



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 23, 2022 – 07:09 pm GMT

PDB ID : 6TBF  
Title : Structure of a beta galactosidase with inhibitor  
Authors : Offen, W.; Davies, G.  
Deposited on : 2019-11-01  
Resolution : 1.50 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.26  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.26

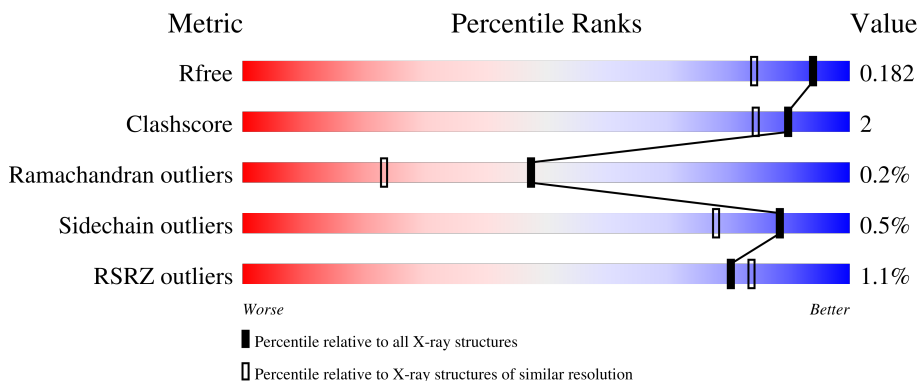
# 1 Overall quality at a glance

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.



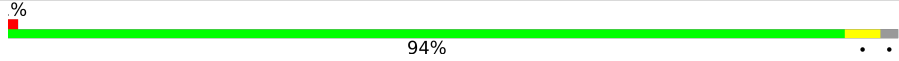
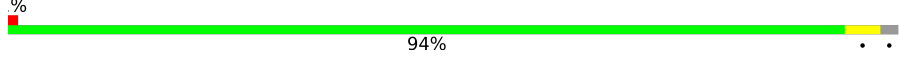
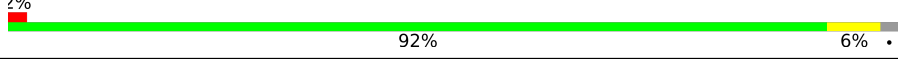
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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  $\leq 5\%$ . 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	550	 94%
1	B	550	 93%
1	C	550	 94%
1	D	550	 93%
1	E	550	 92%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	550	 <p>% 94% • •</p>
1	G	550	 <p>% 94% • •</p>
1	H	550	 <p>2% 92% 6% •</p>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 39416 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 Beta-galactosidase, putative, bgl35A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	538	4284	2740	725	803	16	0	8	0
1	B	540	4286	2739	731	801	15	0	6	0
1	C	540	4292	2744	733	799	16	0	8	0
1	D	539	4313	2759	729	809	16	0	11	0
1	E	540	4320	2763	732	808	17	0	9	0
1	F	539	4385	2803	747	819	16	0	17	0
1	G	540	4379	2799	744	819	17	0	16	0
1	H	539	4324	2765	734	809	16	0	12	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	initiating methionine	UNP B3PBE0
A	27	GLY	-	expression tag	UNP B3PBE0
A	28	SER	-	expression tag	UNP B3PBE0
A	29	SER	-	expression tag	UNP B3PBE0
A	30	HIS	-	expression tag	UNP B3PBE0
A	31	HIS	-	expression tag	UNP B3PBE0
A	32	HIS	-	expression tag	UNP B3PBE0
A	33	HIS	-	expression tag	UNP B3PBE0
A	34	HIS	-	expression tag	UNP B3PBE0
A	35	HIS	-	expression tag	UNP B3PBE0
B	26	MET	-	initiating methionine	UNP B3PBE0
B	27	GLY	-	expression tag	UNP B3PBE0
B	28	SER	-	expression tag	UNP B3PBE0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	29	SER	-	expression tag	UNP B3PBE0
B	30	HIS	-	expression tag	UNP B3PBE0
B	31	HIS	-	expression tag	UNP B3PBE0
B	32	HIS	-	expression tag	UNP B3PBE0
B	33	HIS	-	expression tag	UNP B3PBE0
B	34	HIS	-	expression tag	UNP B3PBE0
B	35	HIS	-	expression tag	UNP B3PBE0
C	26	MET	-	initiating methionine	UNP B3PBE0
C	27	GLY	-	expression tag	UNP B3PBE0
C	28	SER	-	expression tag	UNP B3PBE0
C	29	SER	-	expression tag	UNP B3PBE0
C	30	HIS	-	expression tag	UNP B3PBE0
C	31	HIS	-	expression tag	UNP B3PBE0
C	32	HIS	-	expression tag	UNP B3PBE0
C	33	HIS	-	expression tag	UNP B3PBE0
C	34	HIS	-	expression tag	UNP B3PBE0
C	35	HIS	-	expression tag	UNP B3PBE0
D	26	MET	-	initiating methionine	UNP B3PBE0
D	27	GLY	-	expression tag	UNP B3PBE0
D	28	SER	-	expression tag	UNP B3PBE0
D	29	SER	-	expression tag	UNP B3PBE0
D	30	HIS	-	expression tag	UNP B3PBE0
D	31	HIS	-	expression tag	UNP B3PBE0
D	32	HIS	-	expression tag	UNP B3PBE0
D	33	HIS	-	expression tag	UNP B3PBE0
D	34	HIS	-	expression tag	UNP B3PBE0
D	35	HIS	-	expression tag	UNP B3PBE0
E	26	MET	-	initiating methionine	UNP B3PBE0
E	27	GLY	-	expression tag	UNP B3PBE0
E	28	SER	-	expression tag	UNP B3PBE0
E	29	SER	-	expression tag	UNP B3PBE0
E	30	HIS	-	expression tag	UNP B3PBE0
E	31	HIS	-	expression tag	UNP B3PBE0
E	32	HIS	-	expression tag	UNP B3PBE0
E	33	HIS	-	expression tag	UNP B3PBE0
E	34	HIS	-	expression tag	UNP B3PBE0
E	35	HIS	-	expression tag	UNP B3PBE0
F	26	MET	-	initiating methionine	UNP B3PBE0
F	27	GLY	-	expression tag	UNP B3PBE0
F	28	SER	-	expression tag	UNP B3PBE0
F	29	SER	-	expression tag	UNP B3PBE0
F	30	HIS	-	expression tag	UNP B3PBE0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	31	HIS	-	expression tag	UNP B3PBE0
F	32	HIS	-	expression tag	UNP B3PBE0
F	33	HIS	-	expression tag	UNP B3PBE0
F	34	HIS	-	expression tag	UNP B3PBE0
F	35	HIS	-	expression tag	UNP B3PBE0
G	26	MET	-	initiating methionine	UNP B3PBE0
G	27	GLY	-	expression tag	UNP B3PBE0
G	28	SER	-	expression tag	UNP B3PBE0
G	29	SER	-	expression tag	UNP B3PBE0
G	30	HIS	-	expression tag	UNP B3PBE0
G	31	HIS	-	expression tag	UNP B3PBE0
G	32	HIS	-	expression tag	UNP B3PBE0
G	33	HIS	-	expression tag	UNP B3PBE0
G	34	HIS	-	expression tag	UNP B3PBE0
G	35	HIS	-	expression tag	UNP B3PBE0
H	26	MET	-	initiating methionine	UNP B3PBE0
H	27	GLY	-	expression tag	UNP B3PBE0
H	28	SER	-	expression tag	UNP B3PBE0
H	29	SER	-	expression tag	UNP B3PBE0
H	30	HIS	-	expression tag	UNP B3PBE0
H	31	HIS	-	expression tag	UNP B3PBE0
H	32	HIS	-	expression tag	UNP B3PBE0
H	33	HIS	-	expression tag	UNP B3PBE0
H	34	HIS	-	expression tag	UNP B3PBE0
H	35	HIS	-	expression tag	UNP B3PBE0

