
1:B:111:ASP:HB2	6:B:498:HOH:O	2.13	0.49
1:D:174:LEU:HD11	1:D:194:ARG:HD3	1.93	0.49
1:E:131:VAL:HG11	1:E:191:LEU:HD21	1.94	0.48
1:B:82:ASP:HB3	6:B:426:HOH:O	2.13	0.48
1:A:137:HIS:CE1	1:A:216:ARG:HD2	2.49	0.48
1:B:174:LEU:HD11	1:B:194:ARG:HD3	1.96	0.48
1:B:179:CYS:SG	1:B:190:GLU:HG2	2.53	0.48
1:A:168:TRP:CH2	1:A:191:LEU:HD22	2.48	0.47
1:A:187:LEU:HG	1:A:188:PRO:HD2	1.97	0.47
1:D:4:ALA:O	1:D:204:PRO:HG2	2.13	0.47
1:A:168:TRP:HB2	1:A:218:TRP:HB2	1.97	0.47
1:E:137:HIS:CE1	1:E:216:ARG:HD2	2.50	0.47
1:C:168:TRP:HB2	1:C:218:TRP:HB2	1.96	0.47
1:F:168:TRP:HB2	1:F:218:TRP:HB2	1.97	0.47
1:E:168:TRP:HB2	1:E:218:TRP:HB2	1.97	0.46
1:A:187:LEU:HD23	1:A:189:ALA:N	2.26	0.46
1:C:269:VAL:HG21	1:C:273:VAL:HG22	1.97	0.46
1:B:12:LEU:HD12	1:B:13:ALA:N	2.31	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:48:ARG:NH1	1:C:48:ARG:HG3	2.30	0.46
1:C:225:LEU:O	1:D:62:PRO:HA	2.16	0.46
1:C:128:GLU:HG2	1:C:223:PRO:HD3	1.97	0.45
1:D:47:THR:H	5:D:306:PEG:C1	2.30	0.45
1:A:174:LEU:HD11	1:A:194:ARG:HD3	1.99	0.45
1:D:61:LEU:HB3	1:D:62:PRO:HD3	1.99	0.45
1:B:64:VAL:HG21	1:B:81:GLN:HG3	1.99	0.44
1:B:150:CYS:SG	2:B:301:9CS:N1	2.90	0.44
1:C:84:PRO:HA	1:C:275:HIS:CE1	2.52	0.44
1:A:166:GLU:HB3	1:A:220:ARG:HG3	2.00	0.44
1:A:101:TYR:CG	1:A:262:ARG:HB3	2.52	0.44
1:D:0:ALA:HB2	1:D:166:GLU:OE1	2.17	0.44
1:C:192:GLU:HB3	1:F:65:LEU:HD13	2.00	0.43
1:C:152:ASP:HB3	1:C:232:LEU:HD11	2.01	0.43
1:E:269:VAL:HG21	1:E:273:VAL:HG22	2.01	0.43
1:A:98:PRO:HB3	1:A:262:ARG:NH2	2.33	0.43
1:E:12:LEU:HD23	1:E:14:LEU:HG	2.01	0.43
1:D:80:GLN:NE2	2:D:301:9CS:O7	2.52	0.42
1:A:30:LEU:HD11	1:A:106:ALA:HB1	2.01	0.42
1:D:168:TRP:HB2	1:D:218:TRP:HB2	2.01	0.42
1:C:80[B]:GLN:NE2	1:C:118:SER:OG	2.52	0.42
1:A:188:PRO:HG2	1:E:15:THR:CG2	2.49	0.42
1:D:152:ASP:HB3	1:D:232:LEU:HD11	2.02	0.42
1:F:137:HIS:CE1	1:F:216:ARG:HD2	2.55	0.42
1:C:254:LYS:HA	1:C:268:PHE:CE2	2.55	0.42
1:E:61:LEU:HB3	1:E:62:PRO:HD3	2.00	0.42
1:B:14:LEU:HD12	1:B:44:MET:SD	2.60	0.42
1:F:152:ASP:HB3	1:F:232:LEU:HD11	2.01	0.42
1:D:61:LEU:N	1:D:62:PRO:CD	2.83	0.42
1:F:43:GLY:HA2	6:F:500:HOH:O	2.18	0.42
1:B:12:LEU:HB3	1:B:41:LEU:HD23	2.02	0.42
1:B:137:HIS:CE1	1:B:216:ARG:HD2	2.55	0.41
1:A:61:LEU:HB3	1:A:62:PRO:HD3	2.02	0.41
1:E:168:TRP:CH2	1:E:191:LEU:HD22	2.54	0.41
1:A:155:LEU:HB2	1:A:231:PRO:HG2	2.01	0.41
1:C:113:ARG:HD3	1:C:240:GLU:HA	2.02	0.41
1:B:61:LEU:N	1:B:62:PRO:CD	2.84	0.41
1:F:61:LEU:HB3	1:F:62:PRO:HD3	2.02	0.41
1:A:143:SER:OG	1:D:270:ALA:HB1	2.21	0.41
1:B:128:GLU:HG3	1:B:223:PRO:HD3	2.02	0.41
1:D:30:LEU:HD11	1:D:106:ALA:HB1	2.01	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:148:CYS:HA	1:E:237:HIS:O	2.21	0.41
1:F:131:VAL:HG11	1:F:191:LEU:HD21	2.03	0.41
1:B:134:ASP:OD2	2:B:301:9CS:N1	2.50	0.40
1:D:131:VAL:HG11	1:D:191:LEU:HD21	2.03	0.40
1:E:152:ASP:HB3	1:E:232:LEU:HD11	2.02	0.40
1:A:187:LEU:HG	1:A:188:PRO:CD	2.51	0.40
1:A:61:LEU:N	1:A:62:PRO:CD	2.85	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	A	274/288 (95%)	269 (98%)	5 (2%)	0	100	100
1	B	280/288 (97%)	276 (99%)	4 (1%)	0	100	100
1	C	276/288 (96%)	271 (98%)	5 (2%)	0	100	100
1	D	277/288 (96%)	273 (99%)	4 (1%)	0	100	100
1	E	273/288 (95%)	269 (98%)	4 (2%)	0	100	100
1	F	279/288 (97%)	275 (99%)	4 (1%)	0	100	100
All	All	1659/1728 (96%)	1633 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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

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

The Analysed column shows the number of residues for which the sidechain conformation was



analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	240/248 (97%)	239 (100%)	1 (0%)	91	95
1	B	243/248 (98%)	241 (99%)	2 (1%)	81	89
1	C	241/248 (97%)	240 (100%)	1 (0%)	91	95
1	D	241/248 (97%)	240 (100%)	1 (0%)	91	95
1	E	239/248 (96%)	238 (100%)	1 (0%)	91	95
1	F	243/248 (98%)	241 (99%)	2 (1%)	81	89
All	All	1447/1488 (97%)	1439 (99%)	8 (1%)	86	93

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	133	LEU
1	B	12	LEU
1	B	133	LEU
1	C	133	LEU
1	D	133	LEU
1	E	133	LEU
1	F	133	LEU
1	F	161	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	73	ASN
1	D	80	GLN
1	E	73	ASN

### 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 32 ligands modelled in this entry, 6 are monoatomic - leaving 26 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
4	SO4	A	303	-	4,4,4	0.37	0	6,6,6	0.04	0
4	SO4	B	304	-	4,4,4	0.38	0	6,6,6	0.05	0
4	SO4	E	303	-	4,4,4	0.38	0	6,6,6	0.07	0
4	SO4	E	304	-	4,4,4	0.39	0	6,6,6	0.04	0
4	SO4	D	303	-	4,4,4	0.37	0	6,6,6	0.04	0
4	SO4	F	303	-	4,4,4	0.37	0	6,6,6	0.05	0
4	SO4	E	305	-	4,4,4	0.38	0	6,6,6	0.05	0
2	9CS	A	301	-	35,35,35	0.34	0	46,52,52	0.70	0
4	SO4	A	304	-	4,4,4	0.38	0	6,6,6	0.06	0
4	SO4	C	305	-	4,4,4	0.38	0	6,6,6	0.07	0
4	SO4	D	305	-	4,4,4	0.38	0	6,6,6	0.06	0
4	SO4	F	305	-	4,4,4	0.39	0	6,6,6	0.06	0
4	SO4	C	303	-	4,4,4	0.37	0	6,6,6	0.07	0
4	SO4	B	303	-	4,4,4	0.36	0	6,6,6	0.08	0
2	9CS	B	301	-	35,35,35	0.36	0	46,52,52	0.74	1 (2%)
2	9CS	E	301	-	35,35,35	0.43	0	46,52,52	0.77	0
4	SO4	F	307	-	4,4,4	0.38	0	6,6,6	0.06	0
4	SO4	C	304	-	4,4,4	0.37	0	6,6,6	0.06	0
2	9CS	C	301	-	35,35,35	0.49	0	46,52,52	0.76	0
4	SO4	D	304	-	4,4,4	0.37	0	6,6,6	0.05	0
5	PEG	D	306	-	6,6,6	0.18	0	5,5,5	0.10	0
4	SO4	F	306	-	4,4,4	0.39	0	6,6,6	0.05	0
4	SO4	B	305	-	4,4,4	0.37	0	6,6,6	0.05	0
2	9CS	F	301	-	35,35,35	0.51	1 (2%)	46,52,52	0.63	0
2	9CS	D	301	-	35,35,35	0.43	0	46,52,52	0.54	0
4	SO4	F	304	-	4,4,4	0.38	0	6,6,6	0.07	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
2	9CS	C	301	-	-	2/12/72/72	0/3/3/3
5	PEG	D	306	-	-	1/4/4/4	-
2	9CS	F	301	-	-	4/12/72/72	0/3/3/3
2	9CS	D	301	-	-	3/12/72/72	0/3/3/3
2	9CS	B	301	-	-	3/12/72/72	0/3/3/3
2	9CS	A	301	-	-	2/12/72/72	0/3/3/3
2	9CS	E	301	-	-	2/12/72/72	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	301	9CS	C1-C2	-2.17	1.48	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	301	9CS	O5-C5-C6	2.71	111.05	106.01

