



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 10, 2023 – 03:23 AM EDT

PDB ID : 4HN1  
Title : Crystal Structure of H60N/Y130F double mutant of ChmJ, a 3'-monoepimerase from *Streptomyces bikiniensis* in complex with dTDP  
Authors : Holden, H.M.; Kubiak, R.L.  
Deposited on : 2012-10-18  
Resolution : 1.60 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

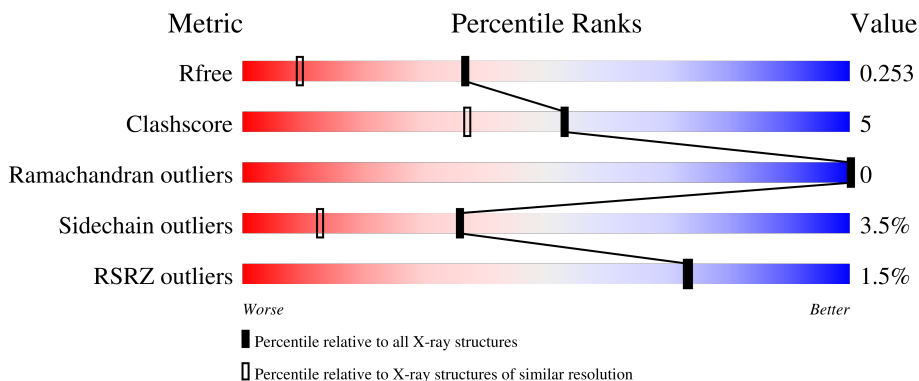
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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	201	 88% 9% ..
1	B	201	 89% 9% ..
1	C	201	 83% 14% .
1	D	201	 85% 14% .

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	TDR	D	303	-	X	-	-

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 6909 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 Putative 3-epimerase in D-allose pathway.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	199	1568	986	281	297	4	0	4	0
1	B	199	1576	991	284	297	4	0	5	0
1	C	201	1585	996	287	298	4	0	2	0
1	D	200	1565	984	284	293	4	0	1	0

There are 28 discrepancies between the modelled and reference sequences:

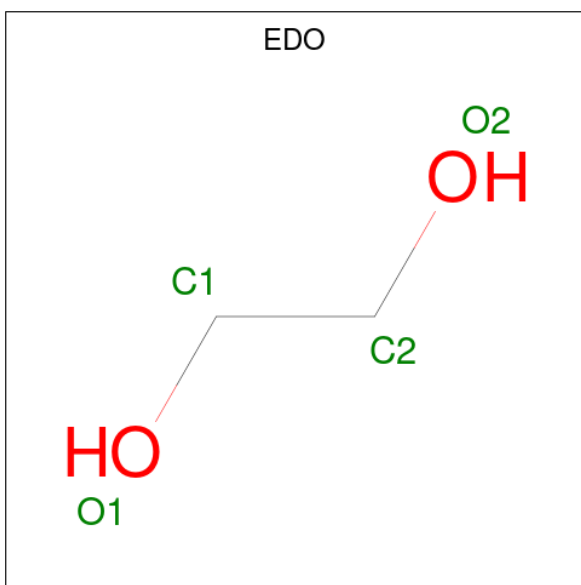
Chain	Residue	Modelled	Actual	Comment	Reference
A	60	ASN	HIS	engineered mutation	UNP Q5SFD1
A	130	PHE	TYR	engineered mutation	UNP Q5SFD1
A	197	LEU	-	expression tag	UNP Q5SFD1
A	198	GLU	-	expression tag	UNP Q5SFD1
A	199	HIS	-	expression tag	UNP Q5SFD1
A	200	HIS	-	expression tag	UNP Q5SFD1
A	201	HIS	-	expression tag	UNP Q5SFD1
B	60	ASN	HIS	engineered mutation	UNP Q5SFD1
B	130	PHE	TYR	engineered mutation	UNP Q5SFD1
B	197	LEU	-	expression tag	UNP Q5SFD1
B	198	GLU	-	expression tag	UNP Q5SFD1
B	199	HIS	-	expression tag	UNP Q5SFD1
B	200	HIS	-	expression tag	UNP Q5SFD1
B	201	HIS	-	expression tag	UNP Q5SFD1
C	60	ASN	HIS	engineered mutation	UNP Q5SFD1
C	130	PHE	TYR	engineered mutation	UNP Q5SFD1
C	197	LEU	-	expression tag	UNP Q5SFD1
C	198	GLU	-	expression tag	UNP Q5SFD1
C	199	HIS	-	expression tag	UNP Q5SFD1
C	200	HIS	-	expression tag	UNP Q5SFD1
C	201	HIS	-	expression tag	UNP Q5SFD1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	60	ASN	HIS	engineered mutation	UNP Q5SFD1
D	130	PHE	TYR	engineered mutation	UNP Q5SFD1
D	197	LEU	-	expression tag	UNP Q5SFD1
D	198	GLU	-	expression tag	UNP Q5SFD1
D	199	HIS	-	expression tag	UNP Q5SFD1
D	200	HIS	-	expression tag	UNP Q5SFD1
D	201	HIS	-	expression tag	UNP Q5SFD1

