



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 10:14 PM EDT

PDB ID : 7RFK  
Title : CamA Adenine Methyltransferase Complexed to Cognate Substrate DNA and Inhibitor Sinefungin  
Authors : Horton, J.R.; Cheng, X.; Zhou, J.  
Deposited on : 2021-07-14  
Resolution : 2.05 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

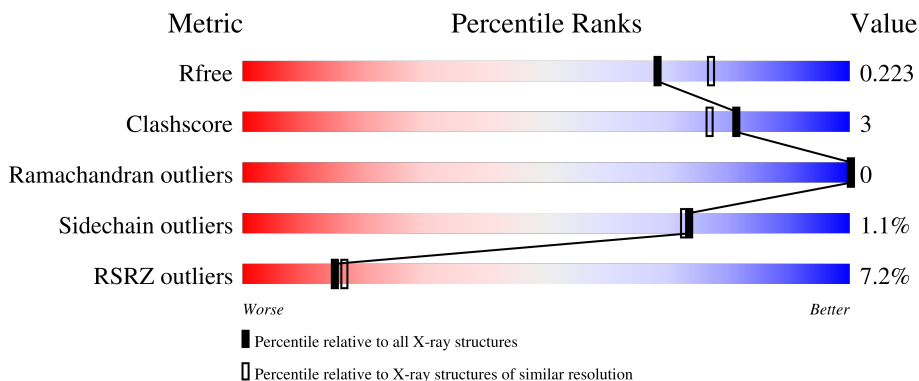
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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.



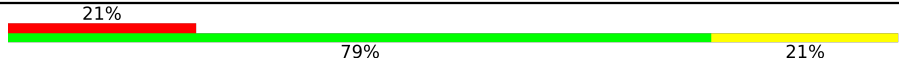
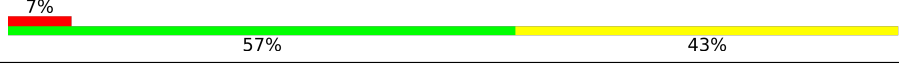
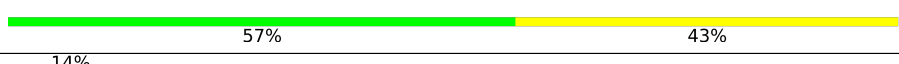

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	578	 3% 93% 5% ..
1	B	578	 4% 93% 6% ..
1	C	578	 13% 90% 7% .
2	E	14	 14% 57% 36% 7%
2	G	14	 100%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	I	14	 21% 79% 21%
3	D	14	 7% 57% 43%
3	F	14	 0% 57% 43%
3	H	14	 14% 71% 29%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 17045 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 Site-specific DNA-methyltransferase (adenine-specific).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	568	Total 4718	C 3064	N 755	O 882	S 17	0	2	0
1	B	575	Total 4790	C 3113	N 770	O 890	S 17	0	4	0
1	C	562	Total 4506	C 2928	N 719	O 842	S 17	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	HIS	-	expression tag	UNP Q183J3
B	0	HIS	-	expression tag	UNP Q183J3
C	0	HIS	-	expression tag	UNP Q183J3

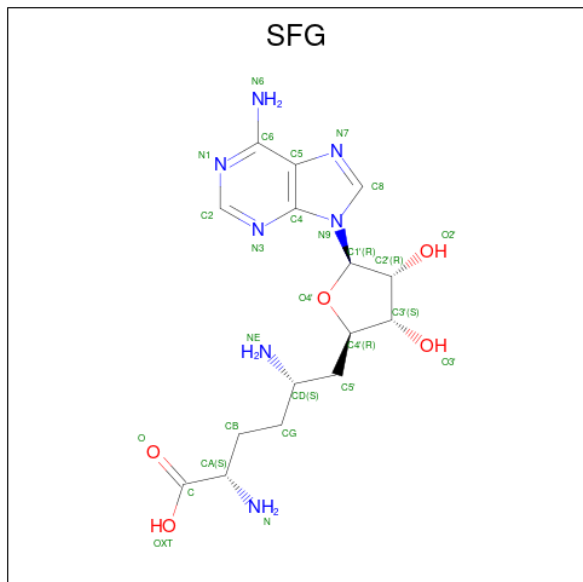
- Molecule 2 is a DNA chain called DNA Strand 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	E	13	Total 269	C 129	N 45	O 82	P 13	0	0	0
2	G	14	Total 287	C 139	N 50	O 85	P 13	0	0	0
2	I	14	Total 287	C 139	N 50	O 85	P 13	0	0	0

- Molecule 3 is a DNA chain called DNA Strand 1.

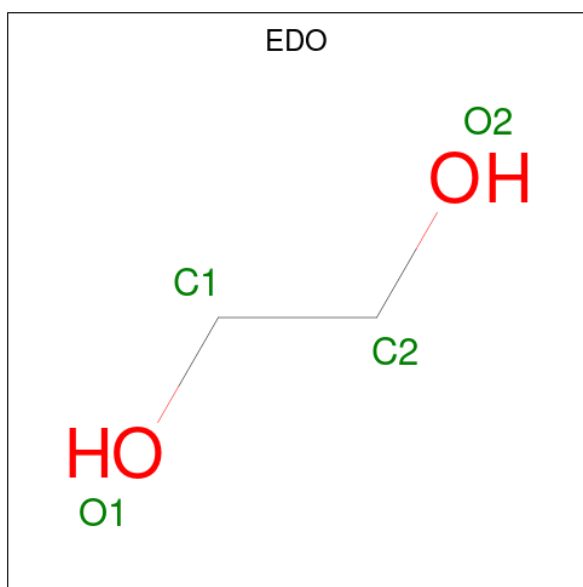
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	14	Total 281	C 136	N 53	O 79	P 13	0	0	0
3	F	14	Total 281	C 136	N 53	O 79	P 13	0	0	0
3	H	14	Total 281	C 136	N 53	O 79	P 13	0	0	0