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	301	9CS	C4-C5-C6-N1
2	B	301	9CS	O5-C5-C6-N1
2	C	301	9CS	C4-C5-C6-N1
2	C	301	9CS	O5-C5-C6-N1
2	D	301	9CS	C4-C5-C6-N1
2	D	301	9CS	O5-C5-C6-N1
2	E	301	9CS	C4-C5-C6-N1
2	F	301	9CS	C4-C5-C6-N1
2	F	301	9CS	O5-C5-C6-N1
2	F	301	9CS	C16-C17-C18-O15
2	A	301	9CS	O12-C17-C18-O15
2	F	301	9CS	O12-C17-C18-O15
2	A	301	9CS	C16-C17-C18-O15
5	D	306	PEG	C4-C3-O2-C2

*Continued on next page...*

*Continued from previous page...*

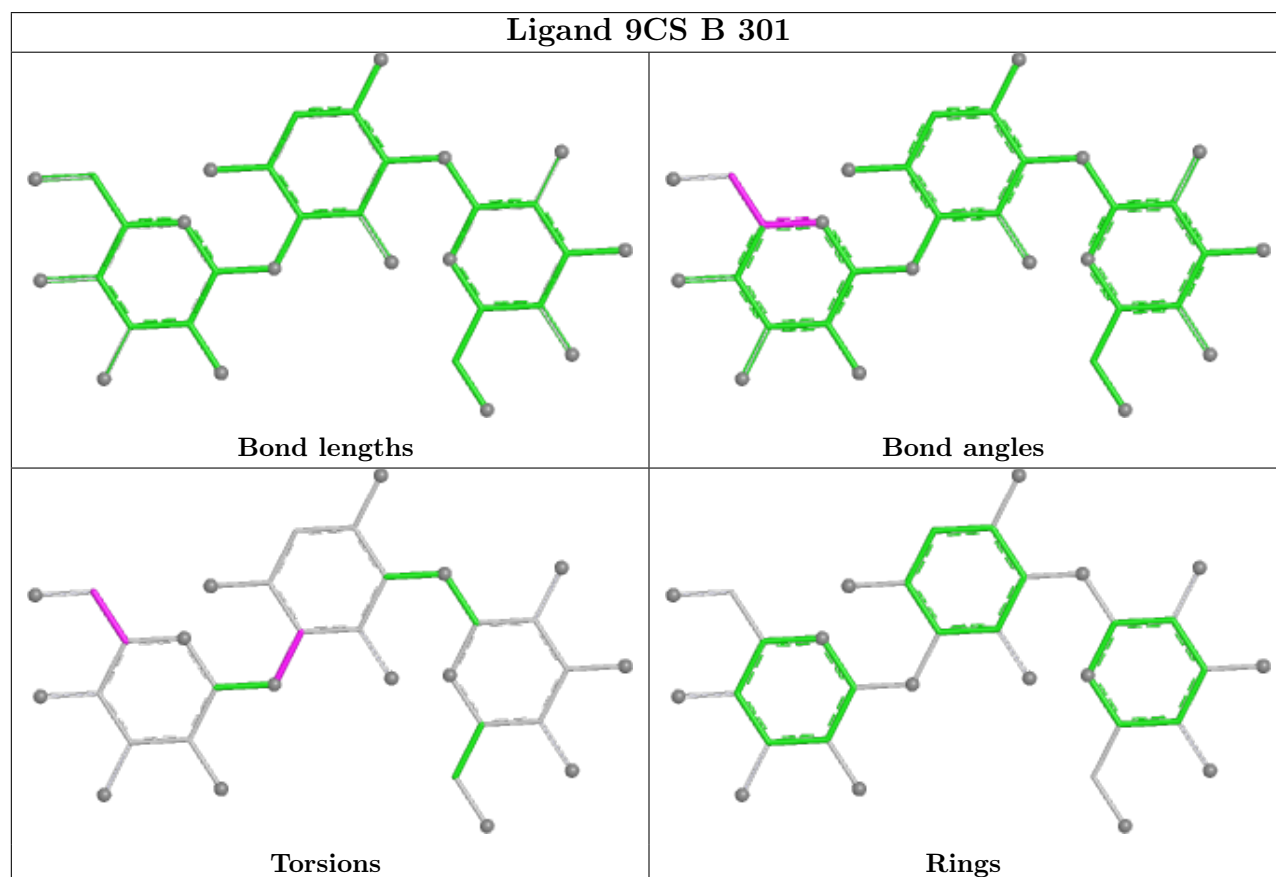
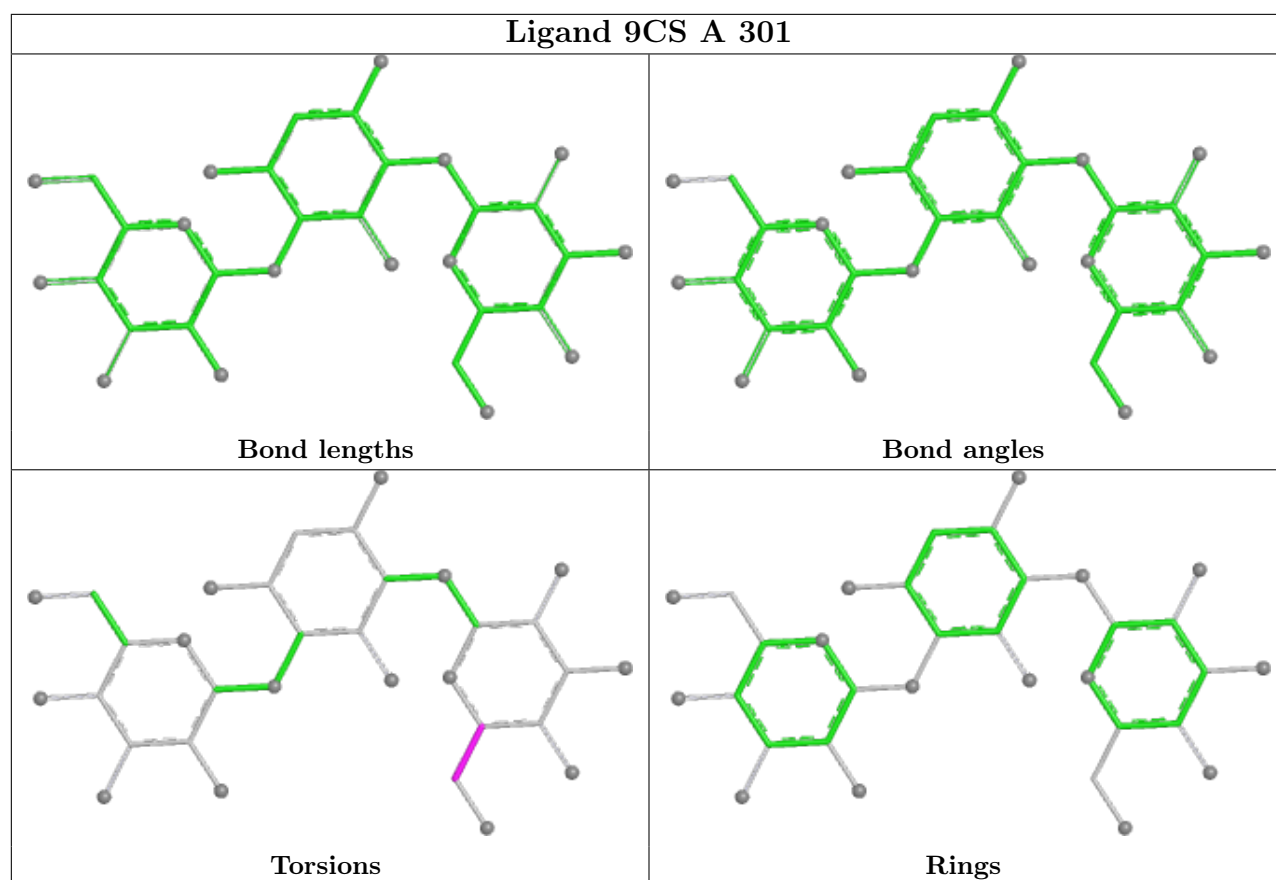
Mol	Chain	Res	Type	Atoms
2	E	301	9CS	O5-C5-C6-N1
2	B	301	9CS	C9-C10-O9-C1
2	D	301	9CS	C16-C17-C18-O15