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



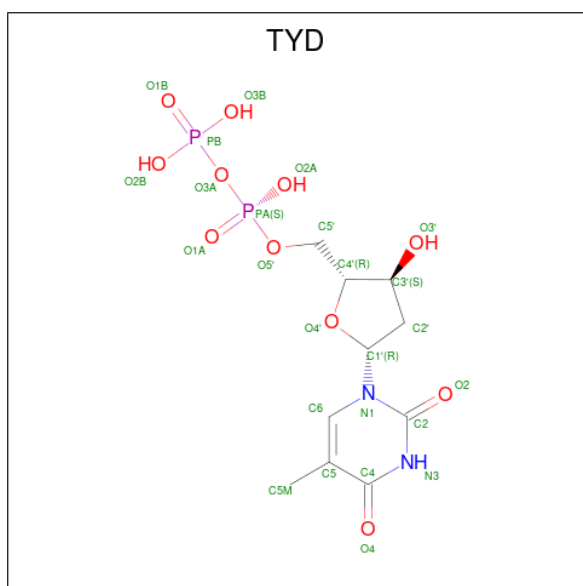
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

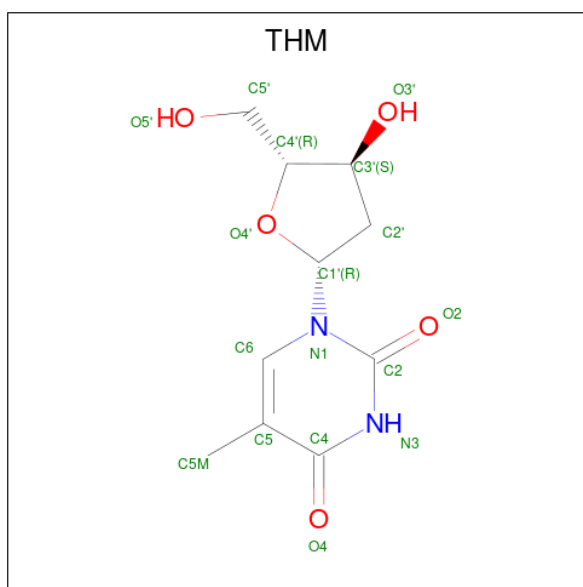
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is THYMIDINE-5'-DIPHOSPHATE (three-letter code: TYD) (formula:  $C_{10}H_{16}N_2O_{11}P_2$ ).