- Molecule 4 is SINEFUNGIN (three-letter code: SFG) (formula:  $C_{15}H_{23}N_7O_5$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	Total	C	N	O	0	0
			27	15	7	5		
4	B	1	Total	C	N	O	0	0
			27	15	7	5		
4	C	1	Total	C	N	O	0	0
			27	15	7	5		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	I	1	Total	C	O	0	0
			4	2	2		
5	I	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	K	0	0
			1	1		
6	B	2	Total	K	0	0
			2	2		

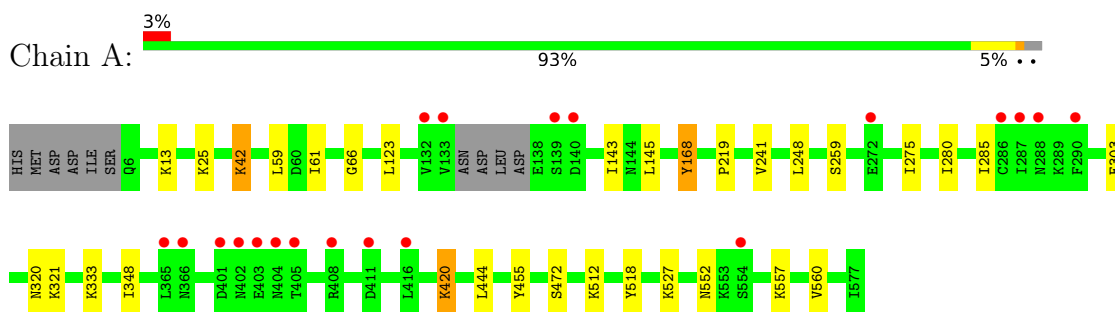
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	349	Total	O	0	0
			349	349		
7	B	365	Total	O	0	0
			365	365		
7	C	242	Total	O	0	0
			242	242		
7	E	31	Total	O	0	0
			31	31		
7	D	41	Total	O	0	0
			41	41		
7	F	40	Total	O	0	0
			40	40		
7	G	38	Total	O	0	0
			38	38		
7	H	32	Total	O	0	0
			32	32		
7	I	27	Total	O	0	0
			27	27		

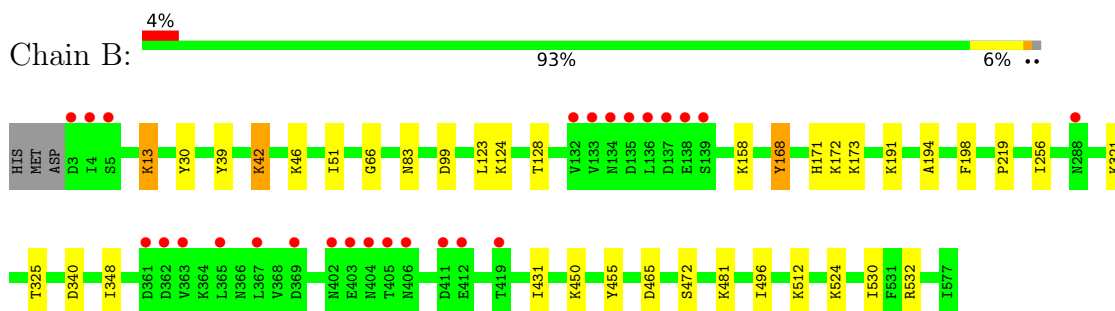
### 3 Residue-property plots [i](#)

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.

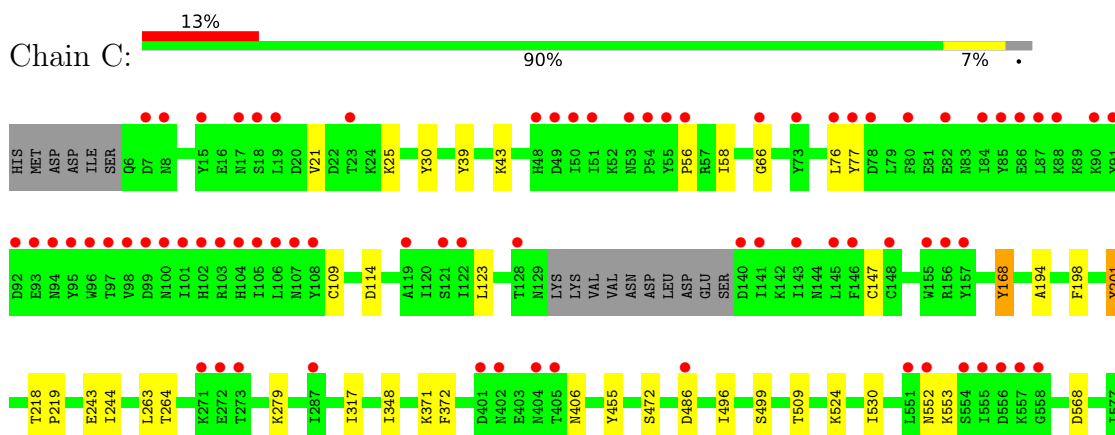
- Molecule 1: Site-specific DNA-methyltransferase (adenine-specific)



- Molecule 1: Site-specific DNA-methyltransferase (adenine-specific)

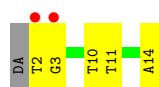


- Molecule 1: Site-specific DNA-methyltransferase (adenine-specific)



- Molecule 2: DNA Strand 2



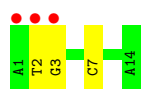
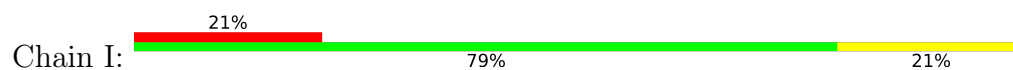


- Molecule 2: DNA Strand 2

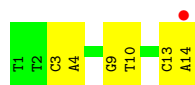


There are no outlier residues recorded for this chain.