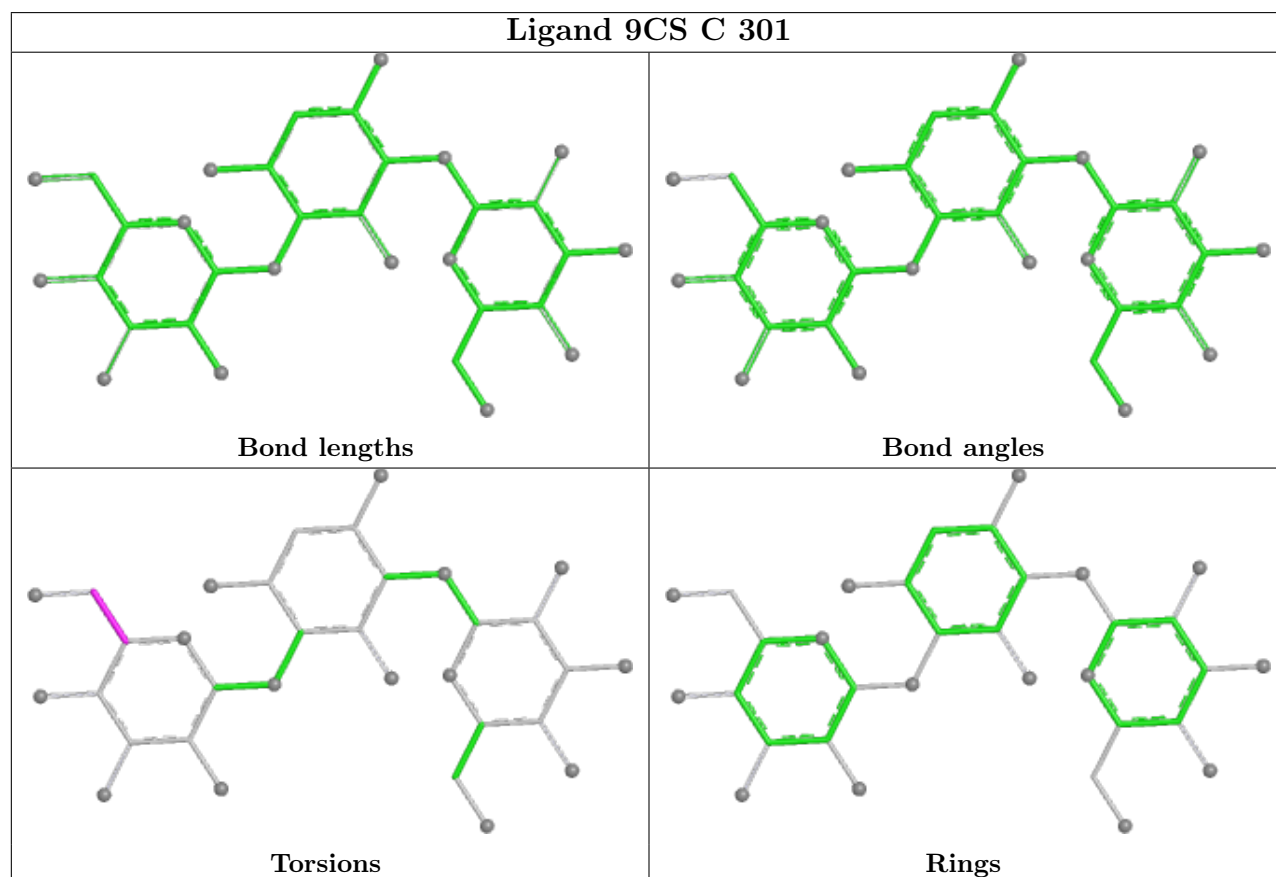
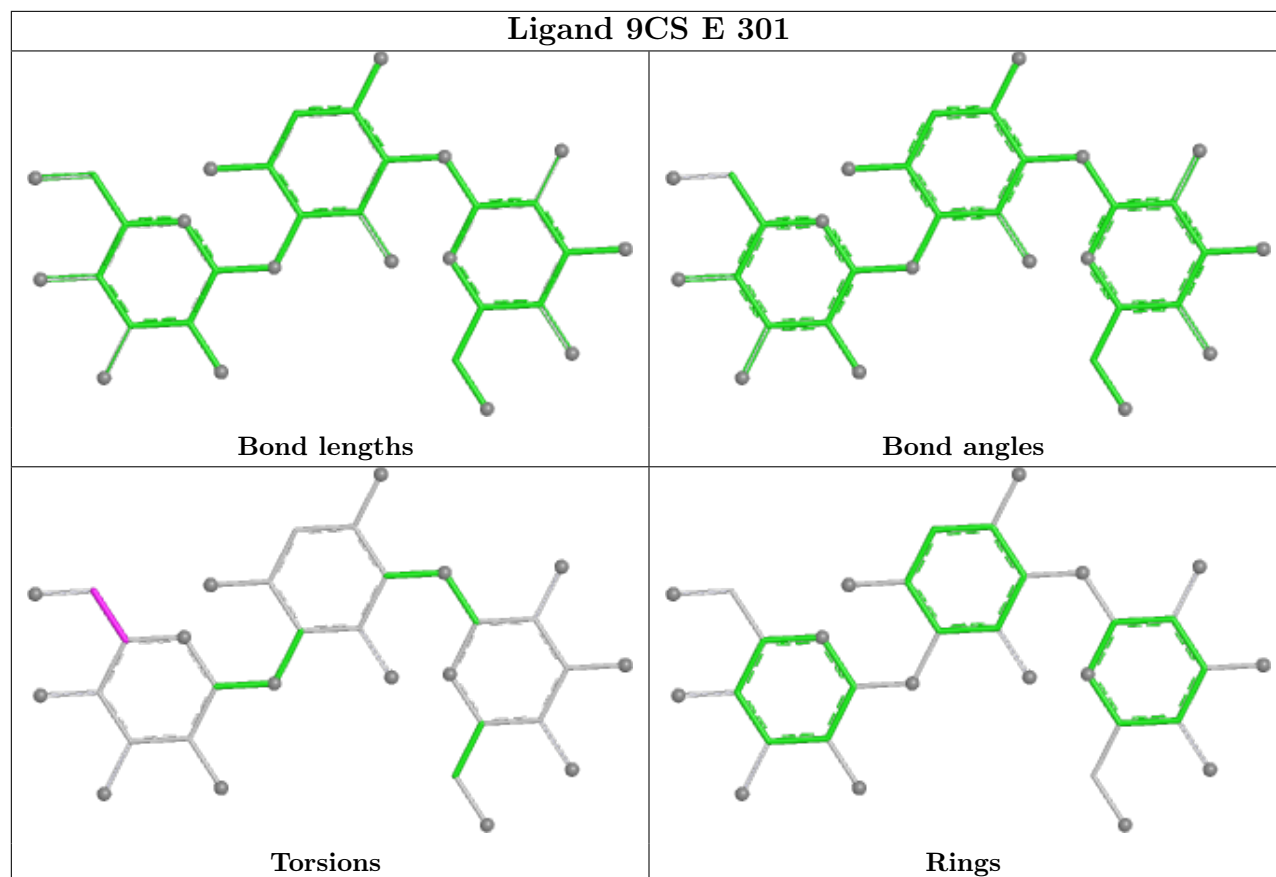
There are no ring outliers.

