wwPDB X-ray Structure Validation Summary Report (i)

Nov 22, 2023 – 01:07 PM JST

PDB ID	:	7WH5
Title	:	Crystal structure of human ClpP in complex with ZG180
Authors	:	Wei, B.Y.; Gan, J.H.; Yang, C.-G.
Deposited on	:	2021-12-29
Resolution	:	2.13 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:

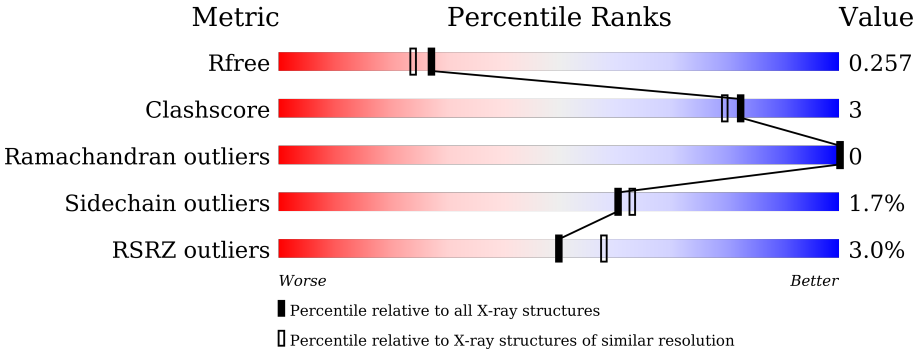
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.13 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	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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	221	2%, 73%, •, 23%
1	B	221	2%, 70%, 8%, 22%
1	C	221	3%, 71%, 6%, 23%
1	D	221	71%, 6%, 23%
1	E	221	2%, 75%, •, 22%
1	F	221	2%, 71%, 8%, 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	221	4% 74% 23%
1	H	221	2% 74% 22%
1	I	221	2% 76% 20%
1	J	221	3% 76% 23%
1	K	221	4% 74% 23%
1	L	221	2% 72% 5% 23%
1	M	221	% 74% 24%
1	N	221	2% 74% 23%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 19072 atoms, of which 0 are hydrogens and 0 are deuteriums.

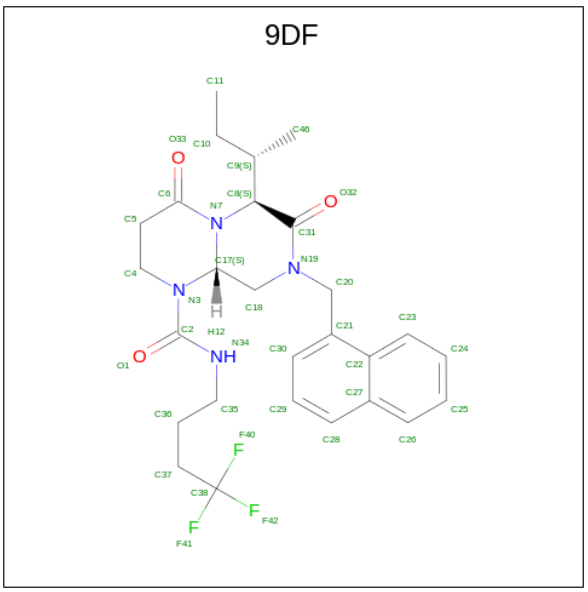
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called ATP-dependent Clp protease proteolytic subunit, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	170	1306	836	221	236	13	0	0	0
1	B	172	1293	830	215	236	12	0	0	0
1	C	170	1302	834	217	238	13	0	0	0
1	D	171	1316	842	221	240	13	0	0	0
1	E	172	1322	846	223	240	13	0	0	0
1	F	173	1301	835	215	238	13	0	0	0
1	G	171	1303	835	219	236	13	0	0	0
1	H	172	1322	848	221	240	13	0	0	0
1	I	177	1343	862	227	242	12	0	0	0
1	J	171	1300	834	220	234	12	0	0	0
1	K	171	1280	821	212	235	12	0	0	0
1	L	170	1295	829	218	235	13	0	0	0
1	M	169	1305	834	222	236	13	0	0	0
1	N	170	1311	838	221	239	13	0	0	0

- Molecule 2 is (6S,9aS)-6-[(2S)-butan-2-yl]-8-(naphthalen-1-ylmethyl)-4,7-bis(oxidanylidene)-N-[4,4,4-tris(fluoranyl)butyl]-3,6,9,9a-tetrahydro-2H-pyrazino[1,2-a]pyrimidine-1-carboxamide (three-letter code: 9DF) (formula: $C_{27}H_{33}F_3N_4O_3$) (labeled as "Ligand of Interest" by

depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	A	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	C	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	C	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	D	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	E	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	F	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	H	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	H	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	I	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	J	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	K	1	37	27	3	4	3	0	0
			Total	C	F	N	O		
2	L	1	37	27	3	4	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	N	1	37	27	3	4	3	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
3	N	1	1	1	0	0

- Molecule 4 is water.

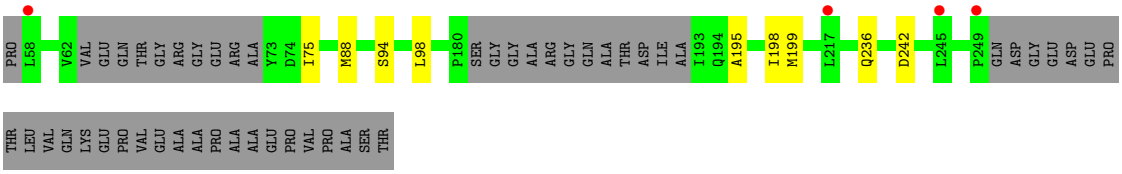
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	14	14	14	0	0
4	B	14	14	14	0	0
4	C	23	23	23	0	0
4	D	36	36	36	0	0
4	E	9	9	9	0	0
4	F	13	13	13	0	0
4	G	13	13	13	0	0
4	H	16	16	16	0	0
4	I	8	8	8	0	0
4	J	5	5	5	0	0
4	K	8	8	8	0	0
4	L	19	19	19	0	0
4	M	37	37	37	0	0
4	N	39	39	39	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

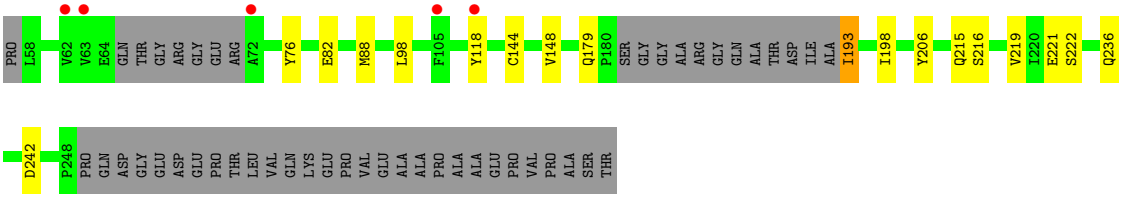
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial

Chain A: 2% 73% 23%



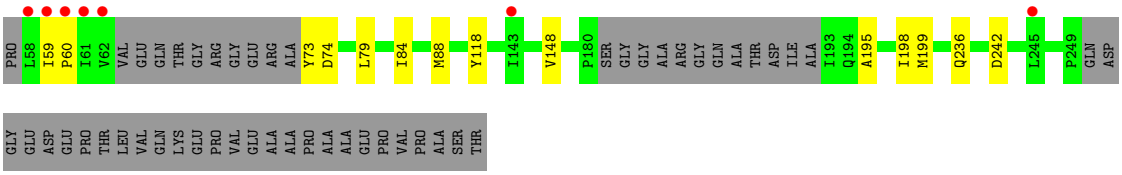
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial

Chain B: 2% 70% 8% 22%



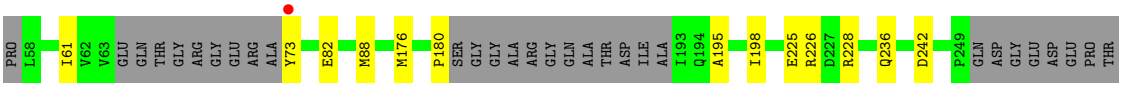
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial

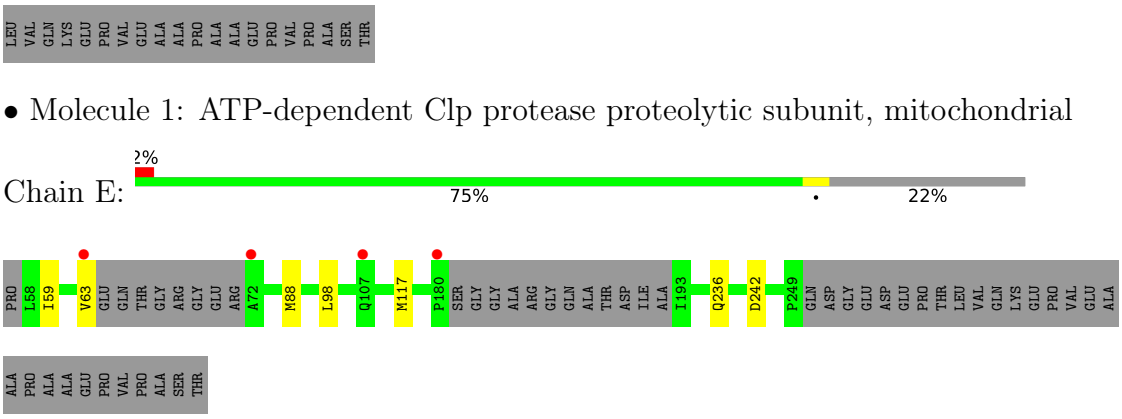
Chain C: 3% 71% 6% 23%

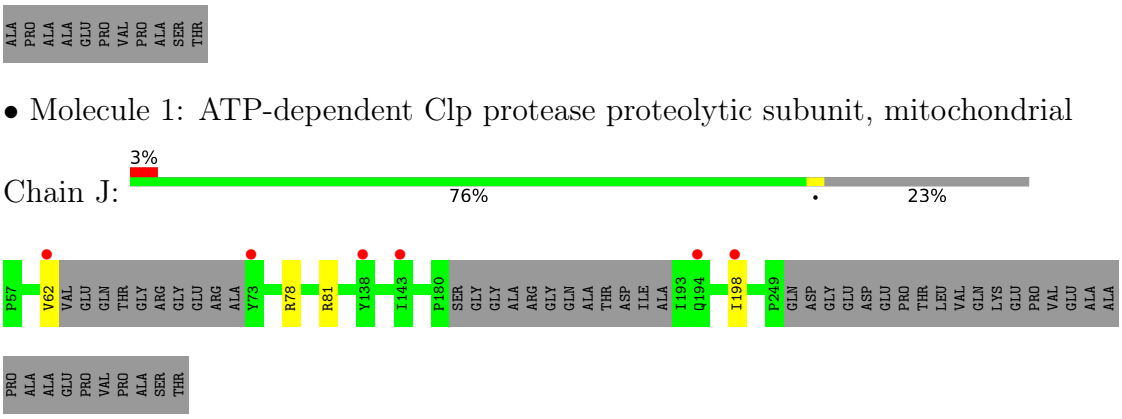


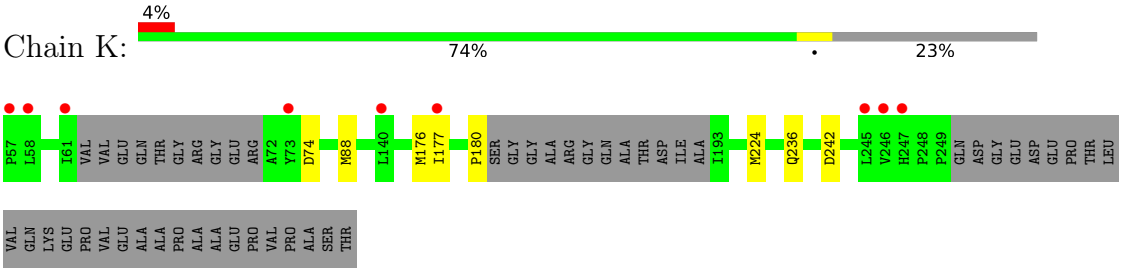
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial

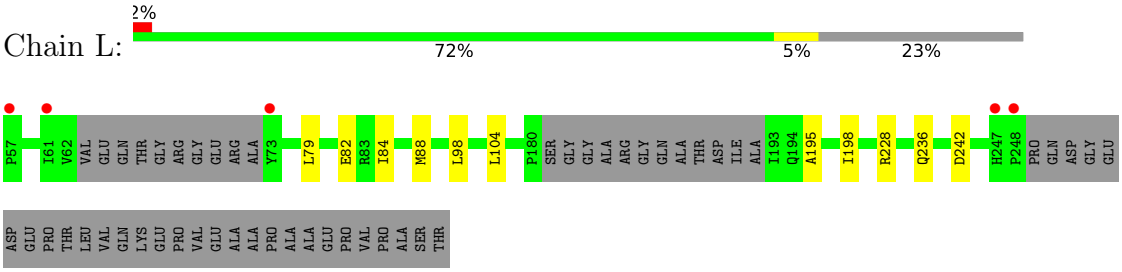
Chain D: 71% 6% 23%

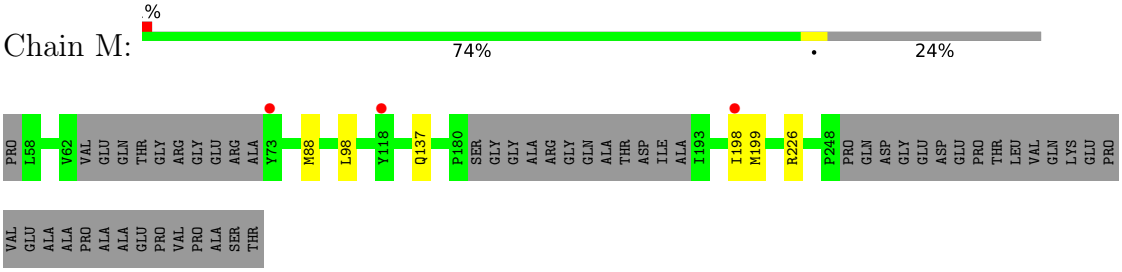


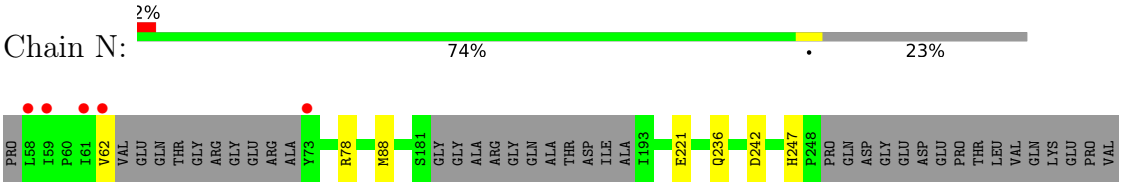

LEU VAL GLN LYS GLU PRO VAL GLU ALA ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain E:
2%
75%
22%
PRO L58 I59 V63 GLU GLN THR GLY ARG GLY GLU ARG A72 M88 L98 Q107 M117 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 Q236 D242 P249 GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN LYS GLU PRO VAL GLU ALA
ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain F:
2%
71%
8%
22%
PRO L58 V63 GLU GLN THR GLY ARG GLY GLU R71 A72 M88 L98 Q107 Y118 V148 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 Q194 A195 I198 L217 M233 Q236 D242 K243 V244 L245 V246 P249 GLN ASP GLY GLU
ASP GLU PRO THR LEU VAL GLN LYS GLU PRO VAL GLU ALA ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain G:
4%
74%
23%
PRO L58 I61 V62 V63 GLU GLN THR GLY ARG GLY GLU ARG ALA Y73 L79 I84 M88 Q107 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 L217 M230 A235 I240 P249 GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN
LYS GLU PRO VAL GLU ALA ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain H:
2%
74%
22%
PRO L58 V63 GLU GLN THR GLY ARG GLY GLU ARG A72 Y73 E82 M88 L98 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 Q194 A195 I198 K203 Q236 D242 P249 GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN LYS GLU
PRO VAL GLU ALA ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain I:
2%
76%
20%
P57 V62 V63 E64 GLN THR GLY ARG G69 E70 R71 A72 R78 L98 W146 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 E197 I198 Q236 D242 P249 GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN LYS GLU PRO VAL GLU ALA


ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR
• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain J:
3%
76%
23%
P57 V62 VAL GLU GLN THR GLY ARG GLY GLU ARG ALA Y73 R78 R81 Y138 I143 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 Q194 I198 P249 GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN LYS GLU PRO VAL GLU ALA ALA
PRO ALA ALA GLU PRO VAL PRO ALA SER THR


• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain K:
4%
74%
23%
P57 L58 I61 VAL VAL GLU GLN THR GLY ARG GLY GLU ARG A72 Y73 D74 M88 L140 M176 I177 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 M224 Q236 D242 L245 V246 H247 P248 P249 GLN ASP GLY GLU ASP GLU PRO THR LEU
VAL GLN LYS GLU PRO VAL GLU ALA ALA PRO ALA ALA ALA GLU PRO VAL PRO ALA SER THR


• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain L:
2%
72%
5%
23%
P57 I61 V62 VAL GLU GLN THR GLY ARG GLY GLU ARG ALA Y73 L79 E82 R83 I84 M88 L98 L104 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 Q194 A195 I198 R228 Q236 D242 H247 P248 PRO GLN ASP GLY GLU
ASP GLU PRO THR LEU VAL GLN LYS GLU PRO VAL GLU ALA ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR


• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain M:
%
74%
24%
PRO L58 V62 VAL GLU GLN THR GLY ARG GLY GLU ARG ALA Y73 M88 L98 Y118 Q137 P180 SER GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 I198 M199 R226 P248 PRO GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN LYS GLU PRO
VAL GLU ALA ALA PRO ALA ALA GLU PRO VAL PRO ALA SER THR


• Molecule 1: ATP-dependent Clp protease proteolytic subunit, mitochondrial
Chain N:
2%
74%
23%
PRO L58 I59 P60 I61 V62 VAL GLU GLN THR GLY ARG GLY GLU ARG ALA Y73 R78 M88 S181 GLY GLY ALA ARG GLY GLN ALA THR ASP ILE ALA I193 E221 Q236 D242 H247 P248 PRO GLN ASP GLY GLU ASP GLU PRO THR LEU VAL GLN LYS GLU PRO VAL

GLU
ALA
ALA
PRO
ALA
ALA
GLU
PRO
VAL
PRO
ALA
SER
THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	111.89Å 97.03Å 131.42Å 90.00° 93.70° 90.00°	Depositor
Resolution (Å)	54.33 – 2.13 54.33 – 2.13	Depositor EDS
% Data completeness (in resolution range)	99.4 (54.33-2.13) 99.5 (54.33-2.13)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.230 , 0.249 0.242 , 0.257	Depositor DCC
R_{free} test set	7770 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	42.3	Xtriage
Anisotropy	0.202	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	19072	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.98% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

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

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 9DF, MG

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	0.54	0/1331	0.71	0/1805
1	B	0.59	0/1317	0.73	0/1791
1	C	0.57	0/1327	0.73	0/1799
1	D	0.64	0/1341	0.76	0/1818
1	E	0.56	0/1347	0.73	0/1827
1	F	0.52	0/1326	0.71	0/1803
1	G	0.54	0/1328	0.72	0/1801
1	H	0.54	0/1347	0.74	0/1827
1	I	0.57	0/1369	0.71	0/1858
1	J	0.51	0/1326	0.68	0/1800
1	K	0.54	0/1306	0.69	0/1776
1	L	0.55	0/1319	0.72	0/1788
1	M	0.58	0/1329	0.75	0/1800
1	N	0.65	0/1335	0.76	0/1807
All	All	0.57	0/18648	0.72	0/25300

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	1306	0	1316	6	0
1	B	1293	0	1278	14	0
1	C	1302	0	1309	12	0
1	D	1316	0	1327	10	0
1	E	1322	0	1334	3	0
1	F	1301	0	1287	17	0
1	G	1303	0	1309	7	0
1	H	1322	0	1341	5	0
1	I	1343	0	1353	7	0
1	J	1300	0	1301	6	0
1	K	1280	0	1255	7	0
1	L	1295	0	1303	6	0
1	M	1305	0	1322	5	0
1	N	1311	0	1330	4	0
2	A	74	0	0	1	0
2	C	74	0	0	1	0
2	D	37	0	0	0	0
2	E	37	0	0	0	0
2	F	37	0	0	0	0
2	H	74	0	0	1	0
2	I	37	0	0	0	0
2	J	37	0	0	0	0
2	K	37	0	0	1	0
2	L	37	0	0	1	0
2	N	37	0	0	0	0
3	N	1	0	0	0	0
4	A	14	0	0	0	0
4	B	14	0	0	2	0
4	C	23	0	0	0	0
4	D	36	0	0	1	0
4	E	9	0	0	0	0
4	F	13	0	0	0	0
4	G	13	0	0	2	0
4	H	16	0	0	0	0
4	I	8	0	0	0	0
4	J	5	0	0	0	0
4	K	8	0	0	0	0
4	L	19	0	0	0	0
4	M	37	0	0	0	0
4	N	39	0	0	0	0
All	All	19072	0	18365	94	0

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

The worst 5 of 94 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:F:198:ILE:HG21	1:J:198:ILE:HD11	1.29	1.10
1:F:198:ILE:HG21	1:J:198:ILE:CD1	1.89	1.02
1:B:118:TYR:CE2	1:B:148:VAL:HG21	2.14	0.82
1:A:94:SER:O	1:A:98:LEU:HD23	1.85	0.76
1:F:198:ILE:CG2	1:J:198:ILE:HD11	2.16	0.72

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	164/221 (74%)	162 (99%)	2 (1%)	0	100	100
1	B	166/221 (75%)	164 (99%)	2 (1%)	0	100	100
1	C	164/221 (74%)	162 (99%)	2 (1%)	0	100	100
1	D	165/221 (75%)	162 (98%)	3 (2%)	0	100	100
1	E	166/221 (75%)	164 (99%)	2 (1%)	0	100	100
1	F	167/221 (76%)	165 (99%)	2 (1%)	0	100	100
1	G	165/221 (75%)	163 (99%)	2 (1%)	0	100	100
1	H	166/221 (75%)	164 (99%)	2 (1%)	0	100	100
1	I	171/221 (77%)	169 (99%)	2 (1%)	0	100	100
1	J	165/221 (75%)	163 (99%)	2 (1%)	0	100	100
1	K	165/221 (75%)	163 (99%)	2 (1%)	0	100	100
1	L	164/221 (74%)	162 (99%)	2 (1%)	0	100	100
1	M	163/221 (74%)	160 (98%)	3 (2%)	0	100	100
1	N	164/221 (74%)	162 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2315/3094 (75%)	2285 (99%)	30 (1%)	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	142/185 (77%)	139 (98%)	3 (2%)	53	54
1	B	136/185 (74%)	131 (96%)	5 (4%)	34	31
1	C	142/185 (77%)	141 (99%)	1 (1%)	84	87
1	D	144/185 (78%)	142 (99%)	2 (1%)	67	70
1	E	144/185 (78%)	141 (98%)	3 (2%)	53	54
1	F	138/185 (75%)	134 (97%)	4 (3%)	42	40
1	G	141/185 (76%)	138 (98%)	3 (2%)	53	54
1	H	145/185 (78%)	142 (98%)	3 (2%)	53	54
1	I	144/185 (78%)	142 (99%)	2 (1%)	67	70
1	J	139/185 (75%)	139 (100%)	0	100	100
1	K	134/185 (72%)	133 (99%)	1 (1%)	84	87
1	L	140/185 (76%)	137 (98%)	3 (2%)	53	54
1	M	143/185 (77%)	141 (99%)	2 (1%)	67	70
1	N	145/185 (78%)	143 (99%)	2 (1%)	67	70
All	All	1977/2590 (76%)	1943 (98%)	34 (2%)	60	63

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	98	LEU
1	L	228	ARG
1	N	88	MET
1	E	98	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	88	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	N	247	HIS
1	K	247	HIS
1	G	207	ASN
1	F	107	GLN
1	I	247	HIS

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 15 ligands modelled in this entry, 1 is monoatomic - leaving 14 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	9DF	H	302	-	38,40,40	0.82	0	50,58,58	1.14	4 (8%)
2	9DF	A	302	-	38,40,40	0.97	2 (5%)	50,58,58	1.23	7 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	9DF	E	301	-	38,40,40	0.82	1 (2%)	50,58,58	1.31	6 (12%)
2	9DF	C	302	-	38,40,40	0.93	2 (5%)	50,58,58	1.18	4 (8%)
2	9DF	N	301	-	38,40,40	0.80	0	50,58,58	1.51	8 (16%)
2	9DF	A	301	-	38,40,40	0.93	3 (7%)	50,58,58	1.12	5 (10%)
2	9DF	J	301	-	38,40,40	0.77	0	50,58,58	1.23	5 (10%)
2	9DF	C	301	-	38,40,40	0.76	1 (2%)	50,58,58	1.25	4 (8%)
2	9DF	H	301	-	38,40,40	0.91	3 (7%)	50,58,58	1.30	8 (16%)
2	9DF	L	301	-	38,40,40	0.75	0	50,58,58	1.19	2 (4%)
2	9DF	K	301	-	38,40,40	0.85	1 (2%)	50,58,58	1.32	9 (18%)
2	9DF	I	301	-	38,40,40	0.77	0	50,58,58	1.21	7 (14%)
2	9DF	F	301	-	38,40,40	0.93	2 (5%)	50,58,58	1.53	7 (14%)
2	9DF	D	301	-	38,40,40	0.83	0	50,58,58	1.44	7 (14%)

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	9DF	H	302	-	-	6/22/55/55	0/4/4/4
2	9DF	A	302	-	-	5/22/55/55	0/4/4/4
2	9DF	E	301	-	-	6/22/55/55	0/4/4/4
2	9DF	C	302	-	-	3/22/55/55	0/4/4/4
2	9DF	N	301	-	-	0/22/55/55	0/4/4/4
2	9DF	A	301	-	-	7/22/55/55	0/4/4/4
2	9DF	J	301	-	-	6/22/55/55	0/4/4/4
2	9DF	C	301	-	-	5/22/55/55	0/4/4/4
2	9DF	H	301	-	-	7/22/55/55	0/4/4/4
2	9DF	L	301	-	-	2/22/55/55	0/4/4/4
2	9DF	K	301	-	-	5/22/55/55	0/4/4/4
2	9DF	I	301	-	-	8/22/55/55	0/4/4/4
2	9DF	F	301	-	-	5/22/55/55	0/4/4/4
2	9DF	D	301	-	-	5/22/55/55	0/4/4/4

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	301	9DF	C18-N19	3.28	1.50	1.46
2	C	302	9DF	C17-N3	3.00	1.50	1.46
2	A	302	9DF	C17-N3	2.90	1.50	1.46
2	A	301	9DF	C18-N19	2.73	1.50	1.46
2	E	301	9DF	C17-N3	2.66	1.49	1.46

The worst 5 of 83 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	301	9DF	C37-C36-C35	-6.14	101.80	112.46
2	N	301	9DF	C37-C36-C35	-5.79	102.40	112.46
2	D	301	9DF	C35-N34-C2	4.18	124.73	120.84
2	L	301	9DF	C10-C9-C8	3.87	118.72	111.48
2	E	301	9DF	C10-C9-C8	3.86	118.70	111.48

There are no chirality outliers.

5 of 70 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	301	9DF	C35-C36-C37-C38
2	E	301	9DF	C36-C37-C38-F42
2	H	302	9DF	O1-C2-N3-C17
2	I	301	9DF	C36-C37-C38-F41
2	I	301	9DF	C11-C10-C9-C46

There are no ring outliers.

