



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2020 – 05:07 PM BST

PDB ID : 5TSX  
Title : HIV-1 CA hexamer with NUP153 peptide - P1 crystal form  
Authors : Skorupka, K.; Ganser-Pornillos, B.K.; Pornillos, O.  
Deposited on : 2016-10-31  
Resolution : 1.90 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](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 : 2.13.1  
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) : 2.13.1

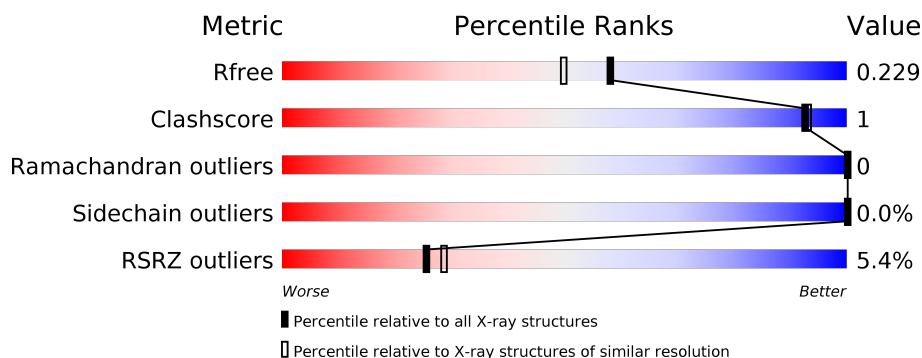
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	231	<div> <div>6%</div> <div> <div></div> <div>92%</div> <div>5%</div> </div> </div>
1	B	231	<div> <div>7%</div> <div> <div></div> <div>90%</div> <div>5%</div> <div>5%</div> </div> </div>
1	C	231	<div> <div>6%</div> <div> <div></div> <div>90%</div> <div>7%</div> </div> </div>
1	D	231	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>7%</div> </div> </div>
1	E	231	<div> <div>4%</div> <div> <div></div> <div>91%</div> <div>7%</div> </div> </div>
1	F	231	<div> <div>6%</div> <div> <div></div> <div>89%</div> <div>7%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	231	<div><div></div><div>6%</div><div>90%</div><div>5%5%</div></div>
1	H	231	<div><div></div><div>2%</div><div>91%</div><div>6%</div></div>
1	I	231	<div><div></div><div>9%</div><div>88%</div><div>10%</div></div>
1	J	231	<div><div></div><div>5%</div><div>87%</div><div>5%8%</div></div>
1	K	231	<div><div></div><div>6%</div><div>90%</div><div>8%</div></div>
1	L	231	<div><div></div><div>2%</div><div>89%</div><div>9%</div></div>
2	M	23	<div><div></div><div>39%</div><div>9%</div><div>52%</div></div>
2	N	23	<div><div></div><div>39%</div><div>9%</div><div>52%</div></div>
2	O	23	<div><div></div><div>4%</div><div>48%</div><div>52%</div></div>
2	P	23	<div><div></div><div>39%</div><div>61%</div></div>
2	R	23	<div><div></div><div>4%</div><div>39%</div><div>61%</div></div>
2	T	23	<div><div></div><div>4%</div><div>30%</div><div>9%</div><div>61%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 43001 atoms, of which 20349 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 HIV-1 CA protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	219	Total	C	H	N	O	S	0	0	0
			3363	1060	1677	297	315	14			
1	B	219	Total	C	H	N	O	S	0	0	0
			3342	1056	1662	295	315	14			
1	C	215	Total	C	H	N	O	S	0	1	0
			3353	1054	1678	292	315	14			
1	D	214	Total	C	H	N	O	S	0	1	0
			3335	1049	1667	291	314	14			
1	E	214	Total	C	H	N	O	S	0	0	0
			3333	1047	1667	291	314	14			
1	F	214	Total	C	H	N	O	S	0	1	0
			3345	1050	1675	294	312	14			
1	G	219	Total	C	H	N	O	S	0	0	0
			3380	1064	1688	297	317	14			
1	H	216	Total	C	H	N	O	S	0	0	0
			3335	1050	1667	293	311	14			
1	I	207	Total	C	H	N	O	S	0	0	0
			3170	1007	1579	272	298	14			
1	J	212	Total	C	H	N	O	S	0	1	0
			3299	1040	1651	288	306	14			
1	K	213	Total	C	H	N	O	S	0	1	0
			3329	1046	1668	290	311	14			
1	L	211	Total	C	H	N	O	S	0	2	0
			3318	1042	1664	290	308	14			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	CYS	ALA	engineered mutation	UNP P12493
A	45	CYS	GLU	engineered mutation	UNP P12493
A	184	ALA	TRP	engineered mutation	UNP P12493
A	185	ALA	MET	engineered mutation	UNP P12493
B	14	CYS	ALA	engineered mutation	UNP P12493

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	45	CYS	GLU	engineered mutation	UNP P12493
B	184	ALA	TRP	engineered mutation	UNP P12493
B	185	ALA	MET	engineered mutation	UNP P12493
C	14	CYS	ALA	engineered mutation	UNP P12493
C	45	CYS	GLU	engineered mutation	UNP P12493
C	184	ALA	TRP	engineered mutation	UNP P12493
C	185	ALA	MET	engineered mutation	UNP P12493
D	14	CYS	ALA	engineered mutation	UNP P12493
D	45	CYS	GLU	engineered mutation	UNP P12493
D	184	ALA	TRP	engineered mutation	UNP P12493
D	185	ALA	MET	engineered mutation	UNP P12493
E	14	CYS	ALA	engineered mutation	UNP P12493
E	45	CYS	GLU	engineered mutation	UNP P12493
E	184	ALA	TRP	engineered mutation	UNP P12493
E	185	ALA	MET	engineered mutation	UNP P12493
F	14	CYS	ALA	engineered mutation	UNP P12493
F	45	CYS	GLU	engineered mutation	UNP P12493
F	184	ALA	TRP	engineered mutation	UNP P12493
F	185	ALA	MET	engineered mutation	UNP P12493
G	14	CYS	ALA	engineered mutation	UNP P12493
G	45	CYS	GLU	engineered mutation	UNP P12493
G	184	ALA	TRP	engineered mutation	UNP P12493
G	185	ALA	MET	engineered mutation	UNP P12493
H	14	CYS	ALA	engineered mutation	UNP P12493
H	45	CYS	GLU	engineered mutation	UNP P12493
H	184	ALA	TRP	engineered mutation	UNP P12493
H	185	ALA	MET	engineered mutation	UNP P12493
I	14	CYS	ALA	engineered mutation	UNP P12493
I	45	CYS	GLU	engineered mutation	UNP P12493
I	184	ALA	TRP	engineered mutation	UNP P12493
I	185	ALA	MET	engineered mutation	UNP P12493
J	14	CYS	ALA	engineered mutation	UNP P12493
J	45	CYS	GLU	engineered mutation	UNP P12493
J	184	ALA	TRP	engineered mutation	UNP P12493
J	185	ALA	MET	engineered mutation	UNP P12493
K	14	CYS	ALA	engineered mutation	UNP P12493
K	45	CYS	GLU	engineered mutation	UNP P12493
K	184	ALA	TRP	engineered mutation	UNP P12493
K	185	ALA	MET	engineered mutation	UNP P12493
L	14	CYS	ALA	engineered mutation	UNP P12493
L	45	CYS	GLU	engineered mutation	UNP P12493
L	184	ALA	TRP	engineered mutation	UNP P12493