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

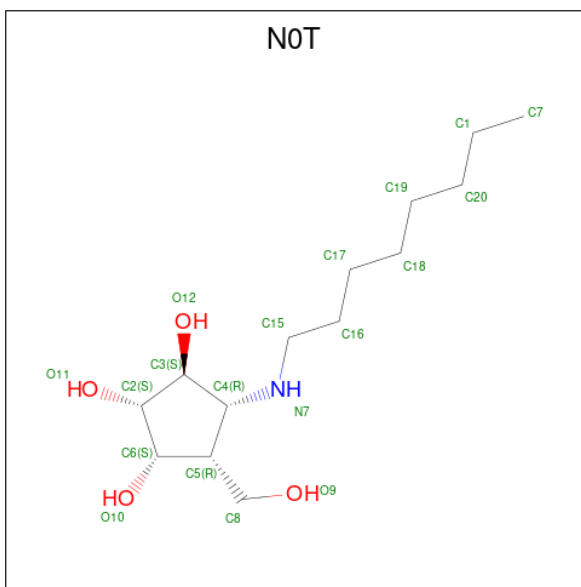
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	4	Total Na 4 4	0	0
2	B	2	Total Na 2 2	0	0
2	C	4	Total Na 4 4	0	0
2	D	5	Total Na 5 5	0	0
2	E	6	Total Na 6 6	0	0
2	F	4	Total Na 4 4	0	0
2	G	5	Total Na 5 5	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	5	Total	Na	0	0
			5	5		

- Molecule 3 is (1 {S},2 {S},3 {S},4 {R},5 {R})-4-(hydroxymethyl)-5-(octylamino)cyclopentane-1,2,3-triol (three-letter code: N0T) (formula: C<sub>14</sub>H<sub>29</sub>NO<sub>4</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			19	14	1	4		
3	B	1	Total	C	N	O	0	0
			19	14	1	4		
3	C	1	Total	C	N	O	0	0
			19	14	1	4		
3	D	1	Total	C	N	O	0	0
			19	14	1	4		
3	E	1	Total	C	N	O	0	0
			19	14	1	4		
3	F	1	Total	C	N	O	0	0
			19	14	1	4		
3	G	1	Total	C	N	O	0	0
			19	14	1	4		
3	H	1	Total	C	N	O	0	0
			19	14	1	4		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



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



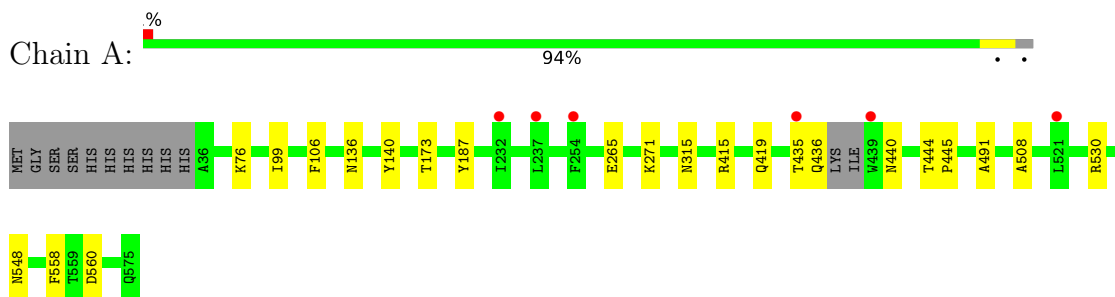
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	444	Total O 445 445	0	1
5	B	528	Total O 529 529	0	1
5	C	532	Total O 534 534	0	2
5	D	640	Total O 643 643	0	3
5	E	653	Total O 654 654	0	1
5	F	622	Total O 624 624	0	2
5	G	681	Total O 681 681	0	0
5	H	480	Total O 480 480	0	0

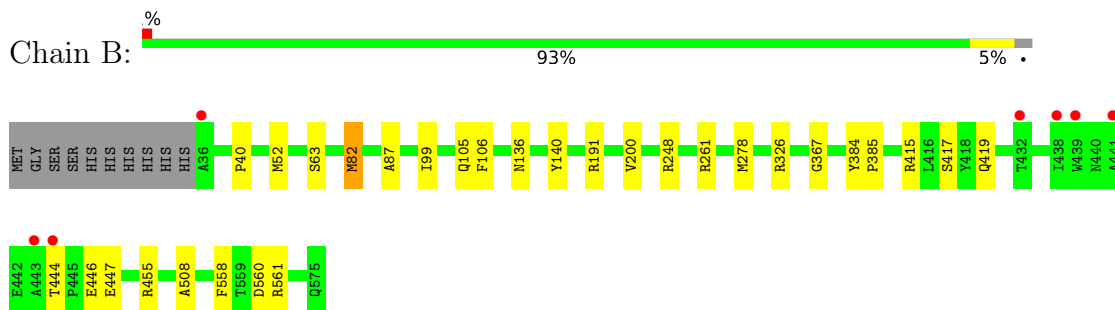
### 3 Residue-property plots [i](#)

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

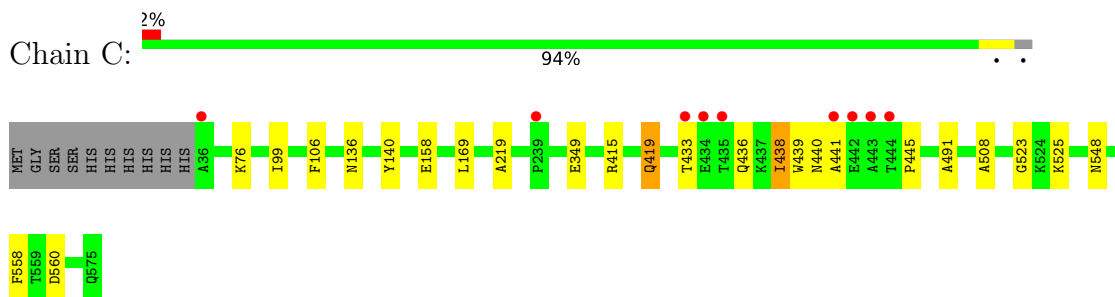
- Molecule 1: Beta-galactosidase, putative, bgl35A



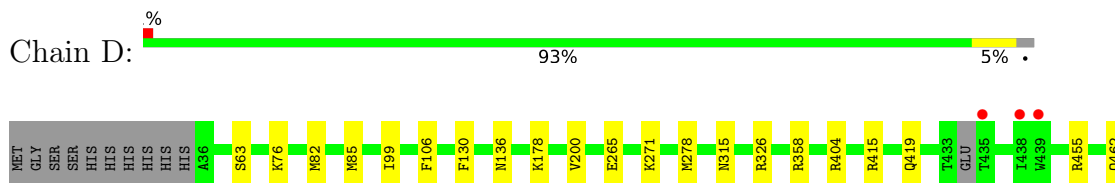
- Molecule 1: Beta-galactosidase, putative, bgl35A



- Molecule 1: Beta-galactosidase, putative, bgl35A

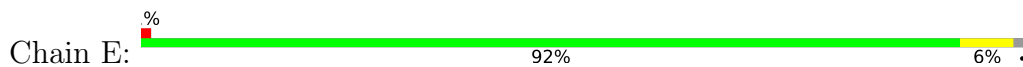


- Molecule 1: Beta-galactosidase, putative, bgl35A

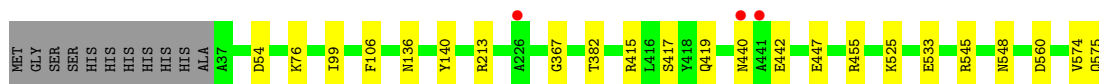
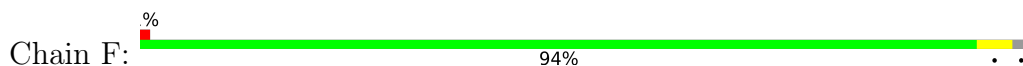




- Molecule 1: Beta-galactosidase, putative, bgl35A



- Molecule 1: Beta-galactosidase, putative, bgl35A



- Molecule 1: Beta-galactosidase, putative, bgl35A



- Molecule 1: Beta-galactosidase, putative, bgl35A



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.42Å 115.63Å 115.94Å 90.35° 90.02° 90.16°	Depositor
Resolution (Å)	115.94 – 1.50 115.94 – 1.50	Depositor EDS
% Data completeness (in resolution range)	96.5 (115.94-1.50) 96.6 (115.94-1.50)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, $R_{free}$	0.128 , 0.178 0.135 , 0.182	Depositor DCC
$R_{free}$ test set	40201 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.1	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.005 for h,l,-k 0.005 for h,-l,k 0.006 for h,-k,-l 0.000 for -h,k,-l 0.000 for -h,-k,l 0.000 for -h,l,k 0.000 for -h,-l,-k	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	39416	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.55% 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: ACT, N0T, NA