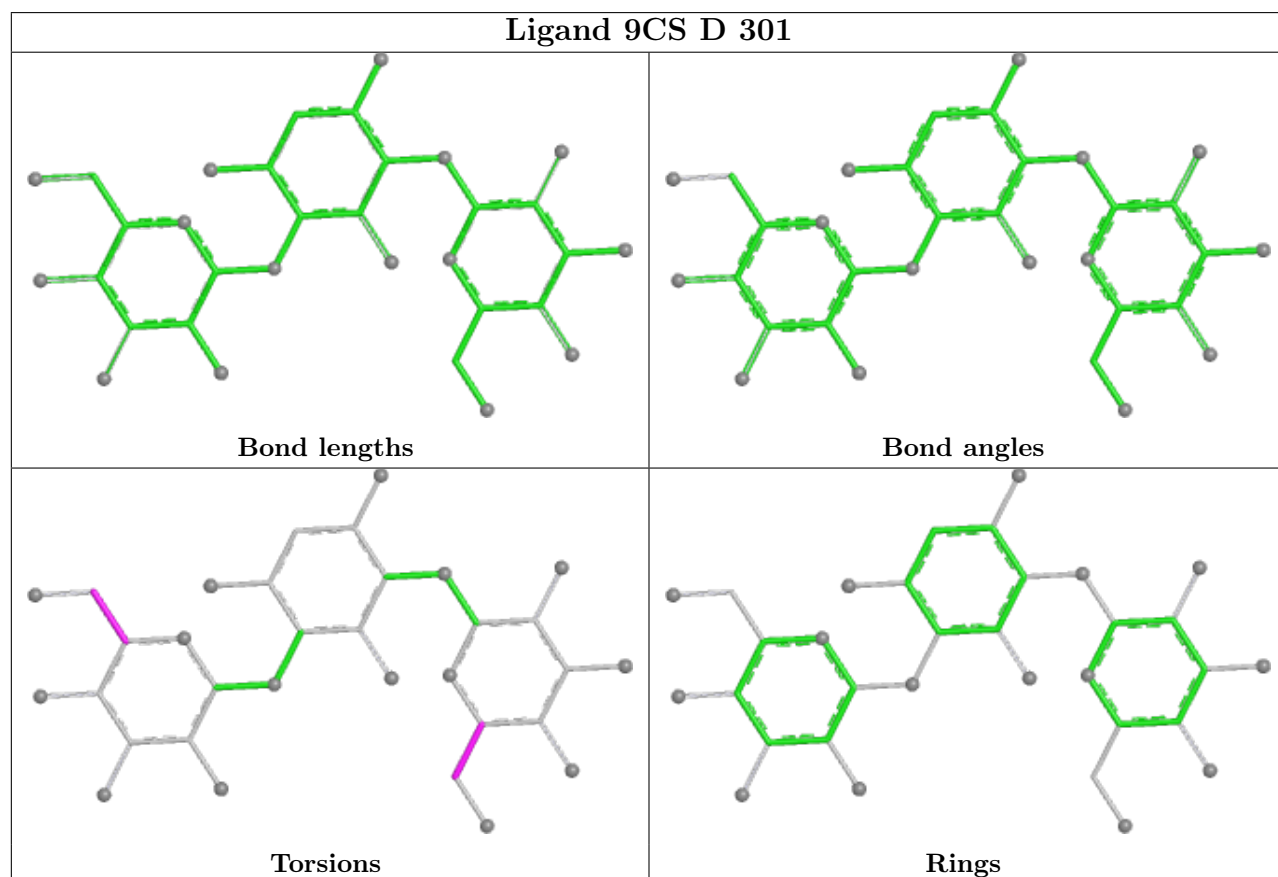
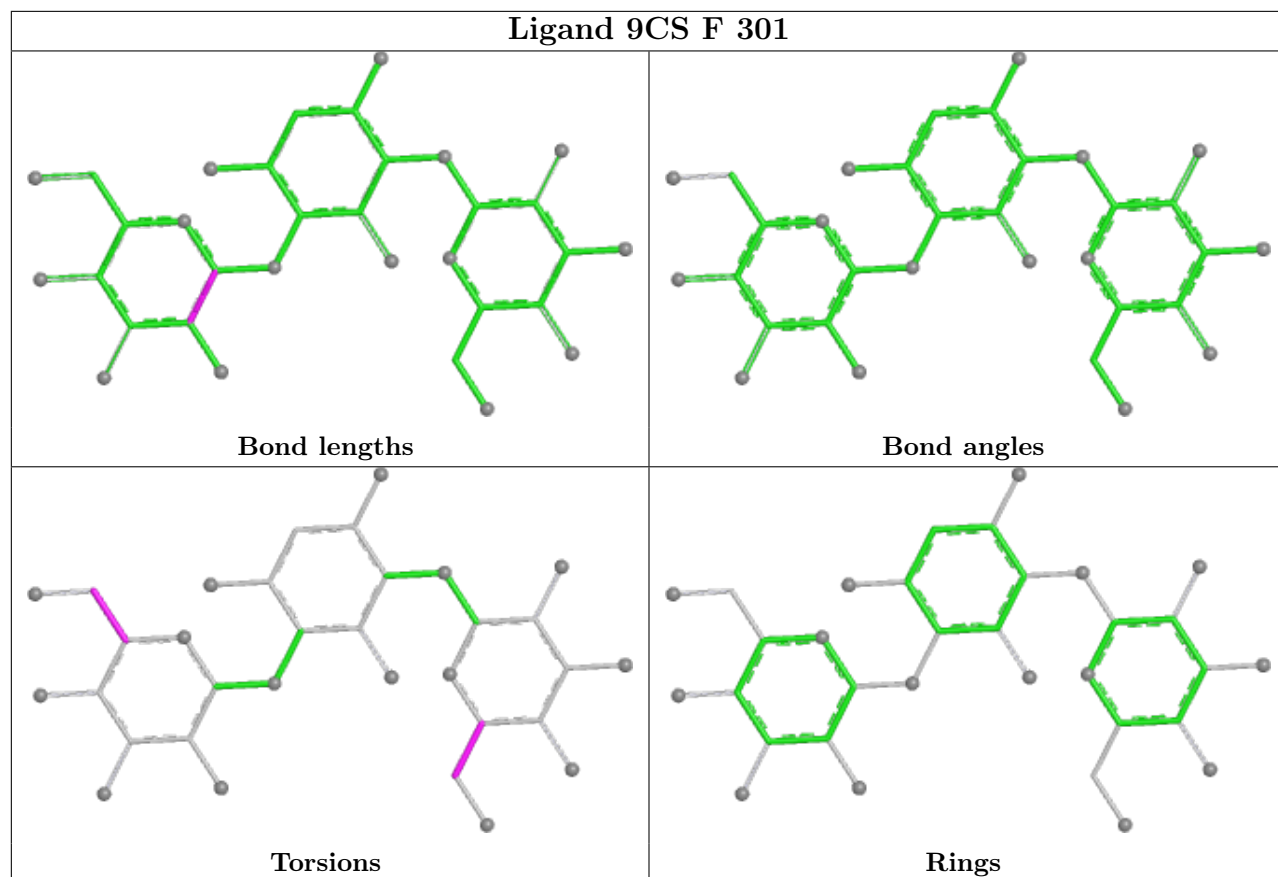
3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	301	9CS	3	0
5	D	306	PEG	1	0
2	D	301	9CS	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	275/288 (95%)	0.92	33 (12%) 4 7	40, 62, 98, 112	1 (0%)
1	B	282/288 (97%)	0.53	9 (3%) 47 59	31, 46, 77, 106	1 (0%)
1	C	277/288 (96%)	0.84	33 (11%) 4 7	39, 62, 106, 141	0
1	D	279/288 (96%)	0.57	10 (3%) 42 55	36, 50, 79, 118	0
1	E	274/288 (95%)	1.53	80 (29%) 0 0	50, 78, 117, 145	0
1	F	283/288 (98%)	0.90	39 (13%) 2 4	36, 63, 103, 124	1 (0%)
All	All	1670/1728 (96%)	0.88	204 (12%) 4 7	31, 59, 103, 145	3 (0%)

All (204) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	274	ASP	8.7
1	E	271	GLY	8.2
1	C	272	ASP	7.0
1	E	179	CYS	6.9
1	C	277	THR	6.7
1	E	46	ARG	6.5
1	E	270	ALA	6.5
1	E	21	LEU	6.2
1	F	281	PRO	6.2
1	E	272	ASP	5.8
1	E	276	LEU	5.8
1	C	273	VAL	5.7
1	E	22	ASP	5.7
1	E	49	LEU	5.5
1	F	24	ALA	5.5
1	F	196	ALA	5.3
1	A	17	ASP	5.3
1	F	68	GLN	5.2
1	D	276	LEU	5.2

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	187	LEU	5.2
1	D	17	ASP	5.1
1	F	282	PHE	5.1
1	C	276	LEU	4.9
1	E	55	THR	4.7
1	F	271	GLY	4.6
1	B	277	THR	4.6
1	E	251	ASP	4.5
1	E	12	LEU	4.5
1	E	259	GLY	4.4
1	E	51	ASP	4.4
1	E	275	HIS	4.4
1	C	12	LEU	4.3
1	E	252	THR	4.3
1	E	18	ASP	4.2
1	B	2	ALA	4.2
1	B	281	PRO	4.2
1	F	178	GLU	4.2
1	A	42	THR	4.1
1	F	17	ASP	4.1
1	A	46	ARG	4.0
1	F	16	PRO	4.0
1	C	274	ASP	4.0
1	F	22	ASP	4.0
1	E	264	THR	4.0
1	F	23	PRO	3.9
1	F	15	THR	3.9
1	E	15	THR	3.9
1	F	284	VAL	3.9
1	A	16	PRO	3.9
1	E	69	ASP	3.9
1	F	187	LEU	3.9
1	A	27	ASP	3.8
1	C	31	ALA	3.8
1	C	24	ALA	3.8
1	E	255	SER	3.8
1	E	53	LEU	3.8
1	C	23	PRO	3.7
1	E	177	ASP	3.7
1	E	261	ASP	3.7
1	F	179	CYS	3.7
1	E	20	THR	3.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	259	GLY	3.7
1	F	277	THR	3.7
1	E	19	LYS	3.6
1	D	23	PRO	3.6
1	C	68	GLN	3.6
1	E	273	VAL	3.6
1	A	255	SER	3.5
1	E	250	PRO	3.5
1	B	283	ALA	3.5
1	C	252	THR	3.4
1	F	273	VAL	3.4
1	C	271	GLY	3.4
1	B	182	GLU	3.4
1	E	56	ALA	3.4
1	C	270	ALA	3.4
1	F	188	PRO	3.3
1	E	48	ARG	3.3
1	A	277	THR	3.3
1	A	3	LEU	3.3
1	B	273	VAL	3.3
1	C	17	ASP	3.3
1	E	95	LEU	3.3
1	F	189	ALA	3.3
1	E	85	VAL	3.3
1	F	283	ALA	3.2
1	E	253	VAL	3.2
1	F	18	ASP	3.1
1	E	256	TRP	3.1
1	A	199	PRO	3.1
1	A	270	ALA	3.1
1	A	25	SER	3.0
1	E	45	LEU	3.0
1	F	272	ASP	2.9
1	E	178	GLU	2.9
1	E	3	LEU	2.9
1	E	26	LEU	2.9
1	E	24	ALA	2.9
1	A	272	ASP	2.9
1	C	250	PRO	2.9
1	E	269	VAL	2.9
1	B	67	GLN	2.9
1	E	63	GLU	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	279	ASP	2.8
1	A	188	PRO	2.8
1	E	146	PRO	2.8
1	E	207	VAL	2.8
1	C	29	ALA	2.8
1	F	176	PRO	2.8
1	E	91	PHE	2.8
1	F	198	ALA	2.8
1	D	27	ASP	2.8
1	E	27	ASP	2.8
1	F	43	GLY	2.7
1	F	19	LYS	2.7
1	A	189	ALA	2.7
1	E	180	TYR	2.7
1	A	177	ASP	2.7
1	C	14	LEU	2.7
1	F	280	HIS	2.7
1	F	276	LEU	2.7
1	E	87	GLU	2.6
1	E	68	GLN	2.6
1	E	65	LEU	2.6
1	D	18	ASP	2.6
1	E	153	VAL	2.6
1	E	260	SER	2.6
1	A	182	GLU	2.6
1	C	267	HIS	2.6
1	A	48	ARG	2.6
1	A	24	ALA	2.6
1	F	28	ARG	2.6
1	A	269	VAL	2.6
1	A	13	ALA	2.5
1	A	250	PRO	2.5
1	C	26	LEU	2.5
1	C	178	GLU	2.5
1	B	272	ASP	2.5
1	C	21	LEU	2.5
1	C	275	HIS	2.5
1	A	256	TRP	2.5
1	C	110	ALA	2.5
1	C	253	VAL	2.5
1	D	91	PHE	2.5
1	F	25	SER	2.5