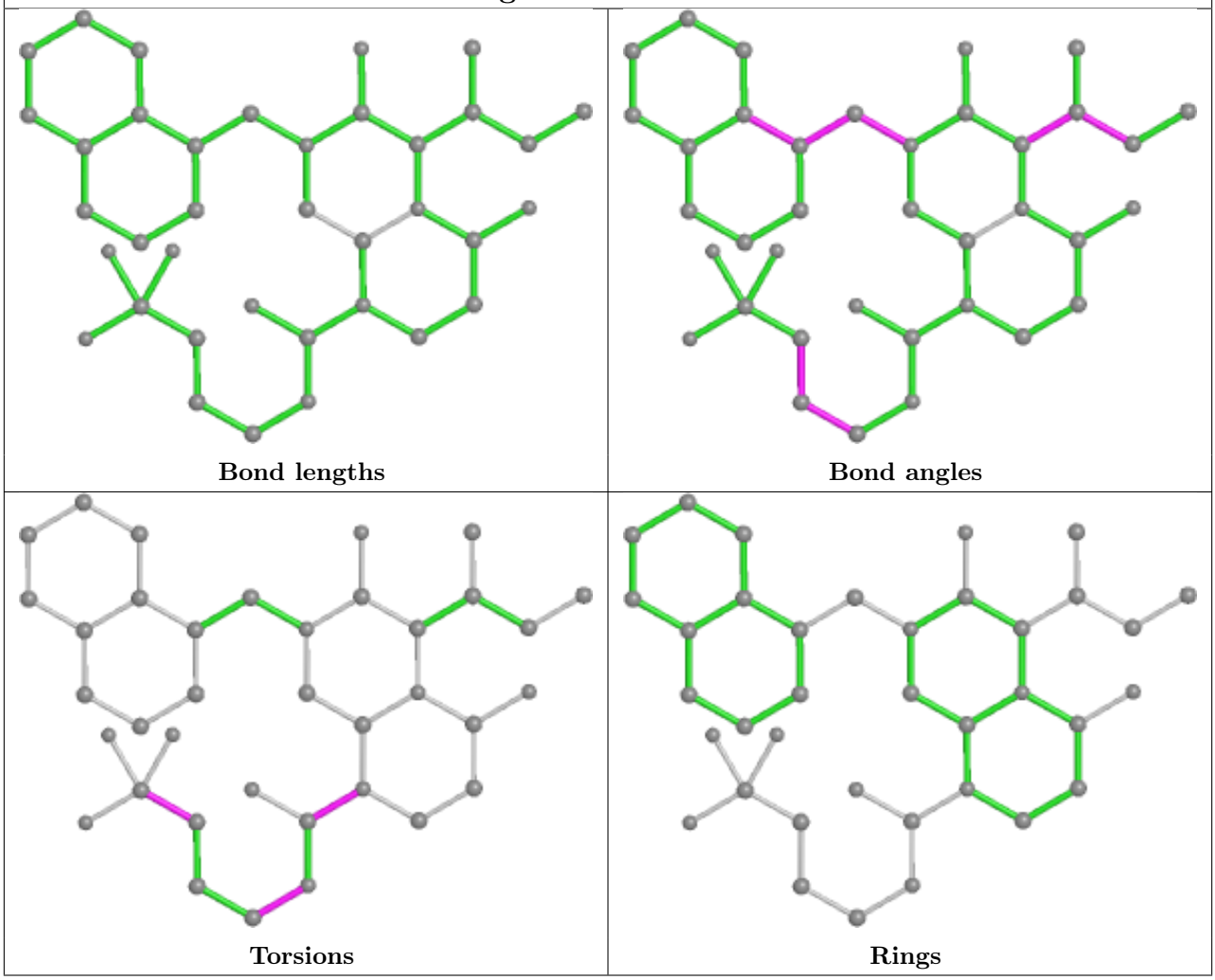
5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	302	9DF	1	0
2	C	301	9DF	1	0
2	H	301	9DF	1	0
2	L	301	9DF	1	0
2	K	301	9DF	1	0

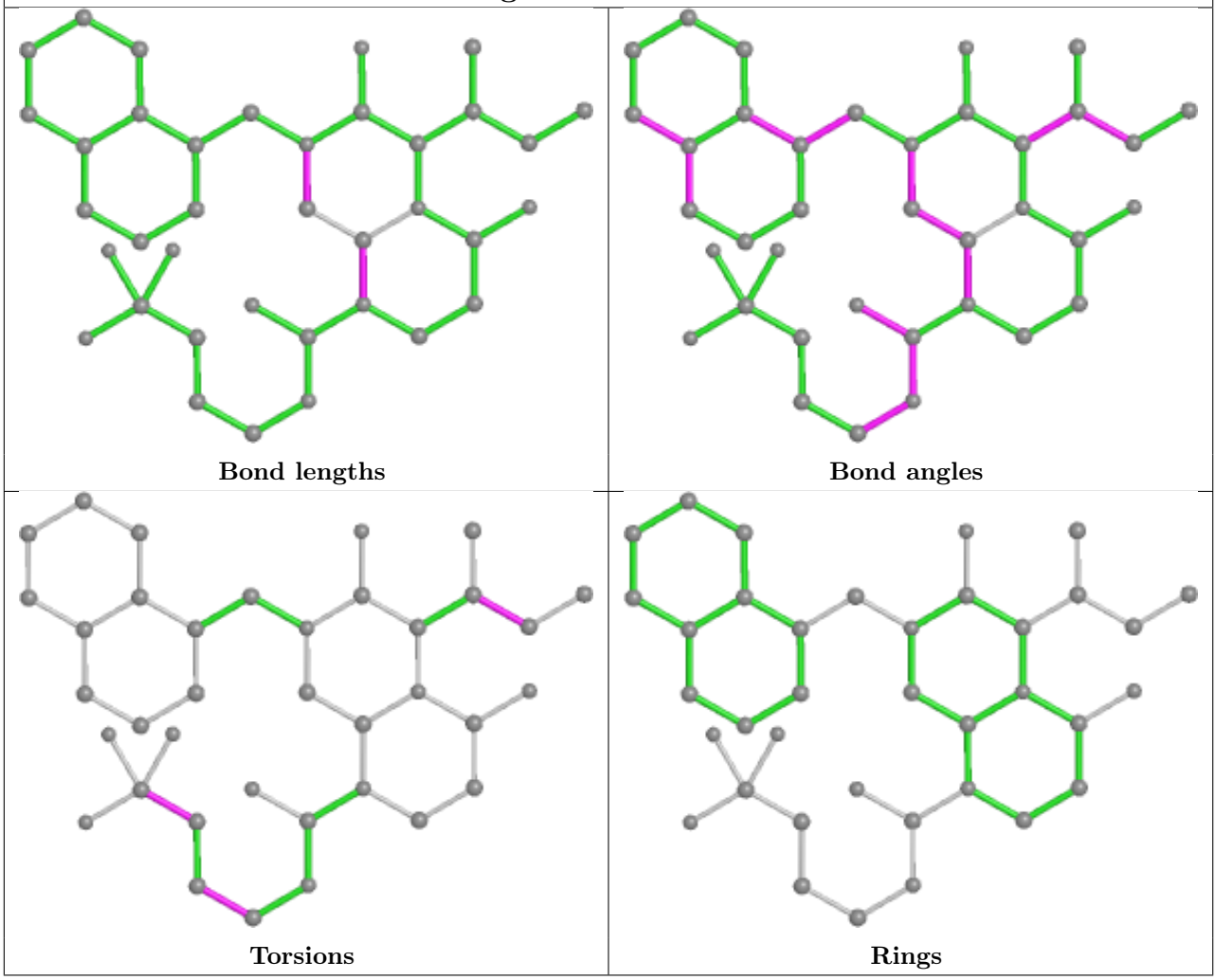
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

