



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 8, 2024 – 02:52 PM EDT

PDB ID : 7HID  
Title : Group deposition for combi-soaks of Chikungunya virus nsP3 macrodomain – Crystal structure of Chikungunya virus nsP3 macrodomain in complex with Z1041785508 and Z50145861 (CHIKV\_MacB-x1487)  
Authors : Aschenbrenner, J.C.; Fairhead, M.; Godoy, A.S.; Balcomb, B.H.; Capkin, E.; Chandran, A.V.; Golding, M.; Koekemoer, L.; Lithgo, R.M.; Marples, P.G.; Ni, X.; Thompson, W.; Tomlinson, C.W.E.; Wild, C.; Winokan, M.; Xavier, M.-A.E.; Fearon, D.; von Delft, F.  
Deposited on : 2024-10-02  
Resolution : 1.50 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
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

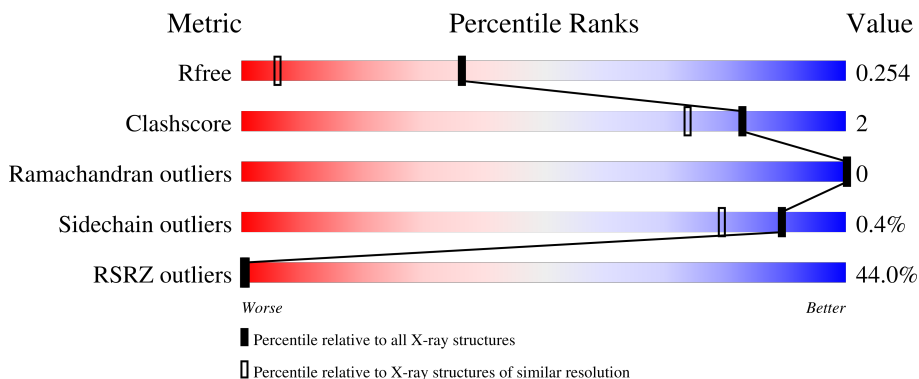
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	3717 (1.50-1.50)
Clashscore	180529	4048 (1.50-1.50)
Ramachandran outliers	177936	3970 (1.50-1.50)
Sidechain outliers	177891	3967 (1.50-1.50)
RSRZ outliers	164620	3718 (1.50-1.50)

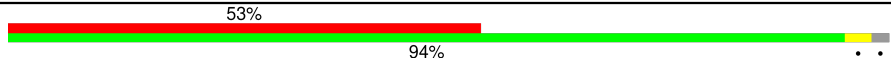
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	163	
1	B	163	
1	C	163	

*Continued on next page...*

Ideal geometry (proteins) : Engh & Huber (2001)  
 Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
 Validation Pipeline (wwPDB-VP) : 2.39

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	D	163	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into segments: a red segment at the beginning, a green segment, a yellow segment, and a grey segment at the end. The text '53%' is positioned above the red segment, and '94%' is positioned below the green segment. Two small black dots are located at the far right end of the bar.</p>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 5881 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 Non-structural protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	163	1262	787	220	245	10	0	2	0
1	B	163	1272	793	222	248	9	0	3	0
1	C	163	1291	804	227	250	10	0	5	0
1	D	159	1246	777	217	243	9	0	3	0

There are 16 discrepancies between the modelled and reference sequences:

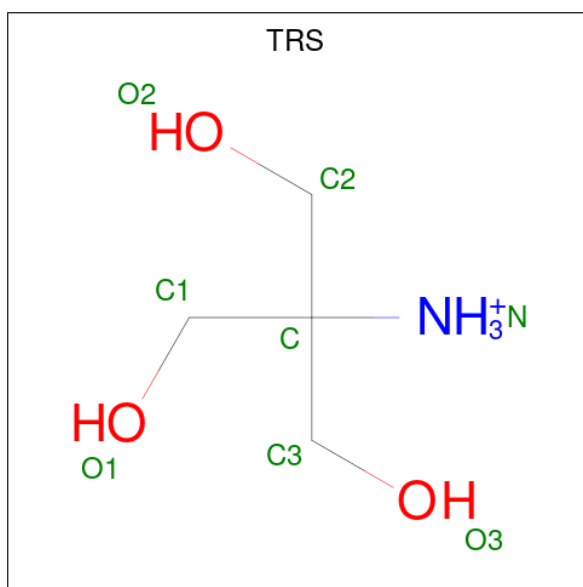
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q8JUX6
A	-1	ALA	-	expression tag	UNP Q8JUX6
A	0	MET	-	expression tag	UNP Q8JUX6
A	77	THR	SER	conflict	UNP Q8JUX6
B	-2	GLY	-	expression tag	UNP Q8JUX6
B	-1	ALA	-	expression tag	UNP Q8JUX6
B	0	MET	-	expression tag	UNP Q8JUX6
B	77	THR	SER	conflict	UNP Q8JUX6
C	-2	GLY	-	expression tag	UNP Q8JUX6
C	-1	ALA	-	expression tag	UNP Q8JUX6
C	0	MET	-	expression tag	UNP Q8JUX6
C	77	THR	SER	conflict	UNP Q8JUX6
D	-2	GLY	-	expression tag	UNP Q8JUX6
D	-1	ALA	-	expression tag	UNP Q8JUX6
D	0	MET	-	expression tag	UNP Q8JUX6
D	77	THR	SER	conflict	UNP Q8JUX6

- Molecule 2 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	C	1	Total	C	O	S	0	0
			4	2	1	1		
2	D	1	Total	C	O	S	0	0
			4	2	1	1		
2	D	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 3 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).

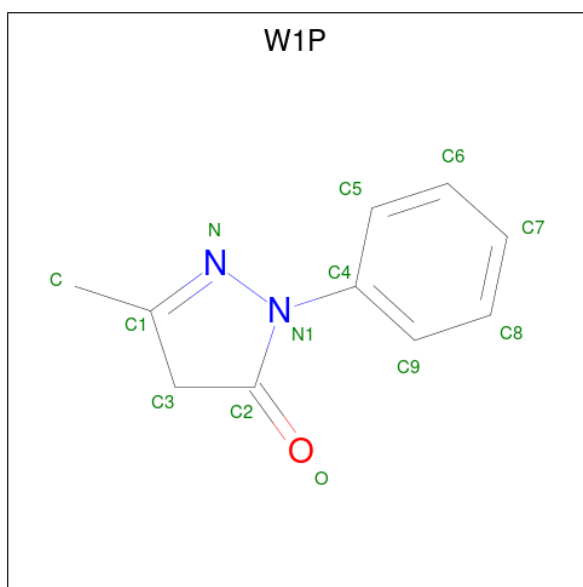


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	8	4	1	3	0	0
3	D	1	8	4	1	3	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

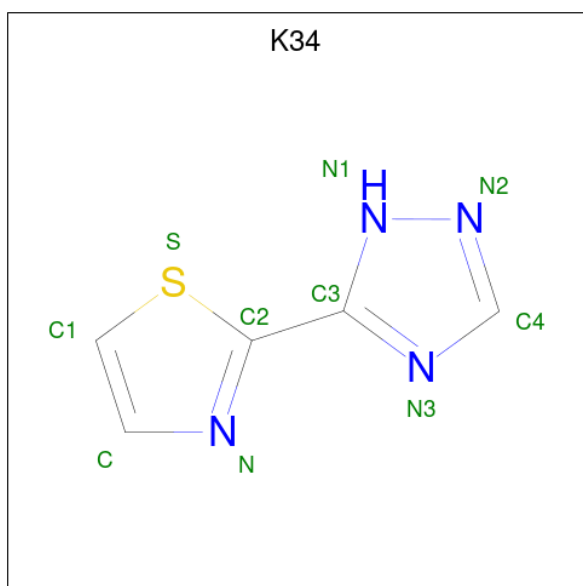
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
4	A	2	2	2	0	0
4	B	1	1	1	0	0
4	C	2	2	2	0	0
4	D	1	1	1	0	0

- Molecule 5 is 5-methyl-2-phenyl-2,4-dihydro-3H-pyrazol-3-one (three-letter code: W1P) (formula: C<sub>10</sub>H<sub>10</sub>N<sub>2</sub>O) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	Total 13	C 10	N 2	O 1	0	0
5	B	1	Total 13	C 10	N 2	O 1	0	0
5	C	1	Total 13	C 10	N 2	O 1	0	0
5	D	1	Total 13	C 10	N 2	O 1	0	0

- Molecule 6 is 5-(1,3-thiazol-2-yl)-1H-1,2,4-triazole (three-letter code: K34) (formula:  $C_5H_4N_4S$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	S	0	0
			10	5	4	1		
6	C	1	Total	C	N	S	0	0
			10	5	4	1		
6	D	1	Total	C	N	S	0	0
			10	5	4	1		

- Molecule 7 is water.