*Continued on next page...*

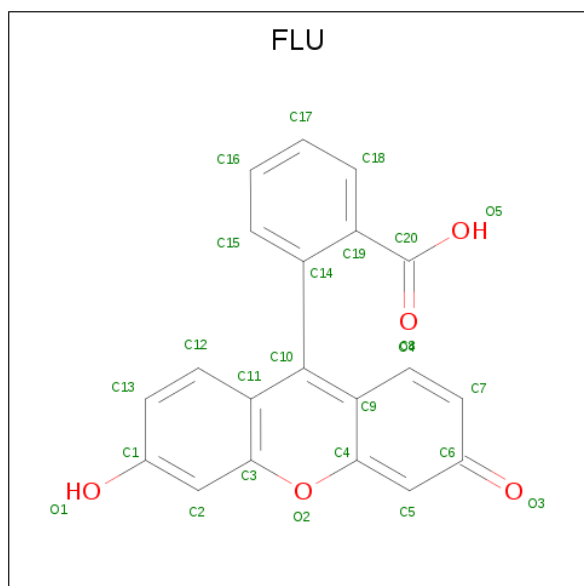
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
L	185	ALA	MET	engineered mutation	UNP P12493

- Molecule 2 is a protein called Nuclear pore complex protein Nup153.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	11	Total	C	H	N	O	0	0	0
			143	49	67	12	15			
2	N	11	Total	C	H	N	O	0	0	0
			143	49	67	12	15			
2	O	11	Total	C	H	N	O	0	0	0
			133	48	60	11	14			
2	P	9	Total	C	H	N	O	0	0	0
			119	42	56	9	12			
2	R	9	Total	C	H	N	O	0	0	0
			119	42	56	9	12			
2	T	9	Total	C	H	N	O	0	0	0
			119	42	56	9	12			

- Molecule 3 is 2-(6-HYDROXY-3-OXO-3H-XANTHEN-9-YL)-BENZOIC ACID (three-letter code: FLU) (formula: C<sub>20</sub>H<sub>12</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	O		0	0
			36	20	11	5			
3	B	1	Total	C	H	O		0	0
			36	20	11	5			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	G	1	Total	C	H	O	0	0
			36	20	11	5		
3	H	1	Total	C	H	O	0	0
			36	20	11	5		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	177	Total	O	0	0
			177	177		
4	B	168	Total	O	0	0
			168	168		
4	C	204	Total	O	0	0
			204	204		
4	D	213	Total	O	0	0
			213	213		
4	E	171	Total	O	0	0
			171	171		
4	F	171	Total	O	0	0
			171	171		
4	G	181	Total	O	0	0
			181	181		
4	H	194	Total	O	0	0
			194	194		
4	I	128	Total	O	0	0
			128	128		
4	J	176	Total	O	0	0
			176	176		
4	K	167	Total	O	0	0
			167	167		
4	L	170	Total	O	0	0
			170	170		
4	M	12	Total	O	0	0
			12	12		
4	N	14	Total	O	0	0
			14	14		
4	O	10	Total	O	0	0
			10	10		
4	P	8	Total	O	0	0
			8	8		
4	R	8	Total	O	0	0
			8	8		

*Continued on next page...*

*Continued from previous page...*

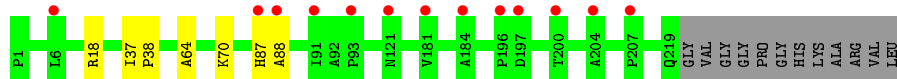
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	T	7	Total	O	0	0
			7	7		



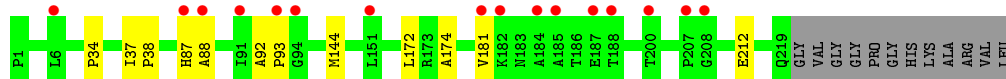
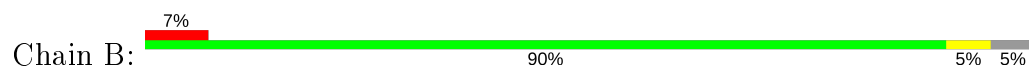
### 3 Residue-property plots [i](#)

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.

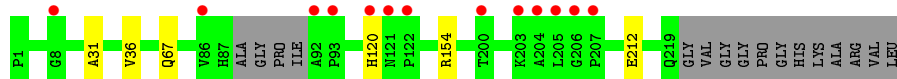
- Molecule 1: HIV-1 CA protein



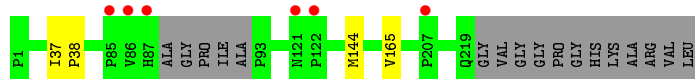
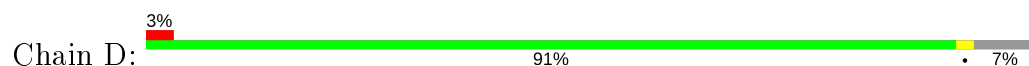
- Molecule 1: HIV-1 CA protein



- Molecule 1: HIV-1 CA protein



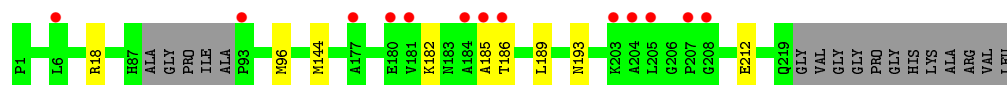
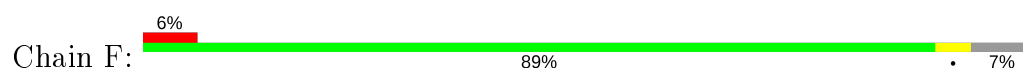
- Molecule 1: HIV-1 CA protein



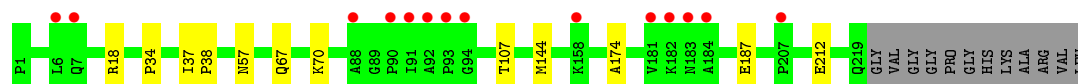
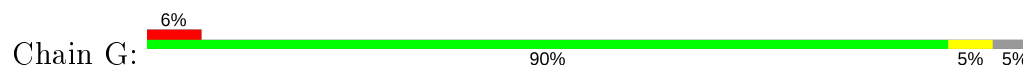
- Molecule 1: HIV-1 CA protein