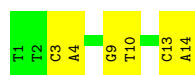
- Molecule 2: DNA Strand 2



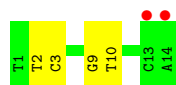
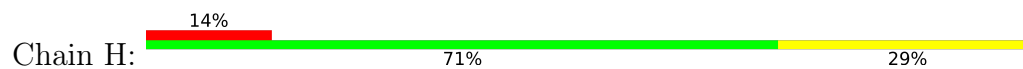
- Molecule 3: DNA Strand 1



- Molecule 3: DNA Strand 1



- Molecule 3: DNA Strand 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.16Å 160.45Å 231.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.29 – 2.05 47.29 – 2.05	Depositor EDS
% Data completeness (in resolution range)	92.7 (47.29-2.05) 92.7 (47.29-2.05)	Depositor EDS
$R_{merge}$	0.28	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.20 (at 2.05Å)	Xtrriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, $R_{free}$	0.189 , 0.225 0.188 , 0.223	Depositor DCC
$R_{free}$ test set	1992 reflections (1.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.7	Xtrriage
Anisotropy	0.356	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 53.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17045	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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: K, SFG, EDO

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	0.31	0/4819	0.52	0/6483
1	B	0.32	0/4895	0.53	0/6581
1	C	0.29	0/4597	0.49	0/6206
2	E	0.67	0/300	1.00	0/462
2	G	0.79	0/321	1.01	0/495
2	I	0.67	0/321	1.00	0/495
3	D	0.72	0/315	0.93	0/483
3	F	0.70	0/315	0.88	0/483
3	H	0.68	0/315	0.86	0/483
All	All	0.38	0/16198	0.59	0/22171

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	4718	0	4695	19	0
1	B	4790	0	4802	22	0
1	C	4506	0	4354	25	0
2	E	269	0	150	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	G	287	0	162	0	0
2	I	287	0	162	2	0
3	D	281	0	159	4	0
3	F	281	0	159	3	0
3	H	281	0	159	2	0
4	A	27	0	22	0	0
4	B	27	0	22	1	0
4	C	27	0	22	1	0
5	A	28	0	40	0	0
5	B	40	0	60	3	0
5	C	12	0	18	1	0
5	E	4	0	6	2	0
5	F	4	0	6	0	0
5	I	8	0	12	0	0
6	A	1	0	0	0	0
6	B	2	0	0	0	0
7	A	349	0	0	5	0
7	B	365	0	0	4	0
7	C	242	0	0	1	0
7	D	41	0	0	0	0
7	E	31	0	0	0	0
7	F	40	0	0	0	0
7	G	38	0	0	0	0
7	H	32	0	0	1	0
7	I	27	0	0	1	0
All	All	17045	0	15010	79	0

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

All (79) 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:532[B]:ARG:NH1	7:B:701:HOH:O	2.21	0.74
1:A:59:LEU:HD21	1:A:61:ILE:HG23	1.73	0.69
3:F:9:DG:H2''	3:F:10:DT:H5''	1.76	0.68
1:A:420:LYS:NZ	7:A:701:HOH:O	2.28	0.66
1:A:66:GLY:HA3	1:A:123:LEU:HD13	1.78	0.65
1:B:158:LYS:HG3	5:B:609:EDO:H12	1.78	0.65
1:B:51:ILE:HD11	1:B:83:ASN:HB3	1.77	0.64
1:C:66:GLY:HA3	1:C:123:LEU:HD13	1.80	0.64