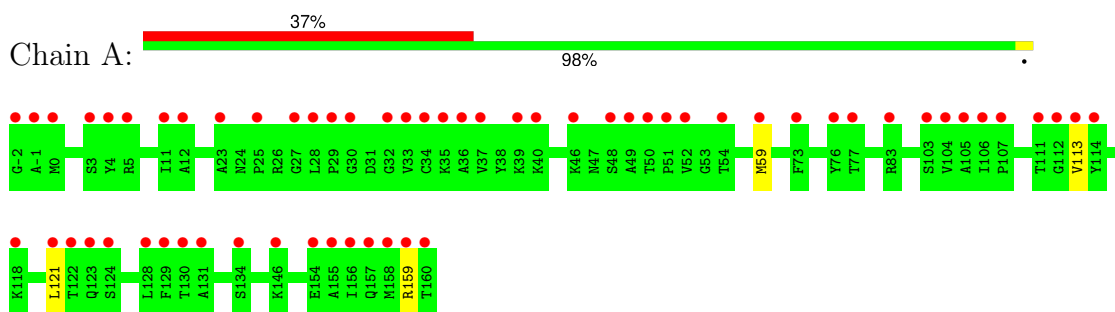
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	192	Total	O	0	0
			192	192		
7	B	165	Total	O	0	0
			165	165		
7	C	172	Total	O	0	0
			172	172		
7	D	137	Total	O	0	0
			137	137		



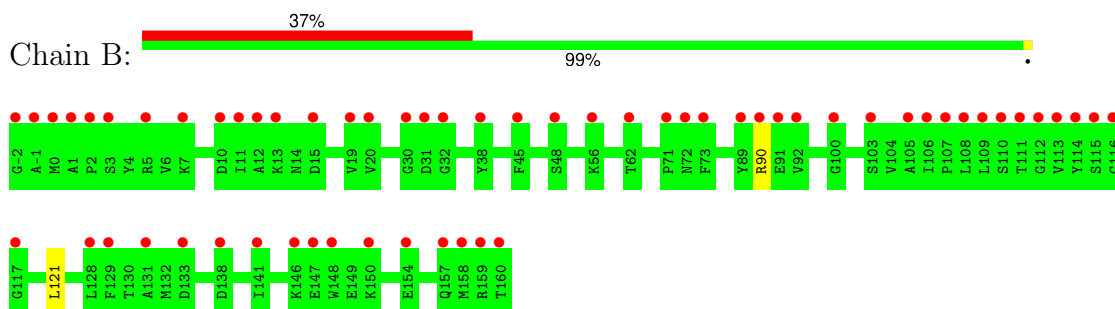
### 3 Residue-property plots [i](#)

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. 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. 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.

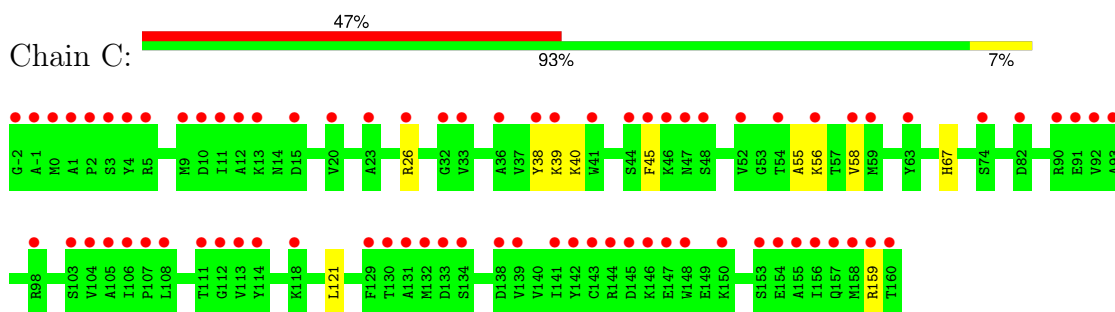
- Molecule 1: Non-structural protein 3



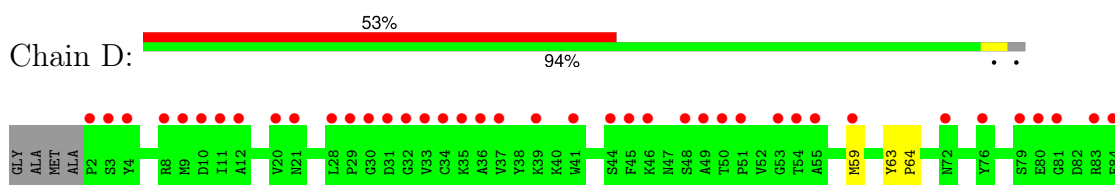
- Molecule 1: Non-structural protein 3

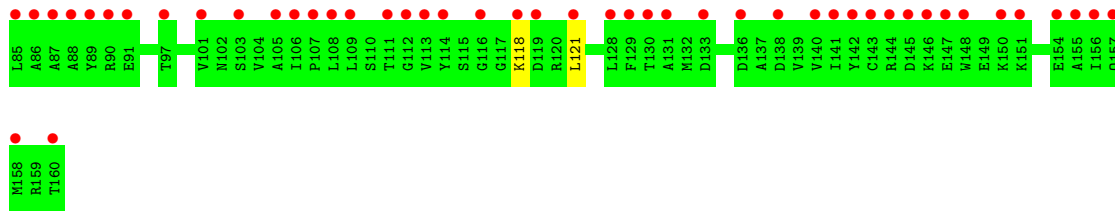


- Molecule 1: Non-structural protein 3



- Molecule 1: Non-structural protein 3





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.43Å 87.43Å 85.49Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.64 – 1.50 34.64 – 1.50	Depositor EDS
% Data completeness (in resolution range)	98.1 (34.64-1.50) 98.2 (34.64-1.50)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.99 (at 1.50Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.190 , 0.222 0.233 , 0.254	Depositor DCC
$R_{free}$ test set	5896 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.4	Xtrriage
Anisotropy	0.010	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 46.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l 0.035 for h,-h-k,-l 0.000 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5881	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.76 % 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 4.3751e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: CL, TRS, W1P, K34, DMS

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	0.74	0/1285	0.79	0/1739
1	B	0.79	0/1295	0.80	0/1752
1	C	0.75	0/1314	0.78	0/1776
1	D	0.74	0/1269	0.76	0/1717
All	All	0.76	0/5163	0.79	0/6984

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	1262	0	1256	5	0
1	B	1272	0	1265	2	0
1	C	1291	0	1286	8	0
1	D	1246	0	1236	4	0
2	A	28	0	42	7	0
2	C	4	0	6	2	0
2	D	8	0	12	1	0
3	A	8	0	12	0	0
3	D	8	0	12	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	C	2	0	0	0	0
4	D	1	0	0	0	0
5	A	13	0	0	0	0
5	B	13	0	0	0	0
5	C	13	0	0	0	0
5	D	13	0	0	0	0
6	B	10	0	0	0	0
6	C	10	0	0	0	0
6	D	10	0	0	0	0
7	A	192	0	0	3	0
7	B	165	0	0	1	0
7	C	172	0	0	5	0
7	D	137	0	0	3	0
All	All	5881	0	5127	24	0

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

All (24) 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:D:118:LYS:NZ	7:D:301:HOH:O	2.07	0.88
2:C:303:DMS:H12	7:C:552:HOH:O	1.80	0.81
1:A:59[A]:MET:HG2	2:A:203:DMS:H11	1.82	0.62
1:A:59[A]:MET:HG2	2:A:203:DMS:C1	2.39	0.52
1:D:59[B]:MET:HG2	7:D:392:HOH:O	2.09	0.52
2:A:201:DMS:H12	7:A:451:HOH:O	2.09	0.51
1:B:121:LEU:HD23	1:B:121:LEU:C	2.31	0.51
2:D:203:DMS:H11	7:D:392:HOH:O	2.11	0.51
1:C:56:LYS:NZ	7:C:402:HOH:O	2.43	0.50
1:C:26:ARG:NH1	7:C:403:HOH:O	2.44	0.49
1:C:39:LYS:HD3	7:C:435:HOH:O	2.13	0.49
1:C:55:ALA:HA	1:C:67:HIS:O	2.13	0.48
1:A:113:VAL:HG21	2:A:208:DMS:H22	1.95	0.48
1:A:121:LEU:C	1:A:121:LEU:HD23	2.36	0.46
1:C:40:LYS:HD3	2:C:303:DMS:H11	1.97	0.46
1:D:121:LEU:HD23	1:D:121:LEU:C	2.37	0.45
1:C:38:TYR:HA	1:C:45:PHE:CE1	2.52	0.45
1:C:58:VAL:HG22	7:C:457:HOH:O	2.18	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:90:ARG:HD2	7:B:434:HOH:O	2.19	0.43
2:A:211:DMS:S	7:A:456:HOH:O	2.61	0.42
1:C:121:LEU:HD23	1:C:121:LEU:C	2.40	0.42
2:A:201:DMS:C1	7:A:451:HOH:O	2.66	0.42
1:D:63:TYR:HA	1:D:64:PRO:HD3	1.92	0.41
1:A:113:VAL:HG21	2:A:208:DMS:C2	2.49	0.41