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	E	267	HIS	2.5
1	E	187	LEU	2.5
1	E	25	SER	2.4
1	F	67	GLN	2.4
1	E	257	VAL	2.4
1	A	252	THR	2.4
1	D	15	THR	2.4
1	E	101	TYR	2.4
1	A	197	VAL	2.4
1	F	177	ASP	2.4
1	A	28	ARG	2.4
1	C	9	GLU	2.4
1	E	86	ARG	2.4
1	E	239	THR	2.4
1	E	240	GLU	2.3
1	A	8	GLY	2.3
1	D	259	GLY	2.3
1	F	128	GLU	2.3
1	A	99	VAL	2.3
1	E	67	GLN	2.3
1	F	195	ARG	2.3
1	F	193	ARG	2.3
1	C	15	THR	2.3
1	E	43	GLY	2.3
1	E	6	PRO	2.3
1	D	187	LEU	2.2
1	E	89	LEU	2.2
1	C	268	PHE	2.2
1	E	242	PHE	2.2
1	E	138	LEU	2.2
1	E	217	LEU	2.2
1	E	16	PRO	2.2
1	E	218	TRP	2.2
1	C	182	GLU	2.2
1	C	254	LYS	2.1
1	F	46	ARG	2.1
1	E	96	LEU	2.1
1	F	21	LEU	2.1
1	E	54	ARG	2.1
1	E	193	ARG	2.1
1	E	5	ALA	2.1
1	E	234	ALA	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	188	PRO	2.1
1	E	102	GLN	2.1
1	A	5	ALA	2.1
1	A	15	THR	2.1
1	D	47	THR	2.1
1	A	70	VAL	2.1
1	F	285	ARG	2.1
1	C	67	GLN	2.1
1	C	251	ASP	2.1
1	F	14	LEU	2.1
1	A	69	ASP	2.1
1	C	11	THR	2.1
1	E	28	ARG	2.0
1	E	190	GLU	2.0
1	E	47	THR	2.0
1	E	175	ASN	2.0
1	A	31	ALA	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	SO4	B	305	5/5	0.43	0.40	125,130,139,140	0
4	SO4	C	305	5/5	0.47	0.42	120,133,133,134	0
4	SO4	E	305	5/5	0.68	0.40	144,146,147,148	0
4	SO4	D	305	5/5	0.70	0.26	104,105,109,112	0
4	SO4	F	305	5/5	0.70	0.54	113,126,129,129	0
5	PEG	D	306	7/7	0.71	0.31	78,86,96,96	0

*Continued on next page...*

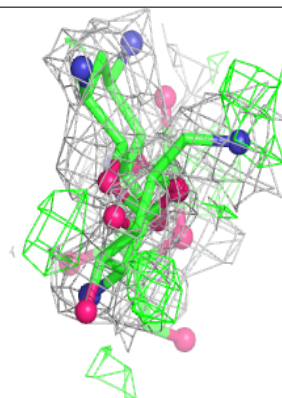
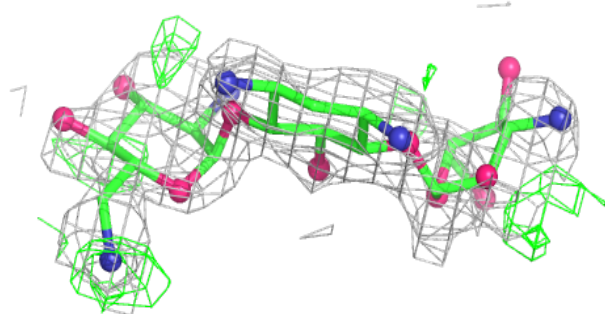
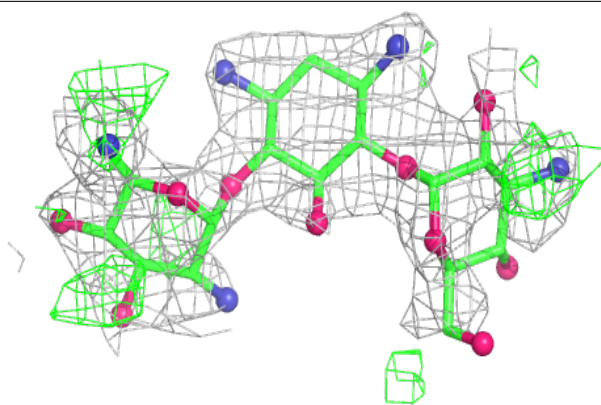
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	9CS	B	301	33/33	0.73	0.35	48,55,68,70	33
4	SO4	F	307	5/5	0.73	0.39	124,125,134,135	0
2	9CS	E	301	33/33	0.73	0.30	57,61,72,73	33
2	9CS	A	301	33/33	0.76	0.57	51,60,65,68	33
2	9CS	C	301	33/33	0.78	0.28	61,65,69,72	33
4	SO4	F	306	5/5	0.78	0.33	138,141,142,144	0
4	SO4	D	304	5/5	0.80	0.26	100,109,113,113	0
2	9CS	D	301	33/33	0.83	0.22	60,68,79,80	7
4	SO4	E	304	5/5	0.85	0.25	109,112,115,116	0
4	SO4	C	304	5/5	0.85	0.20	92,92,101,102	0
4	SO4	E	303	5/5	0.85	0.17	92,93,95,96	0
2	9CS	F	301	33/33	0.89	0.23	37,42,46,48	33
4	SO4	B	304	5/5	0.92	0.40	86,91,93,93	0
4	SO4	F	304	5/5	0.92	0.23	89,89,97,101	0
4	SO4	A	304	5/5	0.92	0.19	93,95,100,102	0
4	SO4	F	303	5/5	0.95	0.16	71,77,82,84	0
4	SO4	D	303	5/5	0.97	0.20	64,70,74,75	0
4	SO4	C	303	5/5	0.97	0.15	56,61,69,69	0
4	SO4	A	303	5/5	0.97	0.17	65,69,74,76	0
4	SO4	B	303	5/5	0.97	0.14	56,59,66,69	0
3	NI	E	302	1/1	0.98	0.06	51,51,51,51	0
3	NI	B	302	1/1	0.99	0.11	30,30,30,30	1
3	NI	C	302	1/1	0.99	0.10	40,40,40,40	0
3	NI	D	302	1/1	0.99	0.10	38,38,38,38	0
3	NI	A	302	1/1	0.99	0.10	44,44,44,44	0
3	NI	F	302	1/1	0.99	0.09	41,41,41,41	0

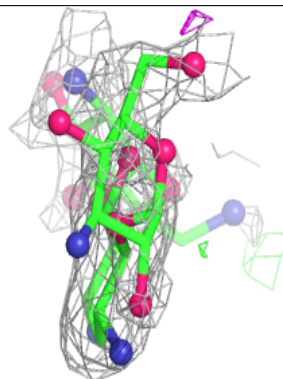
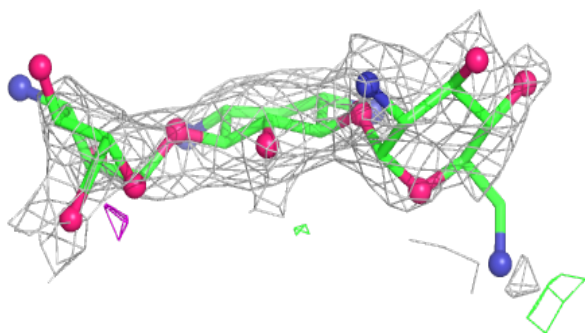
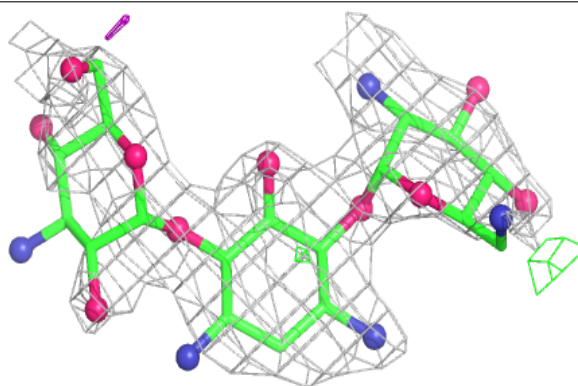
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around 9CS B 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 9CS E 301:**

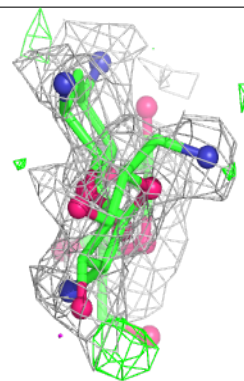
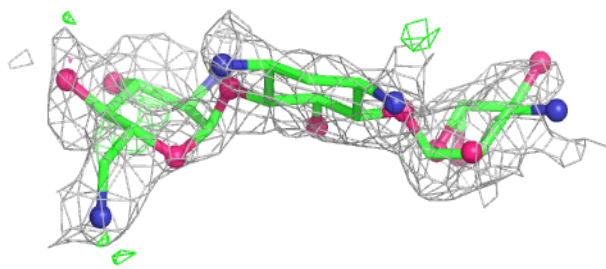
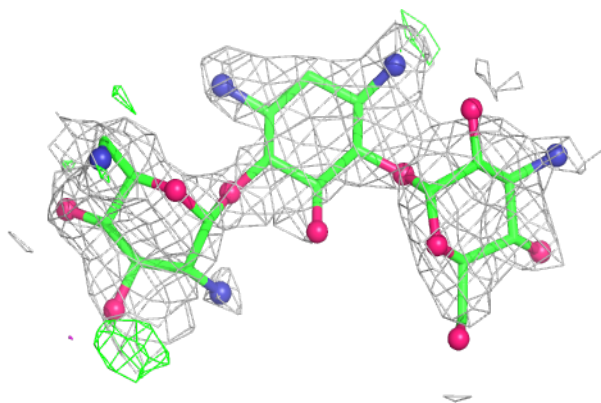
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



