



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 06:04 pm BST

PDB ID : 5TCI  
Title : Crystal structure of tryptophan synthase from *M. tuberculosis* - BRD4592-bound form  
Authors : Michalska, K.; Maltseva, N.; Jedrzejczak, R.; Wellington, S.; Nag, P.P.; Fisher, S.L.; Schreiber, S.L.; Hung, D.T.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2016-09-15  
Resolution : 2.45 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

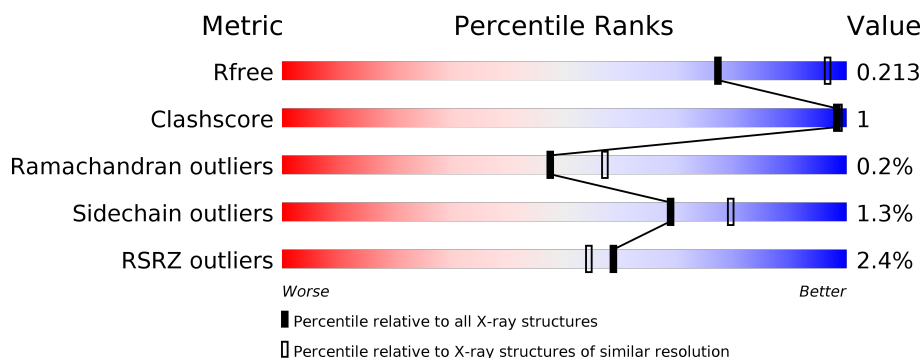
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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 on 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	276	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>%</span> <span>88%</span> <span>10%</span> </div> </div>
1	C	276	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>4%</span> <span>88%</span> <span>10%</span> </div> </div>
1	E	276	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>6%</span> <span>87%</span> <span>11%</span> </div> </div>
1	G	276	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>2%</span> <span>87%</span> <span>11%</span> </div> </div>
2	B	410	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>2%</span> <span>98%</span> <span>.</span> </div> </div>
2	D	410	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span></span> <span>95%</span> <span>..</span> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	410	<div><div></div><div>95%</div><div></div><div></div></div>
2	H	410	<div><div>3%</div><div></div><div>94%</div><div>5%</div><div></div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 20181 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 Tryptophan synthase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	249	Total	C	N	O	S	0	0	0
			1803	1131	323	344	5			
1	G	247	Total	C	N	O	S	0	0	0
			1790	1124	321	340	5			
1	E	245	Total	C	N	O	S	0	0	0
			1776	1115	319	337	5			
1	C	249	Total	C	N	O	S	0	0	0
			1806	1133	323	345	5			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	271	HIS	-	expression tag	UNP P9WFY1
A	272	HIS	-	expression tag	UNP P9WFY1
A	273	HIS	-	expression tag	UNP P9WFY1
A	274	HIS	-	expression tag	UNP P9WFY1
A	275	HIS	-	expression tag	UNP P9WFY1
A	276	HIS	-	expression tag	UNP P9WFY1
G	271	HIS	-	expression tag	UNP P9WFY1
G	272	HIS	-	expression tag	UNP P9WFY1
G	273	HIS	-	expression tag	UNP P9WFY1
G	274	HIS	-	expression tag	UNP P9WFY1
G	275	HIS	-	expression tag	UNP P9WFY1
G	276	HIS	-	expression tag	UNP P9WFY1
E	271	HIS	-	expression tag	UNP P9WFY1
E	272	HIS	-	expression tag	UNP P9WFY1
E	273	HIS	-	expression tag	UNP P9WFY1
E	274	HIS	-	expression tag	UNP P9WFY1
E	275	HIS	-	expression tag	UNP P9WFY1
E	276	HIS	-	expression tag	UNP P9WFY1
C	271	HIS	-	expression tag	UNP P9WFY1
C	272	HIS	-	expression tag	UNP P9WFY1
C	273	HIS	-	expression tag	UNP P9WFY1

*Continued on next page...*

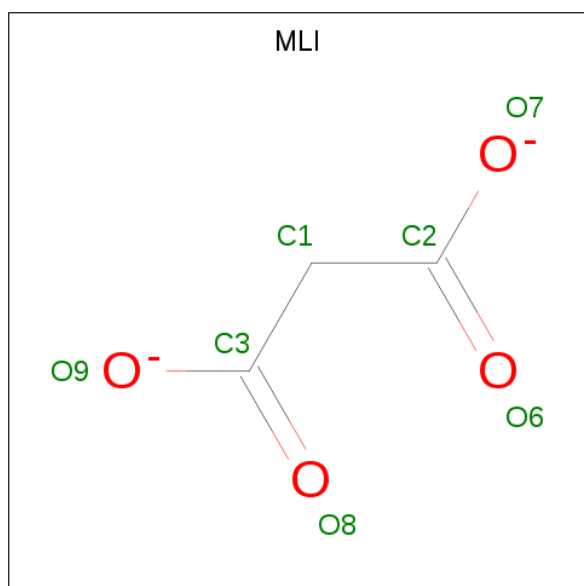
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	274	HIS	-	expression tag	UNP P9WFY1
C	275	HIS	-	expression tag	UNP P9WFY1
C	276	HIS	-	expression tag	UNP P9WFY1

- Molecule 2 is a protein called Tryptophan synthase beta chain.

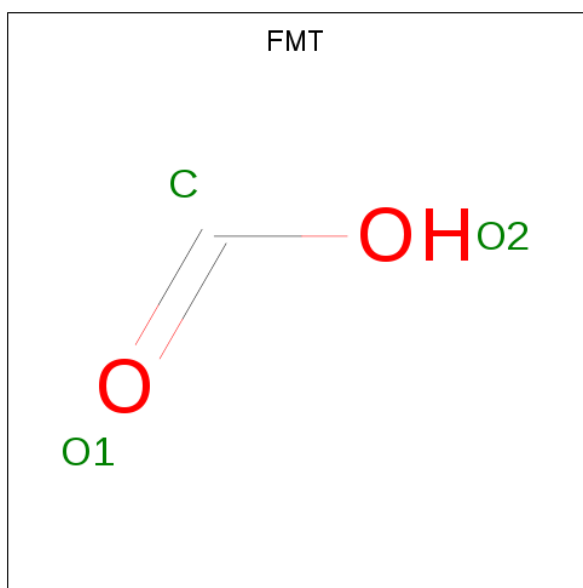
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	405	Total	C	N	O	P	S	0	4	0
			3061	1906	556	585	1	13			
2	H	406	Total	C	N	O	P	S	0	33	0
			3286	2047	595	630	1	13			
2	F	399	Total	C	N	O	P	S	0	0	0
			2987	1865	539	569	1	13			
2	D	398	Total	C	N	O	P	S	0	0	0
			2980	1861	538	567	1	13			

- Molecule 3 is MALONATE ION (three-letter code: MLI) (formula:  $C_3H_2O_4$ ).



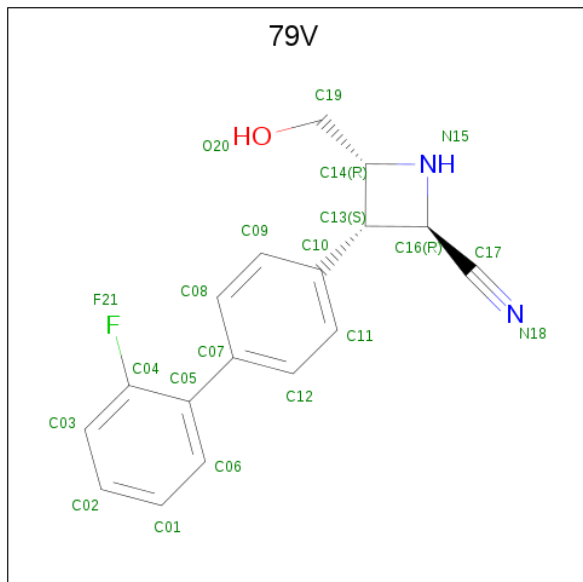
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			7	3	4		
3	G	1	Total	C	O	0	0
			7	3	4		
3	E	1	Total	C	O	0	0
			7	3	4		
3	C	1	Total	C	O	0	0
			7	3	4		