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	163/163 (100%)	162 (99%)	1 (1%)	0	100	100
1	B	164/163 (101%)	164 (100%)	0	0	100	100
1	C	166/163 (102%)	166 (100%)	0	0	100	100
1	D	160/163 (98%)	160 (100%)	0	0	100	100
All	All	653/652 (100%)	652 (100%)	1 (0%)	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	136/134 (102%)	135 (99%)	1 (1%)	81	66
1	B	137/134 (102%)	137 (100%)	0	100	100
1	C	139/134 (104%)	138 (99%)	1 (1%)	81	66
1	D	136/134 (102%)	136 (100%)	0	100	100
All	All	548/536 (102%)	546 (100%)	2 (0%)	89	79

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	159	ARG
1	C	159	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	47	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 25 ligands modelled in this entry, 6 are monoatomic - leaving 19 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	DMS	A	209	-	3,3,3	0.22	0	3,3,3	0.09	0
6	K34	D	206	-	9,11,11	0.98	1 (11%)	5,14,14	6.42	4 (80%)
3	TRS	A	202	-	7,7,7	0.23	0	9,9,9	0.30	0
5	W1P	A	207	-	12,14,14	0.26	0	15,19,19	0.20	0
2	DMS	A	210	-	3,3,3	0.48	0	3,3,3	0.28	0
2	DMS	A	201	-	3,3,3	0.43	0	3,3,3	0.44	0
2	DMS	A	203	-	3,3,3	0.18	0	3,3,3	0.10	0
5	W1P	D	205	-	12,14,14	0.19	0	15,19,19	0.21	0
6	K34	B	203	-	9,11,11	0.94	1 (11%)	5,14,14	6.38	4 (80%)
2	DMS	A	211	-	3,3,3	0.23	0	3,3,3	0.15	0
2	DMS	A	204	-	3,3,3	0.28	0	3,3,3	0.11	0
2	DMS	D	202	-	3,3,3	0.29	0	3,3,3	0.08	0
2	DMS	D	203	-	3,3,3	0.18	0	3,3,3	0.08	0
6	K34	C	305	-	9,11,11	0.97	1 (11%)	5,14,14	6.43	4 (80%)
2	DMS	A	208	-	3,3,3	0.21	0	3,3,3	0.37	0
2	DMS	C	303	-	3,3,3	0.35	0	3,3,3	0.27	0
3	TRS	D	201	-	7,7,7	0.20	0	9,9,9	0.30	0
5	W1P	C	304	-	12,14,14	0.20	0	15,19,19	0.22	0
5	W1P	B	202	-	12,14,14	0.29	0	15,19,19	0.22	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
5	W1P	D	205	-	-	0/4/16/16	0/2/2/2
3	TRS	D	201	-	-	5/9/9/9	-
5	W1P	C	304	-	-	0/4/16/16	0/2/2/2
6	K34	B	203	-	-	0/1/4/4	0/2/2/2
6	K34	C	305	-	-	0/1/4/4	0/2/2/2
5	W1P	B	202	-	-	0/4/16/16	0/2/2/2
6	K34	D	206	-	-	0/1/4/4	0/2/2/2
3	TRS	A	202	-	-	0/9/9/9	-
5	W1P	A	207	-	-	0/4/16/16	0/2/2/2

All (3) bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	206	K34	C4-N2	2.38	1.37	1.32
6	C	305	K34	C4-N2	2.33	1.37	1.32
6	B	203	K34	C4-N2	2.17	1.36	1.32

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	206	K34	C4-N3-C3	13.14	106.45	101.63
6	C	305	K34	C4-N3-C3	13.08	106.43	101.63
6	B	203	K34	C4-N3-C3	12.99	106.39	101.63
6	B	203	K34	N3-C3-N1	-4.83	110.71	114.72
6	C	305	K34	N3-C3-N1	-4.77	110.77	114.72
6	D	206	K34	N3-C3-N1	-4.71	110.82	114.72
6	C	305	K34	C2-C3-N1	2.69	126.64	124.11
6	D	206	K34	C2-C3-N1	2.44	126.40	124.11
6	B	203	K34	C2-C3-N1	2.43	126.39	124.11
6	B	203	K34	C-N-C2	2.19	110.04	104.39
6	D	206	K34	C-N-C2	2.17	109.99	104.39
6	C	305	K34	C-N-C2	2.17	109.98	104.39

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	201	TRS	C1-C-C3-O3
3	D	201	TRS	N-C-C3-O3
3	D	201	TRS	C3-C-C2-O2
3	D	201	TRS	N-C-C2-O2
3	D	201	TRS	C2-C-C3-O3

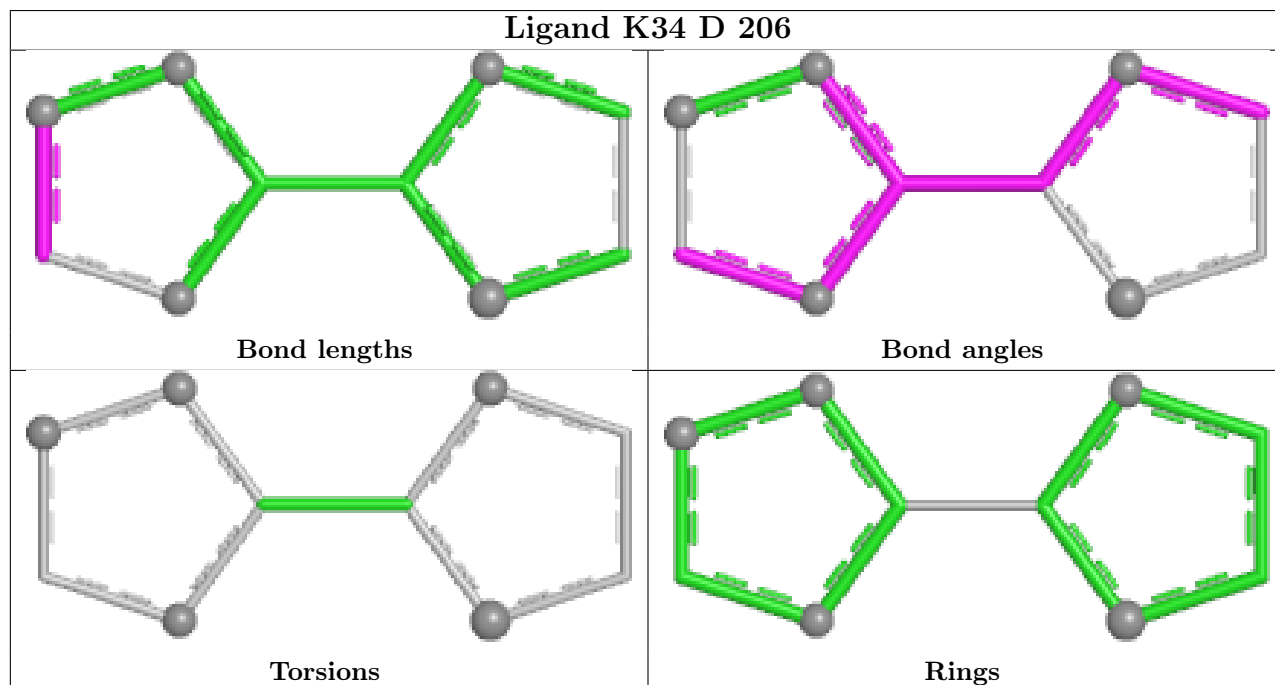
There are no ring outliers.