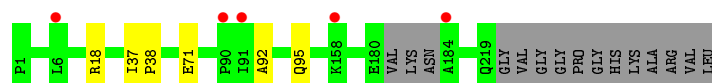
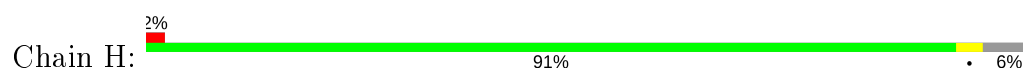
- Molecule 1: HIV-1 CA protein



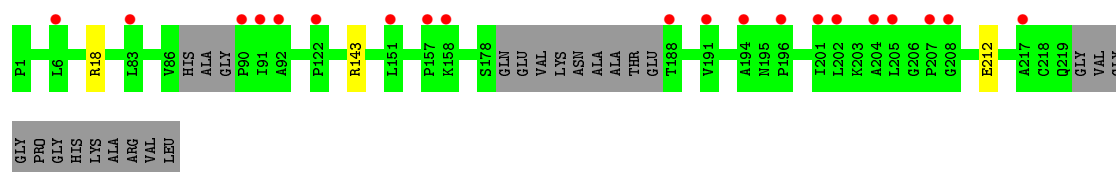
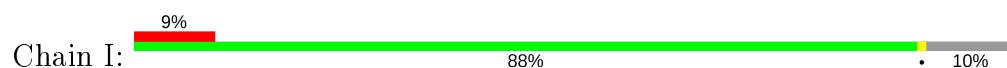
- Molecule 1: HIV-1 CA protein



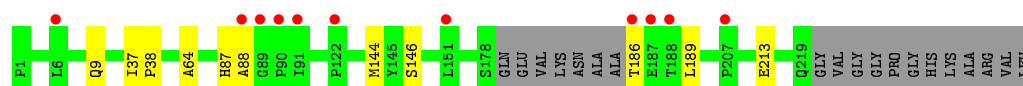
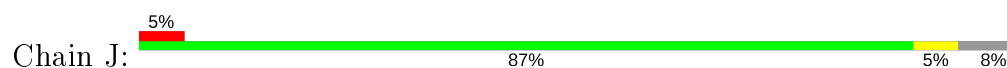
- Molecule 1: HIV-1 CA protein



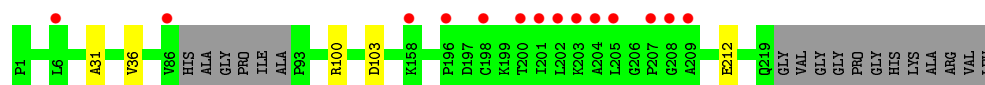
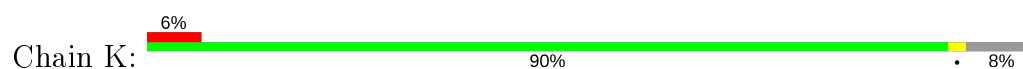
- Molecule 1: HIV-1 CA protein



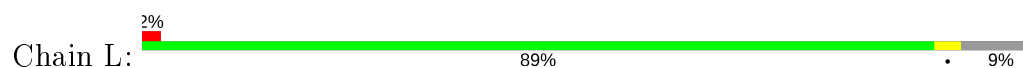
- Molecule 1: HIV-1 CA protein



- Molecule 1: HIV-1 CA protein

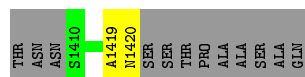


- Molecule 1: HIV-1 CA protein

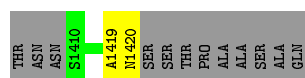
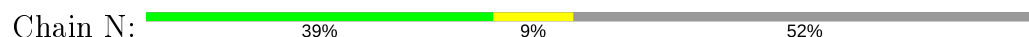




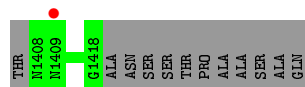
- Molecule 2: Nuclear pore complex protein Nup153



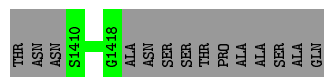
- Molecule 2: Nuclear pore complex protein Nup153



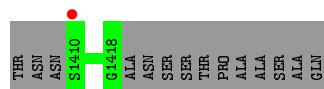
- Molecule 2: Nuclear pore complex protein Nup153



- Molecule 2: Nuclear pore complex protein Nup153



- Molecule 2: Nuclear pore complex protein Nup153



- Molecule 2: Nuclear pore complex protein Nup153



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.33Å 91.49Å 150.05Å 89.54° 90.49° 96.08°	Depositor
Resolution (Å)	45.48 – 1.90 48.27 – 1.90	Depositor EDS
% Data completeness (in resolution range)	85.9 (45.48-1.90) 86.0 (48.27-1.90)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 1.90Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.185 , 0.228 0.187 , 0.229	Depositor DCC
$R_{free}$ test set	10136 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.2	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.016 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	43001	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 71.26 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.7091e-06. 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: FLU

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	0.38	0/1723	0.50	0/2344
1	B	0.38	0/1717	0.50	0/2339
1	C	0.38	0/1712	0.51	0/2325
1	D	0.39	0/1705	0.51	0/2314
1	E	0.36	0/1700	0.49	0/2307
1	F	0.36	0/1707	0.48	0/2316
1	G	0.36	0/1729	0.49	0/2352
1	H	0.35	0/1704	0.53	0/2316
1	I	0.36	0/1625	0.48	0/2207
1	J	0.40	0/1687	0.50	0/2292
1	K	0.36	0/1698	0.51	0/2305
1	L	0.37	0/1692	0.50	0/2294
2	M	0.40	0/78	0.51	0/105
2	N	0.36	0/78	0.58	0/105
2	O	0.41	0/75	0.52	0/101
2	P	0.40	0/65	0.51	0/87
2	R	0.42	0/65	0.49	0/87
2	T	0.43	0/65	0.53	0/87
All	All	0.37	0/20825	0.50	0/28283

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	1686	1677	1677	5	0
1	B	1680	1662	1662	8	0
1	C	1675	1678	1680	5	0
1	D	1668	1667	1669	3	0
1	E	1666	1667	1667	4	0
1	F	1670	1675	1676	8	0
1	G	1692	1688	1688	8	0
1	H	1668	1667	1666	5	0
1	I	1591	1579	1579	3	0
1	J	1648	1651	1652	8	0
1	K	1661	1668	1670	3	0
1	L	1654	1664	1666	5	0
2	M	76	67	67	2	0
2	N	76	67	67	2	0
2	O	73	60	60	0	0
2	P	63	56	56	0	0
2	R	63	56	56	0	0
2	T	63	56	56	2	0
3	A	25	11	11	0	0
3	B	25	11	11	0	0
3	G	25	11	11	0	0
3	H	25	11	11	0	0
4	A	177	0	0	1	0
4	B	168	0	0	0	0
4	C	204	0	0	3	0
4	D	213	0	0	0	0
4	E	171	0	0	5	0
4	F	171	0	0	3	0
4	G	181	0	0	1	0
4	H	194	0	0	2	0
4	I	128	0	0	1	0
4	J	176	0	0	3	0
4	K	167	0	0	0	0
4	L	170	0	0	1	0
4	M	12	0	0	2	0
4	N	14	0	0	1	0
4	O	10	0	0	0	0
4	P	8	0	0	0	0
4	R	8	0	0	0	0
4	T	7	0	0	0	0
All	All	22652	20349	20358	60	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.