- Molecule 4 is FORMIC ACID (three-letter code: FMT) (formula: CH<sub>2</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			3	1	2		
4	B	1	Total	C	O	0	0
			3	1	2		
4	B	1	Total	C	O	0	0
			3	1	2		
4	B	1	Total	C	O	0	0
			3	1	2		
4	G	1	Total	C	O	0	0
			3	1	2		
4	G	1	Total	C	O	0	0
			3	1	2		
4	H	1	Total	C	O	0	0
			3	1	2		
4	H	1	Total	C	O	0	0
			3	1	2		
4	F	1	Total	C	O	0	0
			3	1	2		
4	F	1	Total	C	O	0	0
			3	1	2		
4	C	1	Total	C	O	0	0
			3	1	2		
4	D	1	Total	C	O	0	0
			3	1	2		
4	D	1	Total	C	O	0	0
			3	1	2		

- Molecule 5 is (2R,3S,4R)-3-(2'-fluoro[1,1'-biphenyl]-4-yl)-4-(hydroxymethyl)azetidine-2-carbonitrile (three-letter code: 79V) (formula: C<sub>17</sub>H<sub>15</sub>FN<sub>2</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	F	N	O	0	0
			21	17	1	2	1		
5	H	1	Total	C	F	N	O	0	1
			42	34	2	4	2		
5	F	1	Total	C	F	N	O	0	0
			21	17	1	2	1		
5	D	1	Total	C	F	N	O	0	0
			21	17	1	2	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	37	Total	O	0	0
			37	37		
6	B	99	Total	O	0	1
			99	99		
6	G	46	Total	O	0	1
			46	46		
6	H	105	Total	O	0	0
			105	105		
6	E	14	Total	O	0	0
			14	14		
6	F	86	Total	O	0	0
			86	86		

Continued on next page...

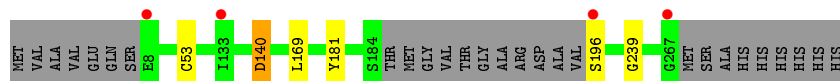
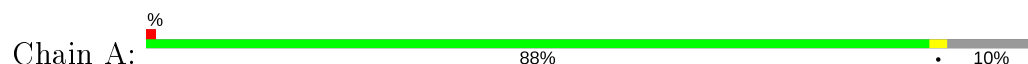
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	32	Total 32	O 32	0	0
6	D	101	Total 101	O 101	0	0

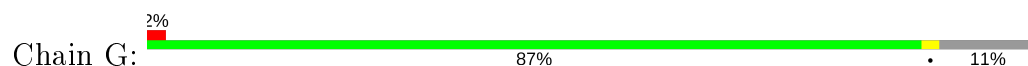
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

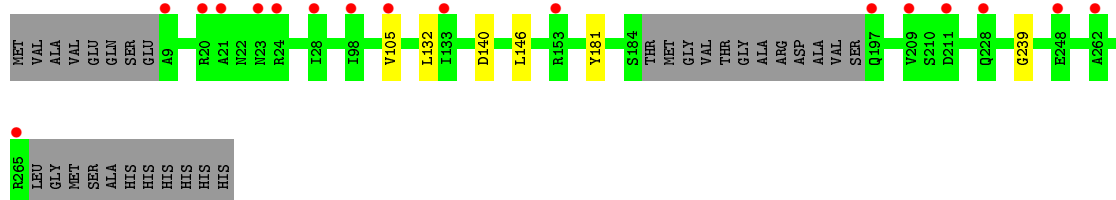
- Molecule 1: Tryptophan synthase alpha chain



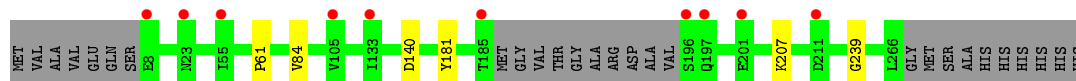
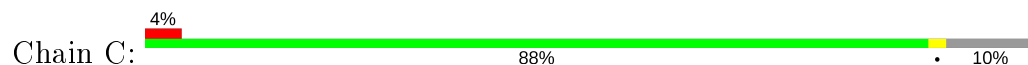
- Molecule 1: Tryptophan synthase alpha chain



- Molecule 1: Tryptophan synthase alpha chain

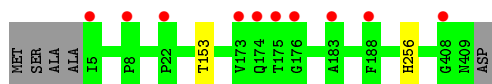


- Molecule 1: Tryptophan synthase alpha chain



- Molecule 2: Tryptophan synthase beta chain





- Molecule 2: Tryptophan synthase beta chain



- Molecule 2: Tryptophan synthase beta chain



- Molecule 2: Tryptophan synthase beta chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.96Å 158.34Å 166.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.45 29.95 – 2.45	Depositor EDS
% Data completeness (in resolution range)	99.5 (30.00-2.45) 99.6 (29.95-2.45)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.45 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.186 , 0.212 0.191 , 0.213	Depositor DCC
$R_{free}$ test set	2558 reflections (1.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.3	Xtriage
Anisotropy	0.331	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 25.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.002 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	20181	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% 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: 79V, FMT, MLI, LLP

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	0.44	0/1832	0.66	0/2500
1	C	0.45	0/1835	0.63	0/2505
1	E	0.44	0/1805	0.62	0/2464
1	G	0.47	0/1819	0.66	0/2483
2	B	0.47	0/3097	0.67	0/4197
2	D	0.47	0/3015	0.66	0/4085
2	F	0.45	0/3022	0.67	0/4095
2	H	0.44	0/3326	0.65	0/4512
All	All	0.45	0/19751	0.66	0/26841

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	1803	0	1822	1	0
1	C	1806	0	1826	1	0
1	E	1776	0	1797	1	0
1	G	1790	0	1813	1	0
2	B	3061	0	2962	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	2980	0	2894	4	0
2	F	2987	0	2901	4	0
2	H	3286	0	3179	12	0
3	A	7	0	2	0	0
3	C	7	0	2	0	0
3	E	7	0	2	0	0
3	G	7	0	2	0	0
4	A	3	0	1	0	0
4	B	9	0	3	0	0
4	C	3	0	1	0	0
4	D	6	0	2	0	0
4	F	6	0	2	0	0
4	G	6	0	2	0	0
4	H	6	0	2	0	0
5	B	21	0	0	0	0
5	D	21	0	0	0	0
5	F	21	0	0	0	0
5	H	42	0	0	1	0
6	A	37	0	0	0	0
6	B	99	0	0	0	0
6	C	32	0	0	0	0
6	D	101	0	0	0	0
6	E	14	0	0	0	0
6	F	86	0	0	0	0
6	G	46	0	0	0	0
6	H	105	0	0	0	0
All	All	20181	0	19215	25	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 1.