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	0.84	2/4397 (0.0%)	0.97	3/5992 (0.1%)
1	B	0.86	2/4400 (0.0%)	0.98	4/6000 (0.1%)
1	C	0.88	4/4405 (0.1%)	0.96	1/6003 (0.0%)
1	D	0.88	0/4429	1.00	6/6033 (0.1%)
1	E	0.88	1/4434 (0.0%)	1.02	7/6039 (0.1%)
1	F	0.89	4/4502 (0.1%)	0.99	2/6123 (0.0%)
1	G	0.88	1/4493 (0.0%)	1.00	4/6115 (0.1%)
1	H	0.87	2/4441 (0.0%)	0.94	5/6049 (0.1%)
All	All	0.87	16/35501 (0.0%)	0.98	32/48354 (0.1%)

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	158	GLU	CD-OE1	9.89	1.36	1.25
1	B	63	SER	CB-OG	-6.34	1.34	1.42
1	C	349	GLU	CD-OE1	6.07	1.32	1.25
1	C	560	ASP	CG-OD2	-6.04	1.11	1.25
1	G	442	GLU	CD-OE2	5.88	1.32	1.25

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	530	ARG	NE-CZ-NH2	-10.66	114.97	120.30
1	H	530	ARG	NE-CZ-NH2	-7.34	116.63	120.30
1	E	455	ARG	NE-CZ-NH1	-7.28	116.66	120.30
1	E	455	ARG	NE-CZ-NH2	7.00	123.80	120.30
1	B	326	ARG	NE-CZ-NH1	-6.64	116.98	120.30

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	4284	0	4096	9	0
1	B	4286	0	4105	25	0
1	C	4292	0	4108	13	0
1	D	4313	0	4130	14	0
1	E	4320	0	4154	22	0
1	F	4385	0	4222	11	0
1	G	4379	0	4204	15	0
1	H	4324	0	4146	19	0
2	A	4	0	0	0	0
2	B	2	0	0	0	0
2	C	4	0	0	0	0
2	D	5	0	0	0	0
2	E	6	0	0	0	0
2	F	4	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
3	A	19	0	0	0	0
3	B	19	0	0	0	0
3	C	19	0	0	0	0
3	D	19	0	0	0	0
3	E	19	0	0	0	0
3	F	19	0	0	0	0
3	G	19	0	0	0	0
3	H	19	0	0	0	0
4	A	4	0	3	0	0
4	B	4	0	3	0	0
4	D	12	0	9	0	0
4	E	8	0	6	1	0
4	F	8	0	6	0	0
4	G	16	0	12	0	0
4	H	4	0	3	0	0
5	A	445	0	0	2	0
5	B	529	0	0	13	0
5	C	534	0	0	0	0
5	D	643	0	0	7	0
5	E	654	0	0	8	0
5	F	624	0	0	6	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	G	681	0	0	5	0
5	H	480	0	0	4	0
All	All	39416	0	33207	122	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.

The worst 5 of 122 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:B:248:ARG:HD2	5:B:1001:HOH:O	1.37	1.19
1:H:436:GLN:O	1:H:440:ASN:HB2	1.51	1.10
1:E:76:LYS:HE2	5:E:991:HOH:O	1.63	0.98
1:B:191[B]:ARG:NH1	5:B:702:HOH:O	1.61	0.92
1:D:462[B]:GLN:NE2	5:D:702:HOH:O	1.95	0.90

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	542/550 (98%)	519 (96%)	21 (4%)	2 (0%)	34	13
1	B	544/550 (99%)	526 (97%)	18 (3%)	0	100	100
1	C	546/550 (99%)	521 (95%)	21 (4%)	4 (1%)	22	6
1	D	546/550 (99%)	527 (96%)	19 (4%)	0	100	100
1	E	547/550 (100%)	528 (96%)	19 (4%)	0	100	100
1	F	555/550 (101%)	534 (96%)	20 (4%)	1 (0%)	47	23
1	G	554/550 (101%)	535 (97%)	18 (3%)	1 (0%)	47	23
1	H	549/550 (100%)	527 (96%)	21 (4%)	1 (0%)	47	23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	4383/4400 (100%)	4217 (96%)	157 (4%)	9 (0%)	47	23

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	445	PRO
1	A	435	THR
1	C	439	TRP
1	F	440	ASN
1	C	491	ALA

### 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	439/461 (95%)	437 (100%)	2 (0%)	88	78
1	B	440/461 (95%)	439 (100%)	1 (0%)	93	86
1	C	438/461 (95%)	435 (99%)	3 (1%)	84	69
1	D	443/461 (96%)	442 (100%)	1 (0%)	93	86
1	E	446/461 (97%)	445 (100%)	1 (0%)	93	86
1	F	452/461 (98%)	449 (99%)	3 (1%)	84	69
1	G	450/461 (98%)	449 (100%)	1 (0%)	93	86
1	H	444/461 (96%)	439 (99%)	5 (1%)	73	53
All	All	3552/3688 (96%)	3535 (100%)	17 (0%)	88	78

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	289	ASN
1	H	442	GLU
1	E	136	ASN
1	F	136	ASN
1	F	419	GLN



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

Mol	Chain	Res	Type
1	C	421	GLN
1	D	45	ASN
1	H	289	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 57 ligands modelled in this entry, 35 are monoatomic - leaving 22 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	ACT	G	607	-	1,3,3	5.12	1 (100%)	0,3,3	-	-
3	N0T	C	603	-	19,19,19	1.22	3 (15%)	21,24,24	1.77	5 (23%)
4	ACT	D	604	-	1,3,3	6.01	1 (100%)	0,3,3	-	-
3	N0T	D	605	-	19,19,19	1.00	1 (5%)	21,24,24	1.52	5 (23%)
4	ACT	A	603	-	1,3,3	4.55	1 (100%)	0,3,3	-	-
3	N0T	H	602	-	19,19,19	1.28	1 (5%)	21,24,24	1.23	3 (14%)
4	ACT	F	606	-	1,3,3	3.02	1 (100%)	0,3,3	-	-
4	ACT	D	607	-	1,3,3	1.74	0	0,3,3	-	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	N0T	E	604	-	19,19,19	1.35	2 (10%)	21,24,24	1.27	2 (9%)
4	ACT	G	604	-	1,3,3	4.39	1 (100%)	0,3,3	-	-
4	ACT	E	605	-	1,3,3	2.23	1 (100%)	0,3,3	-	-
4	ACT	B	603	-	1,3,3	2.61	1 (100%)	0,3,3	-	-
3	N0T	G	603	-	19,19,19	1.67	3 (15%)	21,24,24	1.61	6 (28%)
3	N0T	A	602	-	19,19,19	1.23	1 (5%)	21,24,24	1.25	1 (4%)
4	ACT	H	603	-	1,3,3	7.27	1 (100%)	0,3,3	-	-
4	ACT	D	606	-	1,3,3	1.10	0	0,3,3	-	-
4	ACT	F	605	-	1,3,3	7.37	1 (100%)	0,3,3	-	-
3	N0T	F	604	-	19,19,19	1.35	3 (15%)	21,24,24	1.42	4 (19%)
4	ACT	G	605	-	1,3,3	5.58	1 (100%)	0,3,3	-	-
4	ACT	G	606	-	1,3,3	7.12	1 (100%)	0,3,3	-	-
3	N0T	B	602	-	19,19,19	1.34	2 (10%)	21,24,24	1.37	3 (14%)
4	ACT	E	606	-	1,3,3	3.87	1 (100%)	0,3,3	-	-

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	N0T	C	603	-	-	0/11/31/31	0/1/1/1
3	N0T	F	604	-	-	1/11/31/31	0/1/1/1
3	N0T	H	602	-	-	1/11/31/31	0/1/1/1
3	N0T	D	605	-	-	1/11/31/31	0/1/1/1
3	N0T	A	602	-	-	3/11/31/31	0/1/1/1
3	N0T	E	604	-	-	1/11/31/31	0/1/1/1
3	N0T	G	603	-	-	2/11/31/31	0/1/1/1
3	N0T	B	602	-	-	2/11/31/31	0/1/1/1

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	605	ACT	CH3-C	7.37	1.58	1.48
4	H	603	ACT	CH3-C	7.27	1.58	1.48
4	G	606	ACT	CH3-C	7.12	1.57	1.48
4	D	604	ACT	CH3-C	6.01	1.56	1.48
3	G	603	N0T	C4-N7	5.70	1.57	1.47

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	603	N0T	C5-C6-C2	-4.69	97.77	103.81
3	A	602	N0T	C5-C6-C2	-4.04	98.61	103.81
3	D	605	N0T	C5-C6-C2	-3.51	99.29	103.81
3	E	604	N0T	C5-C6-C2	-3.29	99.57	103.81
3	C	603	N0T	C5-C4-N7	3.21	119.48	112.52