The worst 5 of 60 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:96:MET:SD	4:F:423:HOH:O	2.31	0.87
1:F:182:LYS:O	1:F:186:THR:OG1	2.01	0.78
2:N:1420:ASN:OD1	4:N:1501:HOH:O	2.00	0.78
1:E:18:ARG:NH2	4:E:301:HOH:O	2.18	0.77
1:H:71:GLU:OE1	4:H:401:HOH:O	2.06	0.73

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	217/231 (94%)	214 (99%)	3 (1%)	0	100	100
1	B	217/231 (94%)	214 (99%)	3 (1%)	0	100	100
1	C	212/231 (92%)	211 (100%)	1 (0%)	0	100	100
1	D	211/231 (91%)	210 (100%)	1 (0%)	0	100	100
1	E	210/231 (91%)	209 (100%)	1 (0%)	0	100	100
1	F	211/231 (91%)	206 (98%)	5 (2%)	0	100	100
1	G	217/231 (94%)	213 (98%)	4 (2%)	0	100	100
1	H	212/231 (92%)	211 (100%)	1 (0%)	0	100	100
1	I	201/231 (87%)	199 (99%)	2 (1%)	0	100	100
1	J	209/231 (90%)	208 (100%)	1 (0%)	0	100	100
1	K	210/231 (91%)	208 (99%)	2 (1%)	0	100	100
1	L	207/231 (90%)	205 (99%)	2 (1%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	M	9/23 (39%)	9 (100%)	0	0	100	100
2	N	9/23 (39%)	9 (100%)	0	0	100	100
2	O	9/23 (39%)	8 (89%)	1 (11%)	0	100	100
2	P	7/23 (30%)	7 (100%)	0	0	100	100
2	R	7/23 (30%)	7 (100%)	0	0	100	100
2	T	7/23 (30%)	7 (100%)	0	0	100	100
All	All	2582/2910 (89%)	2555 (99%)	27 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	182/193 (94%)	182 (100%)	0	100	100
1	B	181/193 (94%)	181 (100%)	0	100	100
1	C	184/193 (95%)	184 (100%)	0	100	100
1	D	183/193 (95%)	183 (100%)	0	100	100
1	E	183/193 (95%)	183 (100%)	0	100	100
1	F	183/193 (95%)	183 (100%)	0	100	100
1	G	184/193 (95%)	183 (100%)	1 (0%)	88	89
1	H	181/193 (94%)	181 (100%)	0	100	100
1	I	172/193 (89%)	172 (100%)	0	100	100
1	J	180/193 (93%)	180 (100%)	0	100	100
1	K	183/193 (95%)	183 (100%)	0	100	100
1	L	182/193 (94%)	182 (100%)	0	100	100
2	M	8/17 (47%)	8 (100%)	0	100	100
2	N	8/17 (47%)	8 (100%)	0	100	100
2	O	7/17 (41%)	7 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	P	7/17 (41%)	7 (100%)	0	100	100
2	R	7/17 (41%)	7 (100%)	0	100	100
2	T	7/17 (41%)	7 (100%)	0	100	100
All	All	2222/2418 (92%)	2221 (100%)	1 (0%)	100	100

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	187	GLU

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

Mol	Chain	Res	Type
2	N	1420	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	FLU	A	601	-	24,28,28	2.44	4 (16%)	28,41,41	1.01	1 (3%)
3	FLU	B	301	-	24,28,28	2.49	4 (16%)	28,41,41	1.18	5 (17%)
3	FLU	H	301	-	24,28,28	2.45	5 (20%)	28,41,41	1.12	4 (14%)
3	FLU	G	301	-	24,28,28	2.49	5 (20%)	28,41,41	1.01	2 (7%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	FLU	A	601	-	-	0/4/12/12	0/4/4/4
3	FLU	B	301	-	-	0/4/12/12	0/4/4/4
3	FLU	H	301	-	-	0/4/12/12	0/4/4/4
3	FLU	G	301	-	-	0/4/12/12	0/4/4/4

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	FLU	C14-C10	-8.60	1.40	1.50
3	G	301	FLU	C14-C10	-8.58	1.40	1.50
3	H	301	FLU	C14-C10	-8.57	1.40	1.50
3	B	301	FLU	C14-C10	-8.54	1.40	1.50
3	B	301	FLU	C19-C20	6.78	1.54	1.47

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	301	FLU	O2-C3-C2	2.68	119.25	116.11
3	H	301	FLU	O2-C4-C5	2.66	119.19	116.03
3	B	301	FLU	O2-C4-C5	2.29	118.75	116.03
3	B	301	FLU	C12-C11-C3	2.20	119.05	116.50
3	H	301	FLU	C12-C11-C3	2.20	119.05	116.50

There are no chirality outliers.