All (25) 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)
2:H:186[A]:GLU:OE1	2:H:186[A]:GLU:HA	1.66	0.93
2:H:186[A]:GLU:OE1	2:H:186[A]:GLU:CA	2.31	0.79
5:H:501[A]:79V:C11	5:H:501[A]:79V:C19	2.79	0.61
2:H:190[A]:ASP:OD1	2:H:194[A]:ASN:ND2	2.38	0.57
2:D:173:VAL:HG13	2:D:186:GLU:HG3	1.93	0.51
2:H:170[A]:VAL:O	2:H:170[A]:VAL:HG12	2.13	0.48
2:D:207:GLY:HA2	2:D:294:HIS:O	2.15	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:293:PHE:HB2	2:D:300:LEU:HD22	1.95	0.47
1:E:105:VAL:HB	1:E:132:LEU:HD12	1.97	0.47
2:F:230:ILE:HG21	2:F:238:PRO:HD3	1.97	0.46
2:H:181[A]:LYS:CG	2:H:182[A]:ASP:N	2.81	0.44
1:A:140:ASP:OD1	1:A:140:ASP:N	2.51	0.43
2:H:129:HIS:CE1	2:H:203:GLY:HA2	2.54	0.43
2:F:129:HIS:CE1	2:F:203:GLY:HA2	2.54	0.43
2:H:102:ILE:HA	2:H:105:VAL:HG22	2.00	0.43
2:H:207:GLY:HA2	2:H:294[A]:HIS:O	2.19	0.42
1:C:61:PRO:HG2	1:C:84:VAL:HG11	2.00	0.42
2:H:123:GLU:HG3	2:H:184[B]:ILE:HG12	2.01	0.42
2:H:294[A]:HIS:ND1	2:H:320:TYR:OH	2.52	0.42
2:D:17:GLY:HA3	2:D:209:HIS:HB2	2.01	0.42
1:G:124:LEU:HD22	1:G:132:LEU:HD13	2.02	0.42
2:F:61:LEU:HD12	2:F:110:LEU:HD11	2.03	0.41
2:H:170[A]:VAL:O	2:H:171[A]:VAL:C	2.58	0.41
2:H:100:HIS:NE2	2:H:250:ASN:HB3	2.35	0.41
2:F:121:ILE:HG22	2:F:145:VAL:HB	2.04	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	245/276 (89%)	242 (99%)	2 (1%)	1 (0%)	34	41
1	C	245/276 (89%)	240 (98%)	4 (2%)	1 (0%)	34	41
1	E	241/276 (87%)	238 (99%)	2 (1%)	1 (0%)	34	41
1	G	243/276 (88%)	240 (99%)	2 (1%)	1 (0%)	34	41
2	B	406/410 (99%)	393 (97%)	13 (3%)	0	100	100
2	D	395/410 (96%)	386 (98%)	9 (2%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	396/410 (97%)	388 (98%)	8 (2%)	0	100	100
2	H	436/410 (106%)	423 (97%)	13 (3%)	0	100	100
All	All	2607/2744 (95%)	2550 (98%)	53 (2%)	4 (0%)	47	57

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	239	GLY
1	A	239	GLY
1	G	239	GLY
1	C	239	GLY

### 5.3.2 Protein sidechains ⓘ

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	179/200 (90%)	174 (97%)	5 (3%)	43	56
1	C	180/200 (90%)	177 (98%)	3 (2%)	60	73
1	E	176/200 (88%)	173 (98%)	3 (2%)	60	73
1	G	178/200 (89%)	175 (98%)	3 (2%)	60	73
2	B	301/301 (100%)	299 (99%)	2 (1%)	84	90
2	D	293/301 (97%)	291 (99%)	2 (1%)	84	90
2	F	294/301 (98%)	293 (100%)	1 (0%)	92	95
2	H	323/301 (107%)	314 (97%)	9 (3%)	43	56
All	All	1924/2004 (96%)	1896 (98%)	28 (2%)	69	76

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	CYS
1	A	140	ASP
1	A	169	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	181	TYR
1	A	196	SER
2	B	153	THR
2	B	256	HIS
1	G	20	ARG
1	G	140	ASP
1	G	181	TYR
2	H	186[A]	GLU
2	H	186[B]	GLU
2	H	196[A]	ASP
2	H	196[B]	ASP
2	H	256	HIS
2	H	294[A]	HIS
2	H	294[B]	HIS
2	H	300	LEU
2	H	348	ASP
1	E	140	ASP
1	E	146	LEU
1	E	181	TYR
2	F	300	LEU
1	C	140	ASP
1	C	181	TYR
1	C	207	LYS
2	D	23	SER
2	D	34	LEU

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

Mol	Chain	Res	Type
1	A	231	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	LLP	D	101	2	23,24,25	2.61	6 (26%)	25,32,34	1.56	5 (20%)
2	LLP	B	101	2	23,24,25	2.55	7 (30%)	25,32,34	1.52	5 (20%)
2	LLP	H	101	2	23,24,25	2.57	5 (21%)	25,32,34	1.71	5 (20%)
2	LLP	F	101	2	23,24,25	2.51	6 (26%)	25,32,34	1.63	4 (16%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	LLP	D	101	2	-	4/16/17/19	0/1/1/1
2	LLP	B	101	2	-	4/16/17/19	0/1/1/1
2	LLP	H	101	2	-	5/16/17/19	0/1/1/1
2	LLP	F	101	2	-	4/16/17/19	0/1/1/1

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	101	LLP	C4-C4'	8.12	1.62	1.46
2	F	101	LLP	C4-C4'	8.09	1.62	1.46
2	D	101	LLP	C4-C4'	8.06	1.61	1.46
2	B	101	LLP	C4-C4'	8.05	1.61	1.46
2	D	101	LLP	C4'-NZ	5.29	1.45	1.27
2	B	101	LLP	C4'-NZ	5.26	1.44	1.27
2	F	101	LLP	C4'-NZ	5.19	1.44	1.27
2	H	101	LLP	C4'-NZ	4.64	1.42	1.27
2	H	101	LLP	C4-C5	-4.52	1.36	1.42
2	D	101	LLP	C4-C5	-3.54	1.37	1.42
2	H	101	LLP	C2'-C2	3.53	1.56	1.50
2	B	101	LLP	C4-C5	-3.41	1.37	1.42
2	F	101	LLP	C4-C5	-3.26	1.37	1.42
2	D	101	LLP	C6-N1	3.25	1.41	1.34
2	F	101	LLP	C6-N1	3.05	1.40	1.34
2	B	101	LLP	C6-N1	2.97	1.40	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	101	LLP	C2'-C2	2.95	1.55	1.50
2	H	101	LLP	C6-N1	2.94	1.40	1.34
2	B	101	LLP	C2'-C2	2.83	1.55	1.50
2	F	101	LLP	C2'-C2	2.55	1.54	1.50
2	B	101	LLP	P-OP4	-2.22	1.53	1.60
2	B	101	LLP	C5'-C5	2.20	1.56	1.50
2	D	101	LLP	C5'-C5	2.12	1.56	1.50
2	F	101	LLP	P-OP4	-2.09	1.53	1.60

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	101	LLP	C4-C4'-NZ	-5.44	99.31	124.31
2	D	101	LLP	C4-C4'-NZ	-4.87	101.97	124.31
2	F	101	LLP	C4-C4'-NZ	-4.83	102.15	124.31
2	B	101	LLP	C4-C4'-NZ	-3.74	107.12	124.31
2	B	101	LLP	OP4-C5'-C5	3.61	116.23	109.35
2	D	101	LLP	OP4-C5'-C5	2.68	114.47	109.35
2	F	101	LLP	OP4-C5'-C5	2.37	113.87	109.35
2	F	101	LLP	C3-C4-C4'	-2.34	116.06	120.41
2	D	101	LLP	OP4-P-OP1	2.29	112.89	106.47
2	H	101	LLP	OP4-C5'-C5	2.26	113.66	109.35
2	D	101	LLP	C4-C3-C2	2.25	121.58	120.19
2	B	101	LLP	C3-C4-C5	2.21	119.96	118.26
2	F	101	LLP	C3-C4-C5	2.21	119.96	118.26
2	H	101	LLP	C3-C4-C4'	-2.19	116.33	120.41
2	D	101	LLP	C3-C4-C4'	-2.17	116.37	120.41
2	H	101	LLP	C3-C4-C5	2.14	119.90	118.26
2	B	101	LLP	C2'-C2-C3	-2.05	118.36	120.89
2	B	101	LLP	C2'-C2-N1	2.00	121.58	117.67
2	H	101	LLP	C2'-C2-N1	2.00	121.58	117.67

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	101	LLP	C4-C4'-NZ-CE
2	H	101	LLP	C5'-OP4-P-OP1
2	H	101	LLP	O-C-CA-CB
2	F	101	LLP	C4-C4'-NZ-CE
2	B	101	LLP	CG-CD-CE-NZ
2	H	101	LLP	C5-C4-C4'-NZ

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	D	101	LLP	C3-C4-C4'-NZ
2	H	101	LLP	C3-C4-C4'-NZ
2	D	101	LLP	CD-CE-NZ-C4'
2	F	101	LLP	CD-CE-NZ-C4'
2	B	101	LLP	CD-CE-NZ-C4'
2	D	101	LLP	C5-C4-C4'-NZ
2	F	101	LLP	C3-C4-C4'-NZ
2	D	101	LLP	C4-C4'-NZ-CE
2	H	101	LLP	CD-CE-NZ-C4'
2	F	101	LLP	C5-C4-C4'-NZ
2	B	101	LLP	C3-C4-C4'-NZ