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

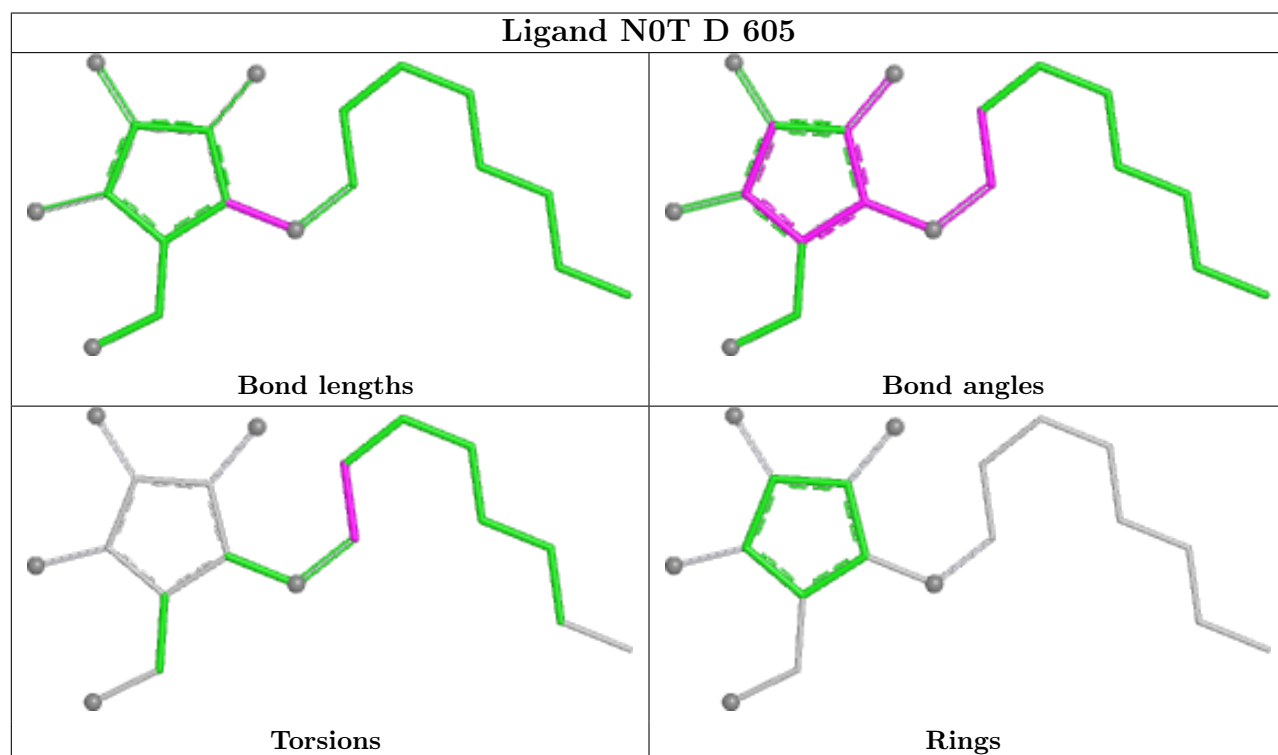
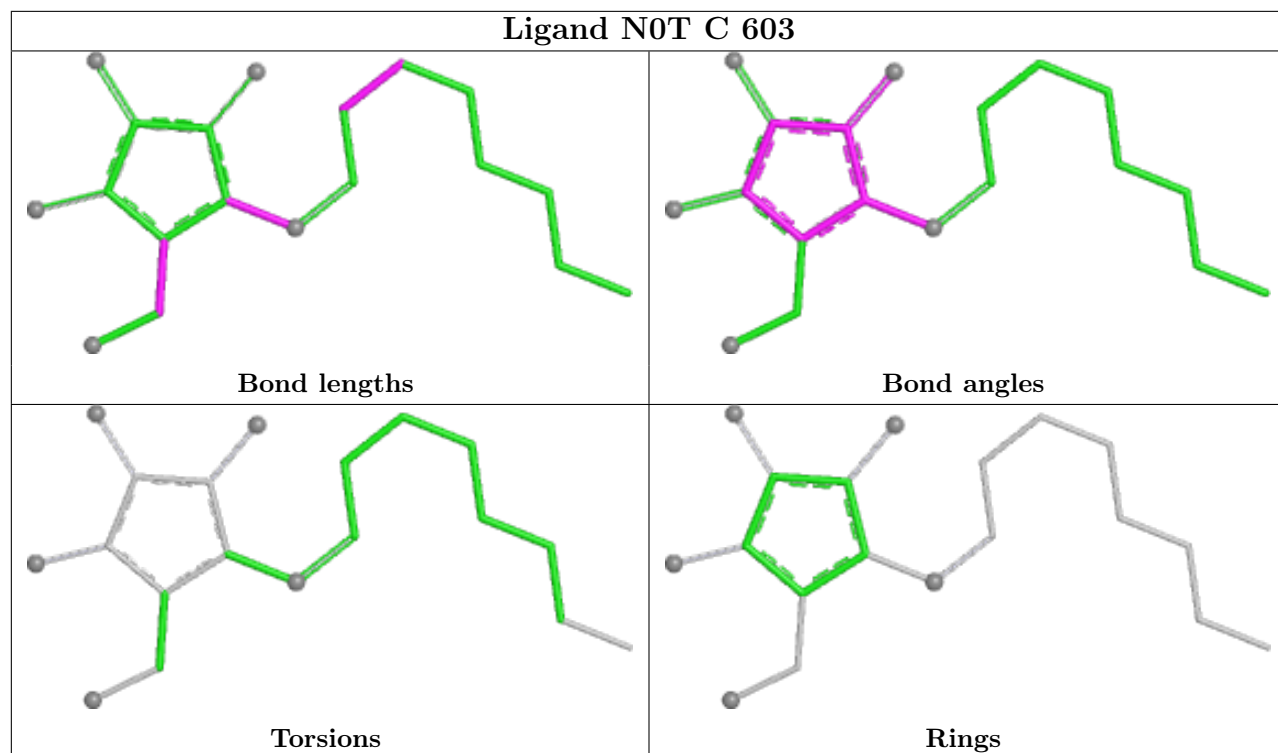
Mol	Chain	Res	Type	Atoms
3	D	605	N0T	N7-C15-C16-C17
3	H	602	N0T	N7-C15-C16-C17
3	F	604	N0T	N7-C15-C16-C17
3	G	603	N0T	N7-C15-C16-C17
3	A	602	N0T	C15-C16-C17-C18