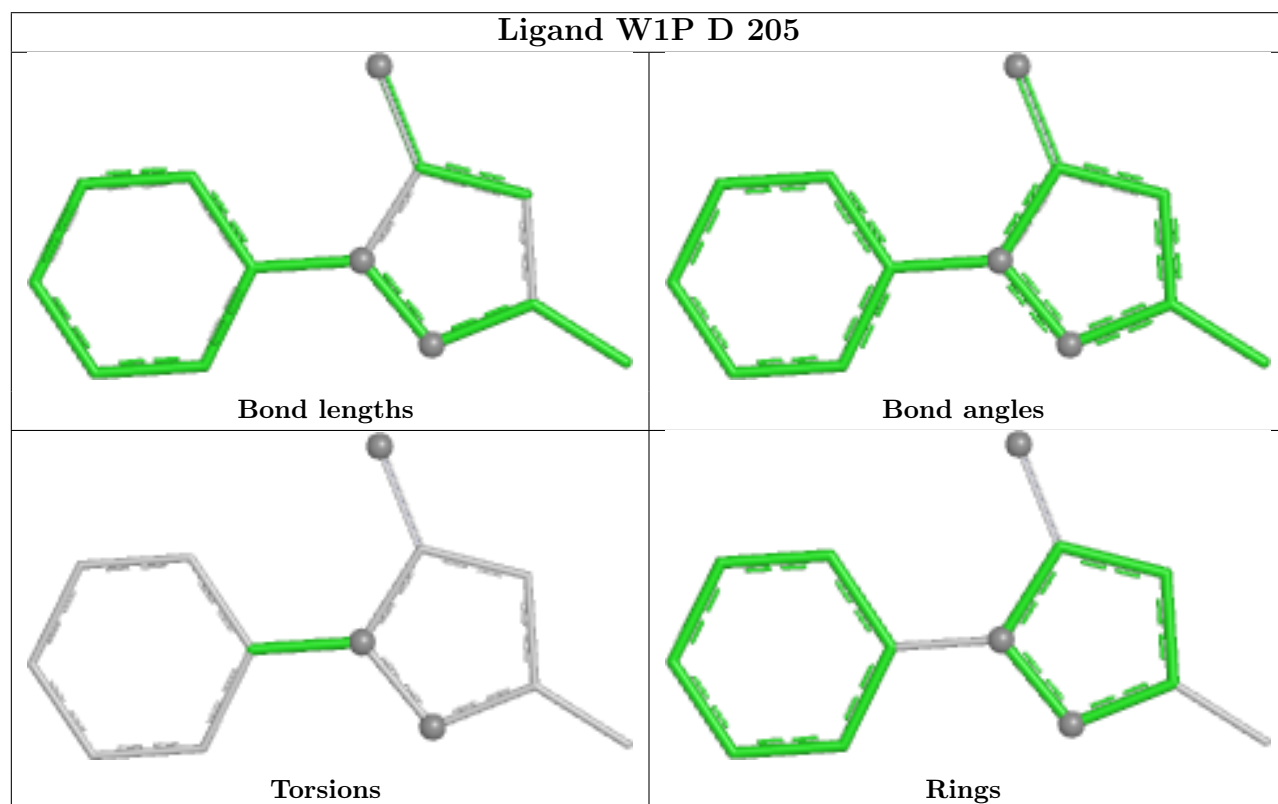
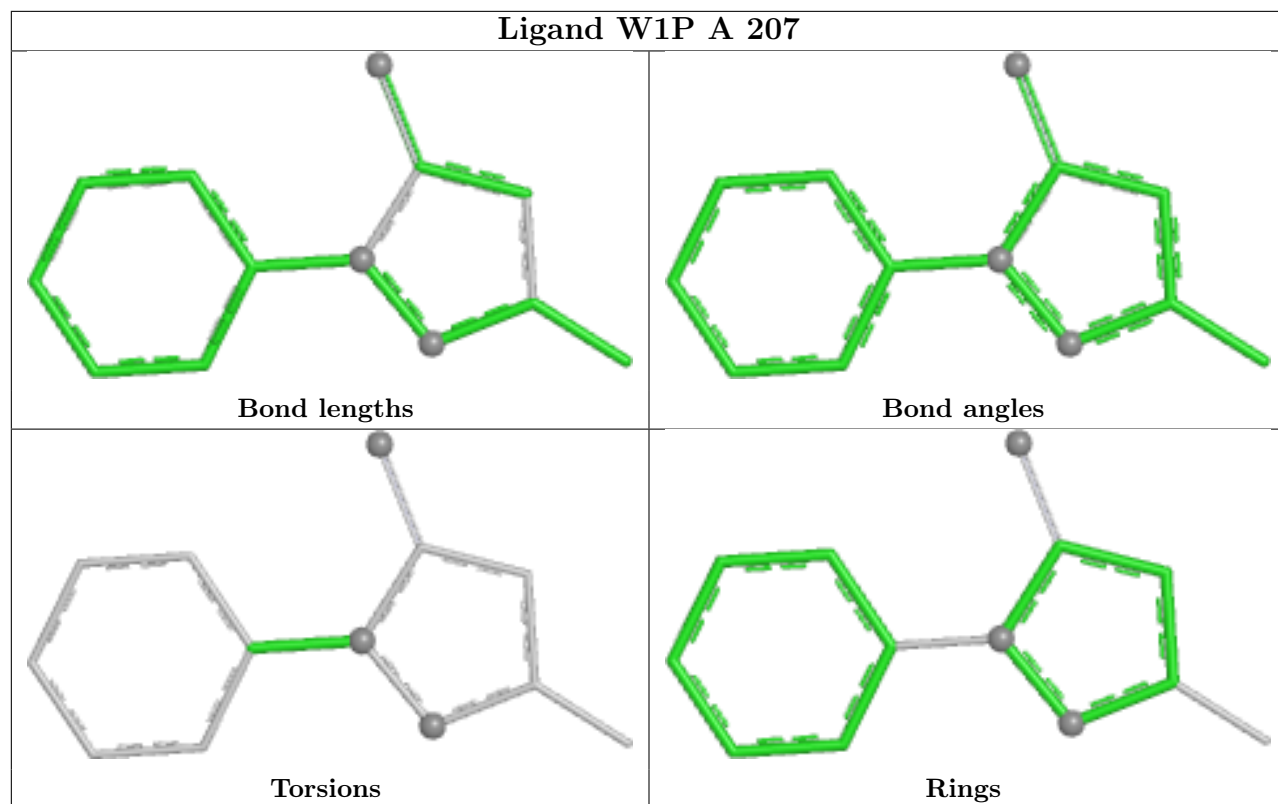
6 monomers are involved in 10 short contacts:

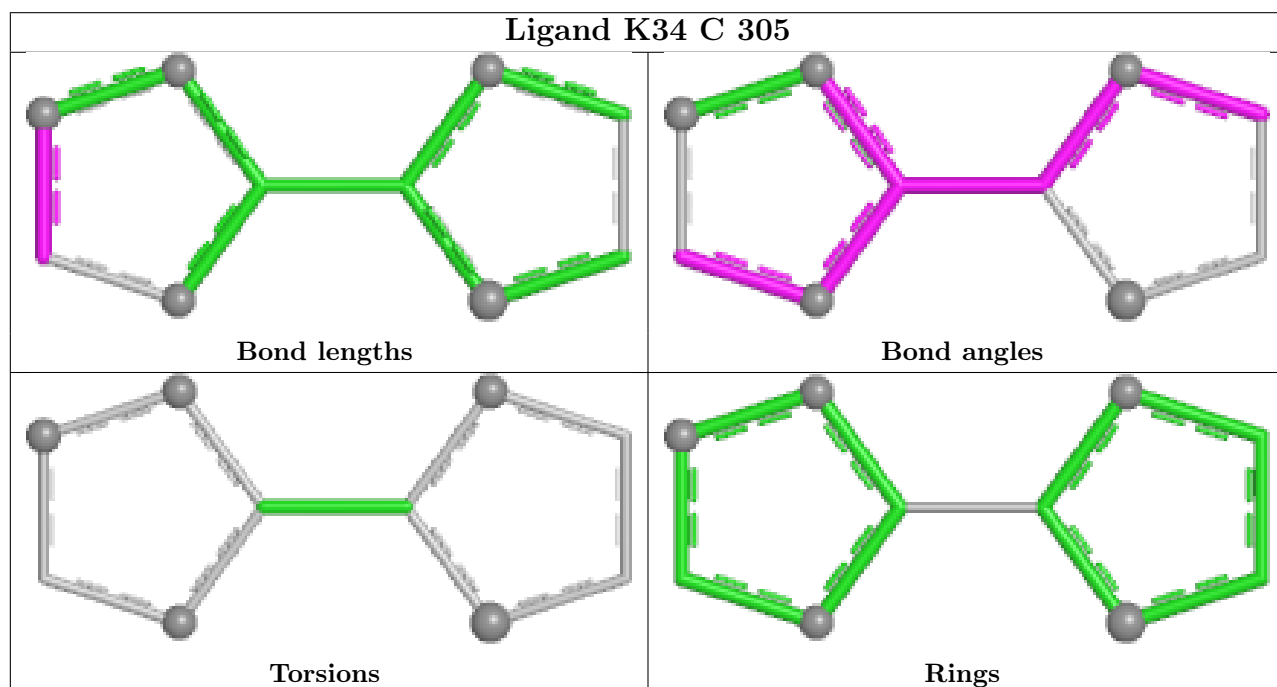
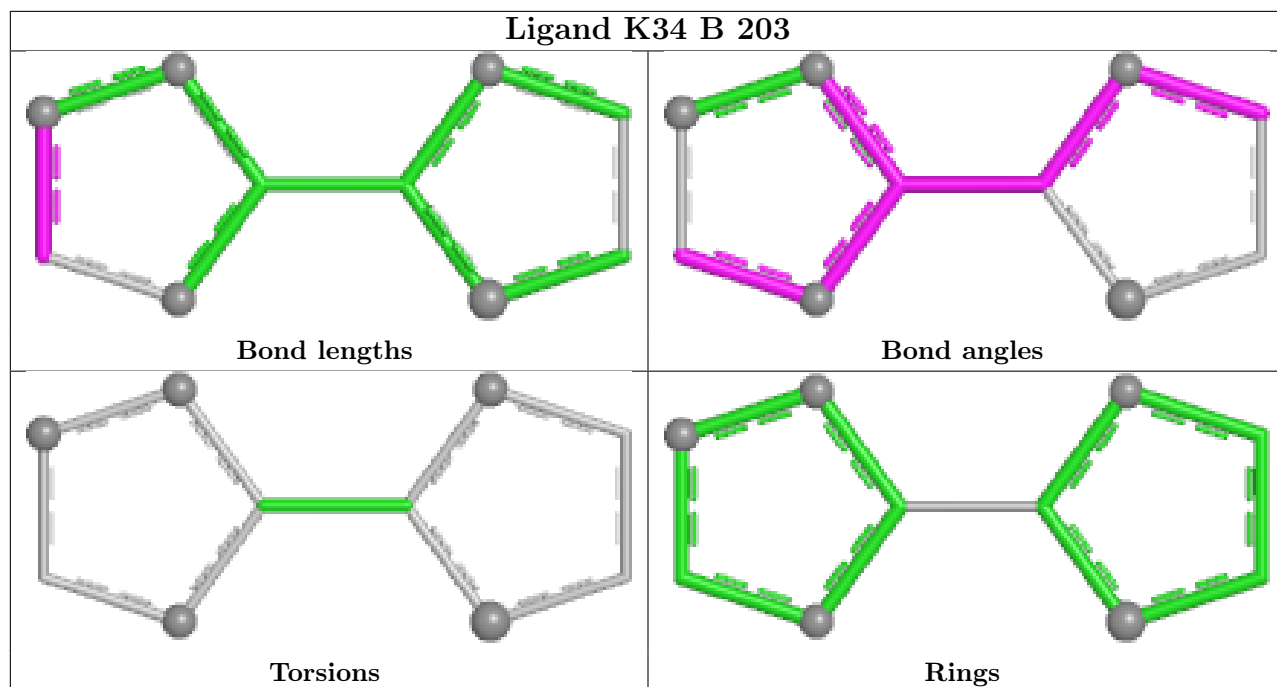
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	DMS	2	0
2	A	203	DMS	2	0
2	A	211	DMS	1	0
2	D	203	DMS	1	0
2	A	208	DMS	2	0
2	C	303	DMS	2	0