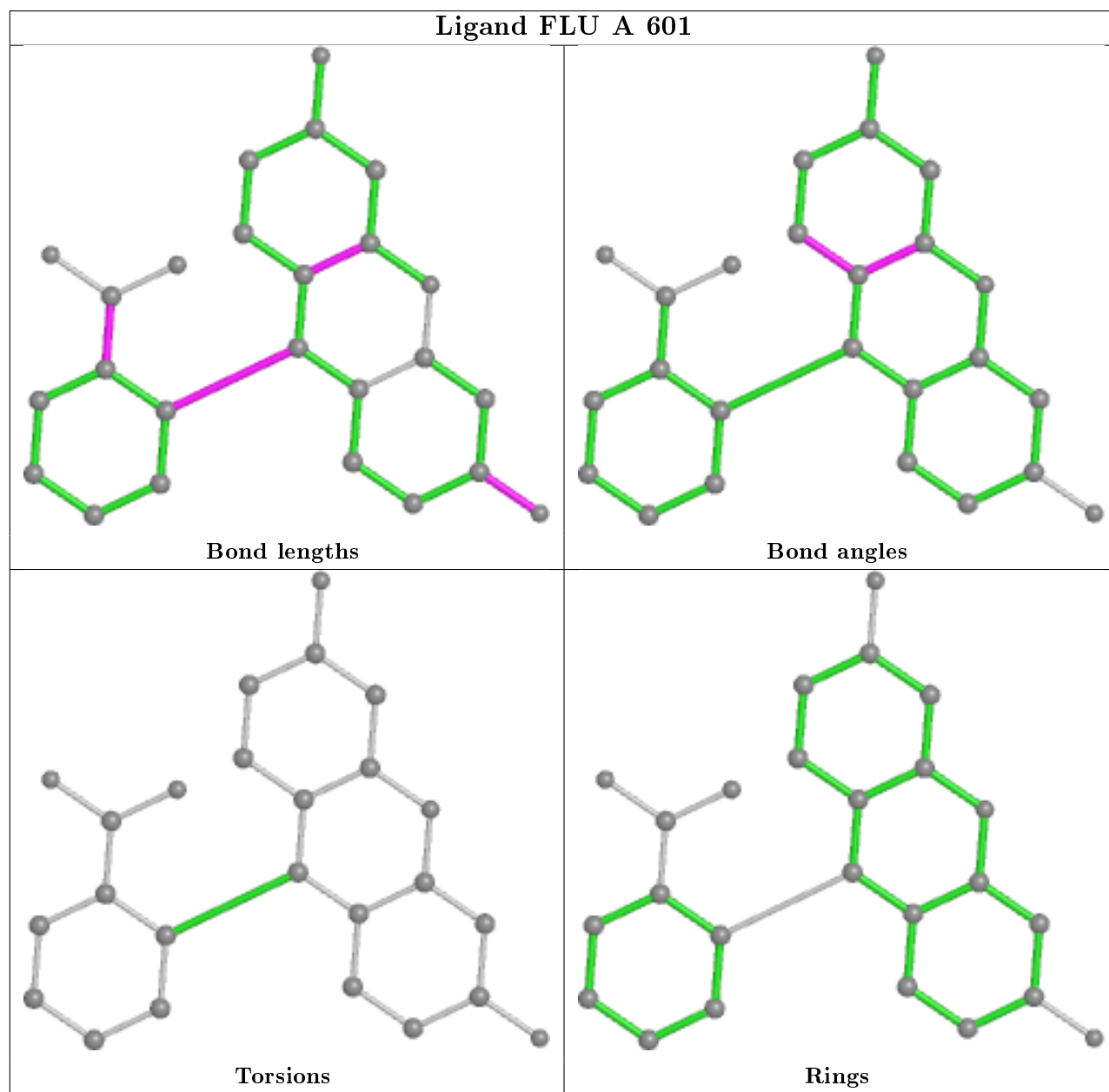
There are no torsion outliers.

There are no ring outliers.

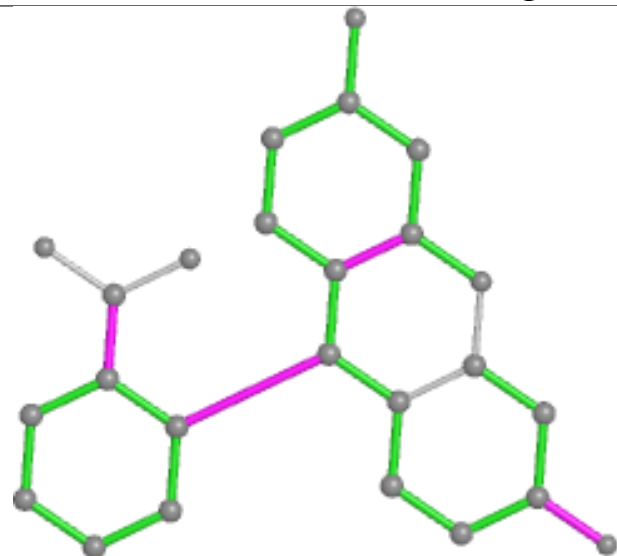
No monomer is involved in short contacts.

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

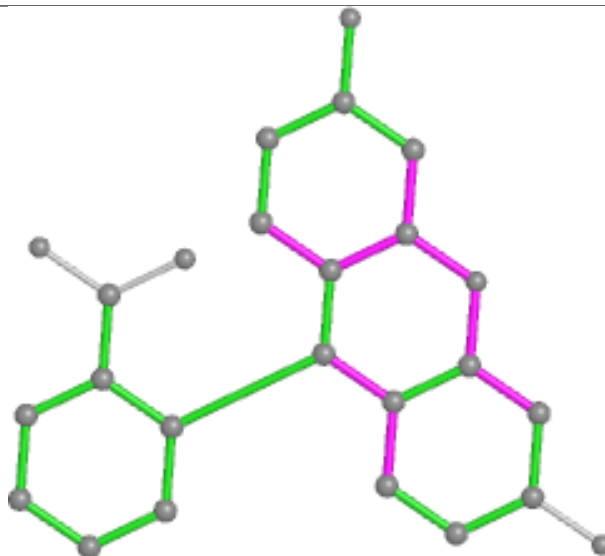
bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 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.