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:527:LYS:NZ	7:A:705:HOH:O	2.33	0.61
1:B:66:GLY:HA3	1:B:123:LEU:HD13	1.83	0.60
1:C:168:TYR:CE1	1:C:219:PRO:HD3	2.37	0.60
1:C:496:ILE:HD11	1:C:530:ILE:HD12	1.85	0.59
1:A:168:TYR:CE1	1:A:219:PRO:HD3	2.39	0.58
2:E:14:DA:H2'	5:E:101:EDO:H22	1.86	0.57
1:C:30:TYR:O	4:C:601:SFG:NE	2.38	0.56
1:A:42:LYS:NZ	7:A:710:HOH:O	2.39	0.55
1:C:21:VAL:O	1:C:25:LYS:HG3	2.07	0.55
1:B:168:TYR:CE1	1:B:219:PRO:HD3	2.42	0.55
1:B:431:ILE:HG22	5:E:101:EDO:H21	1.88	0.54
1:A:280:ILE:HD11	1:A:285:ILE:HG12	1.89	0.54
2:I:7:DC:OP1	7:I:201:HOH:O	2.19	0.51
1:A:512:LYS:HD3	1:A:518:TYR:CE2	2.46	0.50
1:C:279:LYS:NZ	5:C:603:EDO:O1	2.39	0.50
1:B:124:LYS:O	1:B:128:THR:HG23	2.12	0.50
2:E:2:DT:H2''	2:E:3:DG:C8	2.47	0.49
1:B:256:ILE:HD11	5:B:608:EDO:H22	1.94	0.49
3:H:2:DT:H2'	3:H:3:DC:C6	2.47	0.49
1:C:39:TYR:OH	1:C:243:GLU:OE2	2.23	0.48
1:C:244:ILE:HG12	1:C:263:LEU:HD13	1.95	0.48
1:A:420:LYS:NZ	7:A:708:HOH:O	2.36	0.48
1:A:143:ILE:HG22	1:A:145:LEU:HG	1.94	0.48
1:C:486:ASP:OD1	1:C:486:ASP:N	2.44	0.48
1:B:496:ILE:HD11	1:B:530:ILE:HD12	1.96	0.48
1:C:201:TYR:CE2	1:C:218:THR:HG21	2.49	0.48
3:D:9:DG:H2''	3:D:10:DT:H5''	1.95	0.48
1:B:191[B]:LYS:HG2	7:B:961:HOH:O	2.13	0.47
1:C:25:LYS:HE2	7:H:104:HOH:O	2.14	0.47
1:A:348:ILE:O	1:A:472:SER:HA	2.15	0.47
1:B:13:LYS:HG3	5:B:603:EDO:H21	1.97	0.47
1:B:42:LYS:O	1:B:46:LYS:HB2	2.15	0.47
1:C:509:THR:O	1:C:524:LYS:HE3	2.15	0.46
3:D:13:DC:H2''	3:D:14:DA:C8	2.50	0.46
1:A:25:LYS:NZ	3:D:9:DG:OP1	2.49	0.46
1:B:340:ASP:O	1:B:481:LYS:NZ	2.35	0.46
3:F:3:DC:H2'	3:F:4:DA:C8	2.51	0.46
2:I:2:DT:H2''	2:I:3:DG:C8	2.51	0.46
1:B:524:LYS:HD2	7:B:794:HOH:O	2.16	0.45
2:E:10:DT:H2''	2:E:11:DT:H72	1.99	0.45
1:B:30:TYR:O	4:B:601:SFG:NE	2.50	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:333:LYS:NZ	7:A:725:HOH:O	2.50	0.45
1:B:171:HIS:CE1	1:B:172:LYS:HG3	2.52	0.44
1:C:58:ILE:HD12	1:C:76:LEU:HD11	1.98	0.44
1:A:241:VAL:O	1:A:275:ILE:HG12	2.17	0.44
1:B:39:TYR:O	1:B:42:LYS:HG3	2.17	0.44
1:C:406:ASN:ND2	7:C:702:HOH:O	2.30	0.44
1:B:321:LYS:O	1:B:325:THR:HG23	2.18	0.43
1:B:348:ILE:O	1:B:472:SER:HA	2.18	0.43
1:C:317:ILE:HD12	1:C:509:THR:HG22	1.99	0.43
1:B:450:LYS:NZ	1:B:465:ASP:OD2	2.46	0.43
1:C:552:ASN:OD1	1:C:553:LYS:N	2.46	0.43
3:F:13:DC:H2''	3:F:14:DA:C8	2.53	0.43
1:A:13:LYS:HE3	1:A:13:LYS:HB2	1.80	0.43
1:C:43:LYS:HD2	1:C:264:THR:HG21	2.00	0.43
1:B:194:ALA:HB1	1:B:198:PHE:CD2	2.53	0.42
1:C:499:SER:HB2	1:C:568:ASP:OD1	2.19	0.42
1:A:248:LEU:O	1:A:259:SER:HB3	2.19	0.42
1:C:348:ILE:O	1:C:472:SER:HA	2.19	0.42
1:A:557:LYS:O	1:A:560:VAL:HB	2.18	0.42
1:A:303:GLU:OE2	1:A:320:ASN:ND2	2.39	0.42
3:D:3:DC:H2'	3:D:4:DA:C8	2.55	0.42
1:C:56:PRO:O	1:C:109:CYS:HA	2.21	0.41
1:C:371:LYS:HE3	1:C:372:PHE:CZ	2.55	0.41
1:A:321:LYS:HD2	1:A:321:LYS:HA	1.81	0.41
1:B:173:LYS:NZ	7:B:734:HOH:O	2.53	0.41
1:C:244:ILE:HG12	1:C:263:LEU:CD1	2.49	0.41
1:C:114:ASP:O	1:C:147:CYS:HA	2.21	0.41
1:C:194:ALA:HB1	1:C:198:PHE:CD2	2.56	0.41
3:H:9:DG:H2'	3:H:10:DT:H72	2.02	0.41
1:C:168:TYR:CD1	1:C:219:PRO:HD3	2.55	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	566/578 (98%)	551 (97%)	15 (3%)	0	100	100
1	B	577/578 (100%)	556 (96%)	21 (4%)	0	100	100
1	C	558/578 (96%)	541 (97%)	17 (3%)	0	100	100
All	All	1701/1734 (98%)	1648 (97%)	53 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	521/548 (95%)	515 (99%)	6 (1%)	71	70
1	B	530/548 (97%)	524 (99%)	6 (1%)	73	73
1	C	476/548 (87%)	472 (99%)	4 (1%)	81	82
All	All	1527/1644 (93%)	1511 (99%)	16 (1%)	73	75

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	LYS
1	A	168	TYR
1	A	420	LYS
1	A	444	LEU
1	A	455	TYR
1	A	552	ASN
1	B	13	LYS
1	B	42	LYS
1	B	99	ASP
1	B	168	TYR
1	B	455	TYR
1	B	512	LYS
1	C	77	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	168	TYR
1	C	201	TYR
1	C	455	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 30 ligands modelled in this entry, 3 are monoatomic - leaving 27 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	$\# Z  > 2$
5	EDO	B	609	-	3,3,3	0.45	0	2,2,2	0.34	0
5	EDO	B	611	-	3,3,3	0.58	0	2,2,2	0.12	0
5	EDO	A	603	6	3,3,3	0.46	0	2,2,2	0.26	0
5	EDO	B	610	-	3,3,3	0.55	0	2,2,2	0.19	0
5	EDO	C	604	-	3,3,3	0.45	0	2,2,2	0.35	0
5	EDO	A	604	-	3,3,3	0.53	0	2,2,2	0.17	0
5	EDO	A	602	-	3,3,3	0.41	0	2,2,2	0.25	0
5	EDO	B	606	-	3,3,3	0.50	0	2,2,2	0.17	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	EDO	A	608	-	3,3,3	0.48	0	2,2,2	0.33	0
5	EDO	B	608	-	3,3,3	0.48	0	2,2,2	0.26	0
5	EDO	A	605	-	3,3,3	0.47	0	2,2,2	0.35	0
4	SFG	B	601	-	25,29,29	1.00	0	23,42,42	0.94	1 (4%)
5	EDO	C	602	-	3,3,3	0.44	0	2,2,2	0.29	0
5	EDO	I	102	-	3,3,3	0.42	0	2,2,2	0.46	0
5	EDO	B	603	-	3,3,3	0.52	0	2,2,2	0.14	0
4	SFG	C	601	-	25,29,29	0.92	0	23,42,42	0.85	2 (8%)
5	EDO	F	101	-	3,3,3	0.56	0	2,2,2	0.22	0
5	EDO	B	605	-	3,3,3	0.52	0	2,2,2	0.19	0
5	EDO	A	606	-	3,3,3	0.45	0	2,2,2	0.30	0
5	EDO	C	603	-	3,3,3	0.51	0	2,2,2	0.25	0
5	EDO	B	607	-	3,3,3	0.47	0	2,2,2	0.44	0
4	SFG	A	601	-	25,29,29	0.92	0	23,42,42	0.82	1 (4%)
5	EDO	B	602	-	3,3,3	0.40	0	2,2,2	0.36	0
5	EDO	E	101	-	3,3,3	0.56	0	2,2,2	0.14	0
5	EDO	B	604	-	3,3,3	0.51	0	2,2,2	0.32	0
5	EDO	A	607	-	3,3,3	0.44	0	2,2,2	0.37	0
5	EDO	I	101	-	3,3,3	0.50	0	2,2,2	0.30	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	B	609	-	-	0/1/1/1	-
5	EDO	B	611	-	-	0/1/1/1	-
5	EDO	A	603	6	-	0/1/1/1	-
5	EDO	B	610	-	-	1/1/1/1	-
5	EDO	C	604	-	-	0/1/1/1	-
5	EDO	A	604	-	-	1/1/1/1	-
5	EDO	A	602	-	-	0/1/1/1	-
5	EDO	B	606	-	-	0/1/1/1	-
5	EDO	A	608	-	-	0/1/1/1	-
5	EDO	B	608	-	-	0/1/1/1	-
5	EDO	A	605	-	-	0/1/1/1	-
4	SFG	B	601	-	-	3/13/33/33	0/3/3/3
5	EDO	C	602	-	-	0/1/1/1	-
5	EDO	I	102	-	-	0/1/1/1	-
5	EDO	B	603	-	-	0/1/1/1	-
4	SFG	C	601	-	-	1/13/33/33	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	F	101	-	-	1/1/1/1	-
5	EDO	B	605	-	-	0/1/1/1	-
5	EDO	A	606	-	-	0/1/1/1	-
5	EDO	C	603	-	-	0/1/1/1	-
5	EDO	B	607	-	-	0/1/1/1	-
4	SFG	A	601	-	-	4/13/33/33	0/3/3/3
5	EDO	B	602	-	-	0/1/1/1	-
5	EDO	E	101	-	-	1/1/1/1	-
5	EDO	B	604	-	-	1/1/1/1	-
5	EDO	A	607	-	-	0/1/1/1	-
5	EDO	I	101	-	-	0/1/1/1	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	601	SFG	OXT-C-CA	2.41	121.60	113.38
4	C	601	SFG	OXT-C-CA	2.28	121.16	113.38
4	C	601	SFG	OXT-C-O	-2.25	118.98	124.09
4	A	601	SFG	OXT-C-CA	2.19	120.84	113.38