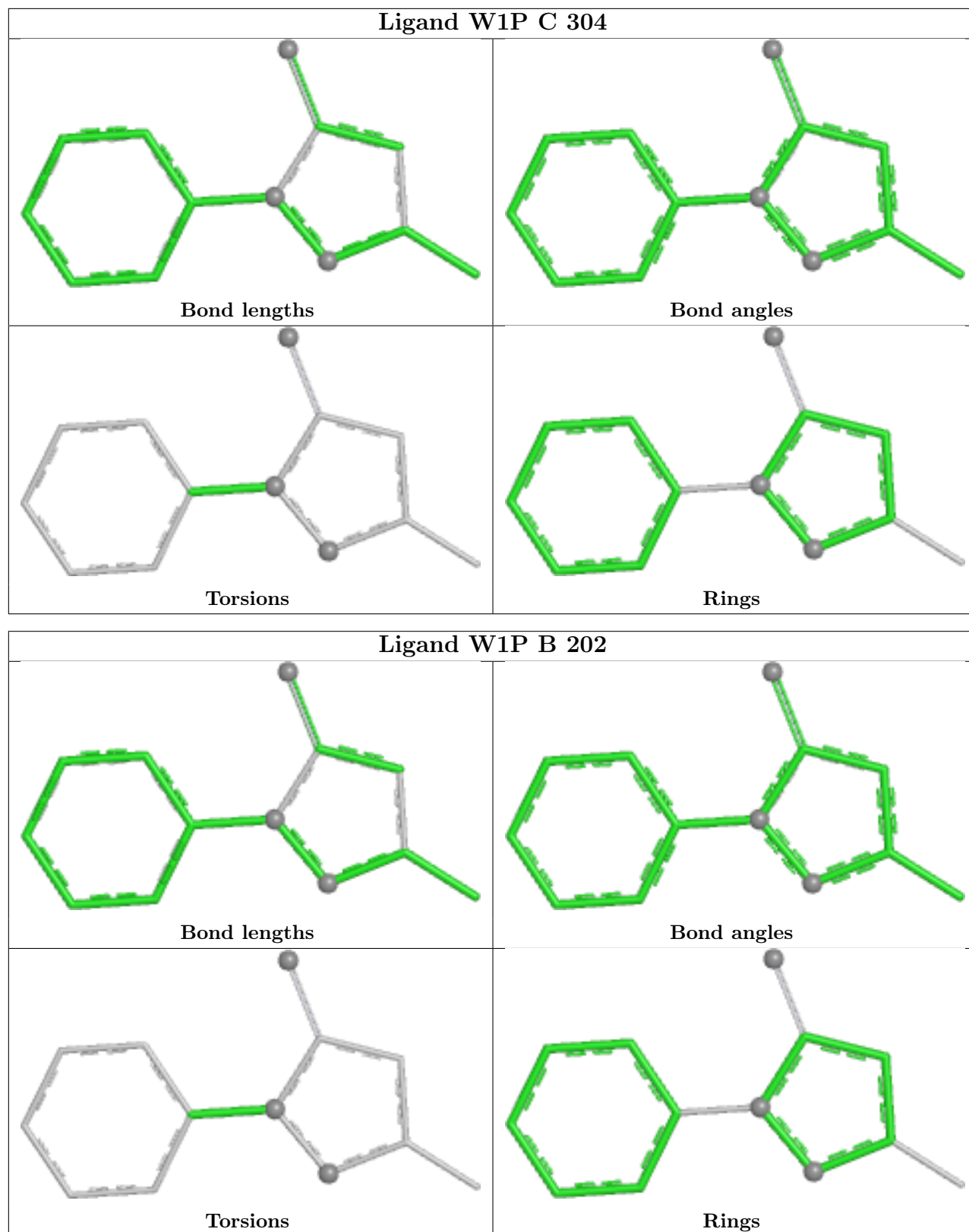
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

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 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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

**Warning:** The R factor obtained from EDS is 0.2544, which does not match the depositor's R factor of 0.18973. Please interpret the results in this section carefully.

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	163/163 (100%)	1.59	61 (37%) 1 0	5, 18, 32, 38	63 (38%)
1	B	163/163 (100%)	2.08	60 (36%) 1 0	5, 19, 32, 43	52 (31%)
1	C	163/163 (100%)	2.37	77 (47%) 0 0	6, 21, 34, 43	59 (36%)
1	D	159/163 (97%)	2.81	87 (54%) 0 0	6, 22, 40, 79	65 (40%)
All	All	648/652 (99%)	2.21	285 (43%) 1 0	5, 20, 34, 79	239 (36%)

All (285) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	160	THR	9.9
1	D	106[A]	ILE	8.6
1	B	106[A]	ILE	8.3
1	B	113	VAL	8.1
1	D	33	VAL	7.9
1	C	106[A]	ILE	7.9
1	B	11	ILE	7.8
1	D	32	GLY	7.5
1	B	-1	ALA	7.4
1	A	-1	ALA	7.4
1	D	131	ALA	7.4
1	D	113	VAL	7.2
1	B	112	GLY	7.1
1	B	105	ALA	6.9
1	C	-1	ALA	6.8
1	D	12	ALA	6.8
1	B	92	VAL	6.7
1	C	139	VAL	6.7
1	D	88	ALA	6.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	128	LEU	6.6
1	C	141	ILE	6.6
1	D	20	VAL	6.6
1	B	129	PHE	6.5
1	D	34	CYS	6.4
1	C	129	PHE	6.4
1	C	108	LEU	6.4
1	C	104	VAL	6.4
1	D	87	ALA	6.4
1	D	128	LEU	6.3
1	B	3	SER	6.3
1	D	11	ILE	6.3
1	B	32	GLY	6.3
1	B	2	PRO	6.2
1	D	29	PRO	6.2
1	D	97	THR	6.2
1	D	141	ILE	6.1
1	C	142	TYR	6.1
1	D	89	TYR	6.1
1	A	113	VAL	6.1
1	C	105	ALA	6.0
1	B	114	TYR	6.0
1	C	13[A]	LYS	6.0
1	B	89	TYR	6.0
1	D	143	CYS	5.9
1	D	142	TYR	5.9
1	B	117	GLY	5.9
1	C	4	TYR	5.8
1	B	1	ALA	5.8
1	C	143	CYS	5.8
1	B	0	MET	5.8
1	C	33	VAL	5.8
1	D	3[A]	SER	5.8
1	C	92	VAL	5.8
1	D	140	VAL	5.8
1	B	30	GLY	5.8
1	B	131	ALA	5.7
1	D	107	PRO	5.7
1	B	73	PHE	5.6
1	C	20	VAL	5.6
1	C	-2	GLY	5.6
1	D	30	GLY	5.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	111	THR	5.6
1	C	107	PRO	5.6
1	A	156	ILE	5.5
1	B	19	VAL	5.5
1	B	146	LYS	5.4
1	C	11	ILE	5.4
1	B	12	ALA	5.4
1	D	36	ALA	5.4
1	C	0	MET	5.4
1	C	90[A]	ARG	5.4
1	D	112	GLY	5.3
1	D	151	LYS	5.3
1	B	141	ILE	5.3
1	D	28	LEU	5.3
1	D	109	LEU	5.3
1	D	111	THR	5.2
1	C	93	ALA	5.2
1	D	150	LYS	5.2
1	D	129	PHE	5.1
1	B	150	LYS	5.1
1	D	114	TYR	5.1
1	D	85	LEU	5.0
1	A	112	GLY	5.0
1	D	86	ALA	5.0
1	A	104	VAL	5.0
1	A	106[A]	ILE	4.9
1	A	160	THR	4.9
1	A	107	PRO	4.9
1	B	20	VAL	4.8
1	B	56[A]	LYS	4.8
1	C	9	MET	4.8
1	C	160	THR	4.8
1	C	59[A]	MET	4.8
1	A	158	MET	4.7
1	B	109	LEU	4.7
1	B	158	MET	4.7
1	A	122	THR	4.7
1	C	130	THR	4.7
1	B	108	LEU	4.6
1	C	148	TRP	4.6
1	D	49	ALA	4.6
1	B	115	SER	4.6