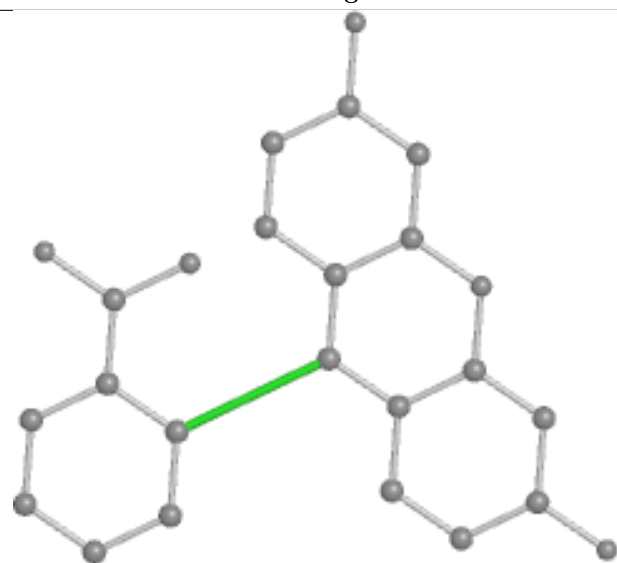
## Ligand FLU B 301



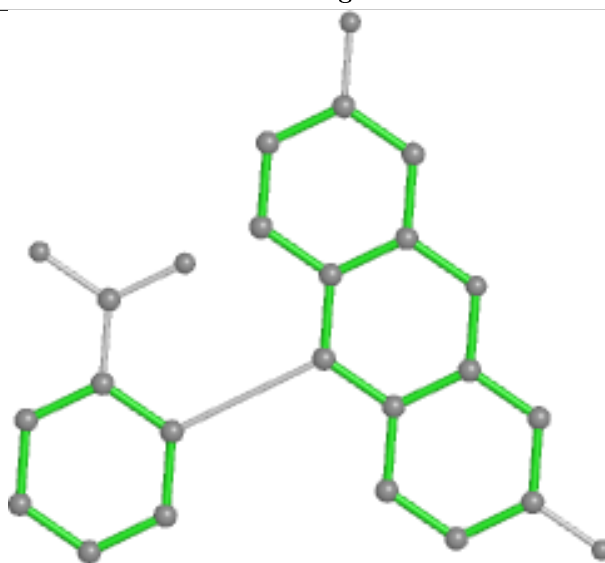
Bond lengths



Bond angles

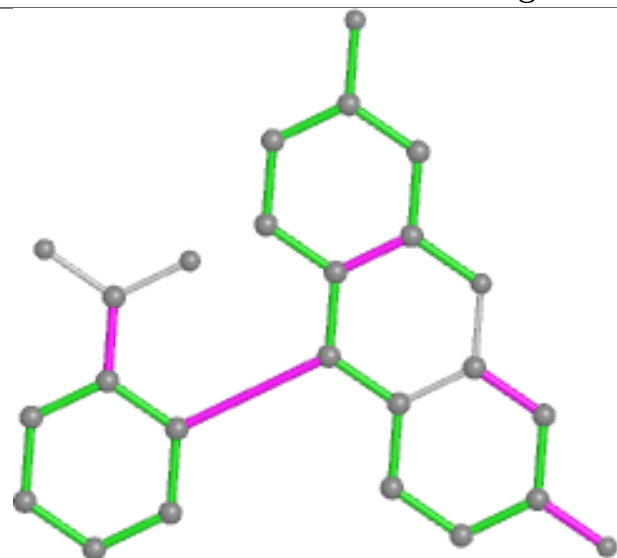


Torsions

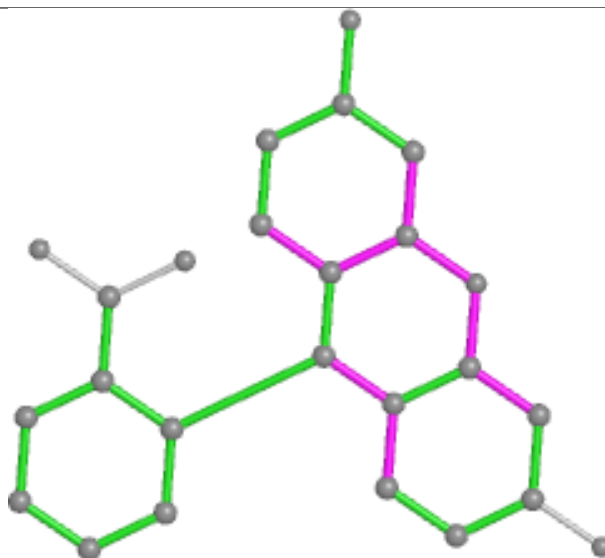


Rings

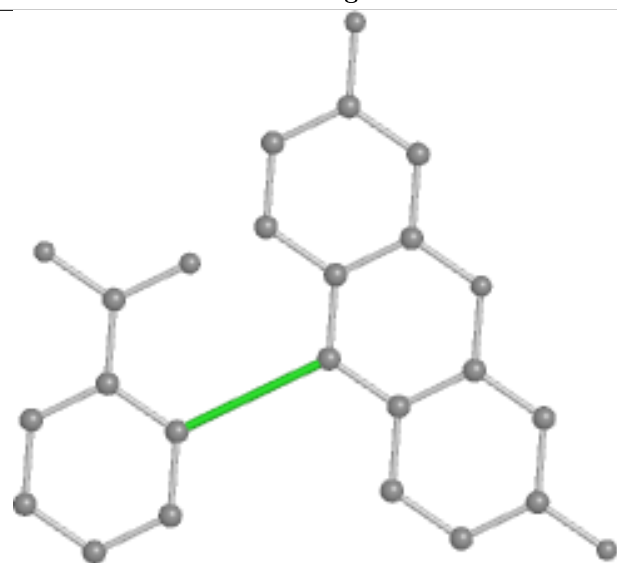
## Ligand FLU H 301



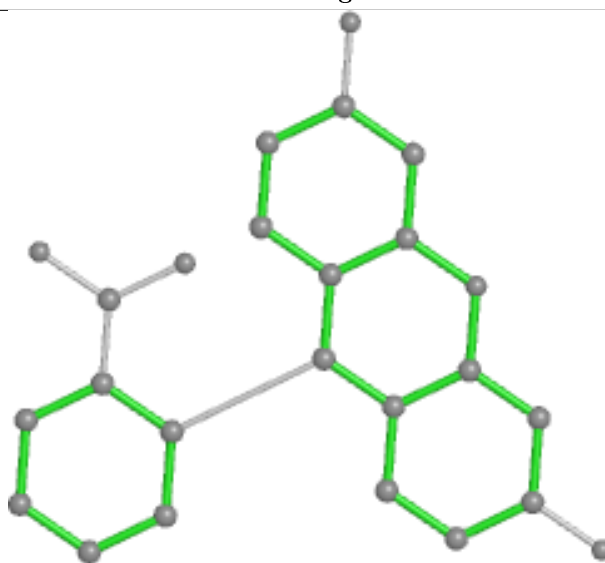
Bond lengths



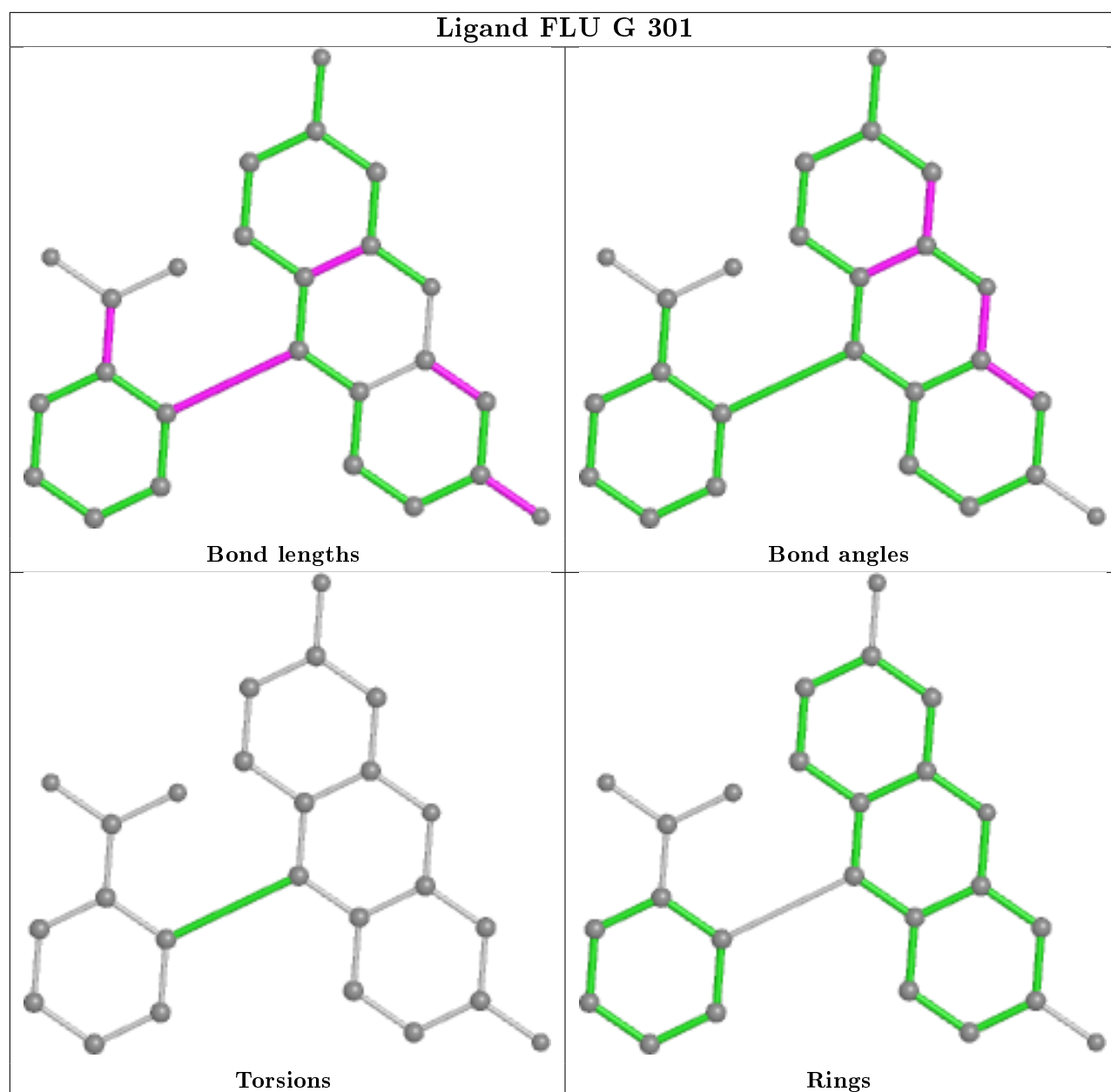
Bond angles



Torsions



Rings



## 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	219/231 (94%)	0.25	13 (5%) 22 25	8, 23, 53, 65	0
1	B	219/231 (94%)	0.38	16 (7%) 15 16	8, 23, 59, 81	0
1	C	215/231 (93%)	0.16	13 (6%) 21 24	8, 20, 57, 74	0
1	D	214/231 (92%)	0.05	6 (2%) 53 56	9, 20, 50, 62	0
1	E	214/231 (92%)	0.17	10 (4%) 31 34	11, 26, 56, 64	0
1	F	214/231 (92%)	0.24	13 (6%) 21 24	10, 25, 62, 73	0
1	G	219/231 (94%)	0.31	14 (6%) 19 22	11, 24, 62, 85	0
1	H	216/231 (93%)	0.09	5 (2%) 60 63	10, 22, 51, 70	0
1	I	207/231 (89%)	0.48	20 (9%) 7 9	12, 29, 62, 79	0
1	J	212/231 (91%)	0.26	11 (5%) 27 30	11, 25, 55, 72	0
1	K	213/231 (92%)	0.18	14 (6%) 18 20	11, 24, 55, 67	0