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

22 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	79V	F	501	-	22,23,23	2.56	8 (36%)	22,32,32	1.01	1 (4%)
3	MLI	G	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	FMT	B	503	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	G	302	-	0,2,2	0.00	-	0,1,1	0.00	-
5	79V	H	501[B]	-	22,23,23	2.60	9 (40%)	22,32,32	1.25	3 (13%)
5	79V	B	501	-	22,23,23	1.96	9 (40%)	22,32,32	0.95	1 (4%)
3	MLI	A	301	-	0,6,6	0.00	-	0,7,7	0.00	-
5	79V	H	501[A]	-	22,23,23	2.22	6 (27%)	22,32,32	1.01	1 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	FMT	H	502	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	H	503	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	G	303	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	D	502	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	F	502	-	0,2,2	0.00	-	0,1,1	0.00	-
3	MLI	C	301	-	0,6,6	0.00	-	0,7,7	0.00	-
3	MLI	E	301	-	0,6,6	0.00	-	0,7,7	0.00	-
4	FMT	B	504	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	D	503	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	F	503	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	C	302	-	0,2,2	0.00	-	0,1,1	0.00	-
5	79V	D	501	-	22,23,23	2.52	6 (27%)	22,32,32	1.08	1 (4%)
4	FMT	B	502	-	0,2,2	0.00	-	0,1,1	0.00	-
4	FMT	A	302	-	0,2,2	0.00	-	0,1,1	0.00	-

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	79V	B	501	-	-	1/9/24/24	0/3/3/3
3	MLI	A	301	-	-	0/0/4/4	-
5	79V	D	501	-	-	1/9/24/24	0/3/3/3
5	79V	H	501[A]	-	-	1/9/24/24	0/3/3/3
5	79V	F	501	-	-	2/9/24/24	0/3/3/3
3	MLI	G	301	-	-	0/0/4/4	-
3	MLI	C	301	-	-	0/0/4/4	-
3	MLI	E	301	-	-	0/0/4/4	-
5	79V	H	501[B]	-	-	9/9/24/24	0/3/3/3

All (38) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	501	79V	C14-N15	8.47	1.56	1.48
5	F	501	79V	C14-N15	7.77	1.55	1.48
5	H	501[B]	79V	C14-N15	7.44	1.55	1.48
5	H	501[A]	79V	C14-N15	6.27	1.53	1.48
5	F	501	79V	C05-C04	4.51	1.45	1.39
5	D	501	79V	C05-C04	4.30	1.45	1.39
5	H	501[B]	79V	C05-C04	4.23	1.45	1.39
5	B	501	79V	C14-N15	4.20	1.52	1.48

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	501[A]	79V	C05-C04	3.42	1.43	1.39
5	D	501	79V	C09-C10	3.14	1.44	1.39
5	H	501[A]	79V	C05-C07	3.01	1.54	1.49
5	F	501	79V	C11-C10	2.88	1.43	1.39
5	H	501[B]	79V	C03-C04	2.86	1.43	1.37
5	B	501	79V	C03-C04	2.82	1.43	1.37
5	H	501[B]	79V	C10-C13	2.72	1.57	1.51
5	H	501[B]	79V	C12-C11	2.71	1.43	1.38
5	H	501[B]	79V	C05-C07	2.68	1.54	1.49
5	F	501	79V	C12-C11	2.65	1.43	1.38
5	F	501	79V	C16-N15	2.59	1.53	1.48
5	H	501[A]	79V	C16-N15	2.56	1.53	1.48
5	F	501	79V	C03-C04	2.56	1.43	1.37
5	H	501[A]	79V	C03-C04	2.55	1.43	1.37
5	B	501	79V	C09-C10	2.54	1.43	1.39
5	H	501[B]	79V	C11-C10	2.53	1.43	1.39
5	H	501[B]	79V	C09-C10	2.52	1.43	1.39
5	H	501[A]	79V	C09-C10	2.46	1.43	1.39
5	F	501	79V	C05-C07	2.45	1.53	1.49
5	B	501	79V	C11-C10	2.43	1.43	1.39
5	D	501	79V	C03-C04	2.40	1.43	1.37
5	B	501	79V	O20-C19	2.40	1.52	1.42
5	B	501	79V	C16-N15	2.39	1.52	1.48
5	B	501	79V	C05-C04	2.33	1.42	1.39
5	F	501	79V	C09-C10	2.31	1.42	1.39
5	D	501	79V	C16-N15	2.29	1.52	1.48
5	H	501[B]	79V	C09-C08	2.14	1.42	1.38
5	B	501	79V	C06-C05	2.14	1.43	1.40
5	B	501	79V	C12-C11	2.10	1.42	1.38
5	D	501	79V	C10-C13	2.01	1.56	1.51

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	H	501[A]	79V	C03-C04-C05	-2.74	119.95	123.30
5	D	501	79V	C03-C04-C05	-2.61	120.11	123.30
5	B	501	79V	C03-C04-C05	-2.47	120.28	123.30
5	H	501[B]	79V	C09-C10-C11	-2.44	115.25	118.29
5	H	501[B]	79V	O20-C19-C14	2.42	116.95	111.09
5	H	501[B]	79V	C03-C04-C05	-2.30	120.49	123.30
5	F	501	79V	C03-C04-C05	-2.26	120.53	123.30

There are no chirality outliers.

All (14) torsion outliers are listed below:

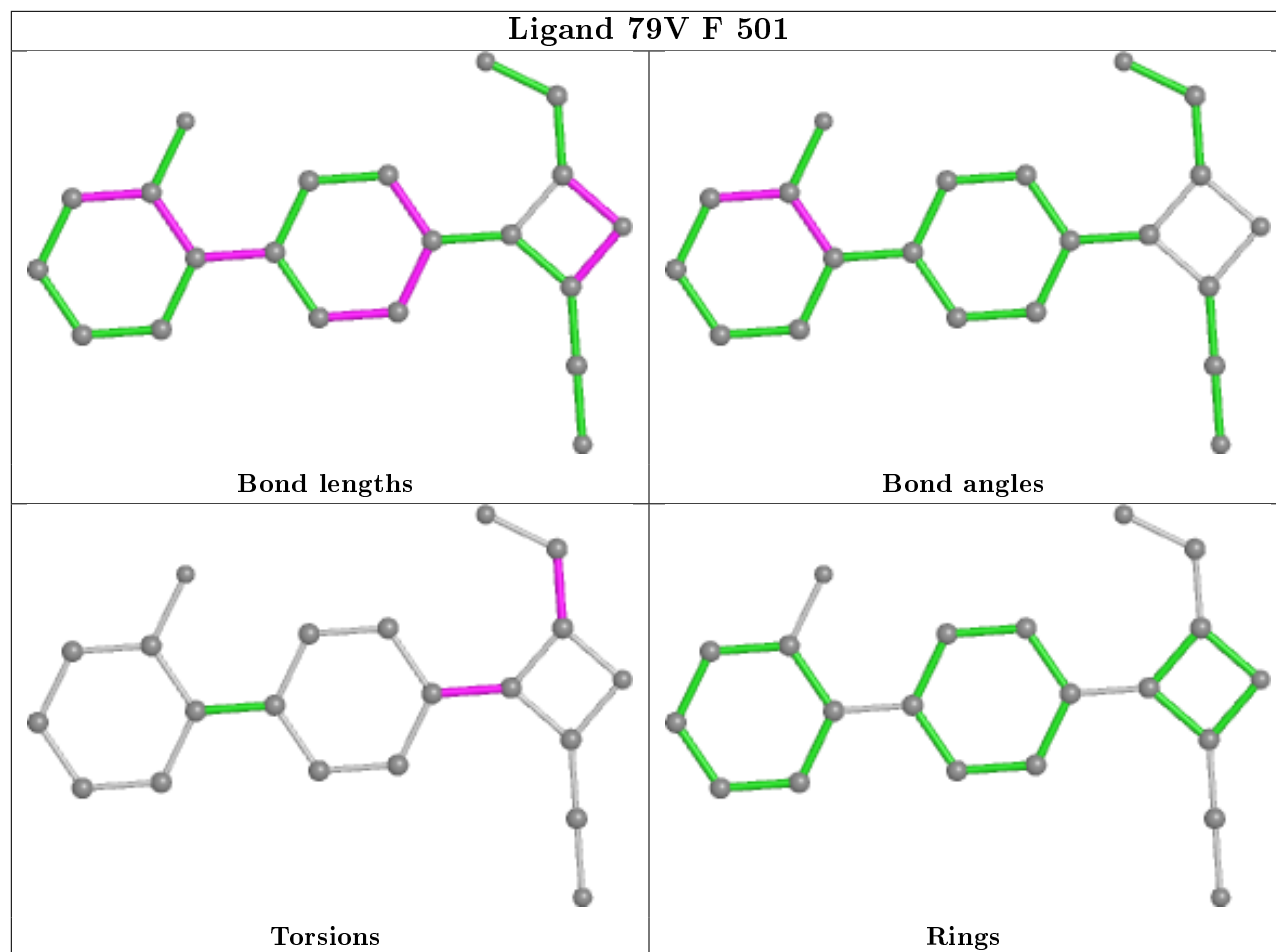
Mol	Chain	Res	Type	Atoms
5	F	501	79V	C13-C14-C19-O20
5	H	501[A]	79V	C13-C14-C19-O20
5	H	501[B]	79V	C04-C05-C07-C12
5	H	501[B]	79V	C04-C05-C07-C08
5	H	501[B]	79V	C13-C14-C19-O20
5	D	501	79V	C13-C14-C19-O20
5	F	501	79V	C09-C10-C13-C16
5	H	501[B]	79V	C06-C05-C07-C08
5	H	501[B]	79V	C11-C10-C13-C16
5	H	501[B]	79V	C11-C10-C13-C14
5	H	501[B]	79V	C09-C10-C13-C16
5	H	501[B]	79V	C09-C10-C13-C14
5	H	501[B]	79V	C06-C05-C07-C12
5	B	501	79V	C13-C14-C19-O20