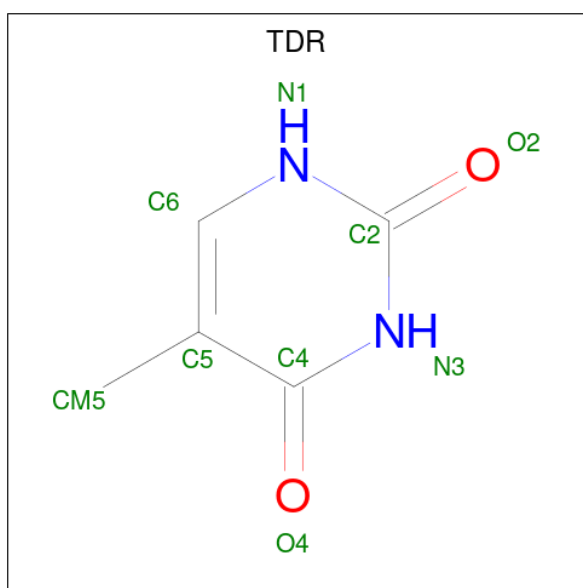
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			25	10	2	11	2		
3	B	1	Total	C	N	O	P	0	0
			25	10	2	11	2		
3	C	1	Total	C	N	O	P	0	0
			25	10	2	11	2		
3	C	1	Total	C	N	O	P	0	0
			25	10	2	11	2		

- Molecule 4 is THYMIDINE (three-letter code: THM) (formula:  $C_{10}H_{14}N_2O_5$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			17	10	2	5		
4	B	1	Total	C	N	O	0	0
			17	10	2	5		

- Molecule 5 is THYMINE (three-letter code: TDR) (formula:  $C_5H_6N_2O_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total	C	N	O	0	0
			9	5	2	2		
5	D	1	Total	C	N	O	0	0
			9	5	2	2		

- Molecule 6 is water.

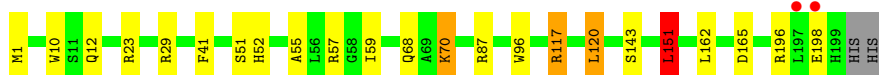
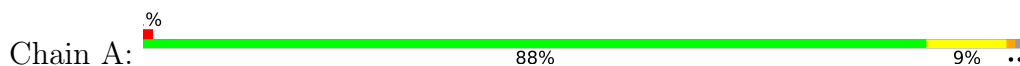
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	143	Total 143	O 143	0	0
6	B	112	Total 112	O 112	0	0
6	C	77	Total 77	O 77	0	0
6	D	87	Total 87	O 87	0	0



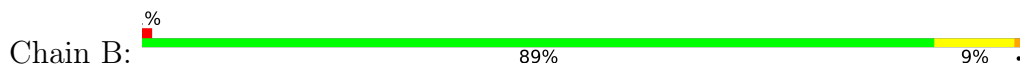
### 3 Residue-property plots

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.

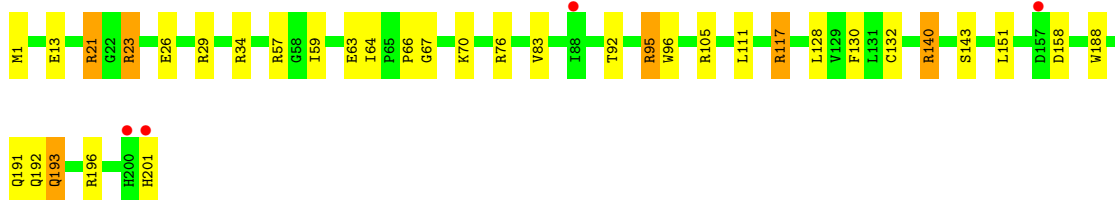
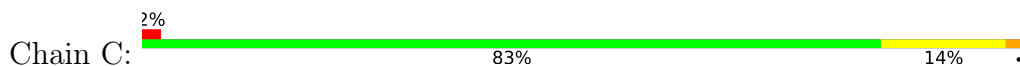
- Molecule 1: Putative 3-epimerase in D-allose pathway