1	L	211/231 (91%)	0.13	5 (2%) 59 62	12, 23, 51, 62	0
2	M	11/23 (47%)	0.23	0 100 100	11, 14, 36, 47	0
2	N	11/23 (47%)	0.13	0 100 100	11, 13, 36, 42	0
2	O	11/23 (47%)	0.52	1 (9%) 9 10	14, 17, 54, 64	0
2	P	9/23 (39%)	-0.06	0 100 100	15, 17, 22, 35	0
2	R	9/23 (39%)	0.08	1 (11%) 5 6	20, 24, 38, 51	0
2	T	9/23 (39%)	1.37	1 (11%) 5 6	25, 35, 58, 69	0
All	All	2633/2910 (90%)	0.23	143 (5%) 25 29	8, 24, 57, 85	0

The worst 5 of 143 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	205	LEU	7.6
1	J	207	PRO	6.7
1	G	184	ALA	6.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	G	91	ILE	6.4
1	B	182	LYS	5.9

## 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.

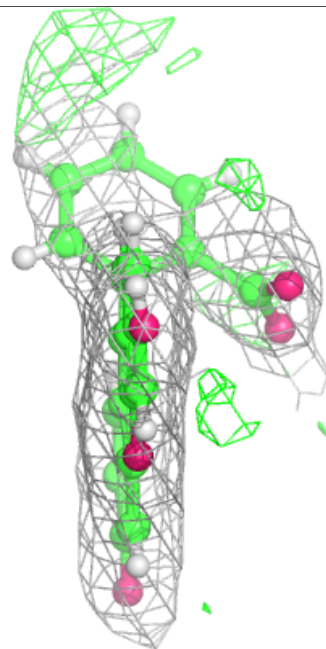
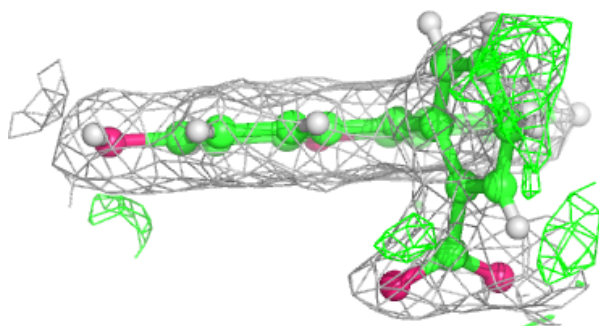
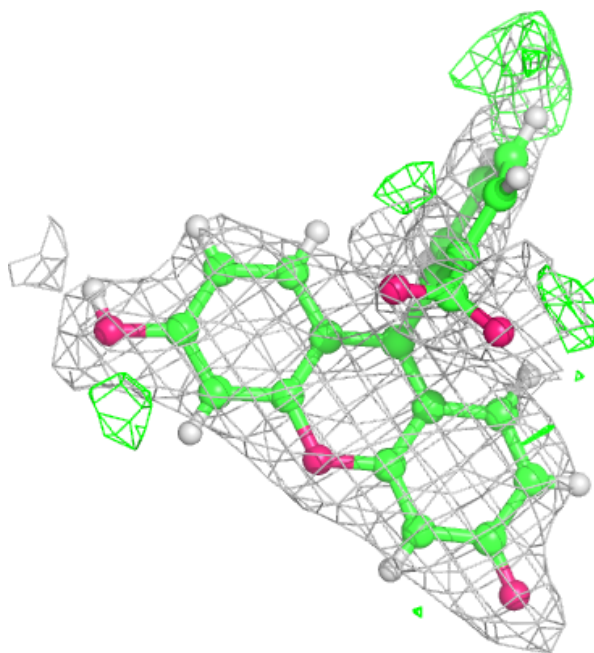
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	FLU	G	301	25/25	0.90	0.14	22,35,59,70	0
3	FLU	H	301	25/25	0.91	0.14	21,34,52,62	0
3	FLU	B	301	25/25	0.92	0.12	23,33,51,55	0
3	FLU	A	601	25/25	0.93	0.11	19,28,40,42	0

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.