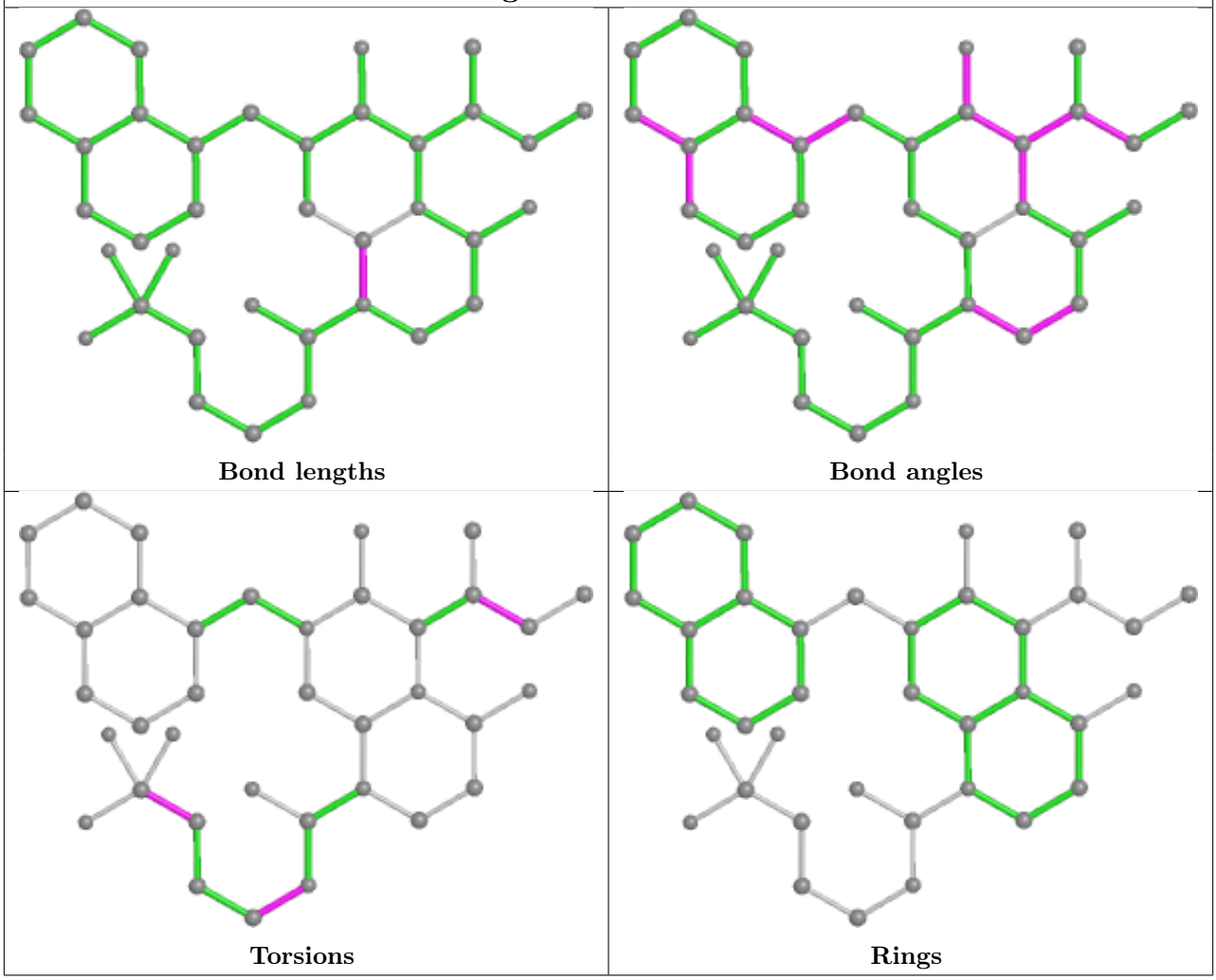
Ligand 9DF H 302



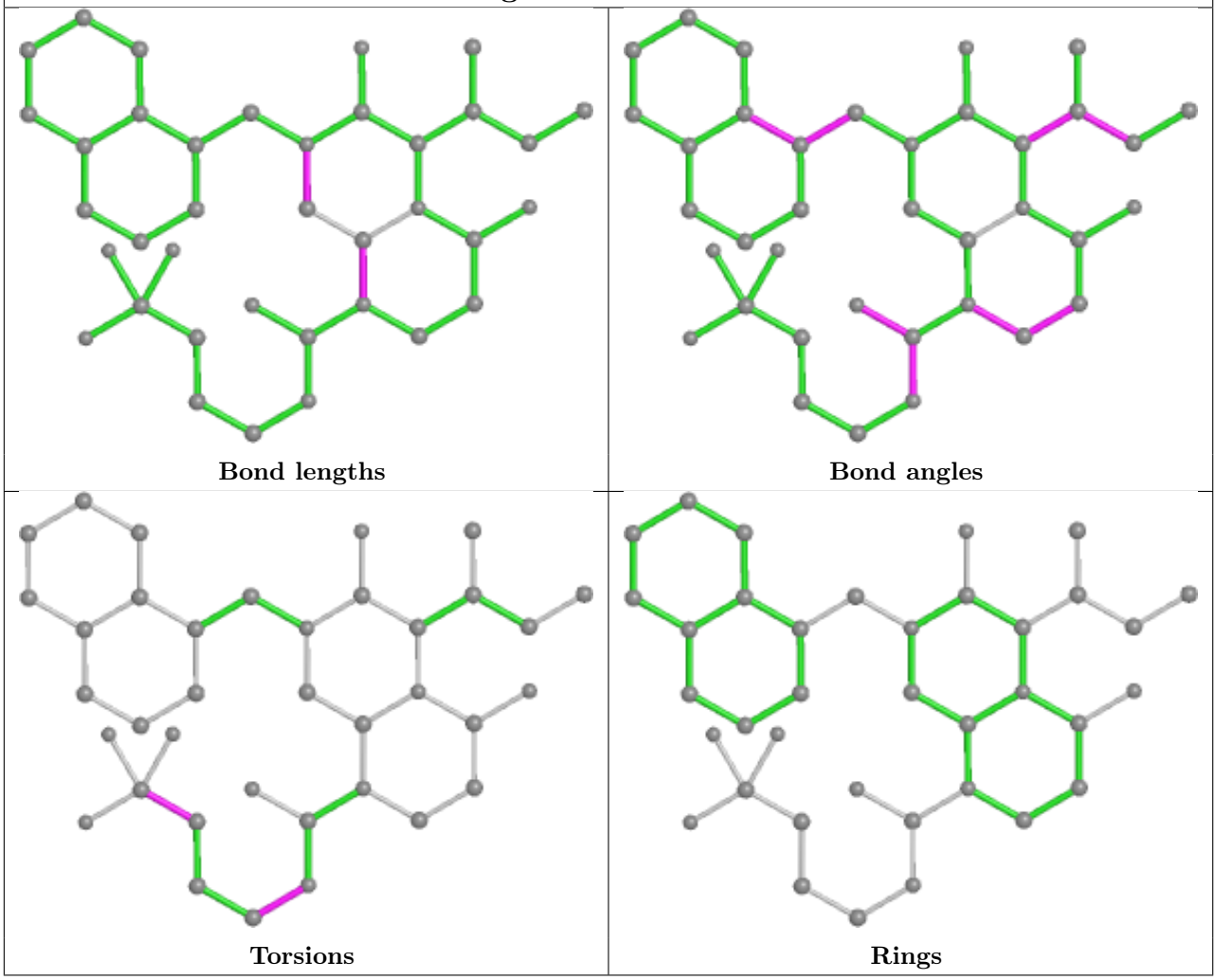
Ligand 9DF A 302



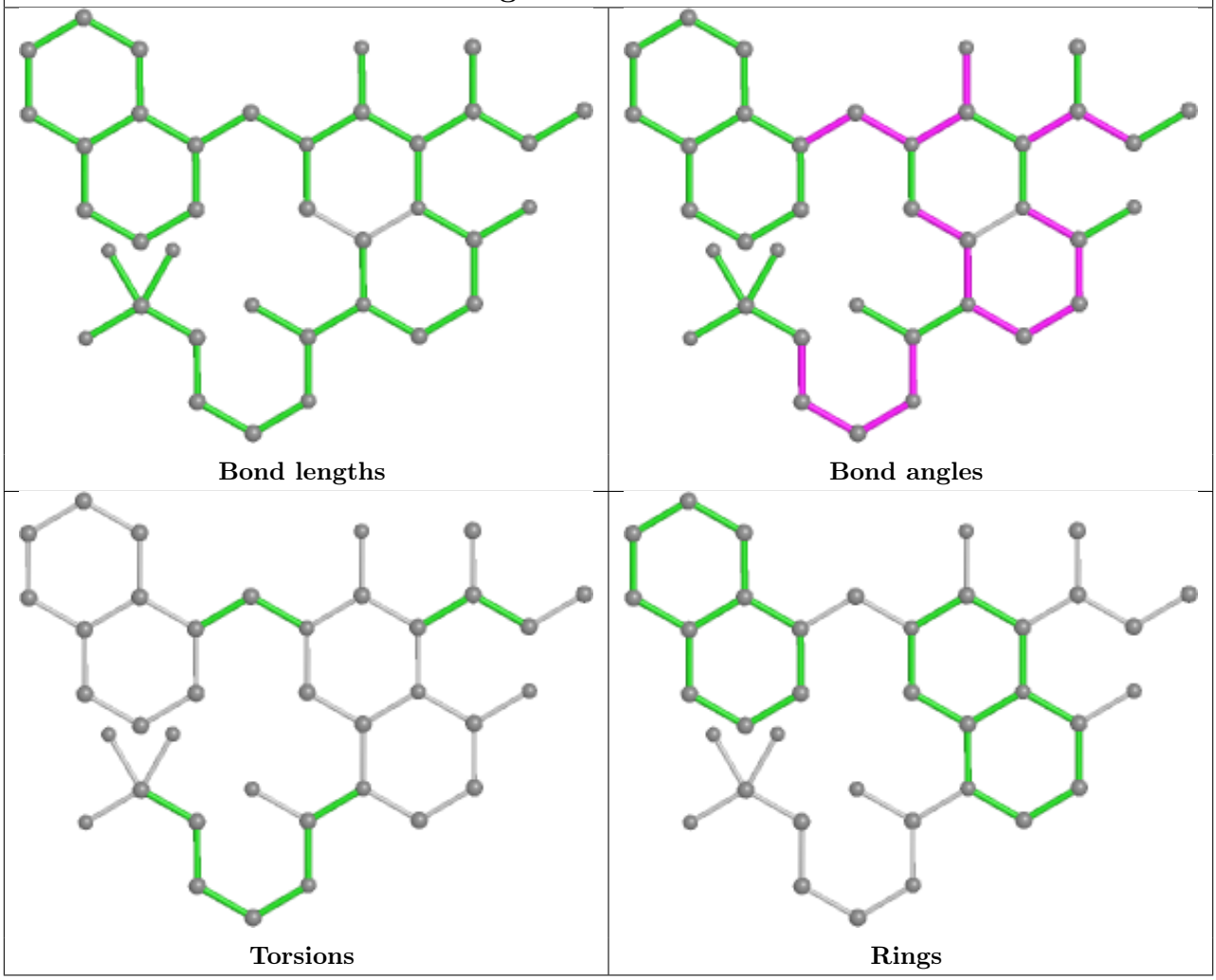
Ligand 9DF E 301



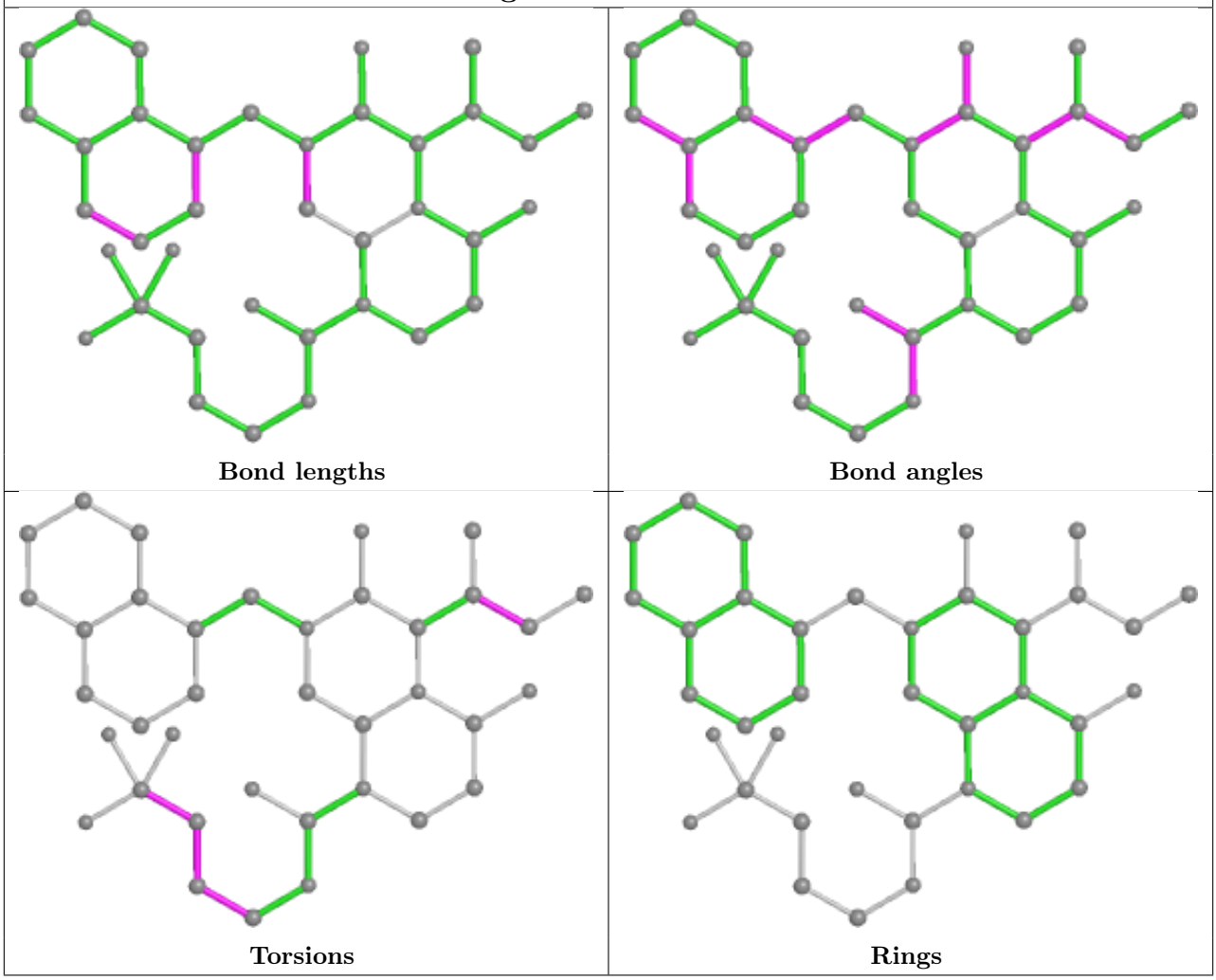
Ligand 9DF C 302



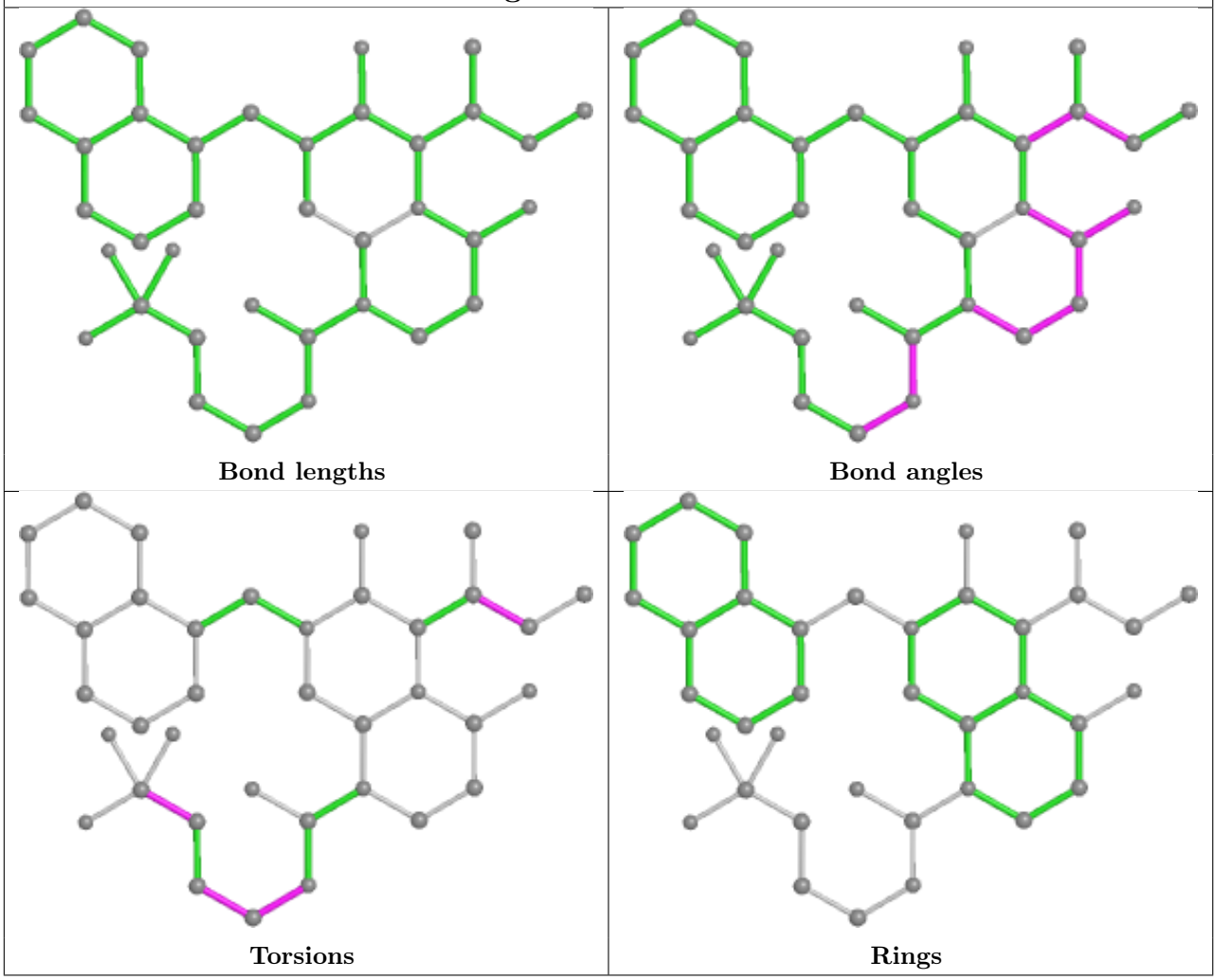
Ligand 9DF N 301



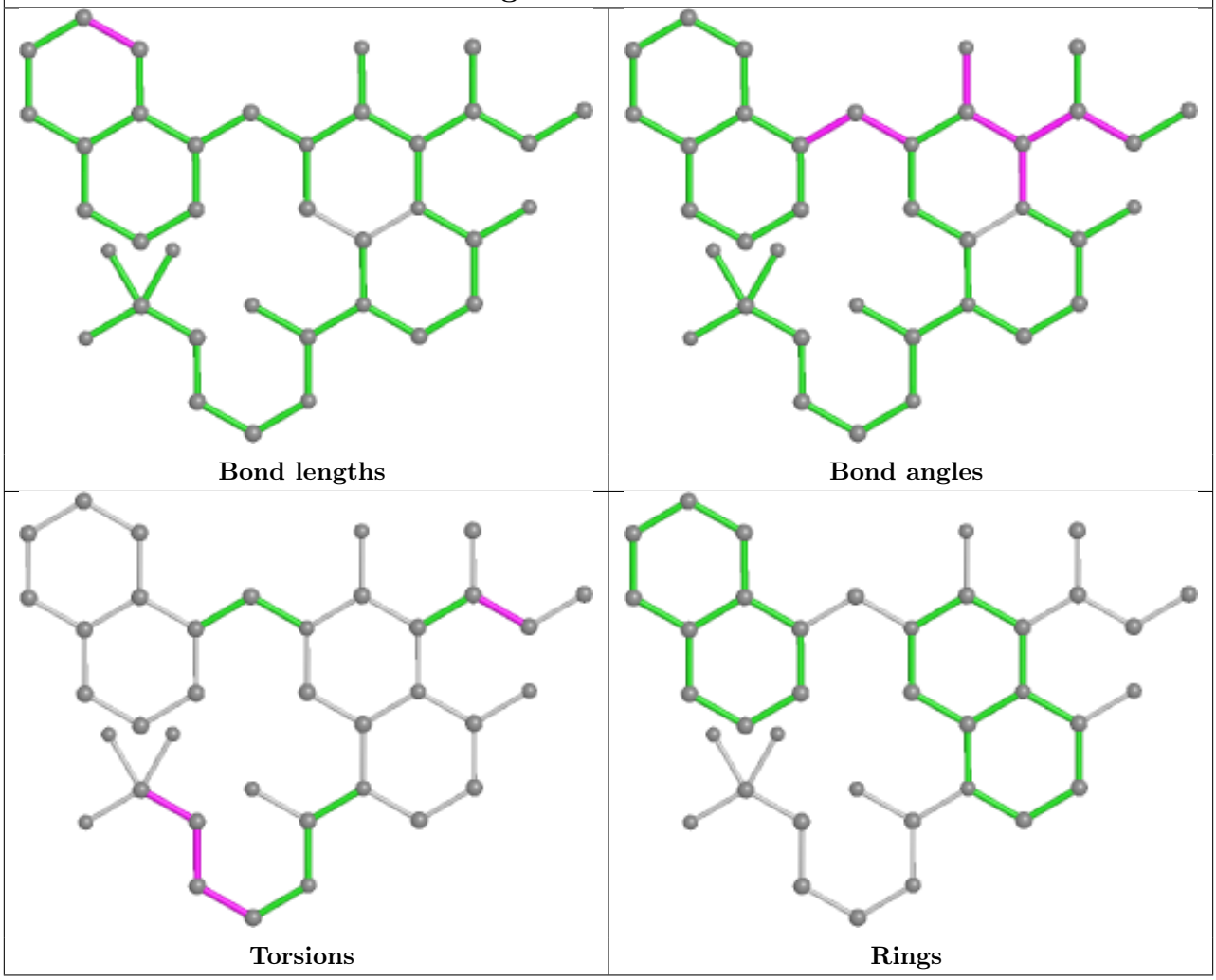
Ligand 9DF A 301



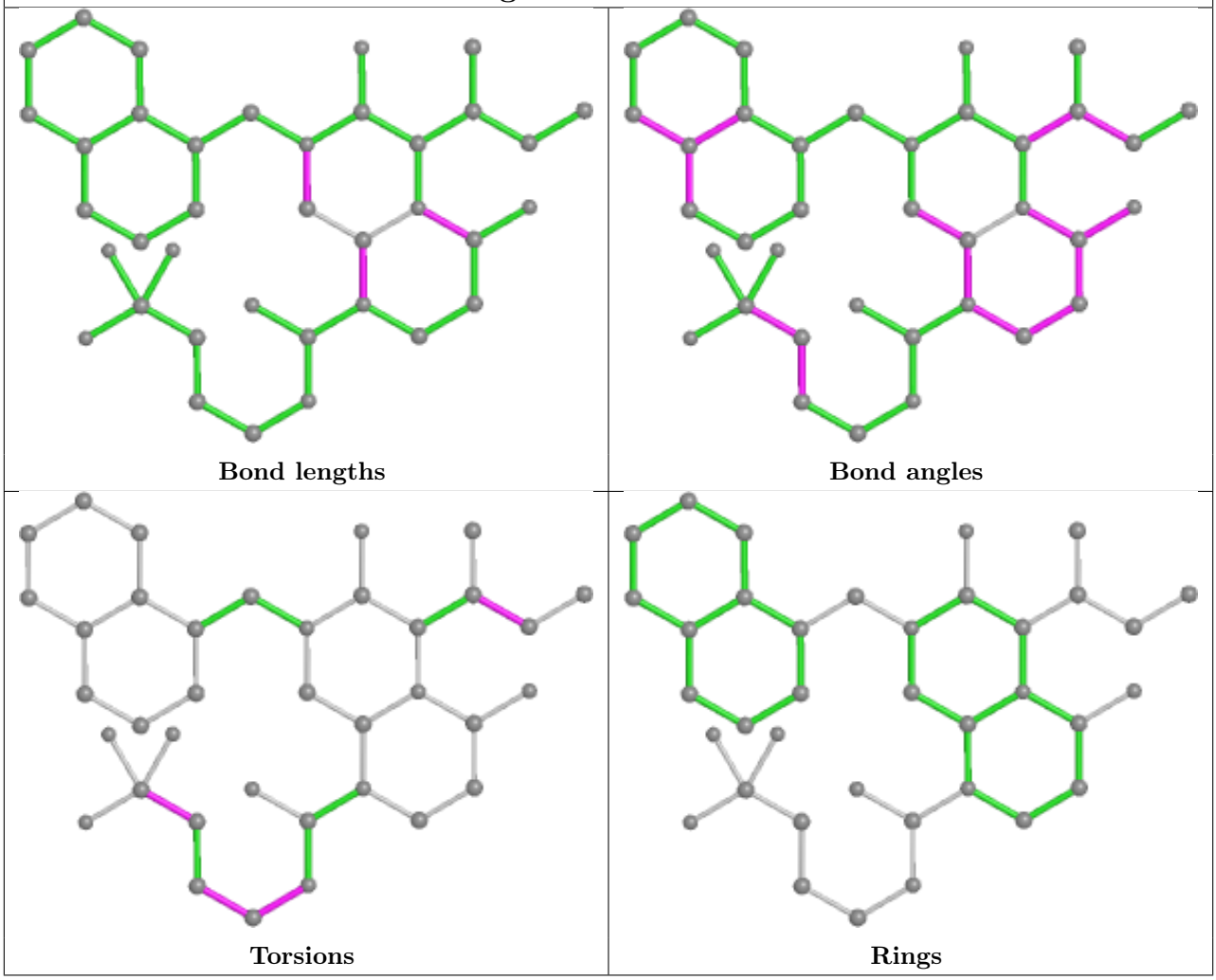
Ligand 9DF J 301



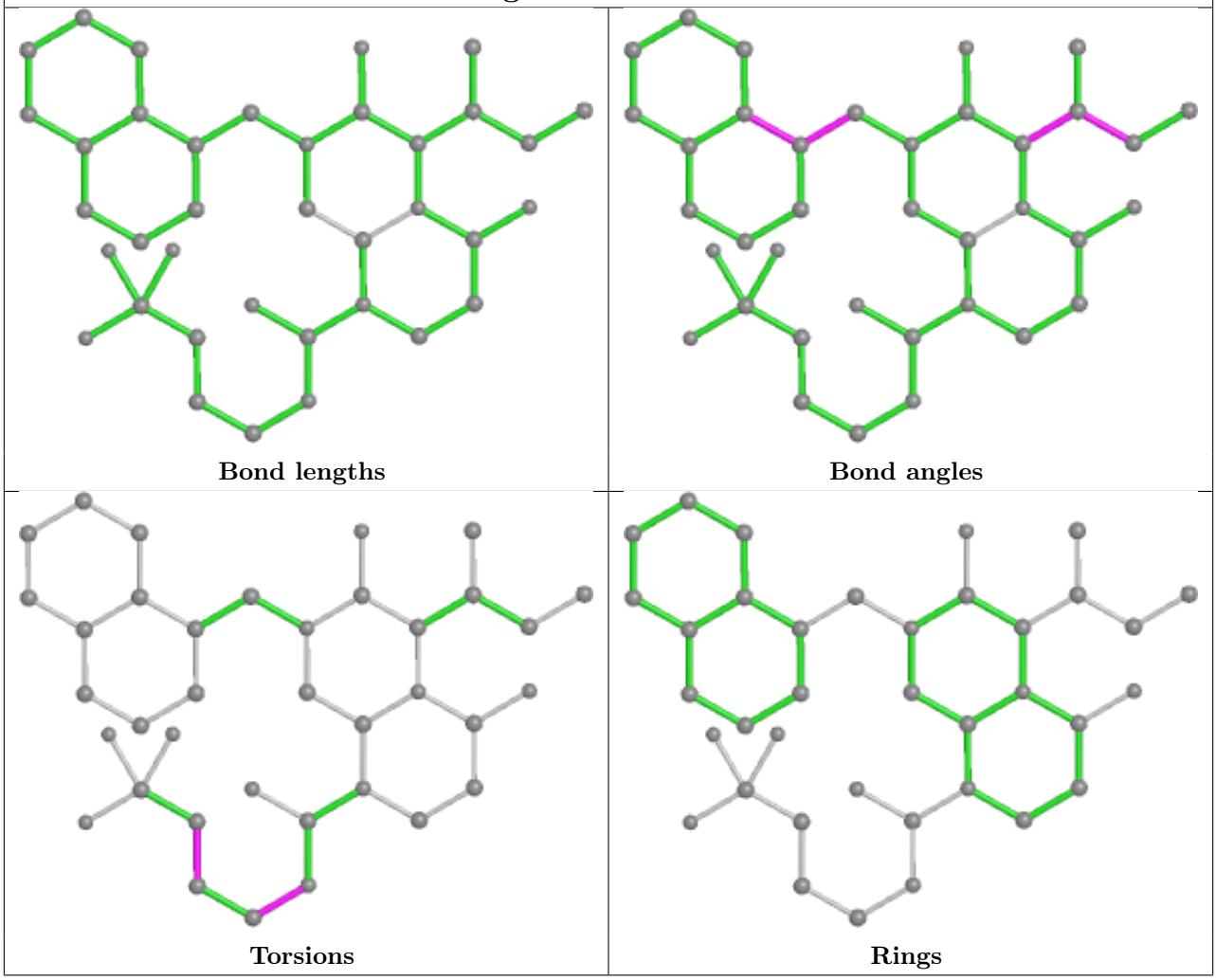
Ligand 9DF C 301



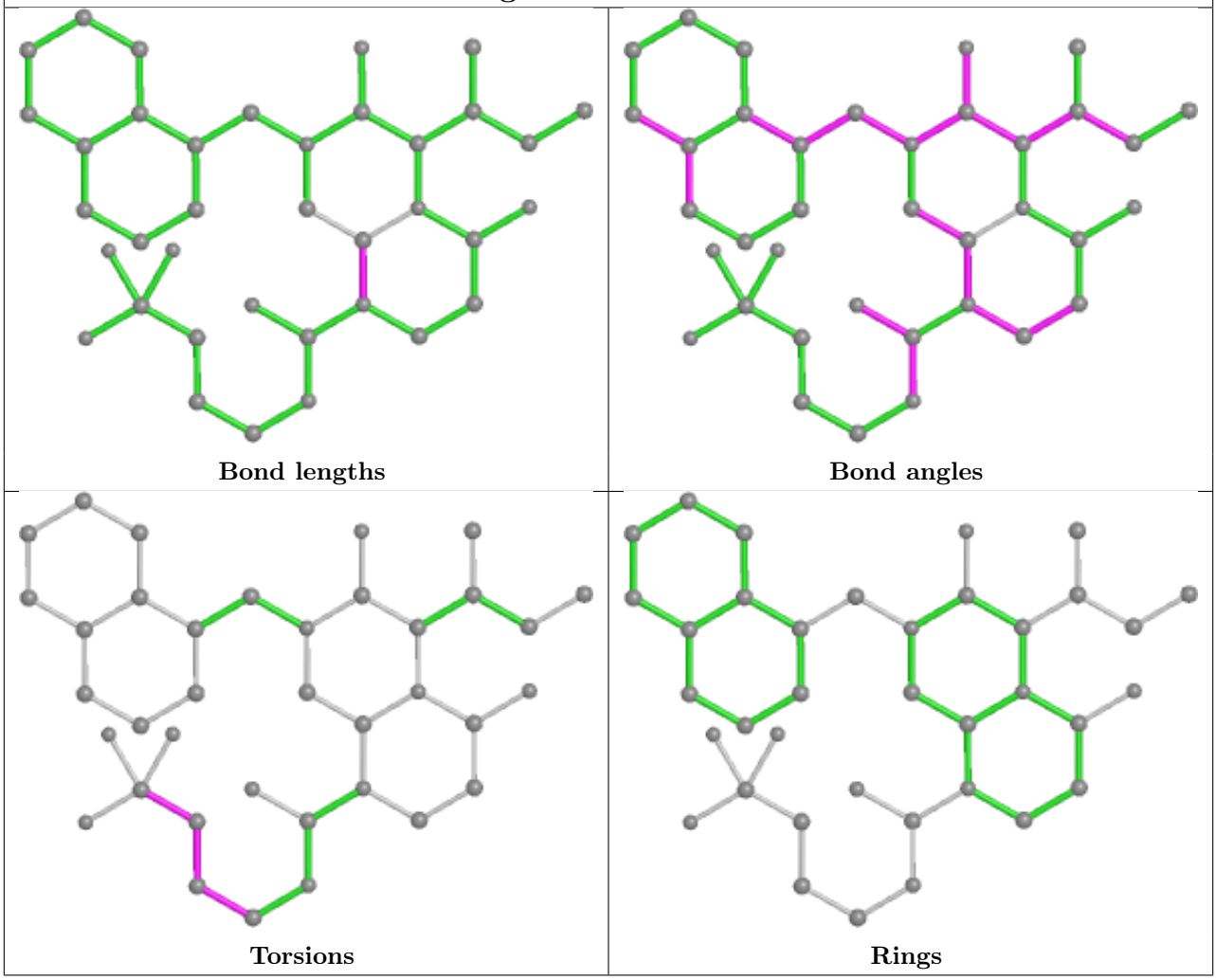
Ligand 9DF H 301



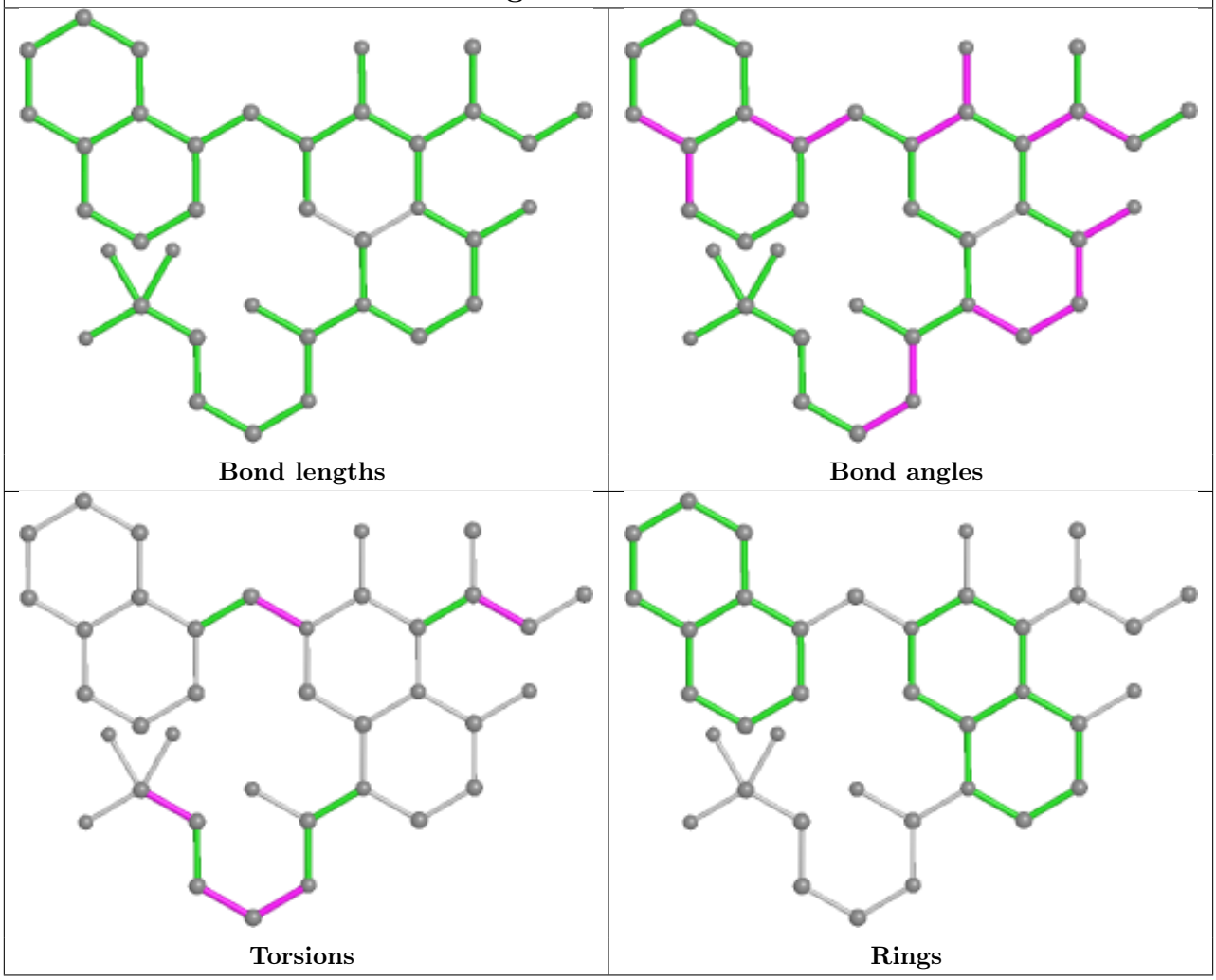
Ligand 9DF L 301



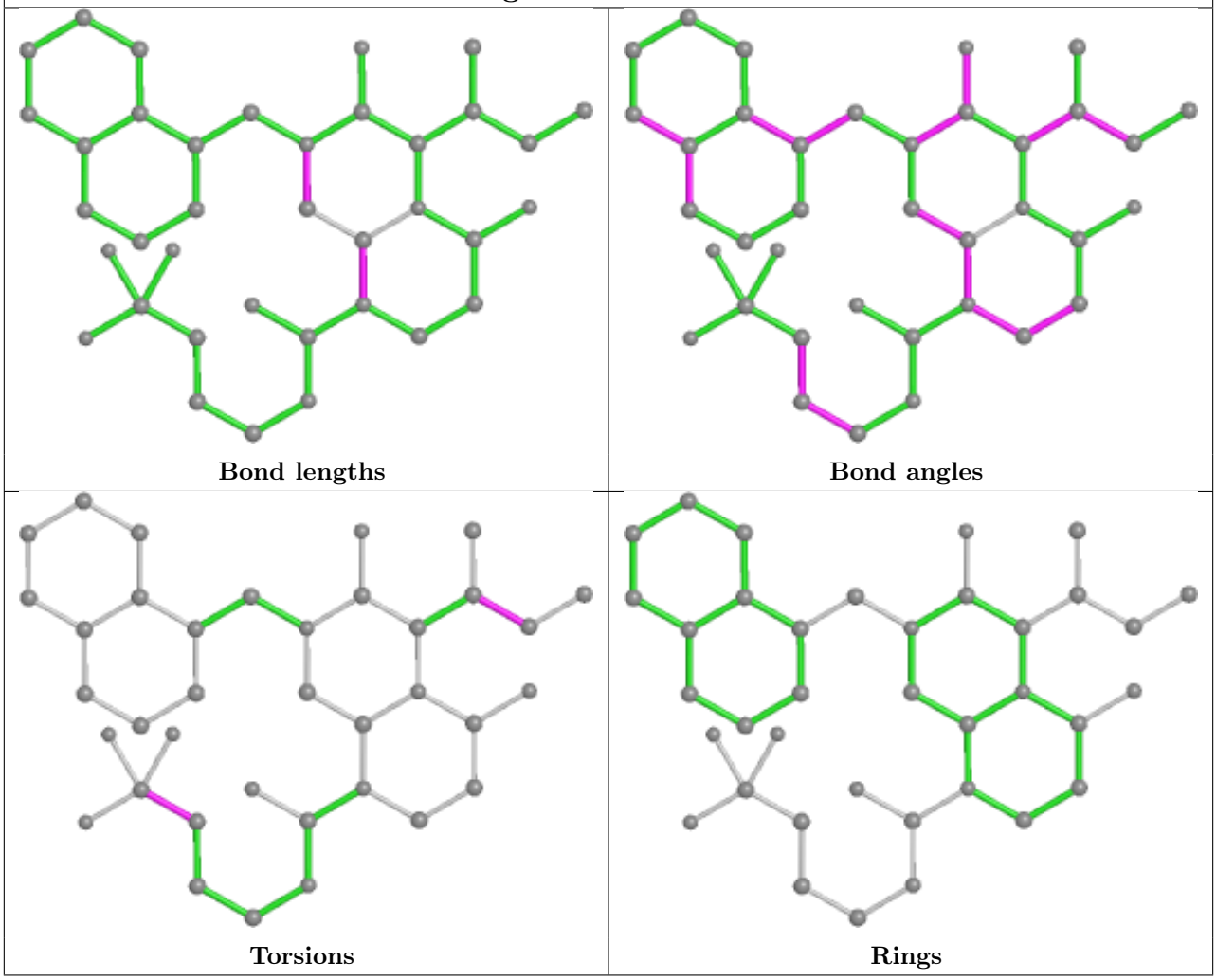
Ligand 9DF K 301



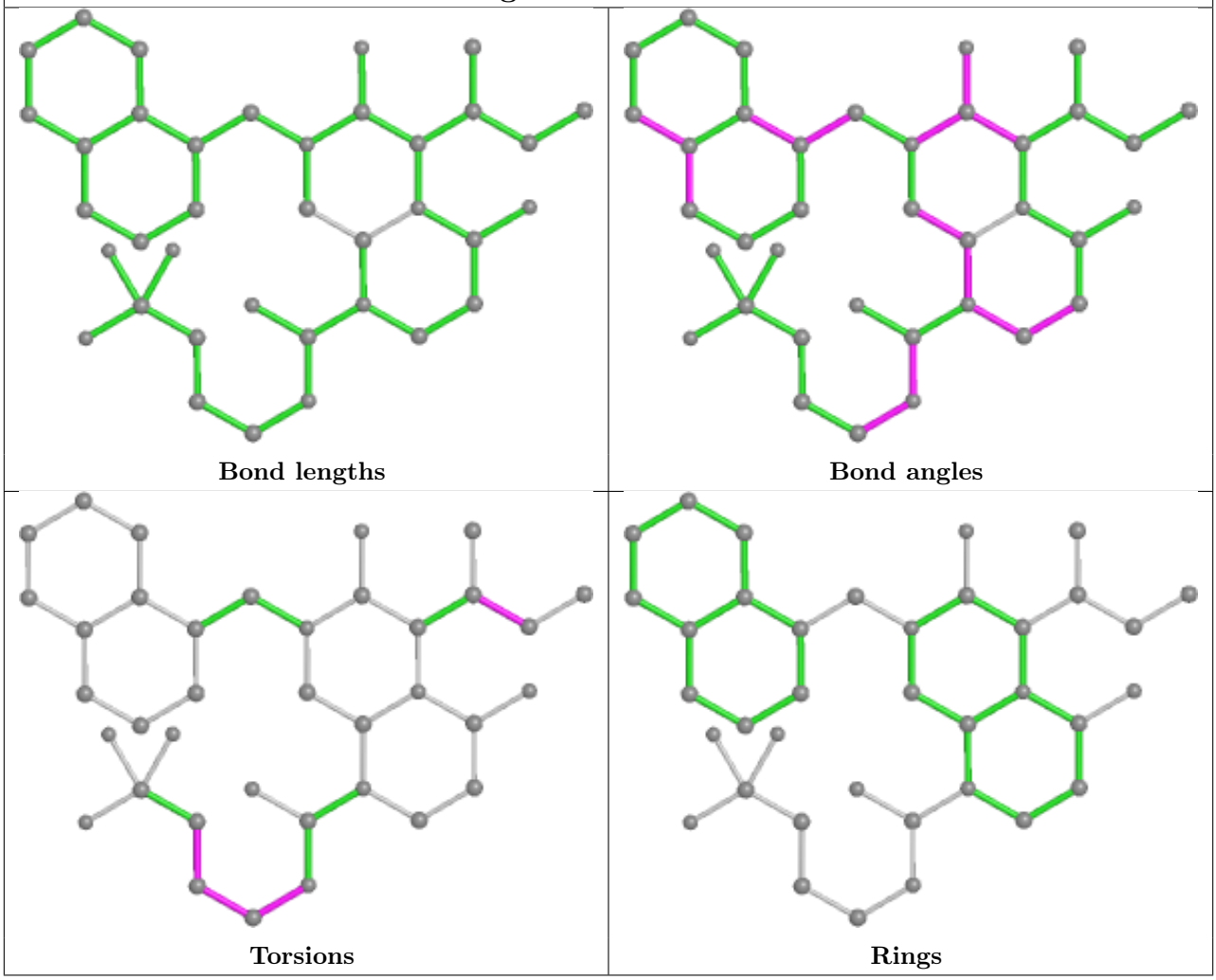
Ligand 9DF I 301



Ligand 9DF F 301



Ligand 9DF D 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 (i)

6.1 Protein, DNA and RNA chains (i)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	170/221 (76%)	0.27	4 (2%)	59	65	36, 51, 79, 114	0