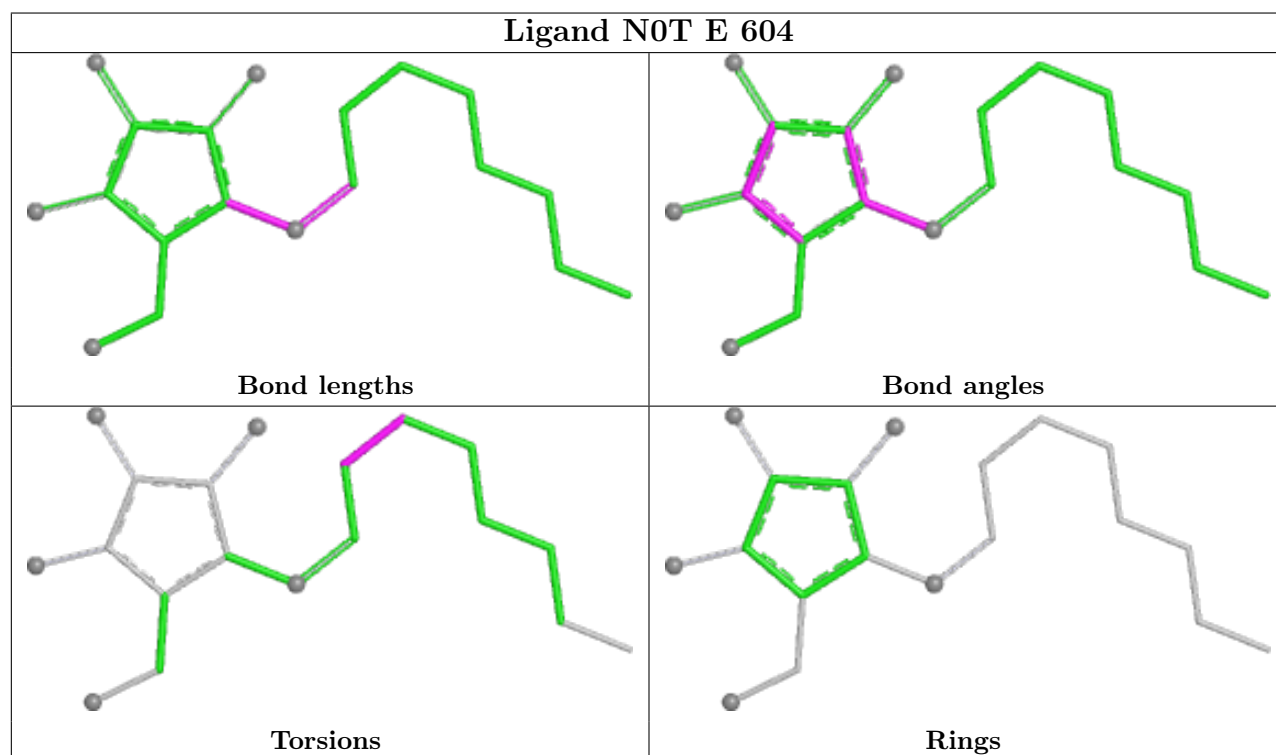
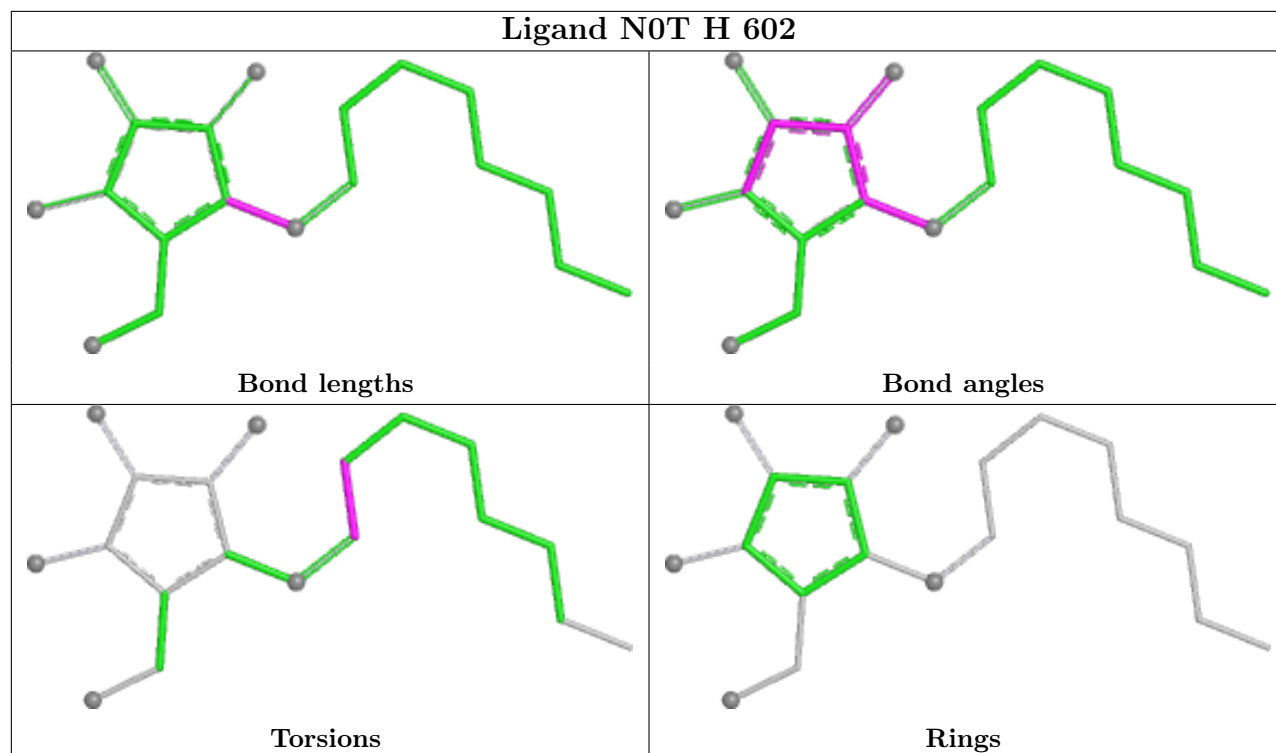
There are no ring outliers.

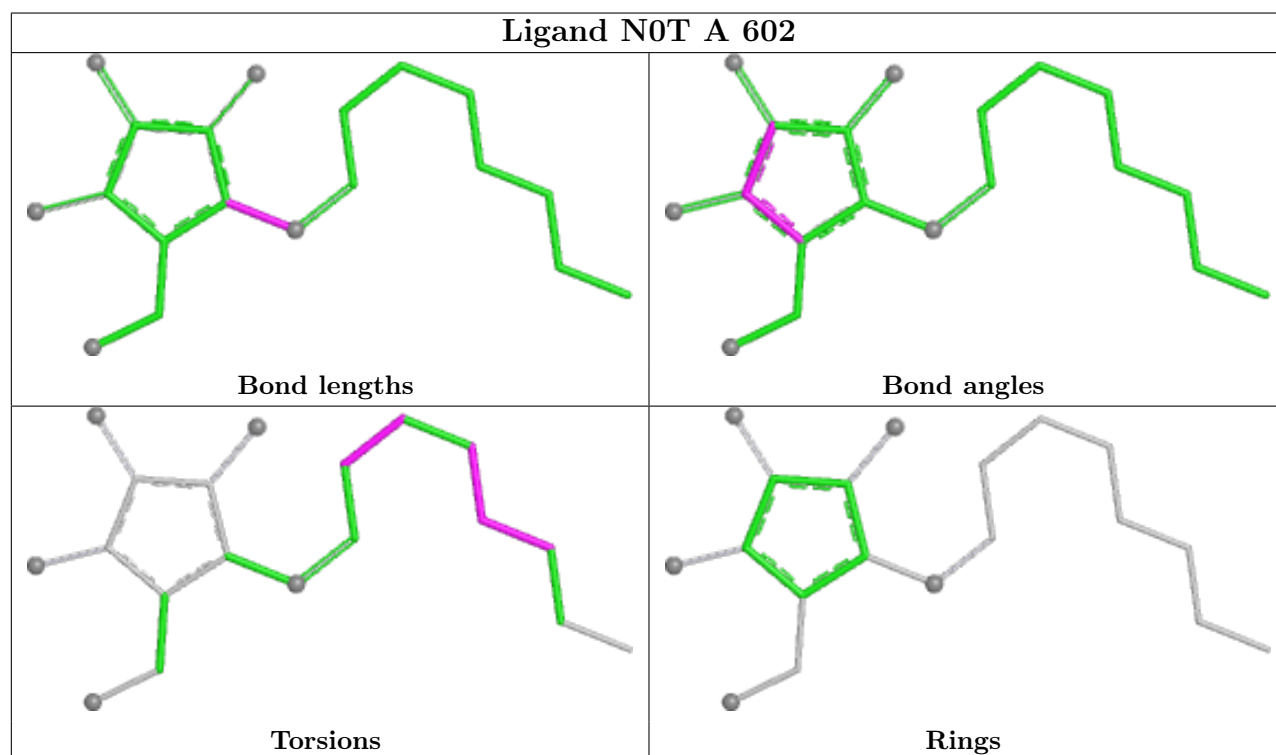
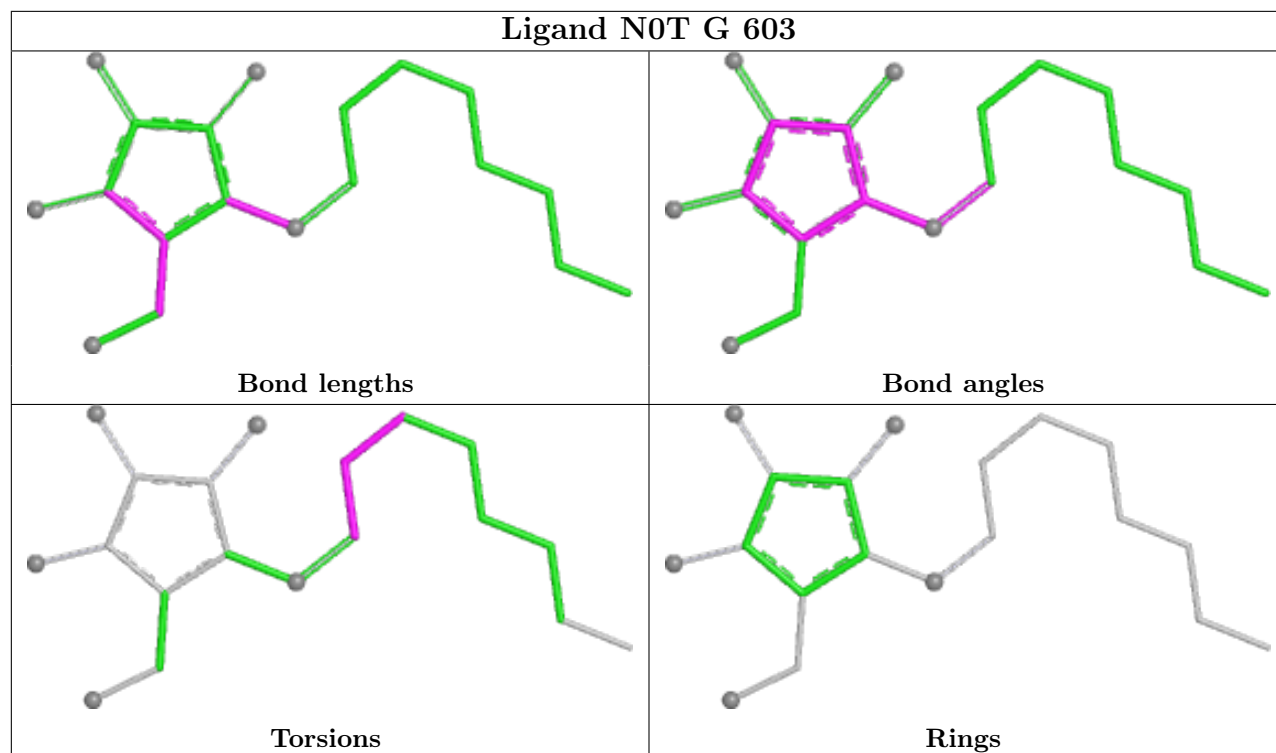
1 monomer is involved in 1 short contact:

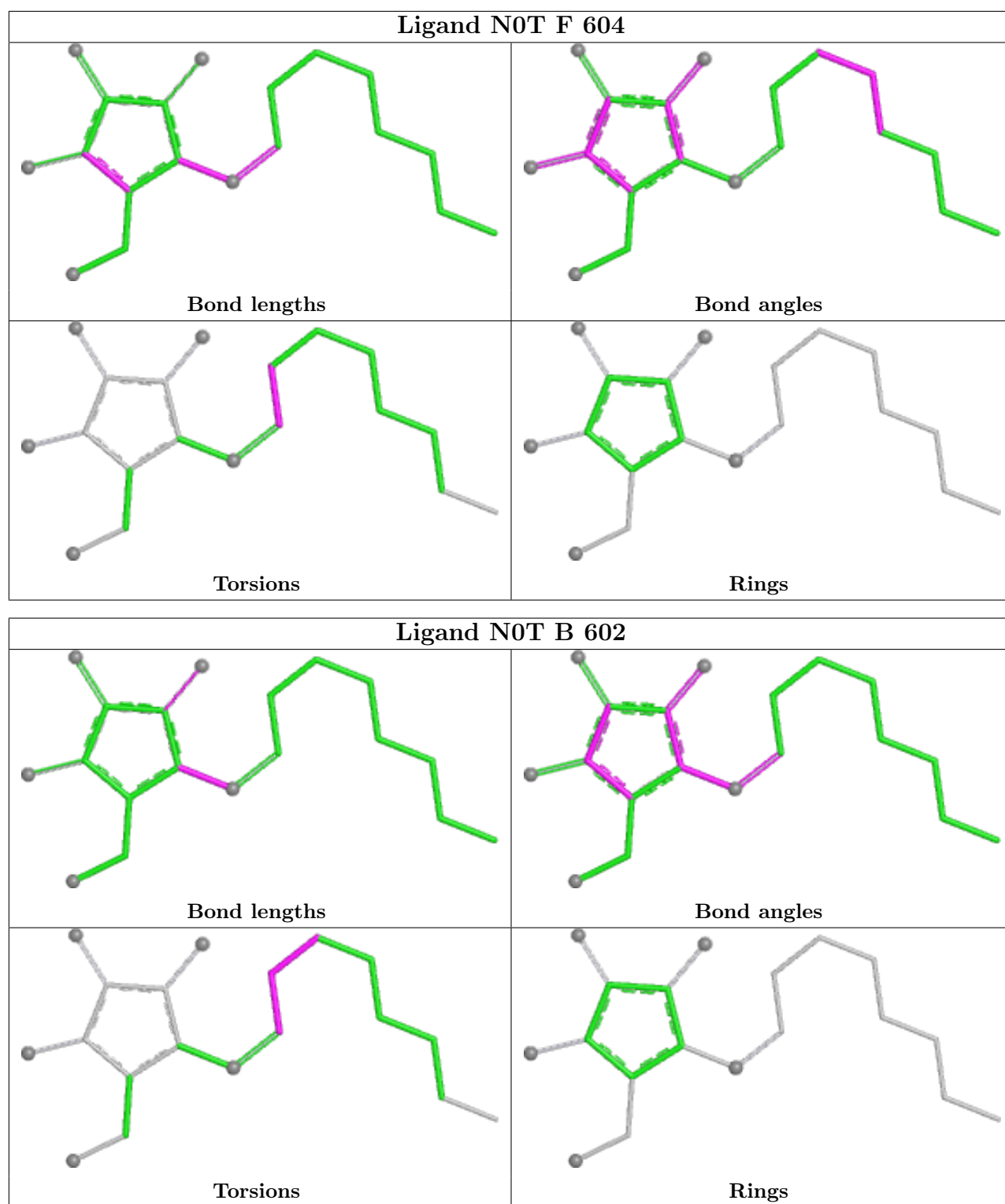
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	606	ACT	1	0

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

There are no chain breaks in this entry.



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

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

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	538/550 (97%)	-0.21	6 (1%) 80 84	18, 27, 51, 86	41 (7%)
1	B	540/550 (98%)	-0.32	7 (1%) 77 81	18, 24, 42, 64	23 (4%)
1	C	540/550 (98%)	-0.26	9 (1%) 70 75	17, 23, 44, 66	29 (5%)
1	D	539/550 (98%)	-0.30	3 (0%) 89 91	14, 20, 39, 68	25 (4%)
1	E	540/550 (98%)	-0.31	7 (1%) 77 81	14, 21, 40, 60	28 (5%)
1	F	539/550 (98%)	-0.28	3 (0%) 89 91	14, 22, 43, 77	24 (4%)
1	G	540/550 (98%)	-0.29	4 (0%) 87 90	14, 20, 40, 76	26 (4%)
1	H	539/550 (98%)	-0.25	9 (1%) 70 75	18, 27, 50, 68	39 (7%)
All	All	4315/4400 (98%)	-0.28	48 (1%) 80 84	14, 23, 45, 86	235 (5%)

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	439	TRP	5.3
1	D	439	TRP	5.2
1	H	444	THR	5.0
1	C	441	ALA	4.5
1	G	435	THR	4.4

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands i

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	ACT	H	603	4/4	0.82	0.12	29,32,34,39	4
4	ACT	F	605	4/4	0.86	0.13	28,28,31,33	4
4	ACT	D	604	4/4	0.86	0.18	28,29,32,38	0
4	ACT	G	606	4/4	0.87	0.14	27,27,30,33	4
4	ACT	E	606	4/4	0.87	0.17	30,30,32,33	4
4	ACT	D	607	4/4	0.88	0.11	26,31,36,39	4
4	ACT	G	607	4/4	0.91	0.09	30,30,33,34	4
4	ACT	F	606	4/4	0.91	0.17	23,27,27,30	4
2	NA	C	602	1/1	0.92	0.08	36,36,36,36	0
4	ACT	B	603	4/4	0.93	0.10	31,34,36,37	4
4	ACT	D	606	4/4	0.93	0.19	28,34,36,37	4
4	ACT	G	605	4/4	0.93	0.15	36,38,44,47	4
4	ACT	E	605	4/4	0.94	0.14	22,23,23,29	4
3	N0T	C	603	19/19	0.94	0.09	15,19,31,38	5
4	ACT	A	603	4/4	0.95	0.07	27,27,30,32	4
4	ACT	G	604	4/4	0.95	0.14	23,24,24,26	4
2	NA	C	601	1/1	0.95	0.10	47,47,47,47	0
3	N0T	D	605	19/19	0.96	0.08	14,15,36,36	0
3	N0T	H	602	19/19	0.96	0.08	17,19,32,37	5
3	N0T	B	602	19/19	0.97	0.07	17,19,38,46	0
2	NA	A	605	1/1	0.97	0.12	36,36,36,36	0
2	NA	E	608	1/1	0.97	0.06	41,41,41,41	0
3	N0T	E	604	19/19	0.97	0.09	14,15,34,40	5
3	N0T	F	604	19/19	0.97	0.09	14,17,36,37	4
3	N0T	G	603	19/19	0.97	0.09	13,15,29,33	5
3	N0T	A	602	19/19	0.97	0.07	17,20,48,49	2
2	NA	D	609	1/1	0.98	0.06	27,27,27,27	0
2	NA	H	601	1/1	0.98	0.09	33,33,33,33	0
2	NA	E	603	1/1	0.99	0.09	22,22,22,22	0
2	NA	E	607	1/1	0.99	0.06	30,30,30,30	0
2	NA	B	604	1/1	0.99	0.07	30,30,30,30	0
2	NA	E	609	1/1	0.99	0.08	22,22,22,22	0
2	NA	F	601	1/1	0.99	0.05	26,26,26,26	0
2	NA	F	607	1/1	0.99	0.06	28,28,28,28	0
2	NA	G	608	1/1	0.99	0.06	30,30,30,30	0
2	NA	G	610	1/1	0.99	0.09	27,27,27,27	0
2	NA	A	601	1/1	0.99	0.15	33,33,33,33	0