*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	157	GLN	4.6
1	D	10	ASP	4.6
1	D	2	PRO	4.5
1	B	116	GLY	4.5
1	A	121	LEU	4.5
1	D	144	ARG	4.5
1	B	71	PRO	4.5
1	D	46	LYS	4.4
1	C	156	ILE	4.4
1	D	108	LEU	4.4
1	C	158	MET	4.4
1	D	148	TRP	4.4
1	A	4	TYR	4.4
1	B	-2	GLY	4.4
1	D	31	ASP	4.4
1	B	107	PRO	4.4
1	C	12	ALA	4.4
1	C	10	ASP	4.3
1	B	103	SER	4.3
1	C	155	ALA	4.3
1	D	35	LYS	4.3
1	A	76	TYR	4.3
1	B	31	ASP	4.3
1	A	11	ILE	4.2
1	A	129	PHE	4.2
1	B	91[A]	GLU	4.2
1	A	37	VAL	4.1
1	A	131	ALA	4.1
1	D	105	ALA	4.1
1	B	13	LYS	4.1
1	D	146	LYS	4.1
1	C	2	PRO	4.0
1	C	3	SER	4.0
1	D	81	GLY	4.0
1	A	111	THR	4.0
1	C	38	TYR	4.0
1	D	53	GLY	4.0
1	A	27	GLY	4.0
1	C	153	SER	4.0
1	D	59[A]	MET	3.9
1	C	138	ASP	3.9
1	B	138	ASP	3.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	0	MET	3.9
1	C	132	MET	3.9
1	D	118	LYS	3.8
1	D	54	THR	3.8
1	C	56	LYS	3.8
1	A	130	THR	3.8
1	D	90	ARG	3.8
1	A	105	ALA	3.8
1	C	146	LYS	3.8
1	A	128	LEU	3.8
1	A	-2	GLY	3.8
1	A	52	VAL	3.8
1	D	119	ASP	3.8
1	C	111	THR	3.7
1	A	103	SER	3.7
1	C	54	THR	3.7
1	B	5	ARG	3.7
1	D	45	PHE	3.7
1	D	9	MET	3.7
1	B	62	THR	3.6
1	B	10	ASP	3.6
1	D	158	MET	3.6
1	C	52	VAL	3.6
1	A	30	GLY	3.6
1	C	112	GLY	3.6
1	C	131	ALA	3.6
1	B	72	ASN	3.6
1	C	113	VAL	3.6
1	C	144	ARG	3.6
1	B	147	GLU	3.6
1	D	21	ASN	3.5
1	B	90	ARG	3.5
1	B	133	ASP	3.5
1	D	103	SER	3.5
1	C	98	ARG	3.5
1	D	39	LYS	3.5
1	A	25	PRO	3.5
1	D	157	GLN	3.5
1	D	76	TYR	3.5
1	D	50	THR	3.5
1	A	28	LEU	3.4
1	C	159	ARG	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	33	VAL	3.4
1	B	154	GLU	3.4
1	A	155	ALA	3.4
1	D	51	PRO	3.3
1	C	154	GLU	3.3
1	B	159	ARG	3.3
1	B	110	SER	3.3
1	C	103	SER	3.3
1	A	40	LYS	3.3
1	B	45	PHE	3.3
1	C	1	ALA	3.3
1	C	91[A]	GLU	3.3
1	C	45	PHE	3.3
1	A	118	LYS	3.2
1	C	118	LYS	3.2
1	C	47	ASN	3.2
1	A	159	ARG	3.2
1	D	48	SER	3.2
1	D	138	ASP	3.2
1	A	32	GLY	3.2
1	D	145	ASP	3.1
1	D	91	GLU	3.1
1	A	51	PRO	3.1
1	A	5	ARG	3.1
1	C	46	LYS	3.1
1	B	157	GLN	3.1
1	B	38	TYR	3.0
1	C	134	SER	3.0
1	D	133	ASP	3.0
1	C	32	GLY	3.0
1	B	15	ASP	3.0
1	A	77	THR	3.0
1	C	26	ARG	3.0
1	B	7	LYS	3.0
1	C	5	ARG	2.9
1	A	36	ALA	2.9
1	A	50	THR	2.9
1	A	157	GLN	2.9
1	C	150	LYS	2.9
1	D	4	TYR	2.8
1	A	83	ARG	2.8
1	A	59[A]	MET	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	D	101	VAL	2.8
1	A	29	PRO	2.8
1	A	23	ALA	2.8
1	A	146	LYS	2.7
1	B	148	TRP	2.7
1	D	37	VAL	2.7
1	C	145	ASP	2.7
1	C	41	TRP	2.7
1	D	41	TRP	2.7
1	C	74	SER	2.6
1	D	44	SER	2.6
1	A	54	THR	2.6
1	D	155	ALA	2.6
1	C	48	SER	2.6
1	D	147	GLU	2.6
1	B	48	SER	2.6
1	A	73	PHE	2.6
1	A	35	LYS	2.6
1	A	34	CYS	2.6
1	A	3	SER	2.6
1	A	134	SER	2.6
1	C	63	TYR	2.5
1	B	160	THR	2.5
1	D	83	ARG	2.5
1	A	114	TYR	2.5
1	D	8	ARG	2.5
1	D	79	SER	2.4
1	A	123	GLN	2.4
1	D	156	ILE	2.4
1	C	44	SER	2.4
1	C	58	VAL	2.4
1	C	114	TYR	2.3
1	A	46	LYS	2.3
1	C	133	ASP	2.3
1	A	154	GLU	2.3
1	D	136	ASP	2.3
1	C	36	ALA	2.3
1	D	55	ALA	2.3
1	A	124	SER	2.3
1	D	84	GLU	2.3
1	D	116	GLY	2.2
1	A	49	ALA	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	100	GLY	2.2
1	D	154	GLU	2.2
1	D	72	ASN	2.2
1	A	12	ALA	2.1
1	D	121	LEU	2.1
1	C	15	ASP	2.1
1	C	82	ASP	2.1
1	D	130	THR	2.1
1	A	48	SER	2.1
1	D	80	GLU	2.1
1	A	39	LYS	2.0
1	C	23	ALA	2.0
1	C	147	GLU	2.0
1	C	39	LYS	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
2	DMS	A	209	4/4	0.56	0.44	42,53,55,56	4
6	K34	B	203	10/10	0.56	0.29	29,31,33,34	10
5	W1P	B	202	13/13	0.62	0.25	28,30,32,32	13
5	W1P	A	207	13/13	0.66	0.24	30,32,34,34	13
3	TRS	D	201	8/8	0.70	0.30	49,52,53,55	8
5	W1P	D	205	13/13	0.71	0.18	28,28,29,29	13
2	DMS	D	203	4/4	0.71	0.28	45,49,51,52	4
6	K34	D	206	10/10	0.72	0.22	31,33,34,35	10
2	DMS	A	204	4/4	0.74	0.26	37,40,41,45	4