1	B	172/221 (77%)	0.31	5 (2%)	51	59	38, 54, 76, 93	0
1	C	170/221 (76%)	0.22	7 (4%)	37	45	32, 47, 73, 107	0
1	D	171/221 (77%)	0.16	1 (0%)	89	91	28, 42, 61, 84	0
1	E	172/221 (77%)	0.29	4 (2%)	60	67	36, 51, 76, 92	0
1	F	173/221 (78%)	0.52	5 (2%)	51	59	38, 54, 77, 93	0
1	G	171/221 (77%)	0.31	8 (4%)	31	38	37, 52, 73, 101	0
1	H	172/221 (77%)	0.24	4 (2%)	60	67	33, 47, 69, 95	0
1	I	177/221 (80%)	0.27	5 (2%)	53	61	37, 54, 85, 102	0
1	J	171/221 (77%)	0.39	6 (3%)	44	51	43, 58, 81, 100	0
1	K	171/221 (77%)	0.41	9 (5%)	26	33	43, 58, 80, 125	0
1	L	170/221 (76%)	0.25	5 (2%)	51	59	35, 51, 78, 120	0
1	M	169/221 (76%)	0.17	3 (1%)	68	74	28, 41, 62, 81	0
1	N	170/221 (76%)	0.28	5 (2%)	51	59	27, 38, 57, 108	0
All	All	2399/3094 (77%)	0.29	71 (2%)	50	58	27, 51, 76, 125	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	58	LEU	7.8
1	A	249	PRO	7.5
1	I	72	ALA	6.3
1	C	62	VAL	6.1
1	N	73	TYR	5.5

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)