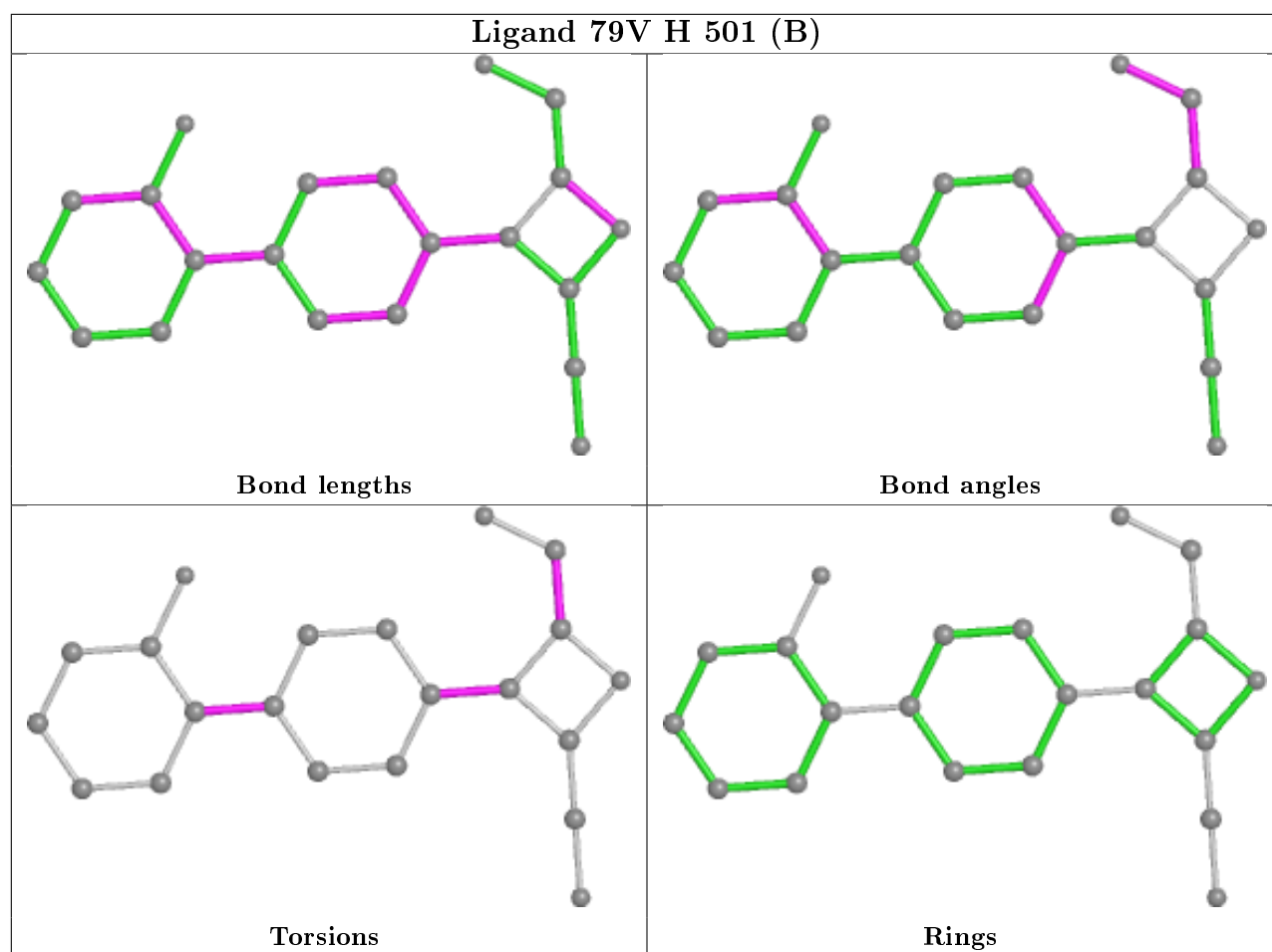
There are no ring outliers.

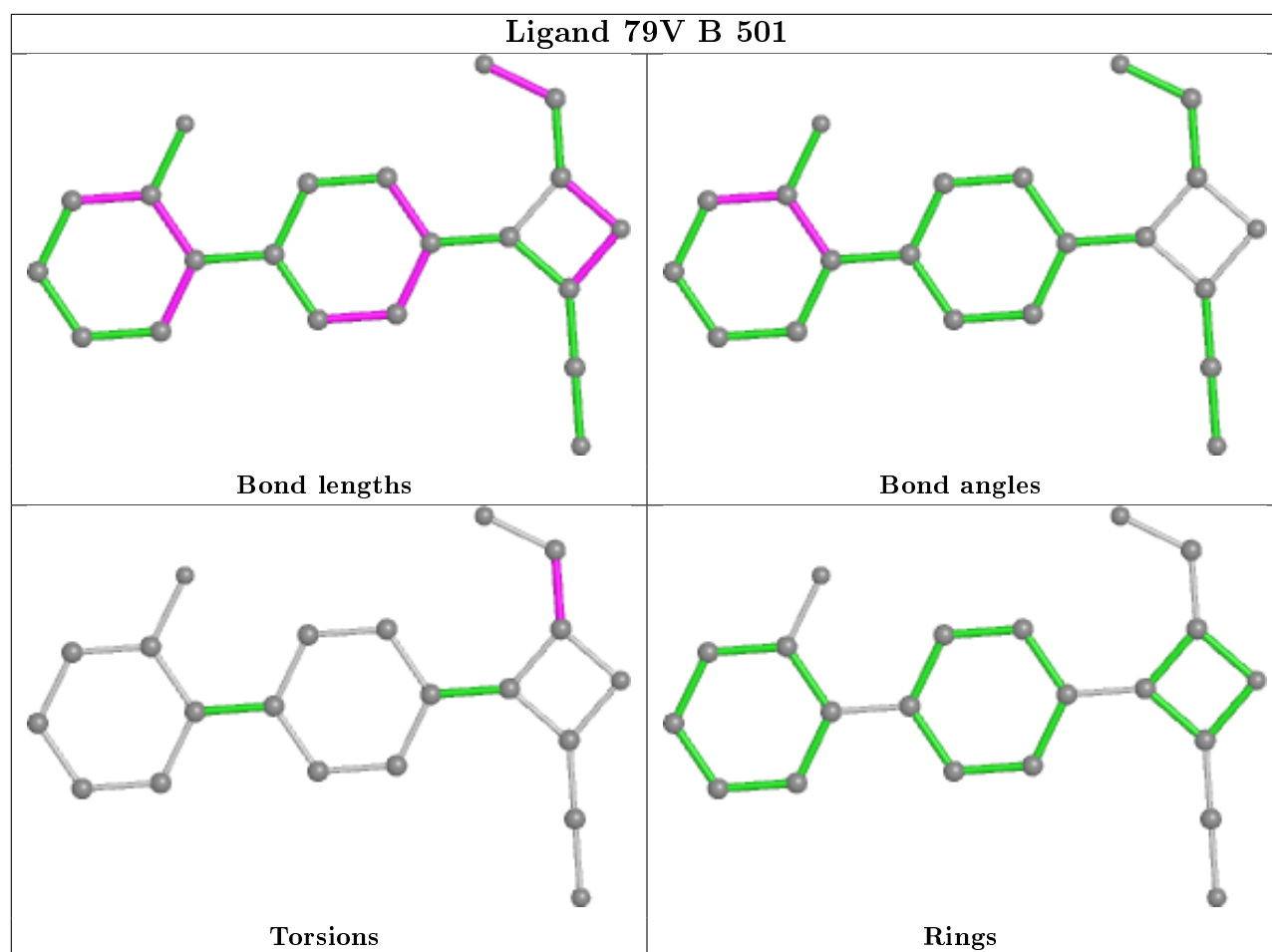
1 monomer is involved in 1 short contact:

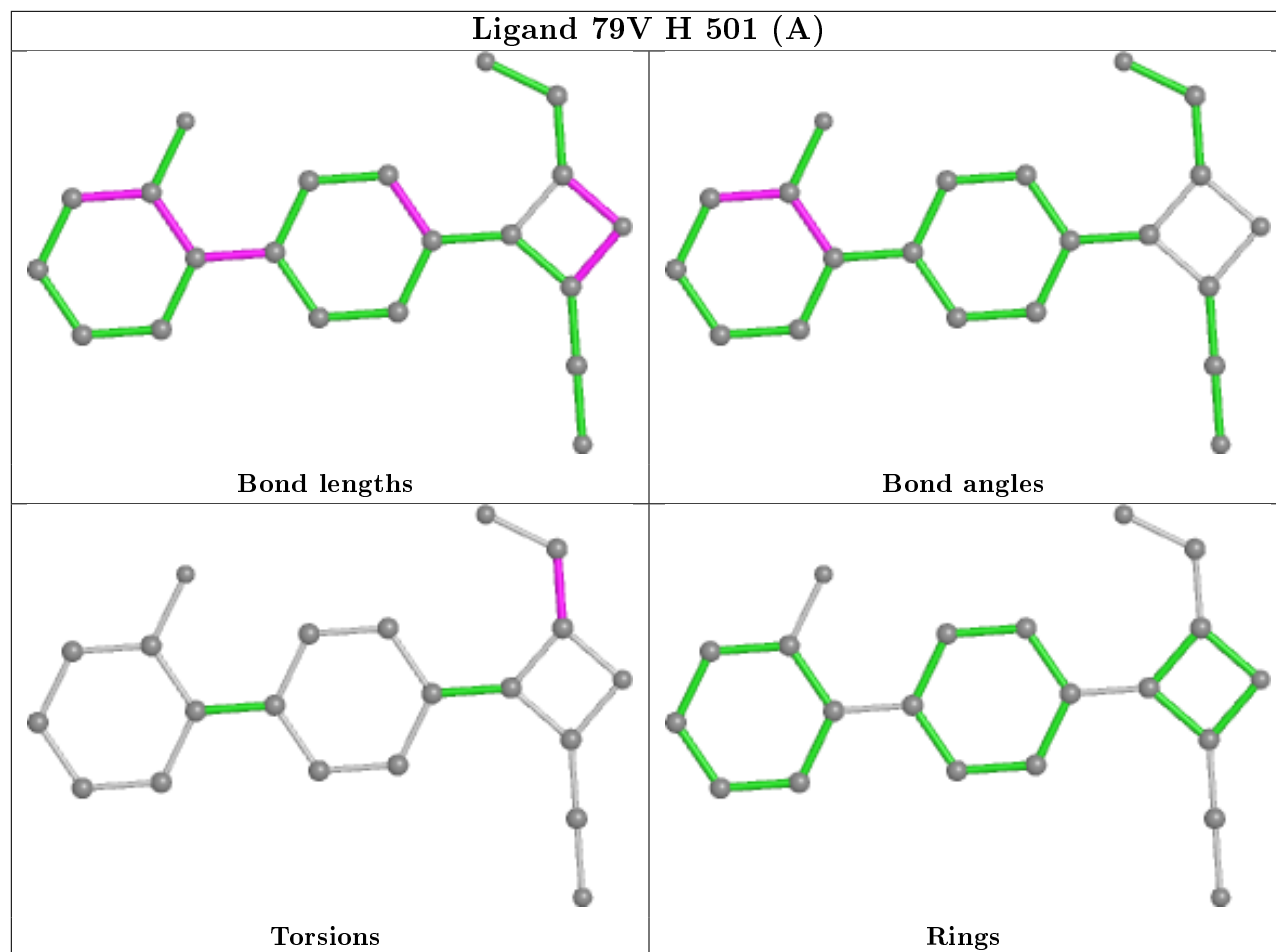
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	H	501[A]	79V	1	0