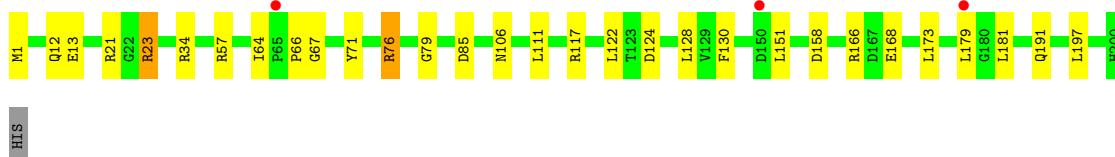
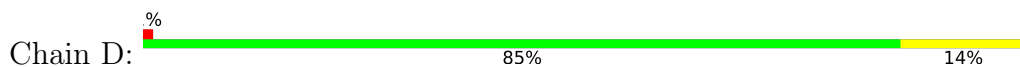
- Molecule 1: Putative 3-epimerase in D-allose pathway



- Molecule 1: Putative 3-epimerase in D-allose pathway



- Molecule 1: Putative 3-epimerase in D-allose pathway



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	141.02Å 141.02Å 116.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.60 28.86 – 1.60	Depositor EDS
% Data completeness (in resolution range)	96.1 (30.00-1.60) 96.1 (28.86-1.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 1.60Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.212 , 0.254 0.212 , 0.253	Depositor DCC
$R_{free}$ test set	7194 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.9	Xtrriage
Anisotropy	0.090	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 30.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.479 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6909	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% 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: TYD, THM, EDO, TDR

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	0.54	0/1623	1.21	7/2215 (0.3%)
1	B	0.53	0/1634	1.19	7/2229 (0.3%)
1	C	0.47	0/1636	1.18	9/2233 (0.4%)
1	D	0.47	0/1612	1.17	14/2201 (0.6%)
All	All	0.50	0/6505	1.18	37/8878 (0.4%)

There are no bond length outliers.

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	21	ARG	NE-CZ-NH2	-12.19	114.21	120.30
1	C	23	ARG	NE-CZ-NH1	9.52	125.06	120.30
1	D	23	ARG	NE-CZ-NH1	9.30	124.95	120.30
1	C	21	ARG	NE-CZ-NH1	8.96	124.78	120.30
1	A	151	LEU	CB-CG-CD2	8.64	125.70	111.00
1	D	34	ARG	NE-CZ-NH2	-8.35	116.13	120.30
1	C	23	ARG	NE-CZ-NH2	-7.71	116.44	120.30
1	D	1	MET	CG-SD-CE	-6.97	89.04	100.20
1	D	23	ARG	NE-CZ-NH2	-6.81	116.89	120.30
1	A	120	LEU	CB-CG-CD2	6.79	122.55	111.00
1	B	87	ARG	NE-CZ-NH2	-6.68	116.96	120.30
1	B	165	ASP	CB-CG-OD1	6.68	124.31	118.30
1	B	87	ARG	NE-CZ-NH1	6.67	123.64	120.30
1	C	1	MET	CG-SD-CE	-6.64	89.58	100.20
1	D	130	PHE	CB-CA-C	-6.55	97.31	110.40
1	B	57	ARG	NE-CZ-NH1	6.34	123.47	120.30
1	B	167	ASP	CB-CG-OD2	6.33	124.00	118.30
1	A	87	ARG	NE-CZ-NH1	6.24	123.42	120.30
1	D	173	LEU	CB-CG-CD2	-6.19	100.48	111.00
1	D	197	LEU	CA-CB-CG	6.16	129.46	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	57	ARG	NE-CZ-NH1	6.06	123.33	120.30
1	C	128	LEU	CA-CB-CG	-5.95	101.63	115.30
1	B	130	PHE	CB-CA-C	-5.83	98.73	110.40
1	A	70	LYS	CD-CE-NZ	-5.83	98.29	111.70
1	C	130	PHE	CB-CA-C	-5.82	98.77	110.40
1	A	87	ARG	NE-CZ-NH2	-5.66	117.47	120.30
1	D	76	ARG	NE-CZ-NH1	5.58	123.09	120.30
1	D	21	ARG	NE-CZ-NH2	-5.57	117.52	120.30
1	A	41	PHE	CB-CG-CD2	5.56	124.69	120.80
1	D	111	LEU	CB-CG-CD2	5.51	120.36	111.00
1	C	111	LEU	CB-CG-CD2	5.46	120.29	111.00
1	D	34	ARG	NE-CZ-NH1	5.40	123.00	120.30
1	D	128	LEU	CA-CB-CG	-5.38	102.93	115.30
1	B	128	LEU	CA-CB-CG	-5.19	103.36	115.30
1	D	124	ASP	CB-CG-OD1	5.14	122.93	118.30
1	D	85	ASP	CB-CG-OD1	5.11	122.90	118.30
1	C	34	ARG	NE-CZ-NH2	-5.05	117.78	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	1568	0	1507	16	0
1	B	1576	0	1520	12	0
1	C	1585	0	1512	21	0
1	D	1565	0	1495	17	0
2	A	16	0	24	3	0
2	B	12	0	18	2	0
2	C	8	0	12	0	0
2	D	8	0	12	2	0
3	A	25	0	13	3	0
3	B	25	0	13	1	0
3	C	50	0	26	3	0
4	A	17	0	14	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	17	0	14	0	0
5	C	9	0	6	0	0
5	D	9	0	6	0	0
6	A	143	0	0	4	0
6	B	112	0	0	4	0
6	C	77	0	0	1	0
6	D	87	0	0	1	0
All	All	6909	0	6192	69	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 5.