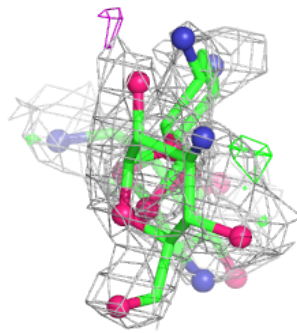
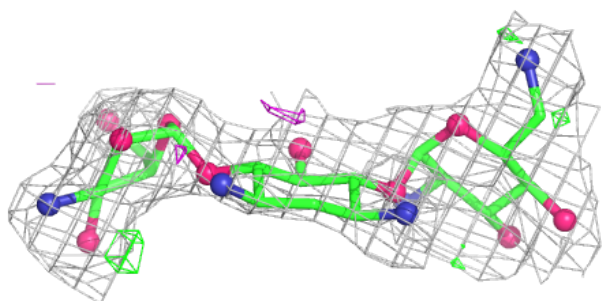
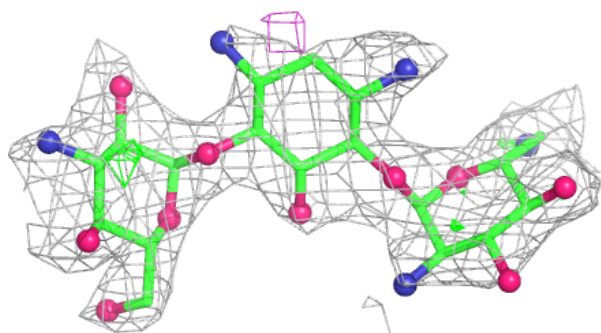


**Electron density around 9CS A 301:**

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

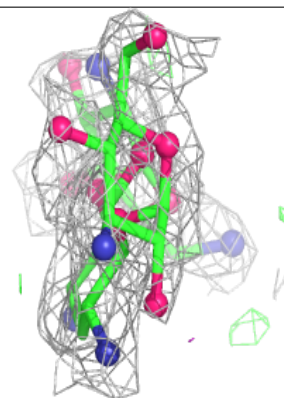
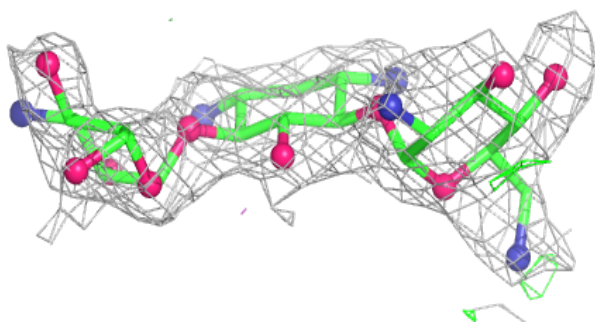
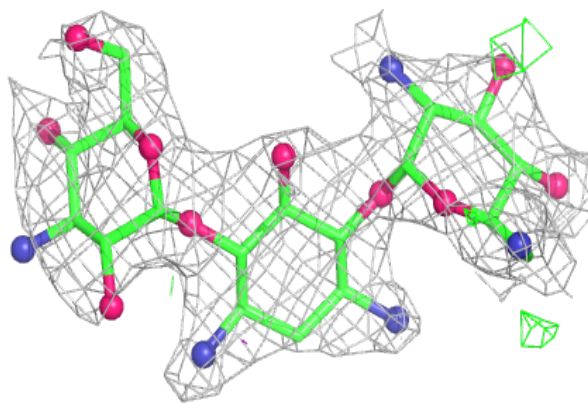
**Electron density around 9CS C 301:**

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

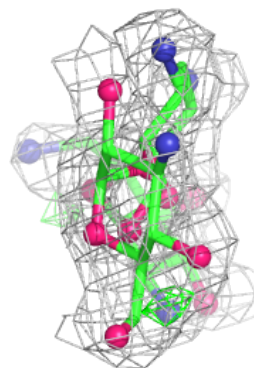
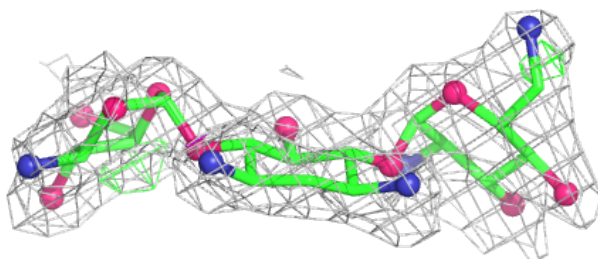
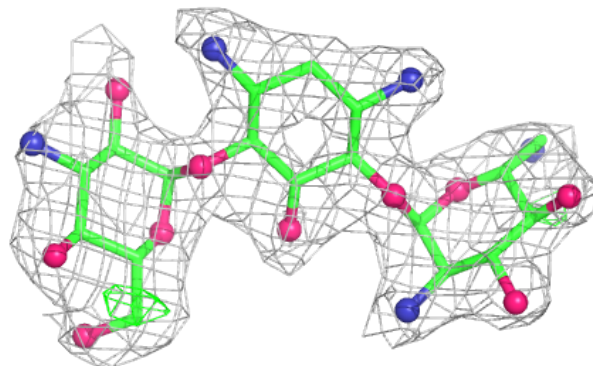