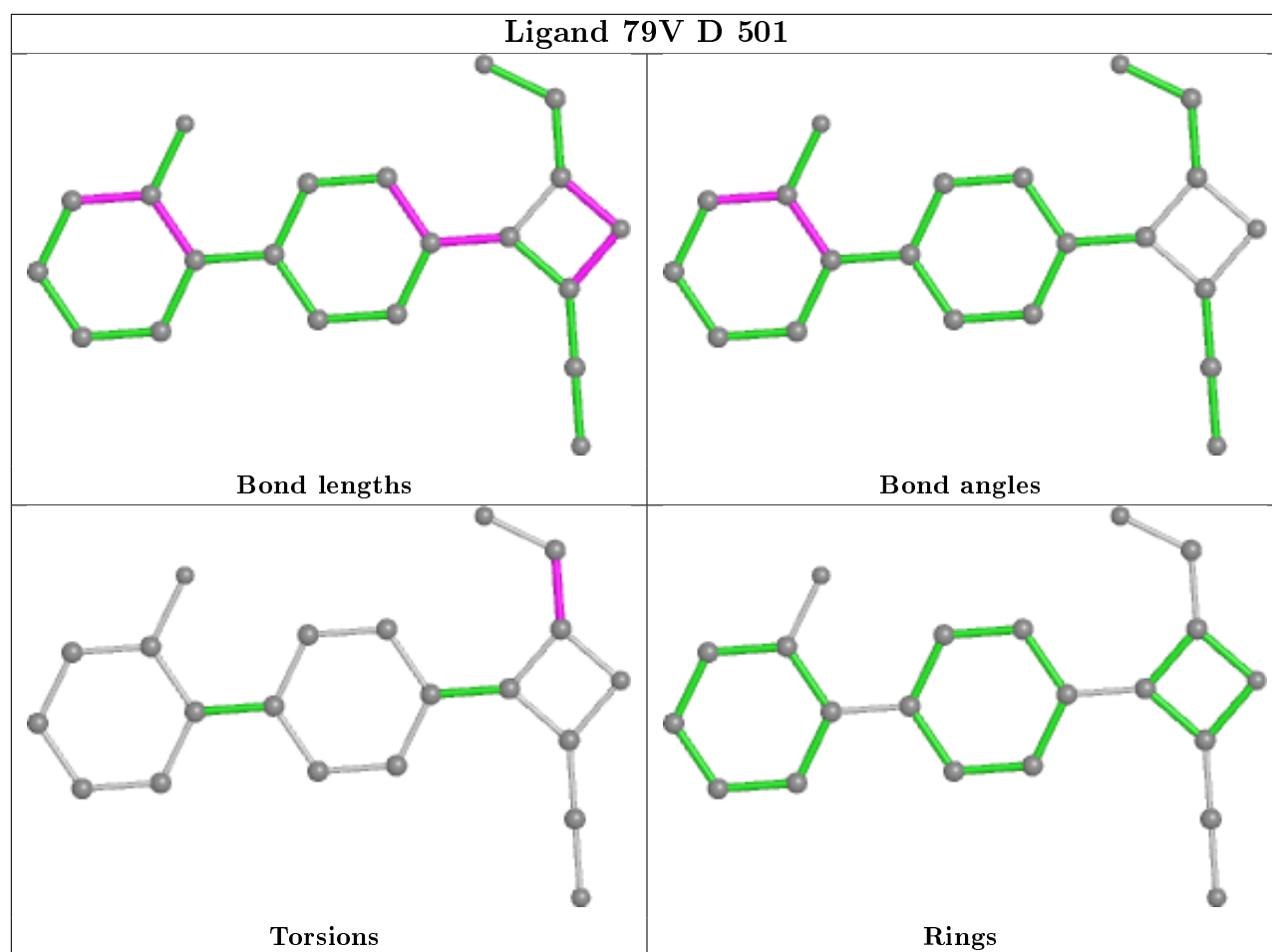
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	249/276 (90%)	-0.28	4 (1%) 72 69	37, 48, 65, 93	0
1	C	249/276 (90%)	-0.14	10 (4%) 38 35	35, 52, 81, 96	0
1	E	245/276 (88%)	0.27	17 (6%) 16 13	44, 73, 103, 118	0
1	G	247/276 (89%)	-0.30	5 (2%) 65 62	34, 48, 63, 83	0
2	B	404/410 (98%)	-0.28	10 (2%) 57 53	30, 39, 70, 90	0
2	D	397/410 (96%)	-0.41	2 (0%) 91 92	29, 38, 53, 67	0
2	F	398/410 (97%)	-0.49	2 (0%) 91 92	29, 39, 56, 81	0
2	H	405/410 (98%)	-0.23	13 (3%) 47 44	28, 38, 68, 98	0
All	All	2594/2744 (94%)	-0.26	63 (2%) 59 54	28, 43, 79, 118	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	173[A]	VAL	5.6
1	E	9	ALA	5.1
1	C	8	GLU	4.6
2	B	183	ALA	4.5
2	H	23	SER	4.4
2	B	174	GLN	4.3
2	H	6	ALA	4.2
2	B	8	PRO	4.0
2	D	22	PRO	3.9
1	A	196	SER	3.8
2	H	172[A]	ALA	3.6
1	C	196	SER	3.5
2	F	22	PRO	3.4
2	H	177[A]	SER	3.4
2	B	173	VAL	3.4
1	A	267	GLY	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	175	THR	3.3
1	G	133	ILE	3.2
1	C	197	GLN	3.2
1	E	211	ASP	3.1
2	H	151	ILE	3.0
1	G	196	SER	3.0
2	F	9	THR	3.0
1	E	98	ILE	2.9
1	E	153	ARG	2.9
1	E	24	ARG	2.8
2	H	22	PRO	2.8
2	H	4	ALA	2.7
2	B	176	GLY	2.7
2	H	176[A]	GLY	2.7
1	A	133	ILE	2.7
1	A	8	GLU	2.7
2	B	408	GLY	2.7
1	E	133	ILE	2.6
1	E	197	GLN	2.6
2	B	188	PHE	2.6
1	C	211	ASP	2.6
1	G	197	GLN	2.5
2	H	8	PRO	2.5
1	C	185	THR	2.5
1	G	23	ASN	2.5
2	B	22	PRO	2.5
2	H	154	ALA	2.4
2	H	179[A]	THR	2.4
1	C	133	ILE	2.3
1	E	23	ASN	2.3
2	B	5	ILE	2.3
1	E	228	GLN	2.3
1	E	28	ILE	2.3
1	E	262	ALA	2.2
1	G	167	GLU	2.2
1	E	20	ARG	2.2
1	E	265	ARG	2.2
1	C	23	ASN	2.2
1	C	201	GLU	2.1
1	E	21	ALA	2.1
1	C	55	ILE	2.1
1	C	105	VAL	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	209	VAL	2.1
1	E	105	VAL	2.0
1	E	248	GLU	2.0
2	H	277	GLU	2.0
2	D	280	ARG	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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	LLP	D	101	24/25	0.98	0.17	30,33,36,37	0
2	LLP	B	101	24/25	0.98	0.16	31,33,36,36	0
2	LLP	H	101	24/25	0.98	0.17	30,32,34,35	0
2	LLP	F	101	24/25	0.98	0.17	30,33,35,36	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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	FMT	D	503	3/3	0.76	0.28	67,67,68,68	0
4	FMT	C	302	3/3	0.78	0.33	60,60,63,64	0
4	FMT	B	503	3/3	0.83	0.13	64,64,68,68	0
4	FMT	H	502	3/3	0.85	0.18	62,62,65,65	0
4	FMT	B	504	3/3	0.85	0.14	53,53,54,56	0
4	FMT	A	302	3/3	0.87	0.34	59,59,61,63	0
5	79V	B	501	21/21	0.88	0.27	38,44,52,55	0
4	FMT	G	302	3/3	0.89	0.28	59,59,61,63	0
3	MLI	E	301	7/7	0.90	0.14	62,68,72,72	0

*Continued on next page...*

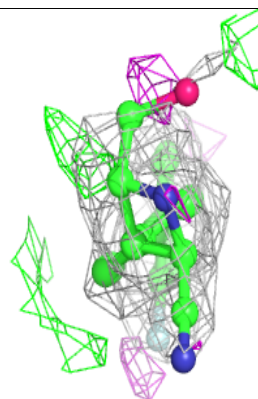
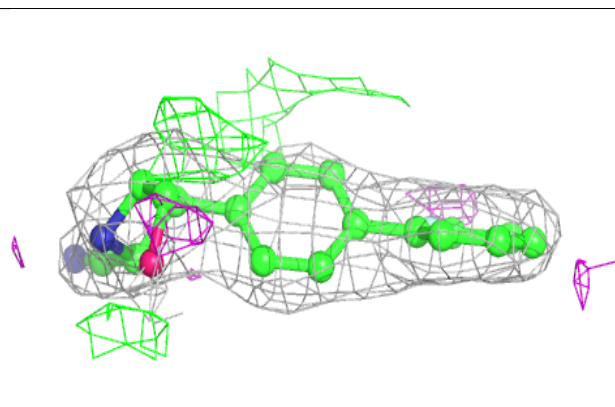
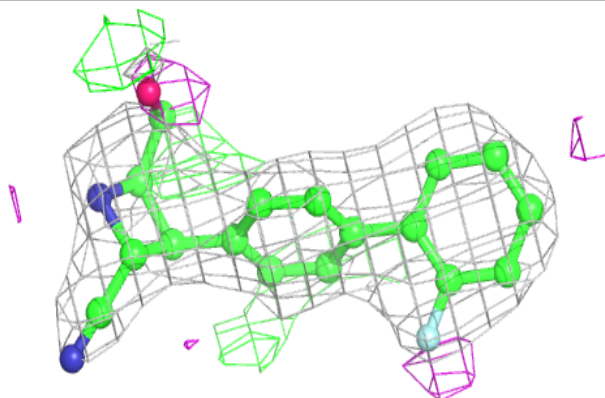
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	79V	H	501[A]	21/21	0.90	0.26	35,36,36,37	21
5	79V	H	501[B]	21/21	0.90	0.26	45,47,54,57	21
4	FMT	D	502	3/3	0.90	0.17	60,60,61,62	0
3	MLI	C	301	7/7	0.90	0.17	70,72,74,74	0
4	FMT	H	503	3/3	0.91	0.14	52,52,52,53	0
4	FMT	B	502	3/3	0.92	0.19	55,55,55,56	0
3	MLI	A	301	7/7	0.93	0.12	50,51,52,53	0
4	FMT	F	502	3/3	0.93	0.10	58,58,58,59	0
3	MLI	G	301	7/7	0.93	0.12	61,66,66,67	0
5	79V	D	501	21/21	0.94	0.16	38,39,41,42	0
4	FMT	G	303	3/3	0.94	0.09	55,55,56,56	0
5	79V	F	501	21/21	0.94	0.14	42,43,46,46	0
4	FMT	F	503	3/3	0.95	0.31	73,73,73,74	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.