*Continued on next page...*

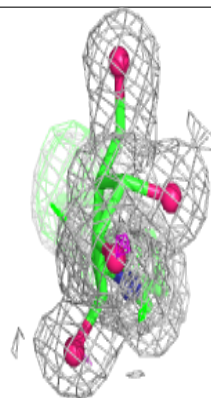
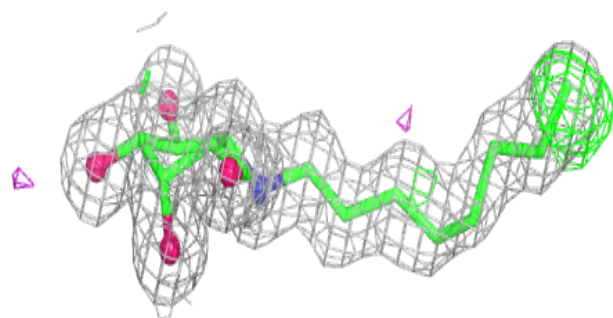
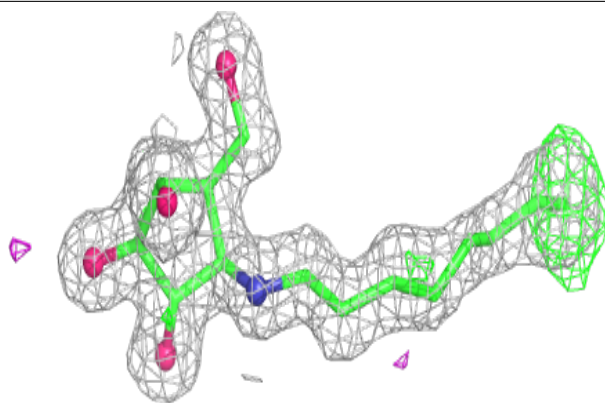
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NA	H	604	1/1	0.99	0.07	26,26,26,26	0
2	NA	H	606	1/1	0.99	0.04	35,35,35,35	1
2	NA	H	607	1/1	0.99	0.12	38,38,38,38	0
2	NA	A	606	1/1	0.99	0.06	34,34,34,34	0
2	NA	D	601	1/1	0.99	0.06	24,24,24,24	0
2	NA	D	602	1/1	0.99	0.07	24,24,24,24	0
2	NA	B	601	1/1	0.99	0.11	35,35,35,35	0
2	NA	E	602	1/1	0.99	0.08	21,21,21,21	0
2	NA	E	601	1/1	1.00	0.05	23,23,23,23	0
2	NA	F	602	1/1	1.00	0.07	25,25,25,25	0
2	NA	F	603	1/1	1.00	0.06	22,22,22,22	0
2	NA	A	604	1/1	1.00	0.08	27,27,27,27	0
2	NA	G	601	1/1	1.00	0.10	28,28,28,28	0
2	NA	G	602	1/1	1.00	0.07	22,22,22,22	0
2	NA	C	604	1/1	1.00	0.05	26,26,26,26	0
2	NA	G	609	1/1	1.00	0.06	24,24,24,24	0
2	NA	D	603	1/1	1.00	0.07	21,21,21,21	0
2	NA	D	608	1/1	1.00	0.06	24,24,24,24	0
2	NA	C	605	1/1	1.00	0.06	29,29,29,29	0
2	NA	H	605	1/1	1.00	0.05	30,30,30,30	0

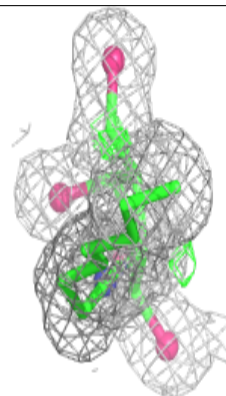
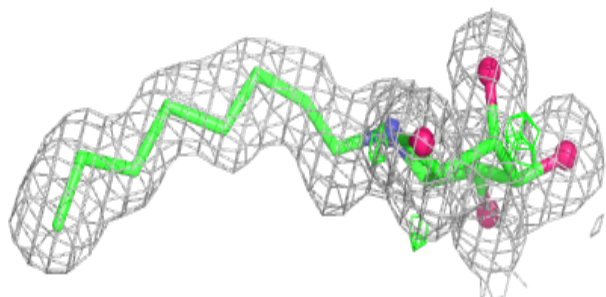
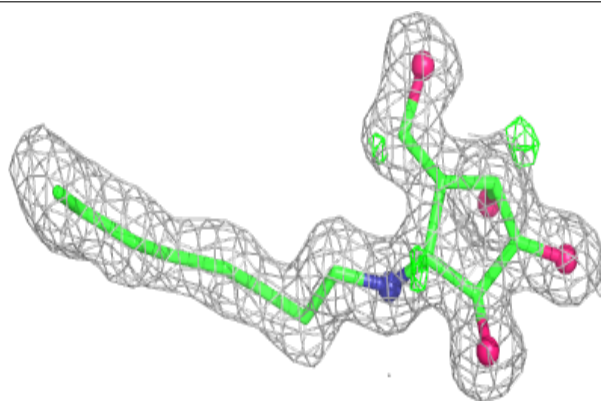
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around N0T C 603:**

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

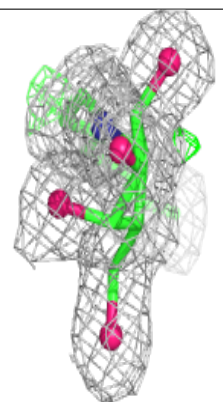
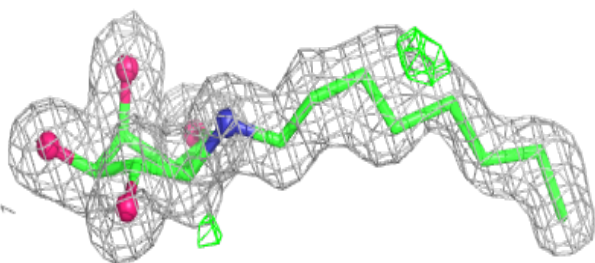
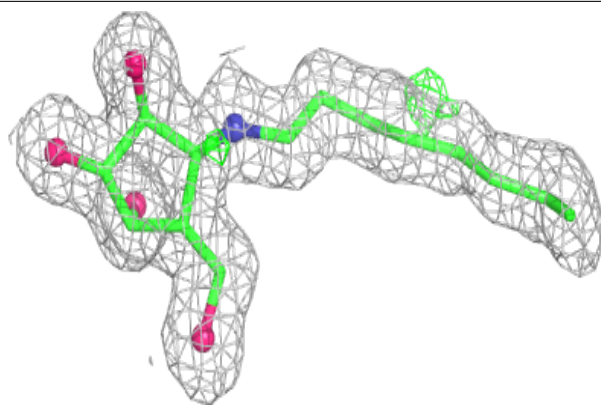
**Electron density around N0T D 605:**