There are no chirality outliers.

All (13) torsion outliers are listed below:

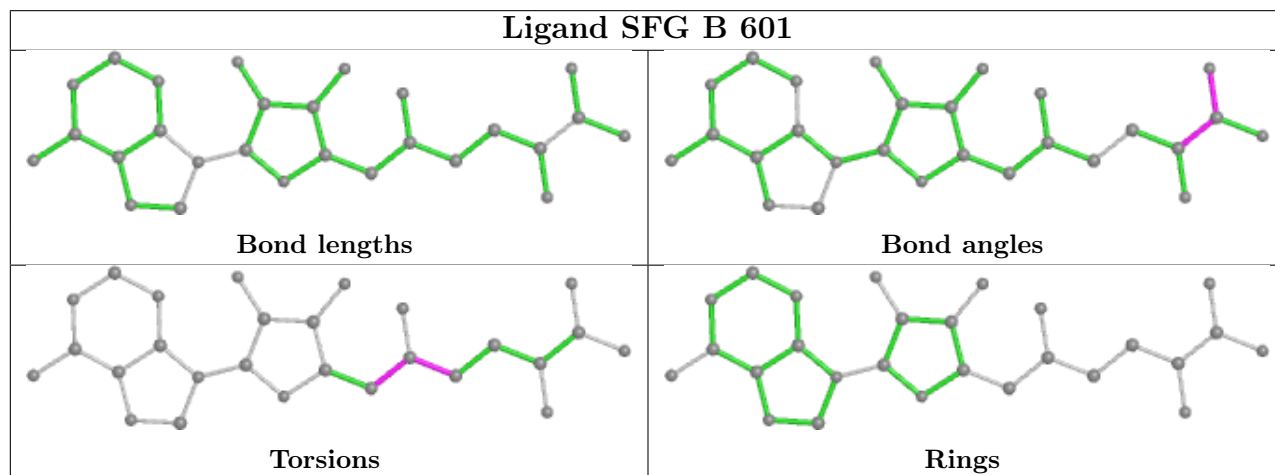
Mol	Chain	Res	Type	Atoms
4	A	601	SFG	N-CA-CB-CG
4	B	601	SFG	NE-CD-CG-CB
4	B	601	SFG	C5'-CD-CG-CB
5	E	101	EDO	O1-C1-C2-O2
4	A	601	SFG	C-CA-CB-CG
4	A	601	SFG	OXT-C-CA-N
4	A	601	SFG	O-C-CA-N
5	B	604	EDO	O1-C1-C2-O2
5	B	610	EDO	O1-C1-C2-O2
5	F	101	EDO	O1-C1-C2-O2
4	B	601	SFG	C4'-C5'-CD-NE
4	C	601	SFG	OXT-C-CA-N
5	A	604	EDO	O1-C1-C2-O2