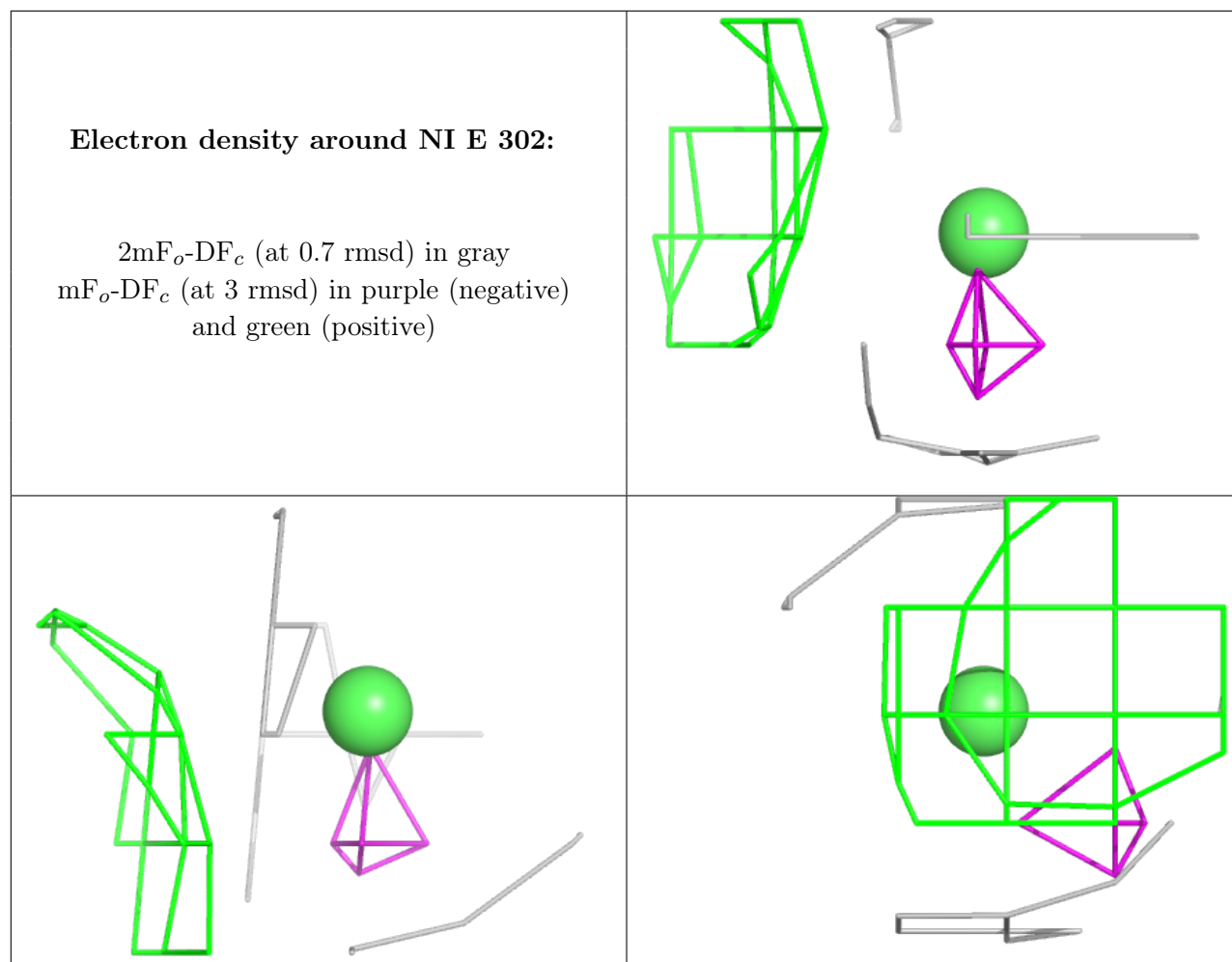
**Electron density around 9CS D 301:**

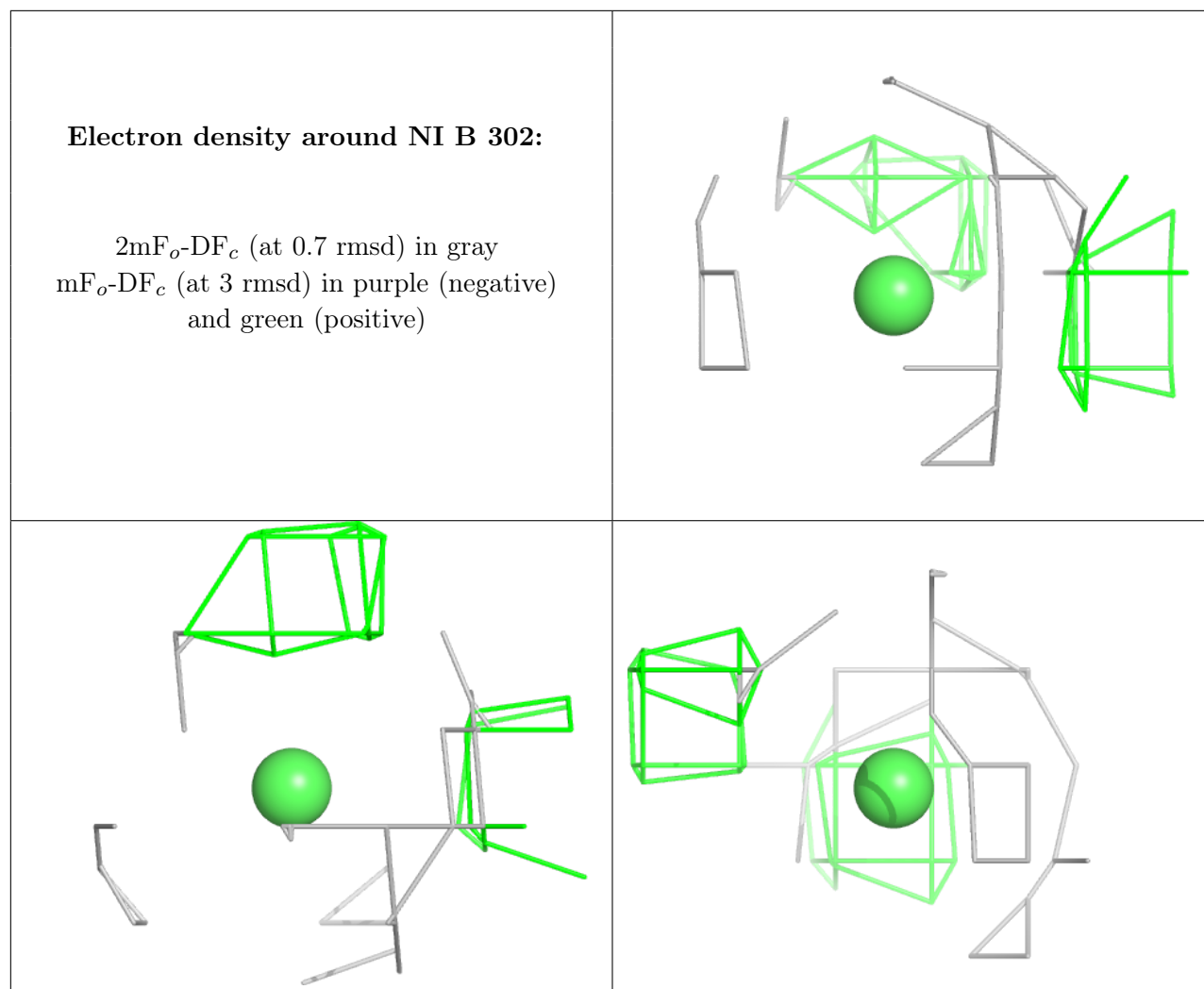
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