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

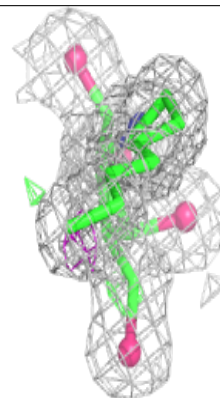
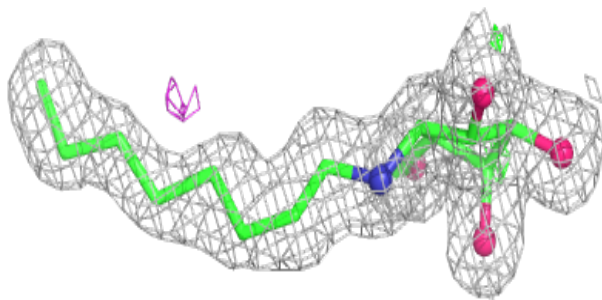
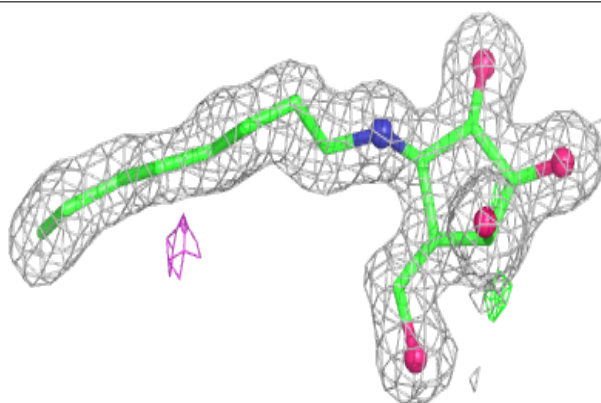


**Electron density around N0T H 602:**

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

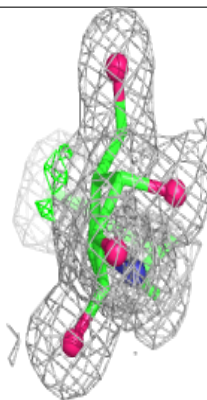
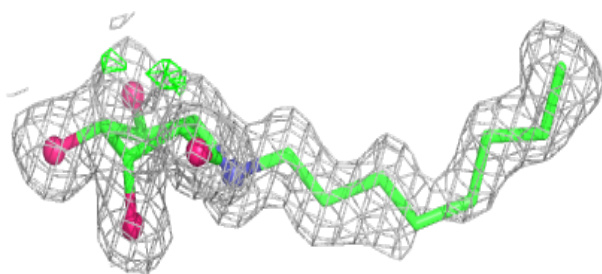
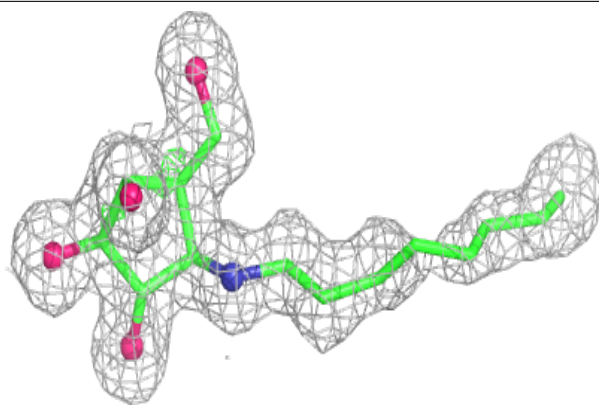
**Electron density around N0T B 602:**

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

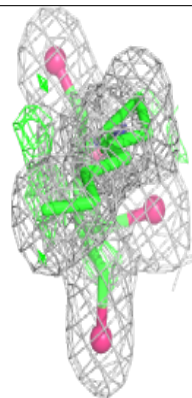
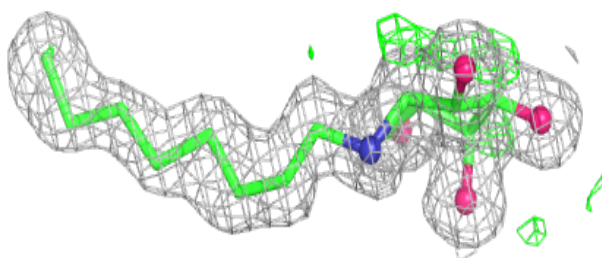
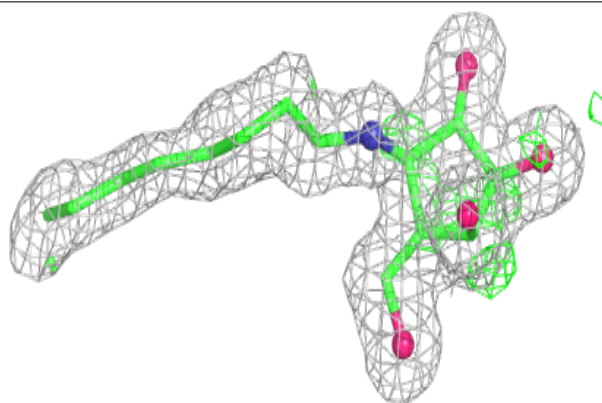


**Electron density around N0T E 604:**

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

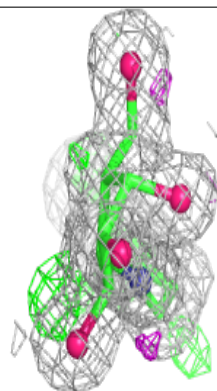
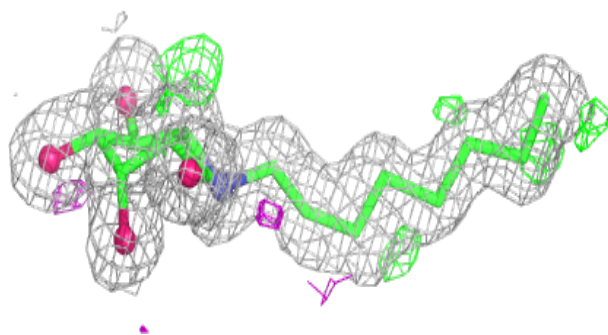
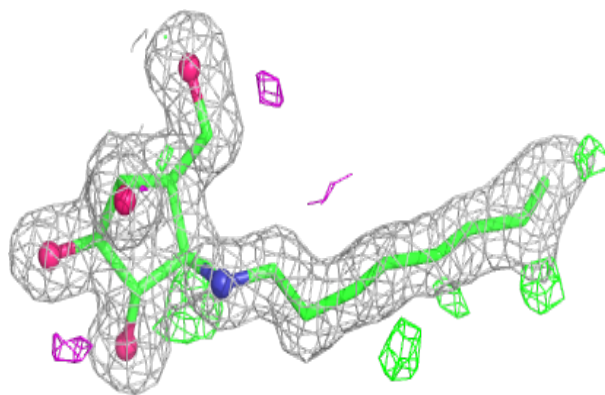
**Electron density around N0T F 604:**

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

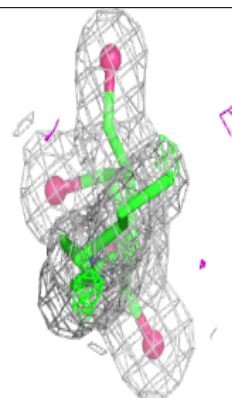
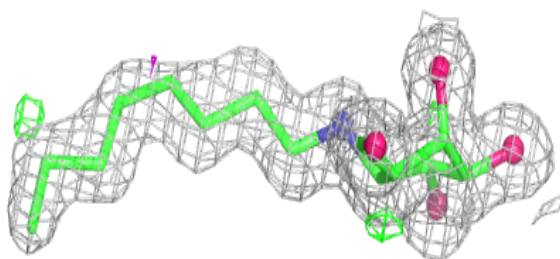
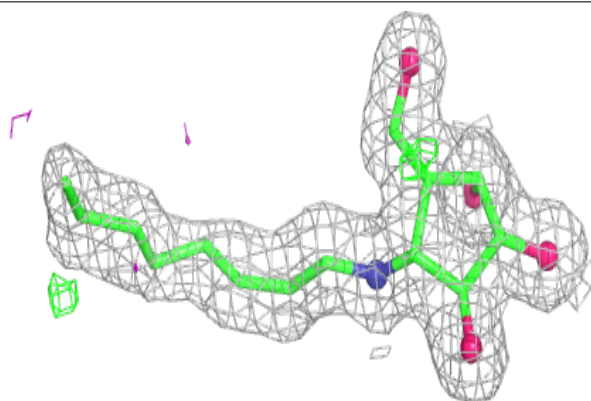


**Electron density around N0T G 603:**

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

**Electron density around N0T A 602:**

$2mF_o-DF_c$  (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.