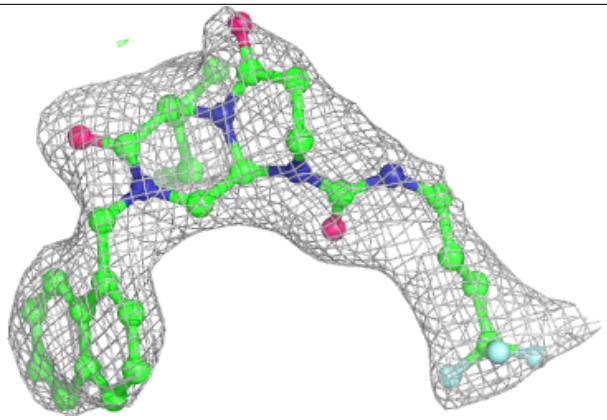
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

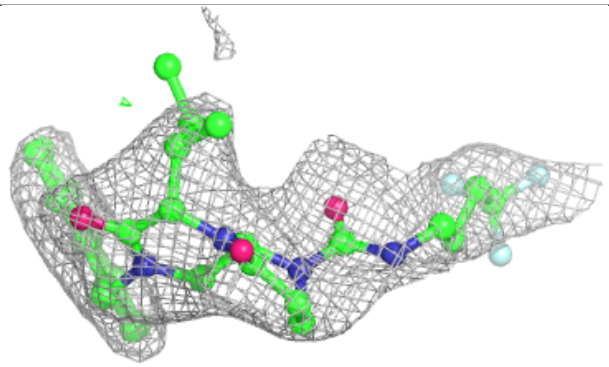
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MG	N	302	1/1	0.81	0.09	57,57,57,57	0
2	9DF	J	301	37/37	0.85	0.20	66,79,111,123	0
2	9DF	L	301	37/37	0.85	0.42	61,87,115,123	0
2	9DF	A	302	37/37	0.85	0.27	74,106,155,158	0
2	9DF	I	301	37/37	0.86	0.21	76,93,117,129	0
2	9DF	K	301	37/37	0.86	0.22	71,77,137,144	0
2	9DF	E	301	37/37	0.88	0.26	68,79,120,125	0
2	9DF	C	301	37/37	0.89	0.18	44,60,87,94	0
2	9DF	C	302	37/37	0.90	0.20	59,70,117,124	0
2	9DF	D	301	37/37	0.90	0.14	57,65,90,95	0
2	9DF	H	302	37/37	0.92	0.13	57,63,92,97	0
2	9DF	A	301	37/37	0.93	0.18	47,65,87,94	0
2	9DF	N	301	37/37	0.94	0.15	40,45,64,79	0
2	9DF	F	301	37/37	0.94	0.18	50,63,85,100	0
2	9DF	H	301	37/37	0.95	0.14	41,52,74,75	0

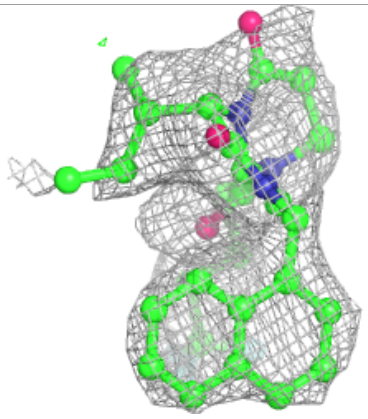
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 9DF J 301:

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

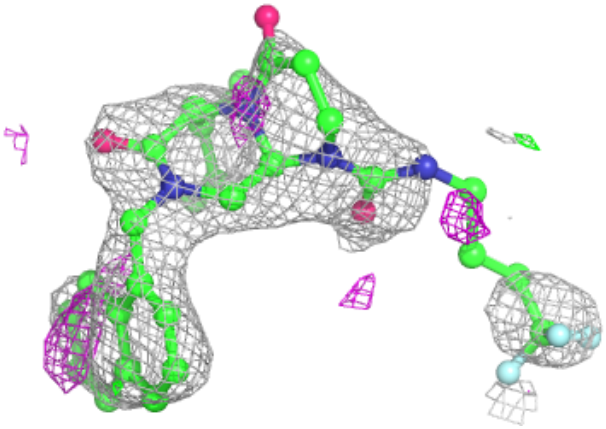


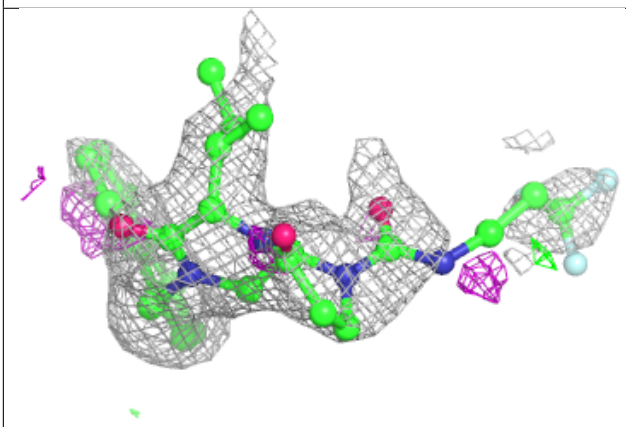


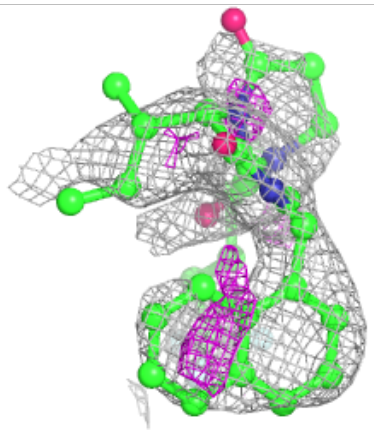


Electron density around 9DF L 301:

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

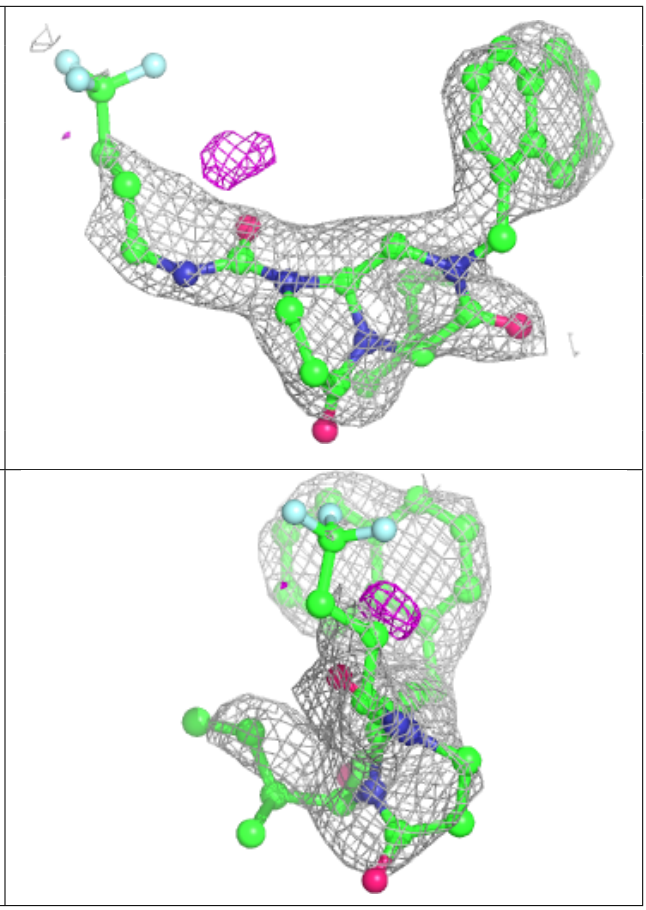






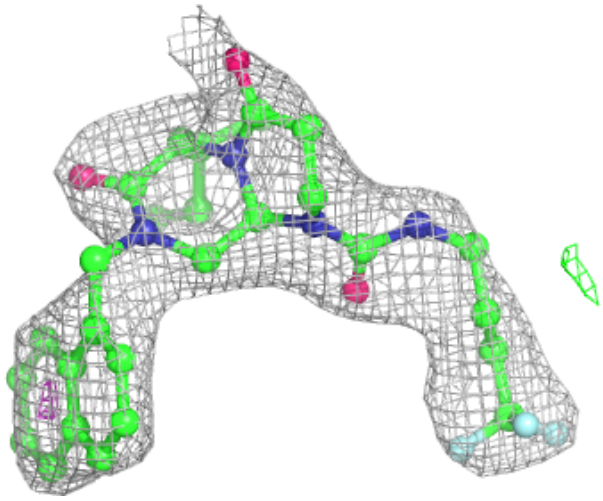
Electron density around 9DF A 302:

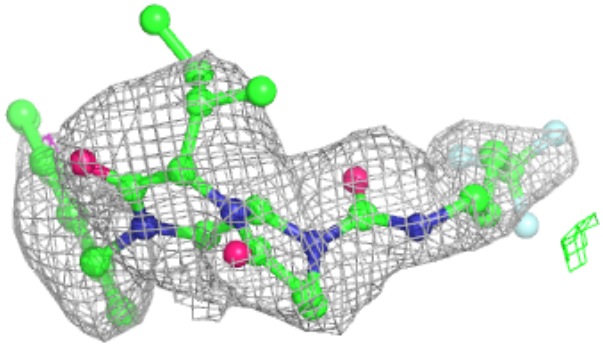
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

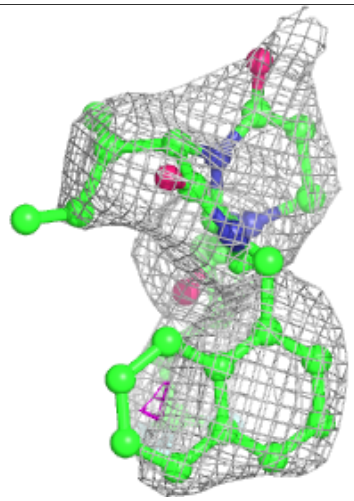


Electron density around 9DF I 301:

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

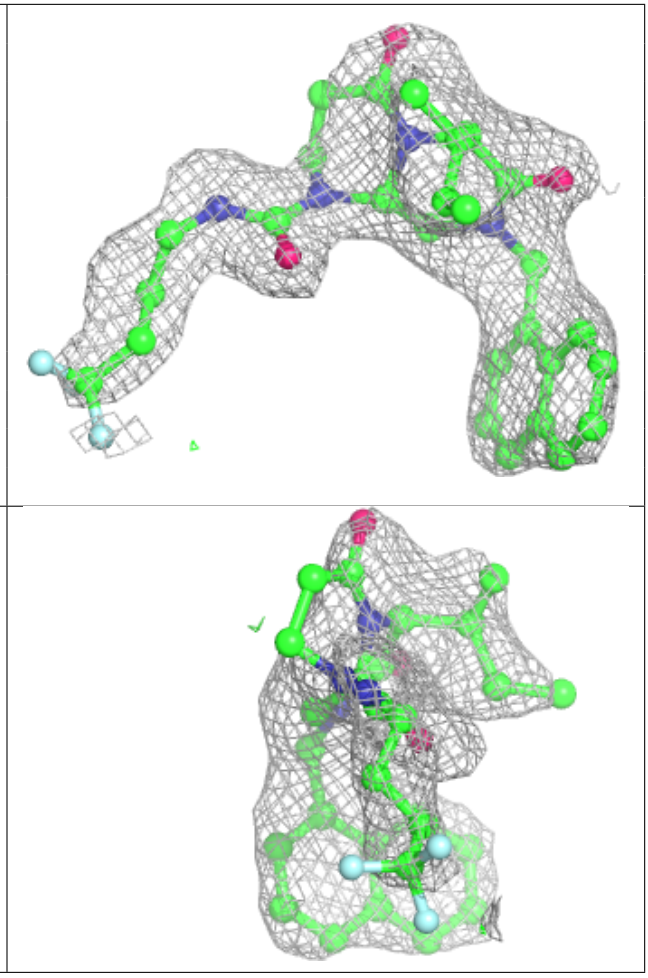






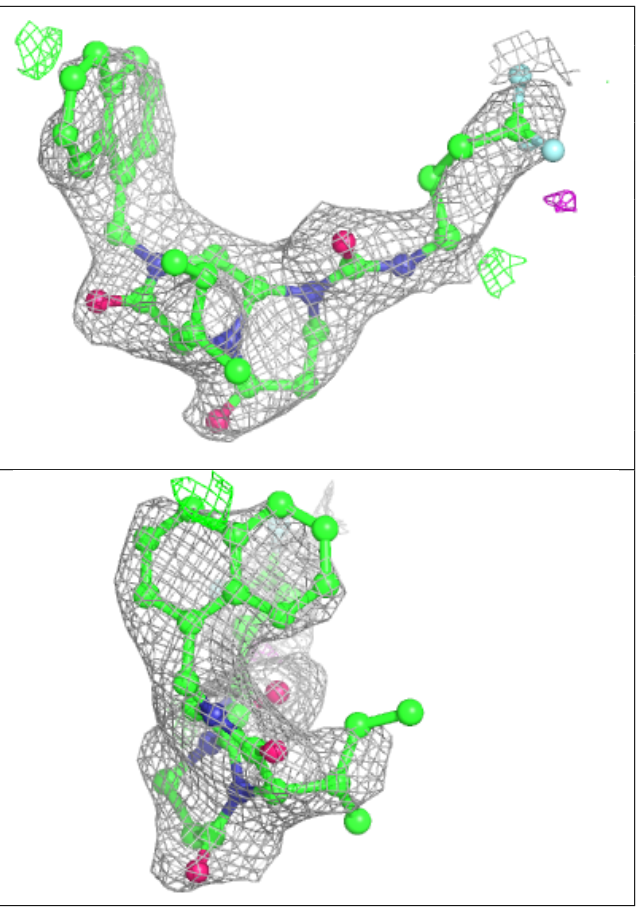
Electron density around 9DF K 301:

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



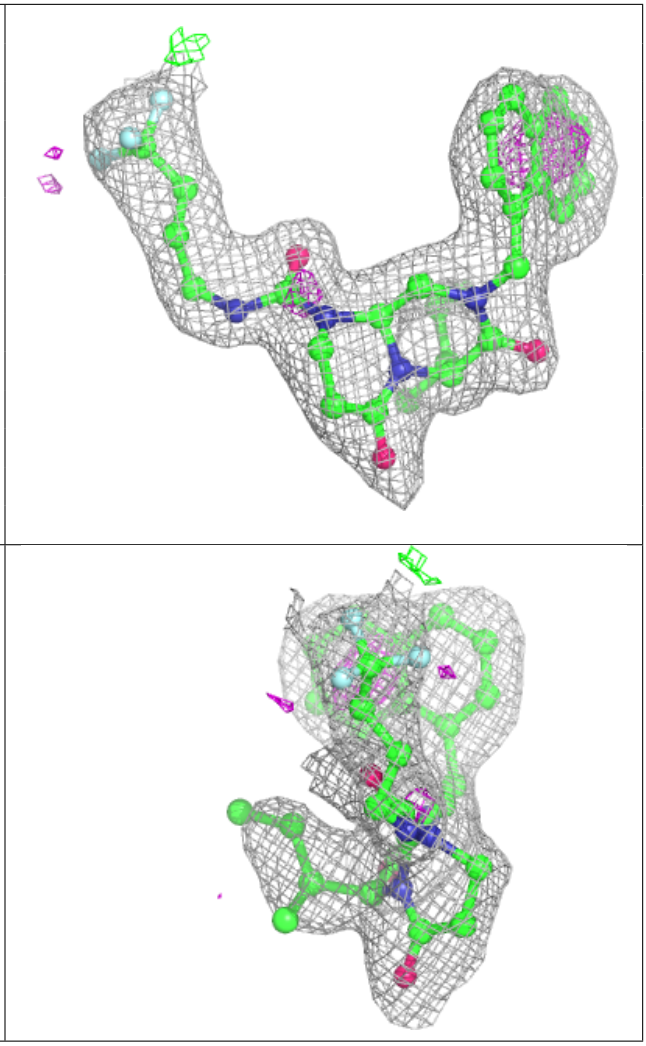
Electron density around 9DF E 301:

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



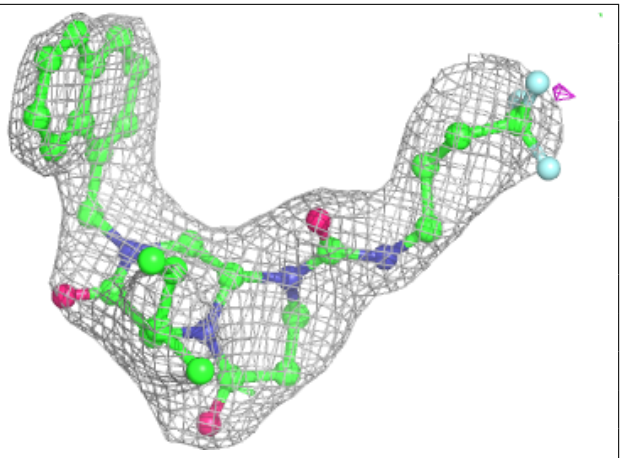
Electron density around 9DF C 301:

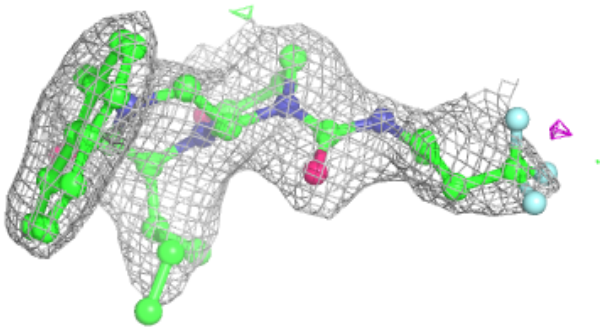
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