*Continued on next page...*

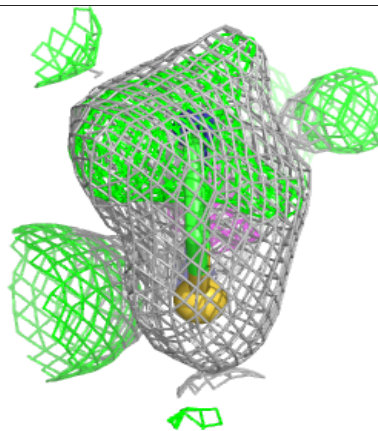
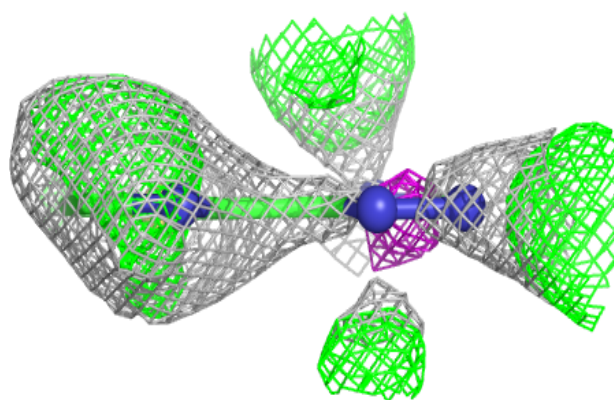
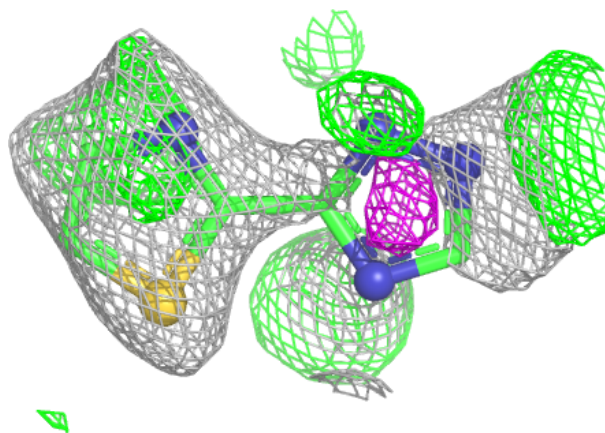
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	DMS	A	208	4/4	0.79	0.23	33,35,36,37	4
2	DMS	D	202	4/4	0.80	0.26	43,44,45,47	4
3	TRS	A	202	8/8	0.80	0.20	41,44,46,47	8
5	W1P	C	304	13/13	0.81	0.16	20,24,25,25	13
2	DMS	A	203	4/4	0.83	0.21	38,41,41,44	4
2	DMS	A	211	4/4	0.88	0.17	58,66,67,71	0
2	DMS	C	303	4/4	0.88	0.16	47,49,52,52	0
2	DMS	A	201	4/4	0.90	0.14	43,48,50,51	0
4	CL	C	302	1/1	0.90	0.14	41,41,41,41	0
2	DMS	A	210	4/4	0.94	0.14	30,34,35,37	0
6	K34	C	305	10/10	0.95	0.09	14,14,15,16	10
4	CL	D	204	1/1	0.97	0.06	38,38,38,38	0
4	CL	C	301	1/1	0.97	0.07	32,32,32,32	0
4	CL	A	206	1/1	0.99	0.10	23,23,23,23	0
4	CL	B	201	1/1	0.99	0.03	24,24,24,24	0
4	CL	A	205	1/1	0.99	0.03	23,23,23,23	0

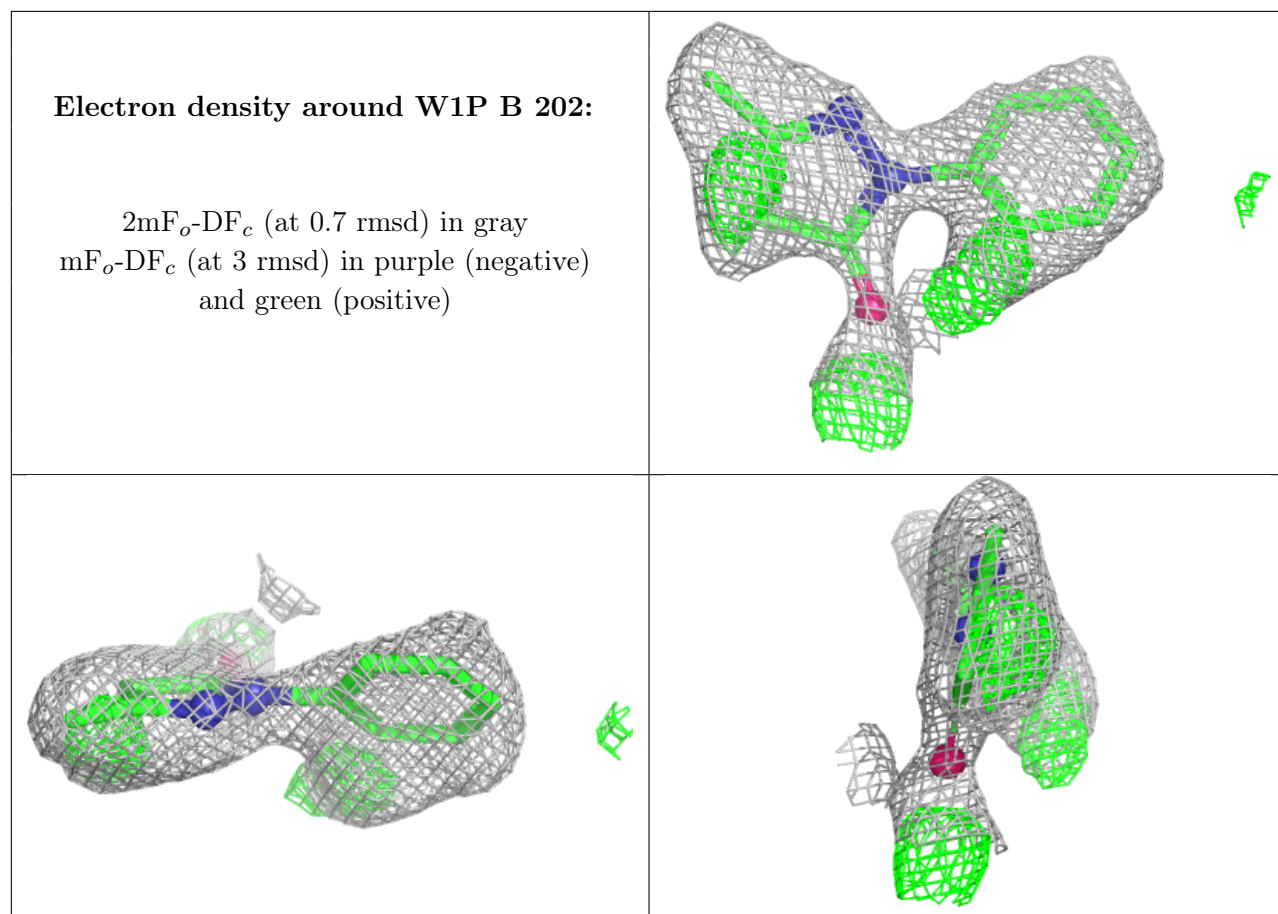
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around K34 B 203:**

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

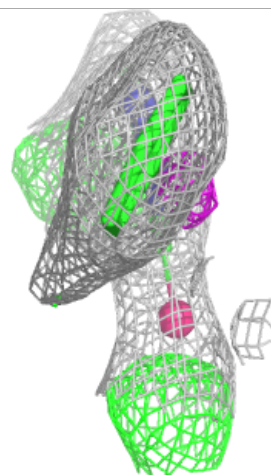
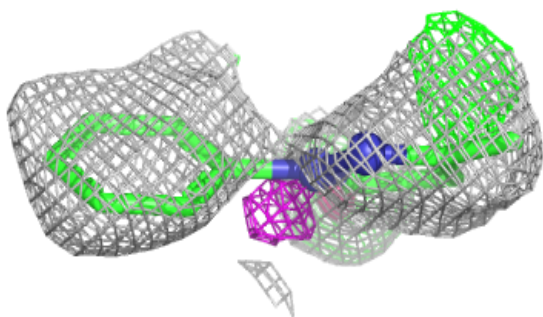
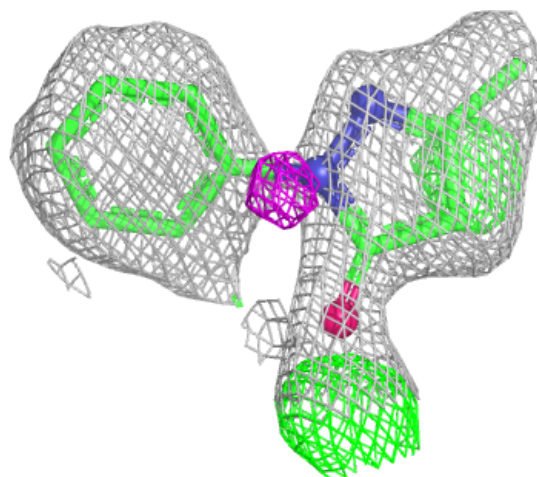