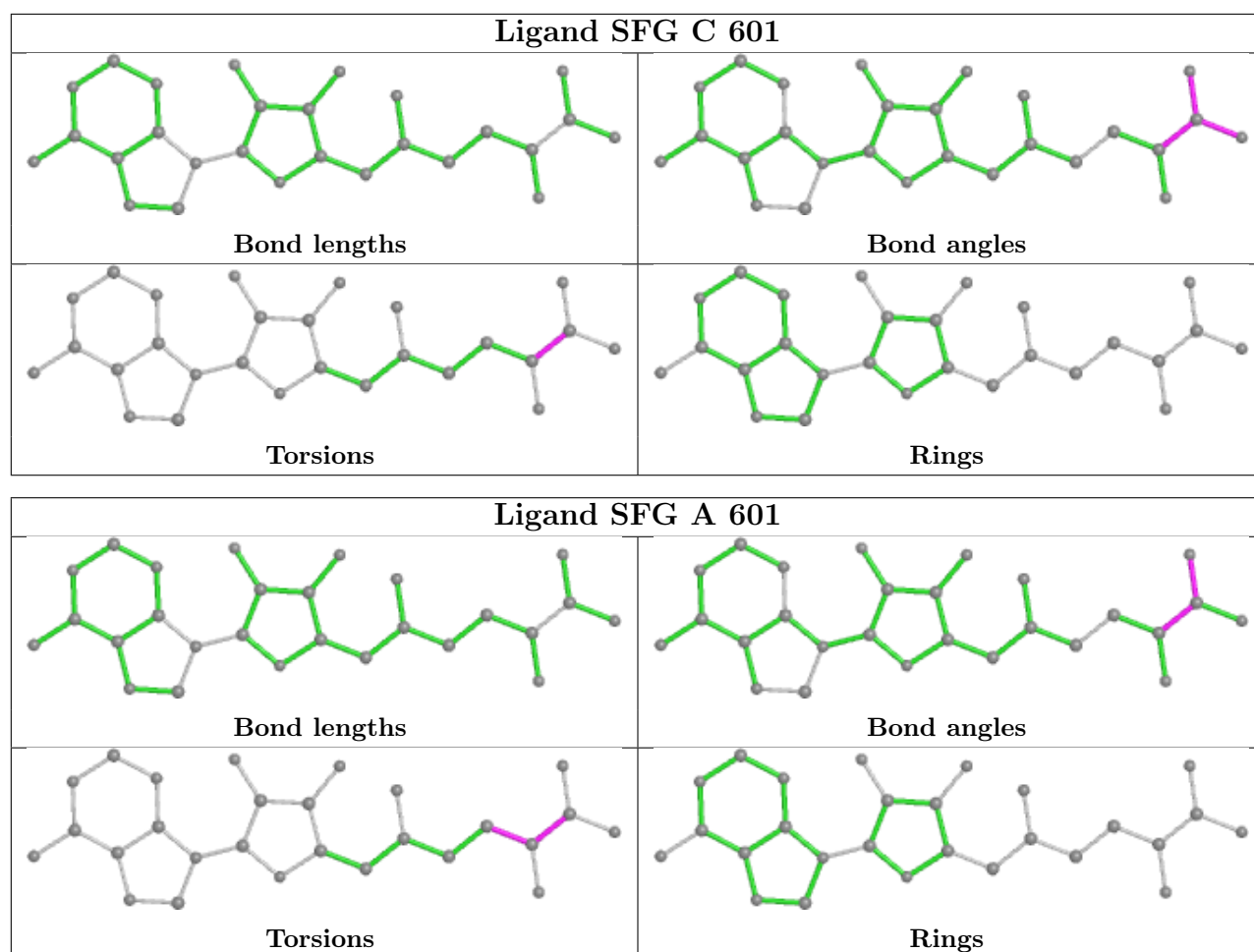
There are no ring outliers.

7 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	609	EDO	1	0
5	B	608	EDO	1	0
4	B	601	SFG	1	0
5	B	603	EDO	1	0
4	C	601	SFG	1	0
5	C	603	EDO	1	0
5	E	101	EDO	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

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

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

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	568/578 (98%)	-0.06	20 (3%) 44 48	21, 38, 68, 116	0
1	B	575/578 (99%)	0.15	26 (4%) 33 35	21, 36, 66, 120	0
1	C	562/578 (97%)	0.52	75 (13%) 3 2	27, 49, 98, 133	0
2	E	13/14 (92%)	0.09	2 (15%) 2 1	27, 31, 100, 120	0
2	G	14/14 (100%)	-0.55	0 100 100	28, 32, 50, 50	0
2	I	14/14 (100%)	0.71	3 (21%) 0 0	32, 38, 134, 147	0
3	D	14/14 (100%)	0.16	1 (7%) 16 17	25, 30, 99, 112	0
3	F	14/14 (100%)	-0.20	0 100 100	23, 30, 56, 59	0
3	H	14/14 (100%)	0.38	2 (14%) 2 2	28, 40, 96, 126	0
All	All	1788/1818 (98%)	0.20	129 (7%) 15 17	21, 40, 87, 147	0

All (129) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	136	LEU	9.6
1	B	135	ASP	8.2
1	C	91	TYR	8.1
1	C	108	TYR	7.0
1	B	134	ASN	6.7
1	C	84	ILE	6.0
1	C	51	ILE	6.0
1	C	94	ASN	6.0
1	B	137	ASP	6.0
3	H	14	DA	5.9
1	A	405	THR	5.6
1	C	555	ILE	5.5
2	I	1	DA	5.5
1	C	99	ASP	4.9
1	B	133	VAL	4.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	92	ASP	4.8
1	C	55	TYR	4.8
1	A	133	VAL	4.7
1	A	288	ASN	4.7
1	C	85	TYR	4.7
1	C	95	TYR	4.5
1	C	50	ILE	4.4
1	C	98	VAL	4.4
1	C	73	TYR	4.4
1	B	402	ASN	4.4
1	C	87	LEU	4.3
1	C	402	ASN	4.3
1	C	96	TRP	4.3
1	C	145	LEU	4.1
1	C	93	GLU	4.1
1	A	404	ASN	4.0
1	C	103	ARG	4.0
1	B	405	THR	4.0
1	C	53	ASN	4.0
1	C	102	HIS	3.9
1	B	132	VAL	3.9
1	C	77	TYR	3.8
1	A	140	ASP	3.8
1	B	139	SER	3.8
1	A	403	GLU	3.7
2	I	2	DT	3.7
1	B	4	ILE	3.7
1	C	141	ILE	3.6
1	C	19	LEU	3.6
1	C	101	ILE	3.5
1	B	367	LEU	3.5
1	A	132	VAL	3.4
1	C	105	ILE	3.4
1	C	97	THR	3.3
1	C	405	THR	3.3
1	C	66	GLY	3.2
1	C	100	ASN	3.2
1	C	90	LYS	3.2
1	C	271	LYS	3.2
1	A	287	ILE	3.2
1	B	3	ASP	3.2
1	C	7	ASP	3.2