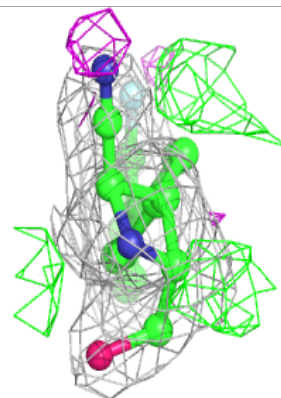
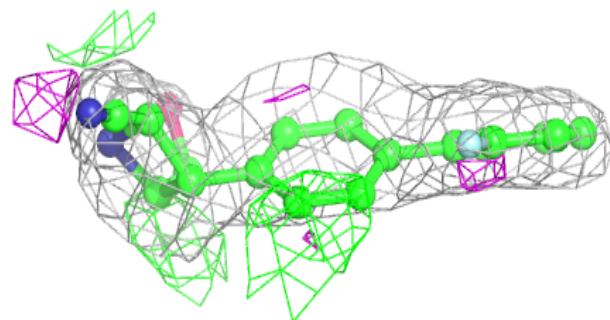
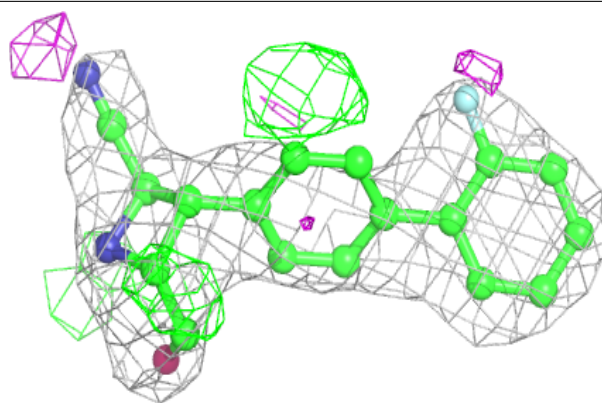
**Electron density around 79V B 501:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

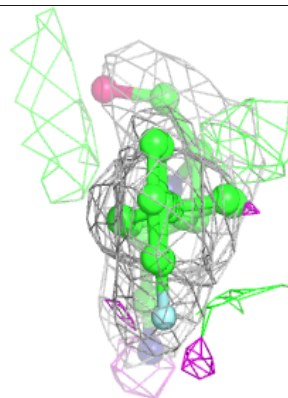
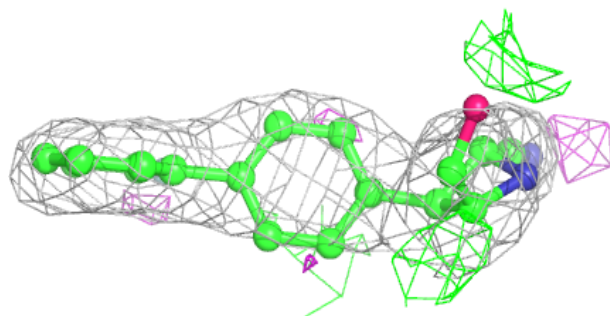
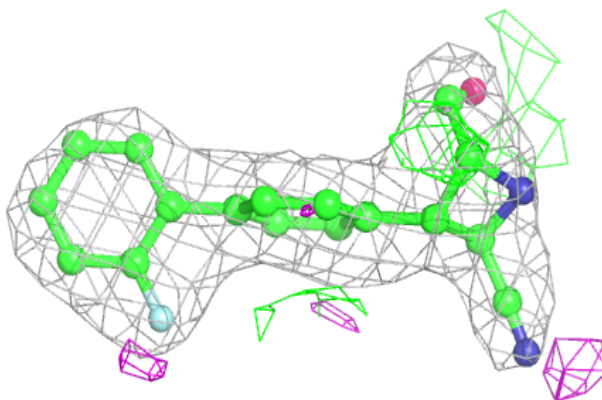


**Electron density around 79V H 501 (A):**

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

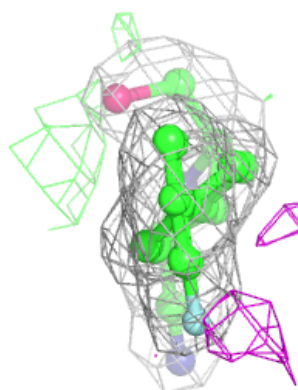
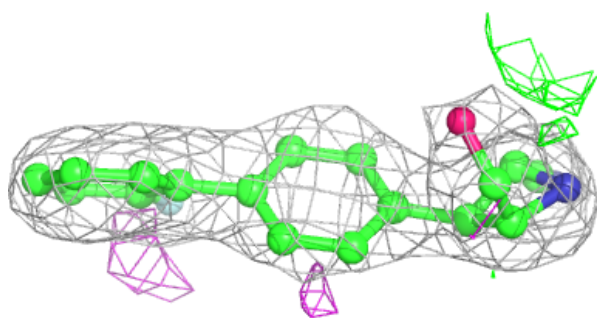
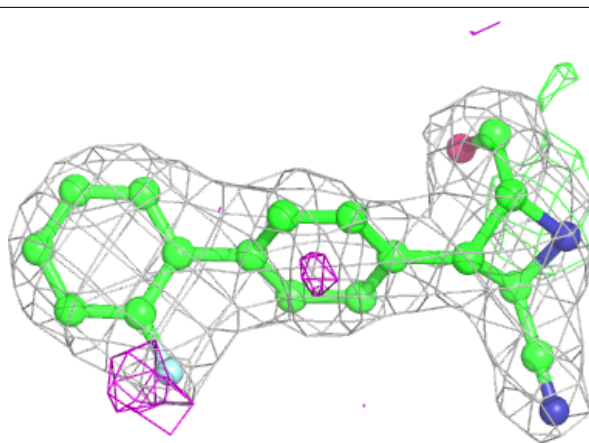
**Electron density around 79V H 501 (B):**

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

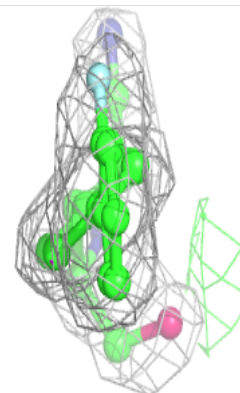
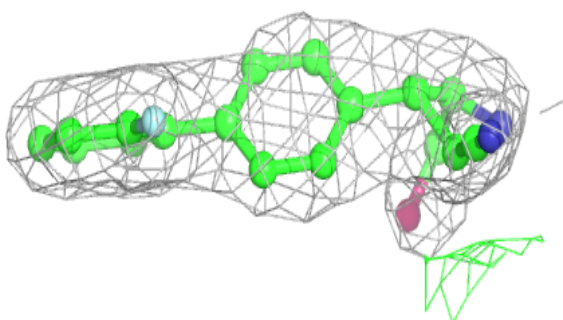
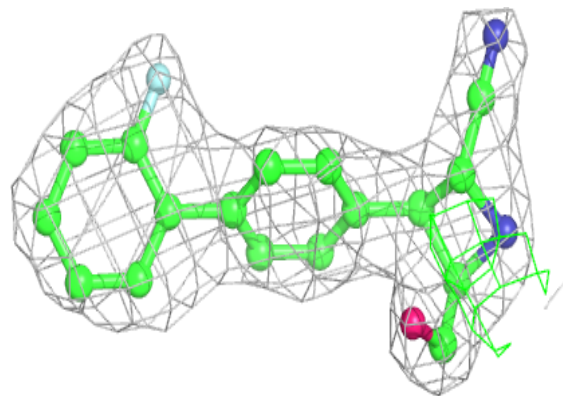


**Electron density around 79V D 501:**

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

**Electron density around 79V F 501:**

$2mF_o-DF_c$  (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.