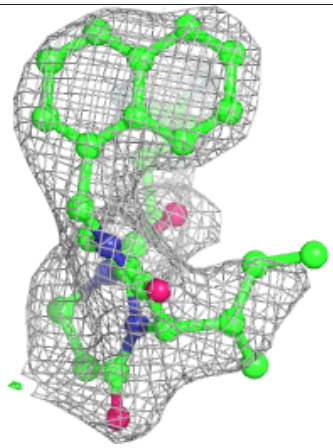


Electron density around 9DF C 302:

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

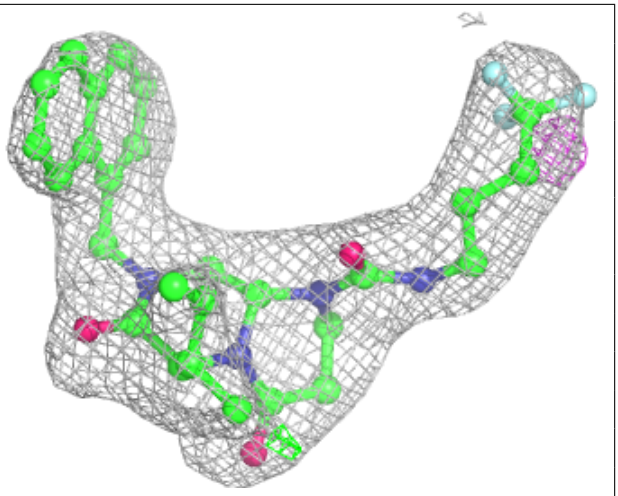


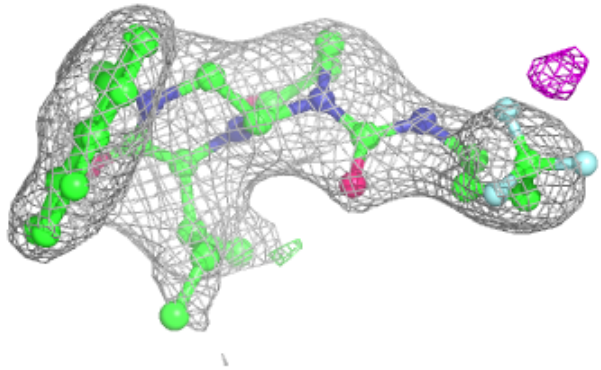


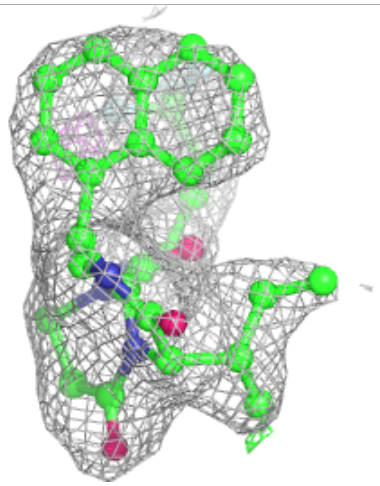


Electron density around 9DF D 301:

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

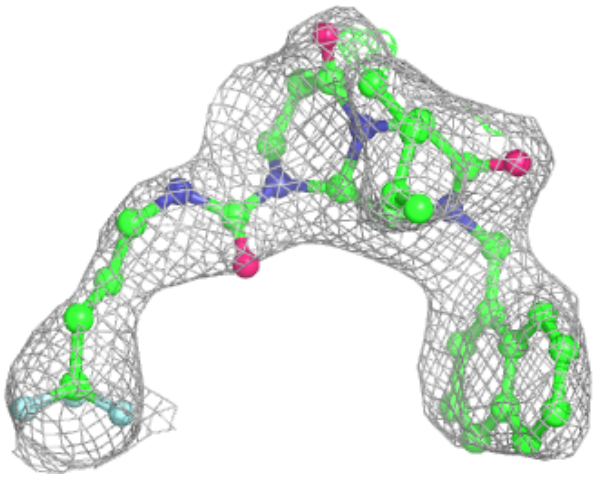


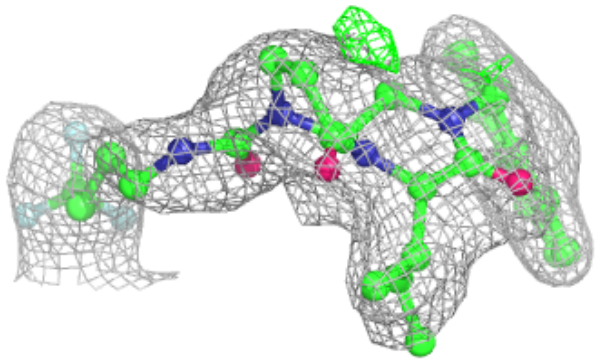


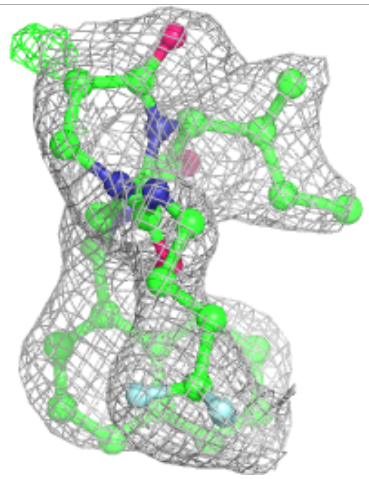


Electron density around 9DF H 302:

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

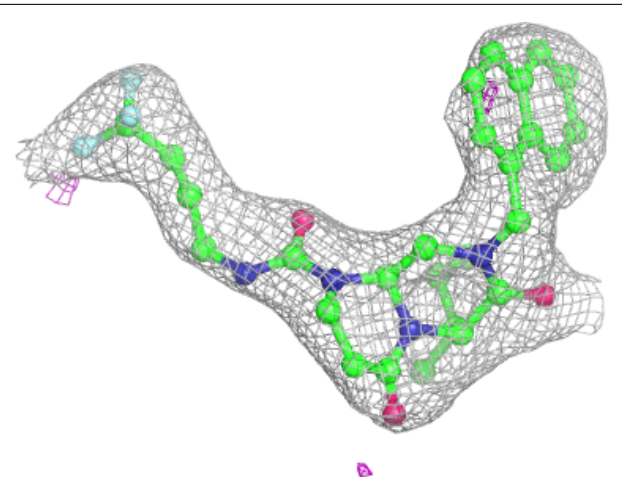


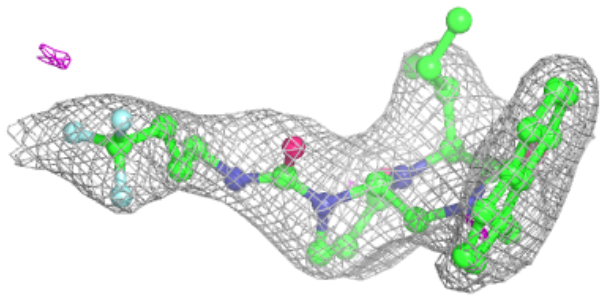


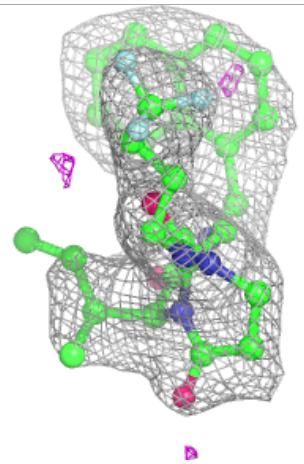


Electron density around 9DF A 301:

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

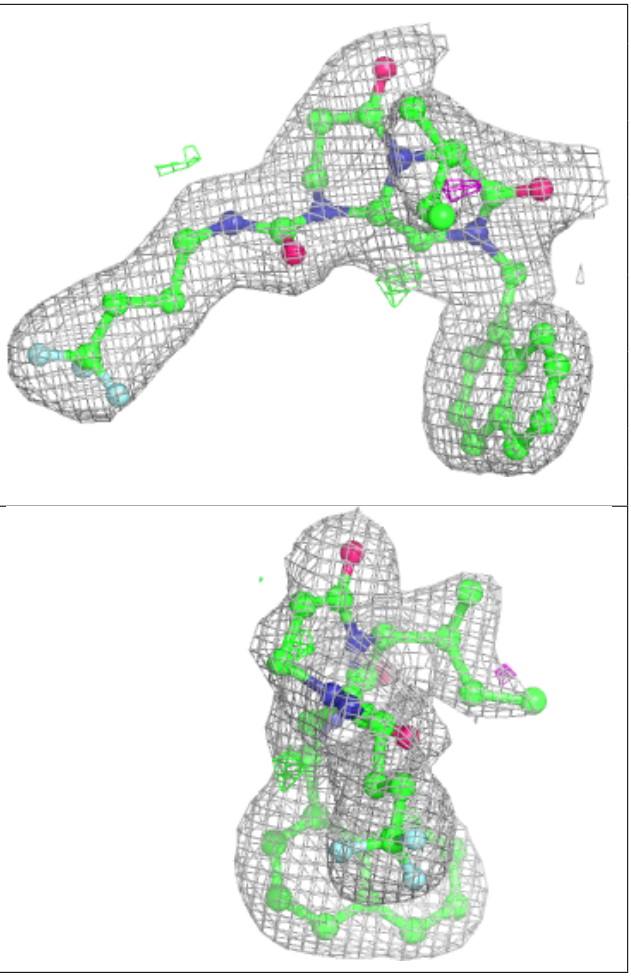






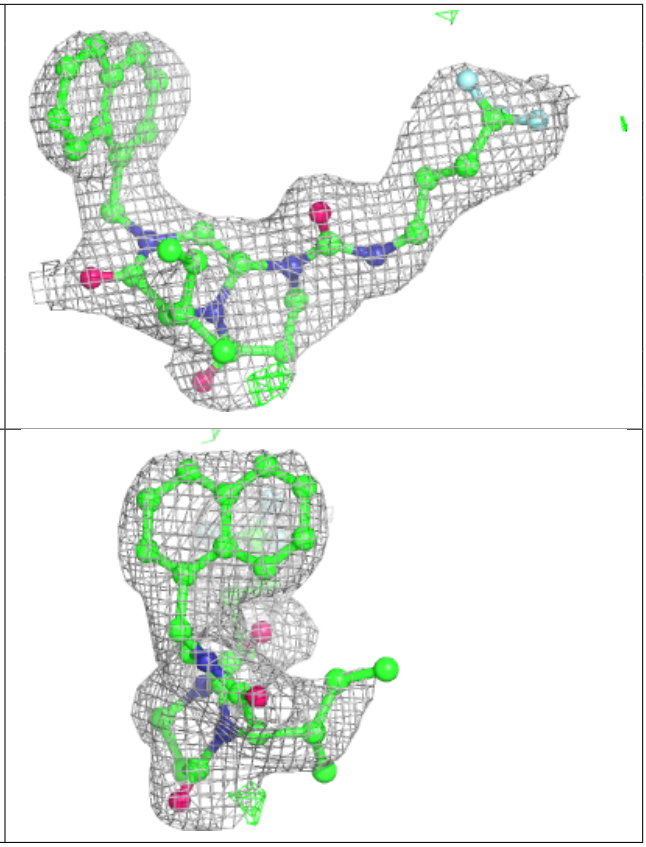
Electron density around 9DF N 301:

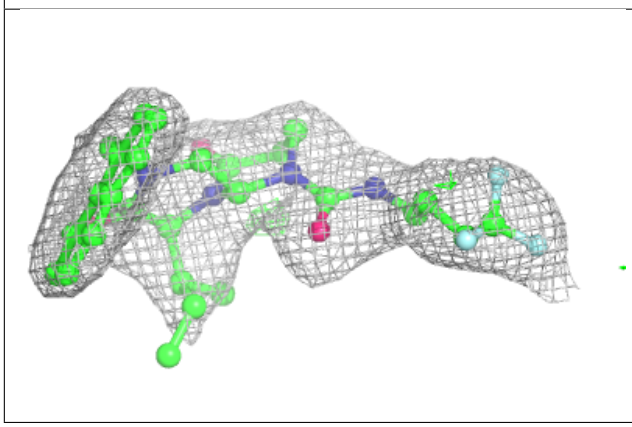
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around 9DF F 301:

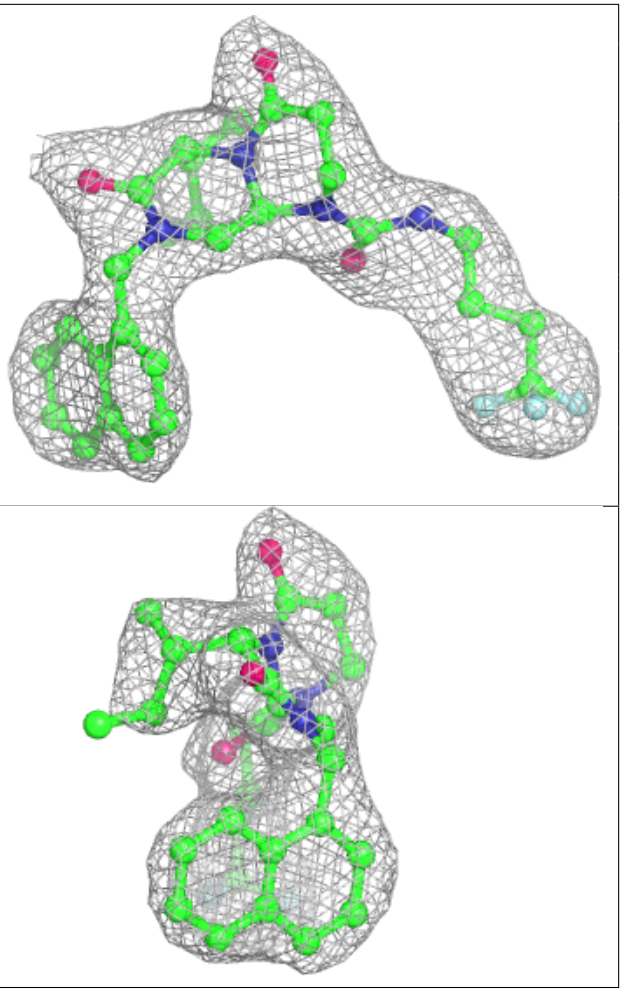
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





Electron density around 9DF H 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.