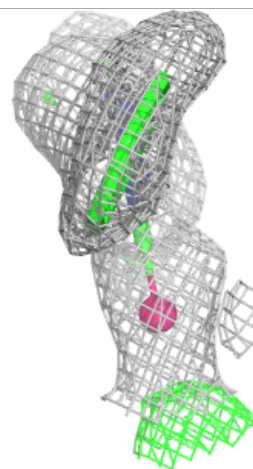
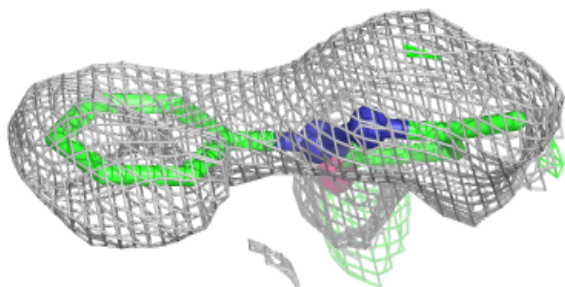
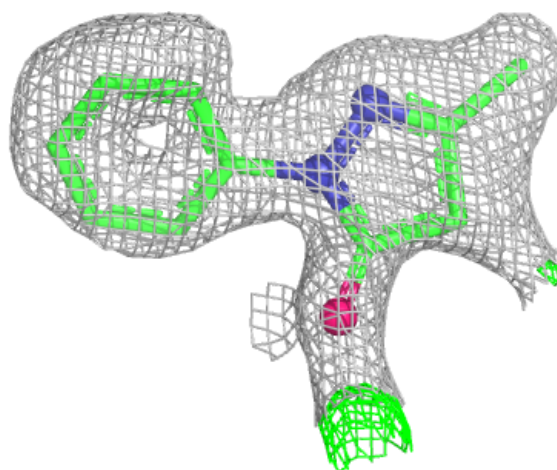
**Electron density around W1P A 207:**

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



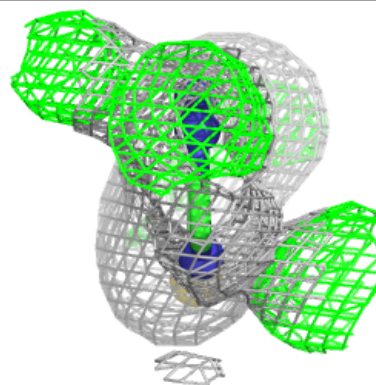
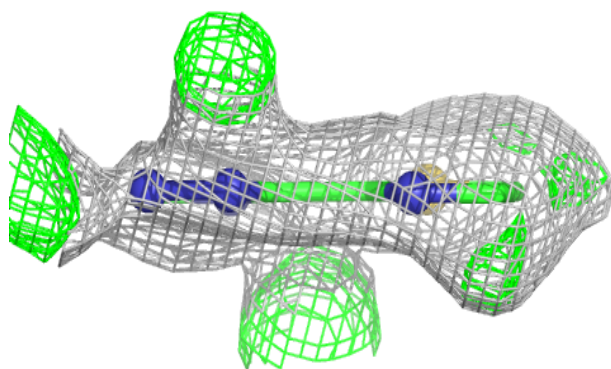
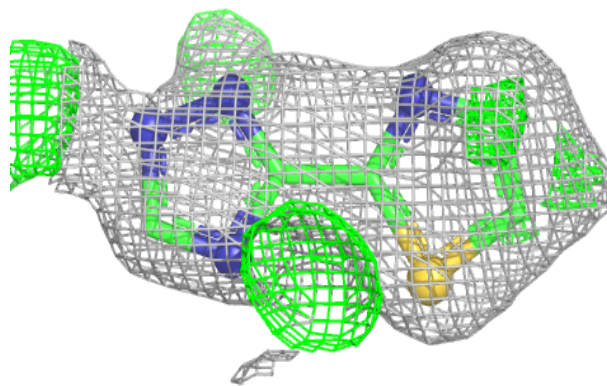
**Electron density around W1P D 205:**

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



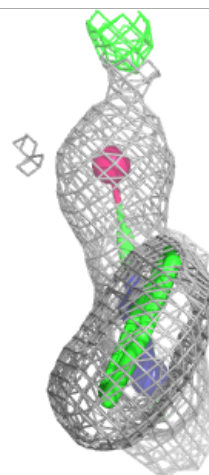
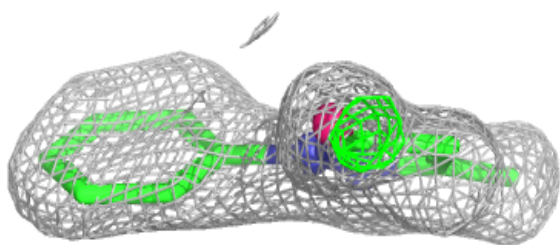
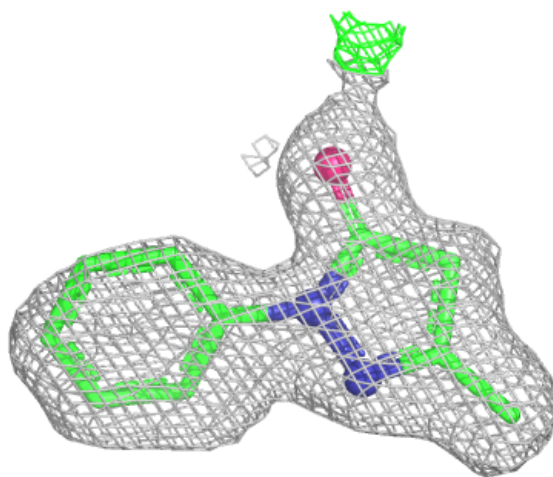
**Electron density around K34 D 206:**

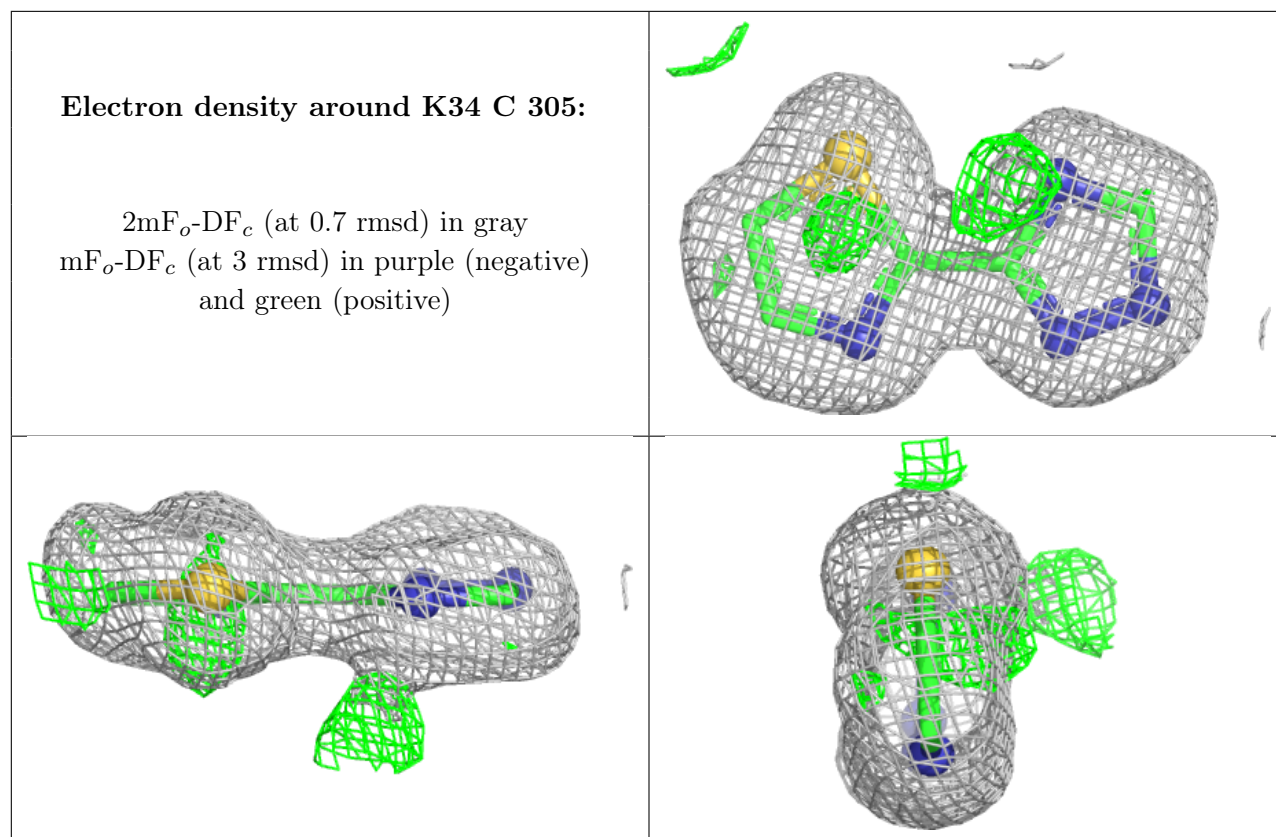
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around W1P C 304:**

$2mF_o-DF_c$  (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.