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	404	ASN	3.1
1	C	272	GLU	3.1
2	I	3	DG	3.1
1	B	138	GLU	3.1
1	B	419	THR	3.1
1	C	86	GLU	3.0
1	C	76	LEU	3.0
1	C	107	ASN	3.0
1	A	286	CYS	3.0
1	C	148	CYS	3.0
1	C	554	SER	3.0
1	B	406	ASN	2.9
1	B	363	VAL	2.9
1	B	361	ASP	2.9
1	B	411	ASP	2.9
1	C	104	HIS	2.9
1	C	23	THR	2.9
1	C	401	ASP	2.9
1	C	143	ILE	2.8
1	C	106	LEU	2.8
1	C	17	ASN	2.8
1	C	49	ASP	2.8
1	C	119	ALA	2.8
1	C	122	ILE	2.8
1	A	402	ASN	2.8
3	D	14	DA	2.8
1	C	82	GLU	2.7
1	C	80	PHE	2.7
1	C	146	PHE	2.7
1	C	121	SER	2.7
1	C	558	GLY	2.7
1	B	403	GLU	2.7
1	B	5	SER	2.6
2	E	2	DT	2.6
1	C	78	ASP	2.6
1	C	552	ASN	2.6
1	C	8	ASN	2.6
1	C	404	ASN	2.6
1	C	157	TYR	2.6
1	C	556	ASP	2.6
1	B	412	GLU	2.6
1	B	365	LEU	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	273	THR	2.5
1	C	54	PRO	2.5
1	B	288	ASN	2.5
1	B	362	ASP	2.5
1	A	366	ASN	2.4
1	A	411	ASP	2.4
1	A	365	LEU	2.4
1	C	88	LYS	2.4
1	A	408	ARG	2.3
1	A	139	SER	2.3
1	C	156	ARG	2.3
2	E	3	DG	2.3
1	C	140	ASP	2.3
1	C	155	TRP	2.2
1	C	486	ASP	2.2
1	C	287	ILE	2.2
1	A	272	GLU	2.2
1	C	557	LYS	2.2
1	C	15	TYR	2.1
1	C	18	SER	2.1
1	B	369	ASP	2.1
1	C	128	THR	2.1
1	A	401	ASP	2.1
1	C	48	HIS	2.1
1	C	551	LEU	2.1
1	A	290	PHE	2.1
1	A	554	SER	2.0
1	C	56	PRO	2.0
1	A	416	LEU	2.0
3	H	13	DC	2.0

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

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

## 6.3 Carbohydrates [\(i\)](#)

There are no monosaccharides in this entry.



## 6.4 Ligands i

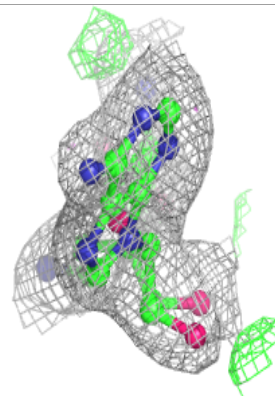
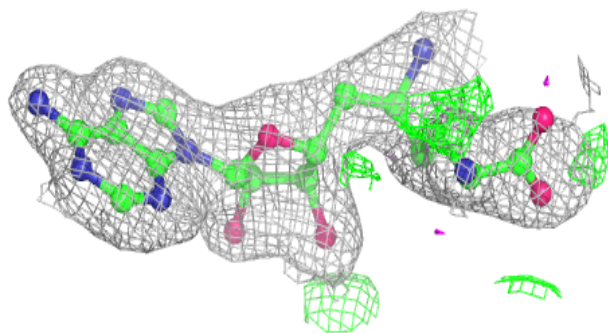
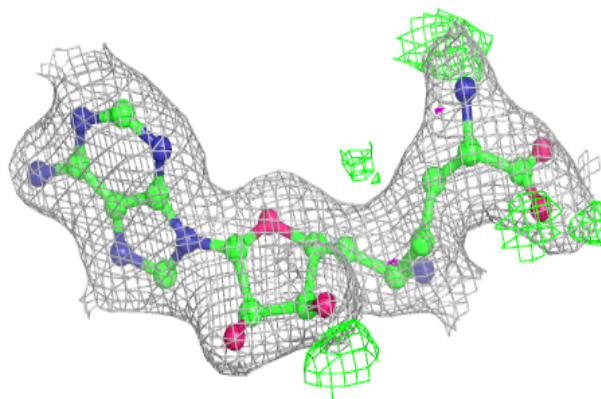
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	EDO	A	603	4/4	0.79	0.18	44,49,56,59	0
5	EDO	A	606	4/4	0.81	0.40	48,53,54,59	0
5	EDO	A	604	4/4	0.84	0.28	45,51,53,58	0
5	EDO	A	605	4/4	0.85	0.19	51,62,62,71	0
5	EDO	E	101	4/4	0.85	0.43	38,47,56,62	0
5	EDO	F	101	4/4	0.85	0.24	36,39,40,45	0
5	EDO	C	603	4/4	0.86	0.16	47,51,55,67	0
5	EDO	B	604	4/4	0.86	0.14	54,57,61,67	0
5	EDO	B	611	4/4	0.86	0.15	35,37,45,47	0
5	EDO	A	607	4/4	0.88	0.17	41,49,51,62	0
5	EDO	B	603	4/4	0.90	0.14	48,53,55,61	0
5	EDO	B	609	4/4	0.90	0.20	48,50,52,52	0
4	SFG	C	601	27/27	0.91	0.11	39,52,63,69	0
5	EDO	A	608	4/4	0.92	0.14	48,50,56,58	0
6	K	B	613	1/1	0.93	0.10	40,40,40,40	0
5	EDO	B	606	4/4	0.94	0.08	44,44,47,48	0
5	EDO	C	604	4/4	0.94	0.16	38,47,59,62	0
4	SFG	A	601	27/27	0.94	0.11	27,35,43,48	0
5	EDO	B	610	4/4	0.94	0.20	35,45,50,66	0
5	EDO	B	605	4/4	0.94	0.16	35,39,40,46	0
5	EDO	I	101	4/4	0.95	0.16	35,44,45,59	0
5	EDO	I	102	4/4	0.95	0.15	49,50,57,66	0
5	EDO	B	607	4/4	0.95	0.12	33,37,42,44	0
4	SFG	B	601	27/27	0.96	0.11	22,28,38,39	0
5	EDO	B	608	4/4	0.96	0.13	42,43,54,70	0
5	EDO	C	602	4/4	0.97	0.08	34,41,44,49	0
5	EDO	B	602	4/4	0.98	0.09	31,32,33,33	0
6	K	A	609	1/1	0.98	0.06	44,44,44,44	0
5	EDO	A	602	4/4	0.98	0.11	32,32,32,34	0
6	K	B	612	1/1	0.99	0.12	41,41,41,41	0