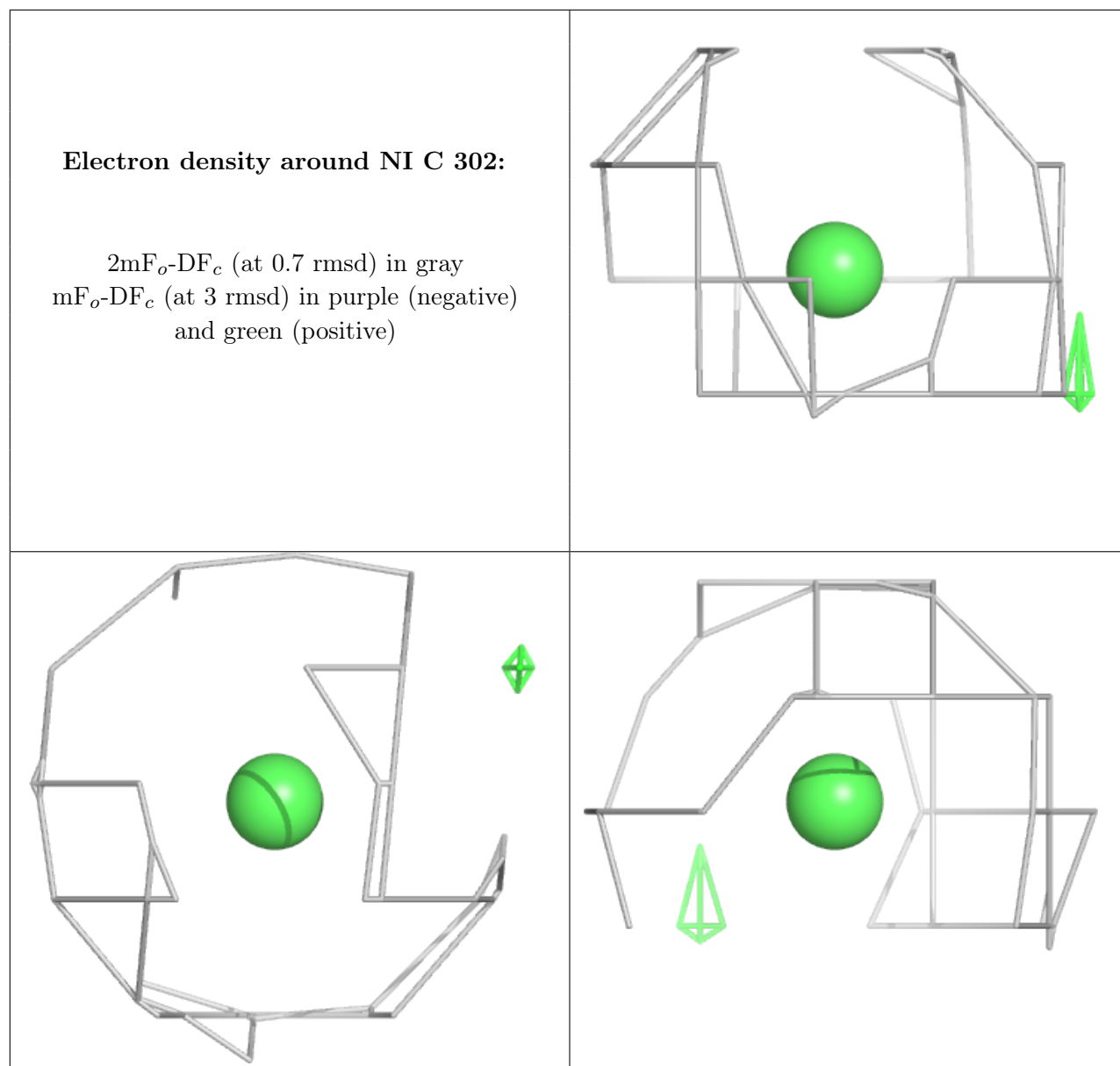
**Electron density around 9CS F 301:**

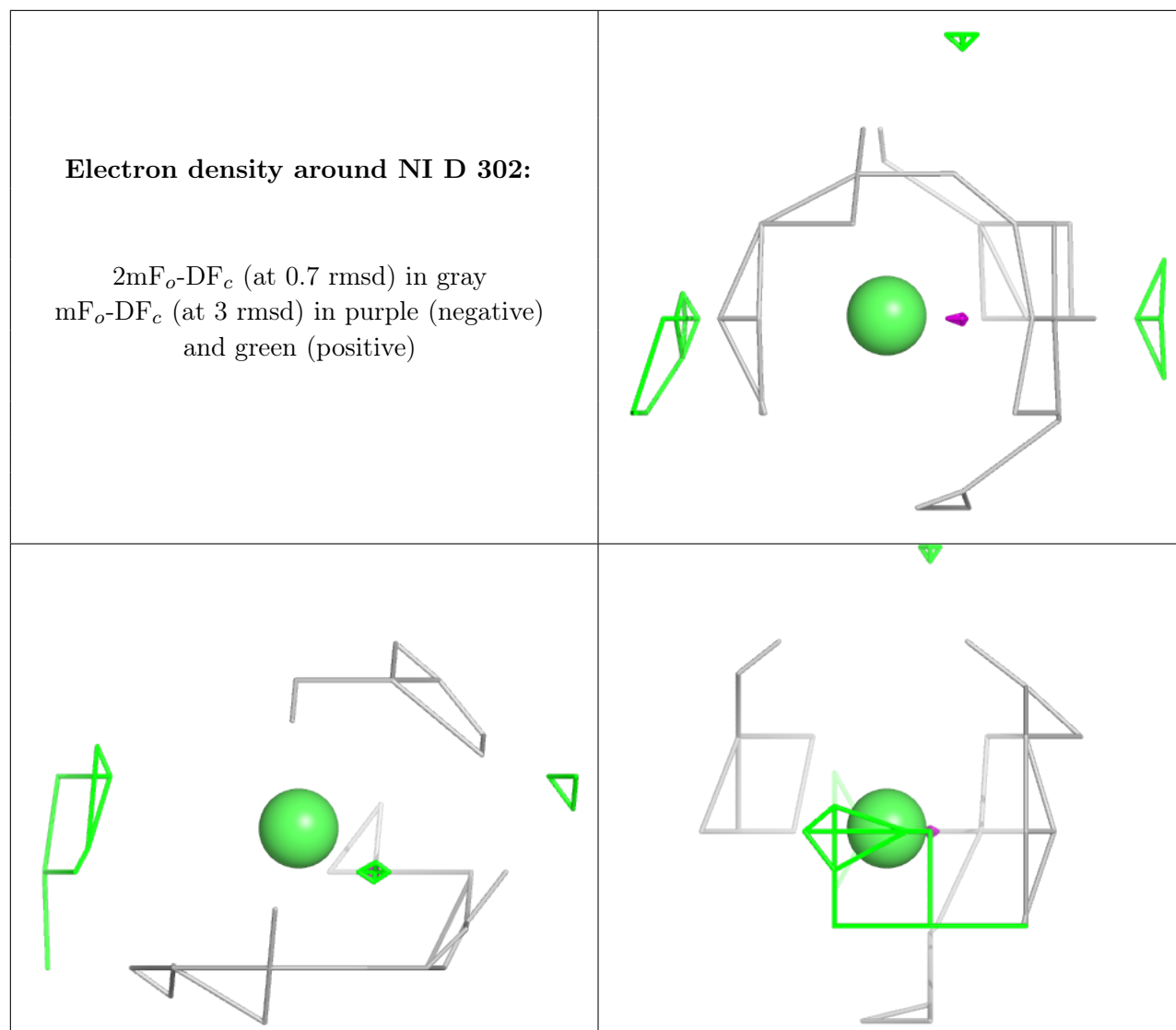
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

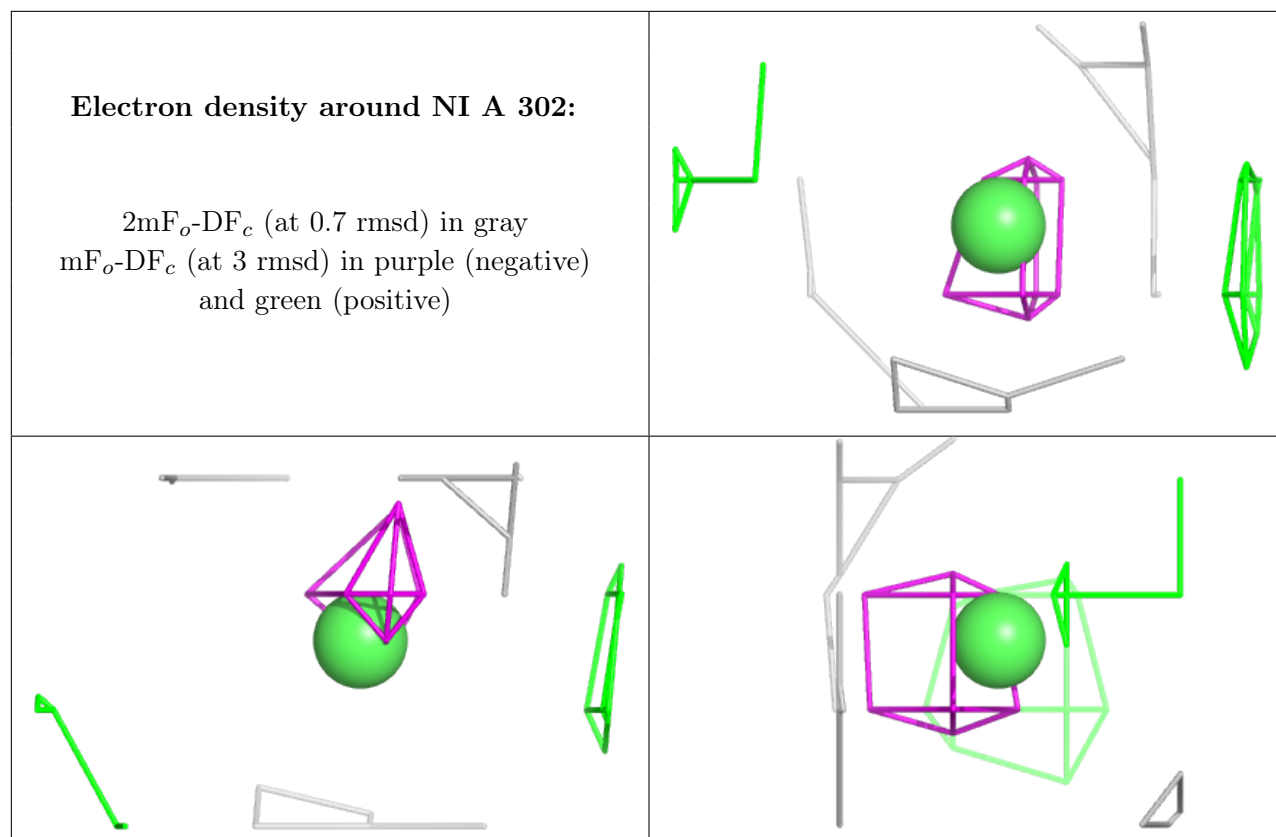


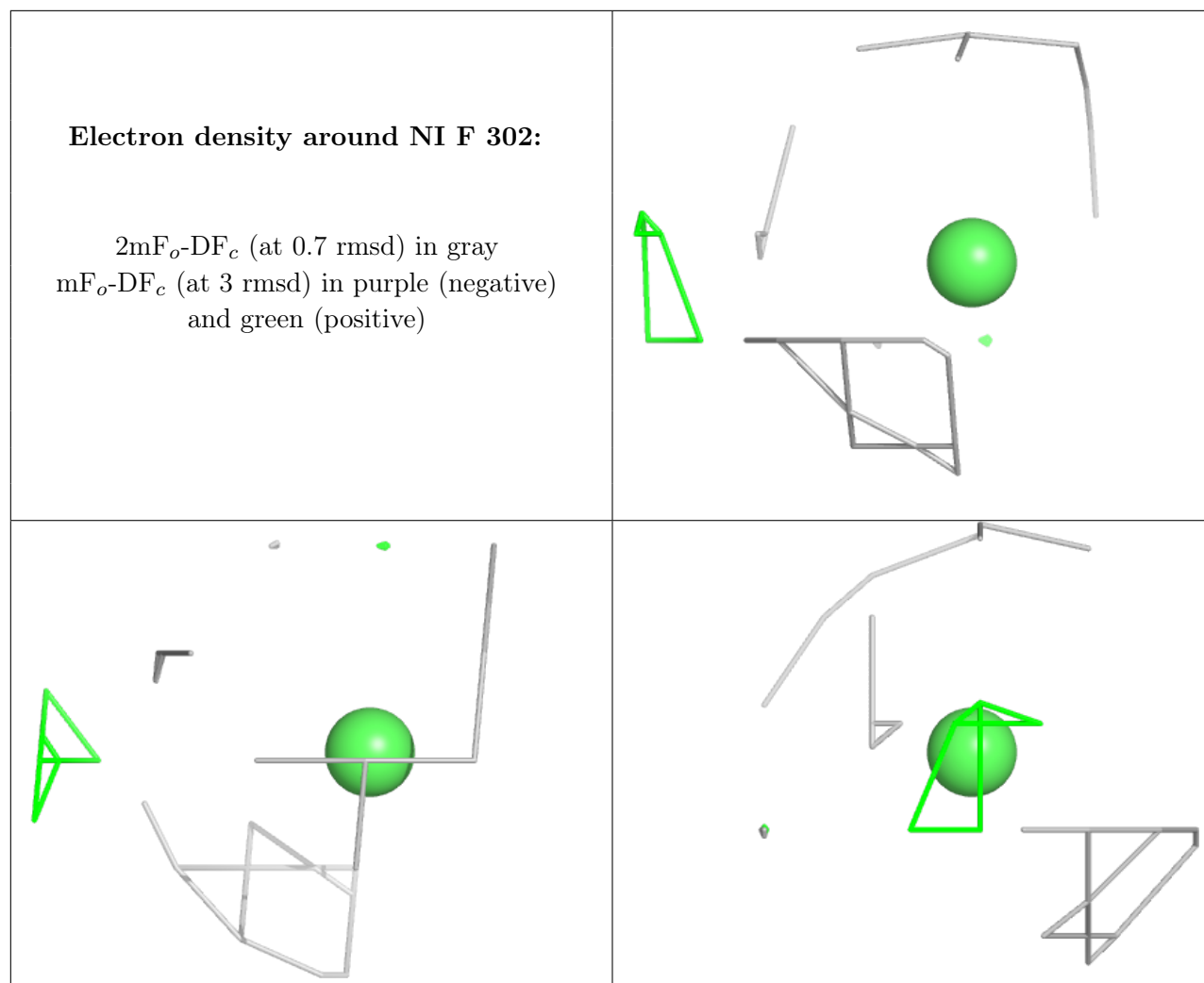












## 6.5 Other polymers [i](#)

There are no such residues in this entry.