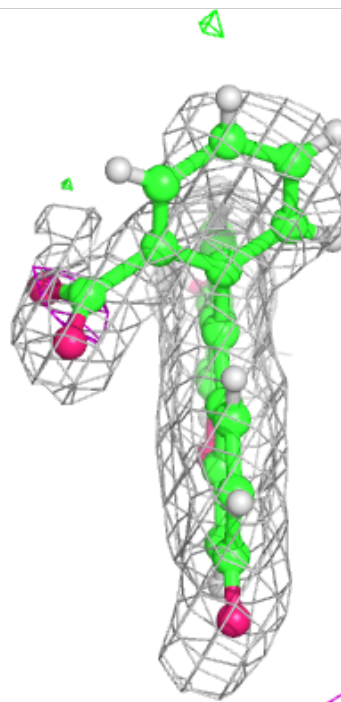
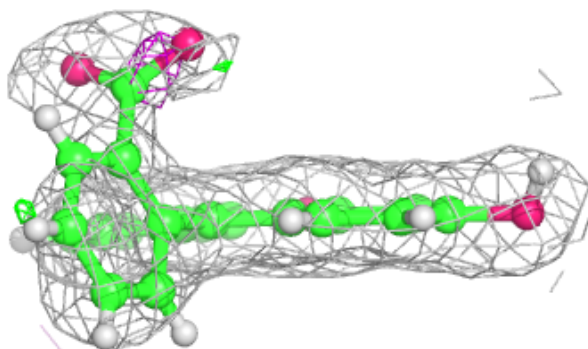
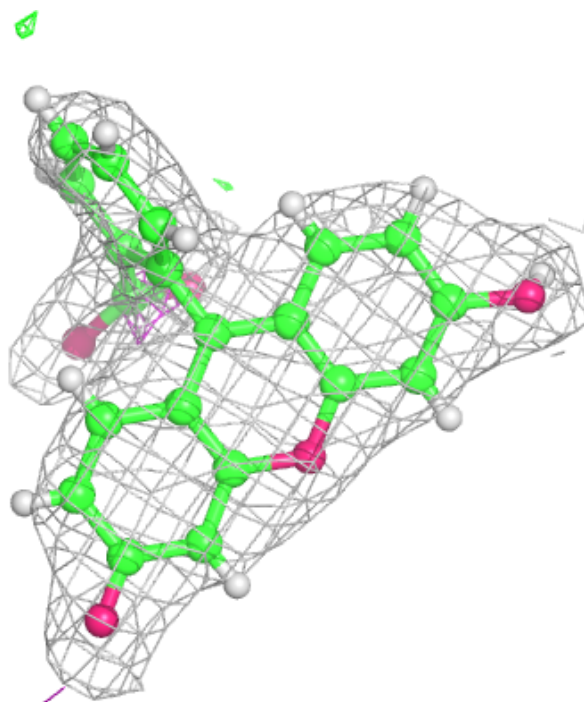
**Electron density around FLU G 301:**

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



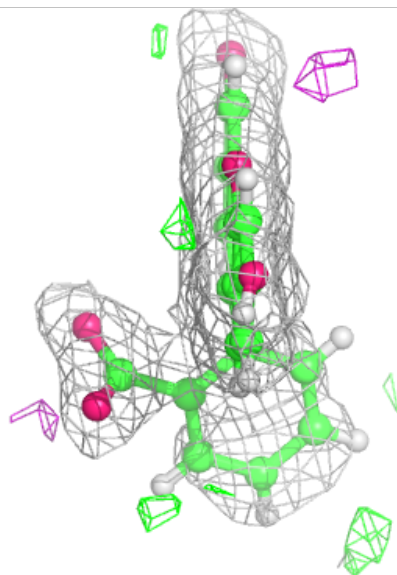
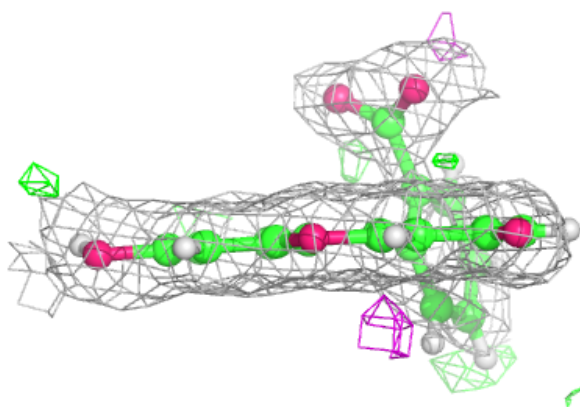
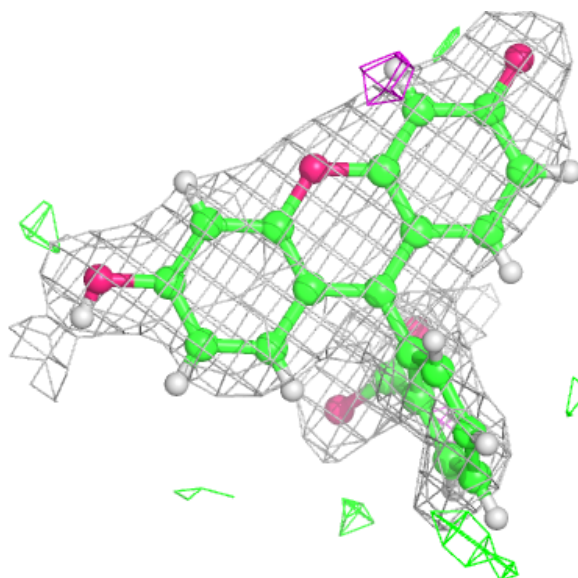
**Electron density around FLU H 301:**

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



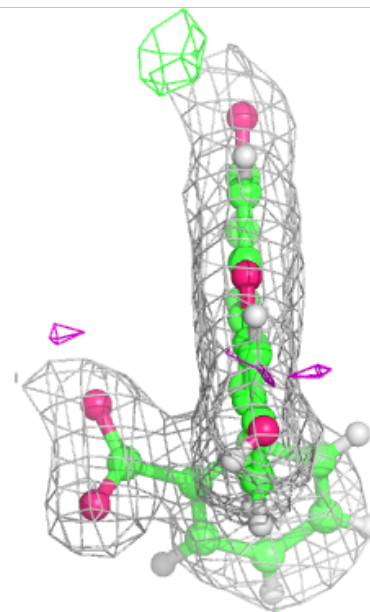
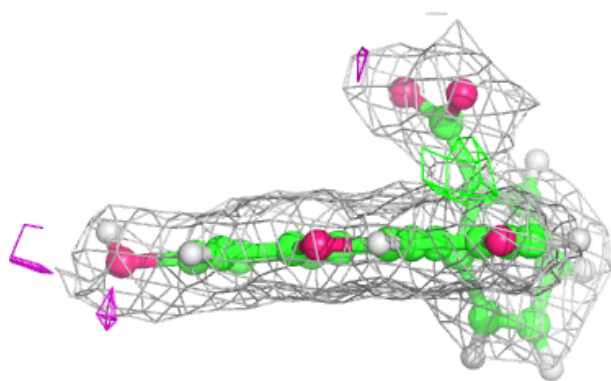
**Electron density around FLU B 301:**

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



**Electron density around FLU A 601:**

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



## 6.5 Other polymers ⓘ

There are no such residues in this entry.