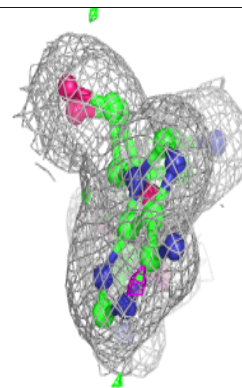
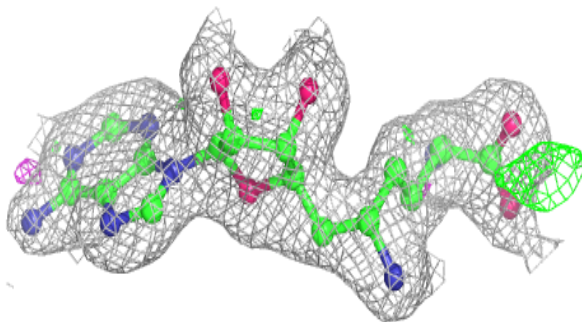
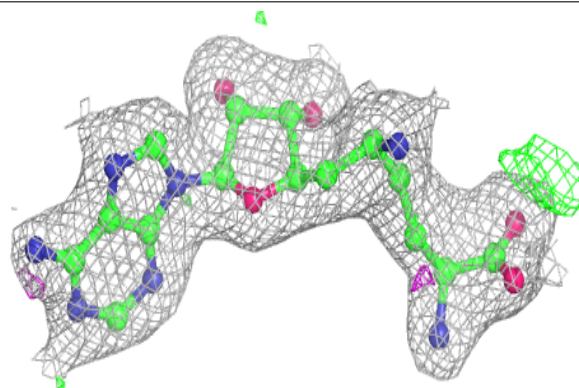
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

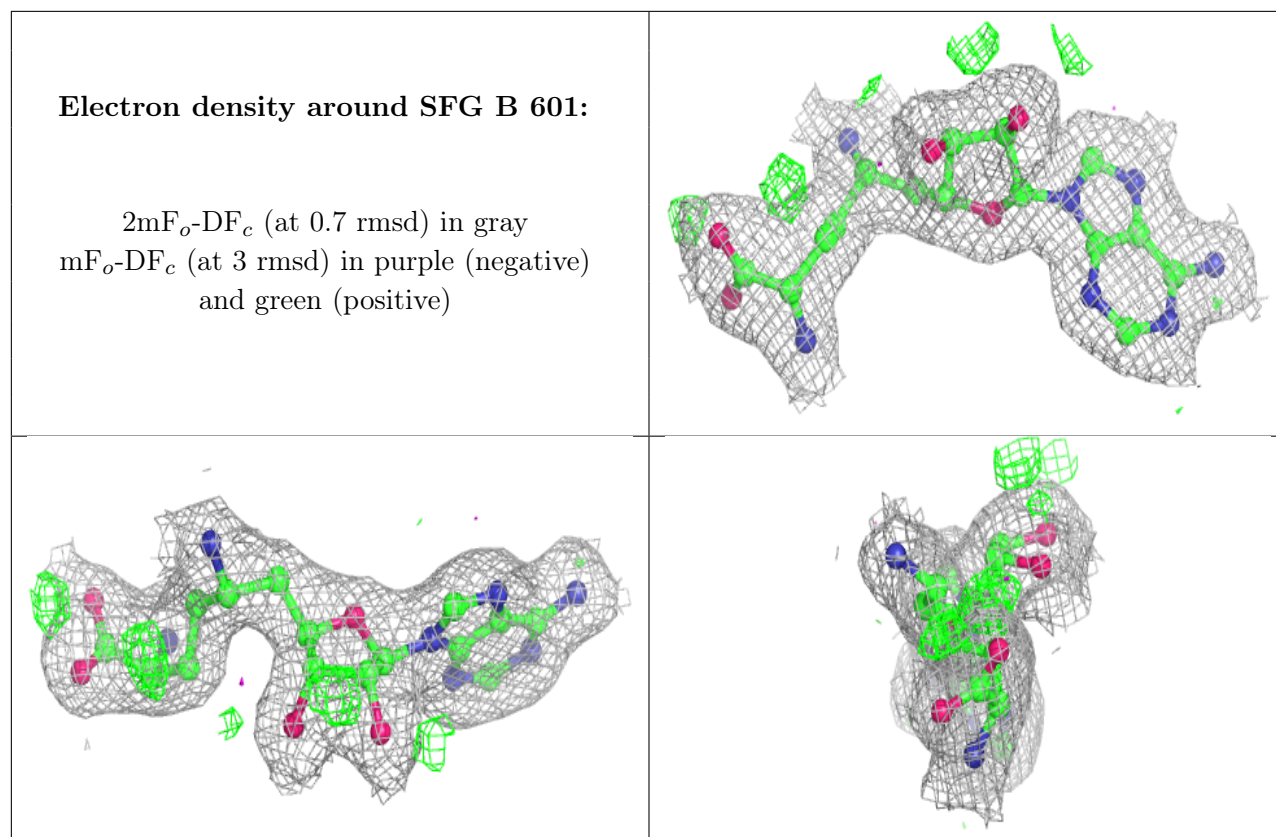
**Electron density around SFG C 601:**

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

**Electron density around SFG 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 [i](#)

There are no such residues in this entry.