All (69) 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:C:105:ARG:HD3	6:C:477:HOH:O	1.52	1.06
1:A:117:ARG:HG3	1:A:117:ARG:HH11	1.23	1.00
3:C:303:TYD:O1B	1:D:57:ARG:NH1	2.10	0.84
1:A:117:ARG:HH11	1:A:117:ARG:CG	1.90	0.84
1:C:26[A]:GLU:OE1	1:C:29:ARG:NE	2.10	0.84
1:C:57:ARG:NH1	3:C:304:TYD:O2B	2.13	0.79
1:D:66:PRO:HA	1:D:191:GLN:HE22	1.48	0.78
1:A:10:TRP:CZ3	2:A:304:EDO:H22	2.18	0.78
1:C:70:LYS:NZ	1:C:132:CYS:SG	2.62	0.73
1:A:52:HIS:HD2	6:A:471:HOH:O	1.73	0.71
1:C:193:GLN:NE2	1:C:196:ARG:HH21	1.87	0.71
1:A:117:ARG:HG3	1:A:117:ARG:NH1	2.00	0.69
1:D:66:PRO:HA	1:D:191:GLN:NE2	2.08	0.69
1:C:26[A]:GLU:OE1	1:C:29:ARG:CD	2.43	0.67
1:D:71:TYR:OH	2:D:301:EDO:H21	1.94	0.67
1:B:52:HIS:ND1	6:B:508:HOH:O	2.28	0.66
1:C:66:PRO:HA	1:C:191:GLN:HE22	1.61	0.66
1:D:23:ARG:HH21	1:D:106:ASN:HD21	1.47	0.61
2:A:302:EDO:O1	3:B:304:TYD:H5'1	2.01	0.61
1:B:92:THR:HB	1:B:95:ARG:HG3	1.83	0.60
1:C:193:GLN:HE21	1:C:196:ARG:HH21	1.51	0.58
1:C:66:PRO:HA	1:C:191:GLN:NE2	2.18	0.58
3:A:305:TYD:O3'	6:A:487:HOH:O	2.06	0.57
1:C:21:ARG:NH1	3:C:303:TYD:O3B	2.27	0.57
1:D:23:ARG:HH21	1:D:106:ASN:ND2	2.03	0.57
3:A:305:TYD:H5'1	2:B:301:EDO:O1	2.05	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:79:GLY:HA2	1:D:122:LEU:HG	1.90	0.53
1:D:168:GLU:OE1	6:D:464:HOH:O	2.18	0.53
1:A:117:ARG:CG	1:A:117:ARG:NH1	2.59	0.52
1:C:70:LYS:HD2	1:C:117:ARG:NH1	2.25	0.52
1:D:71:TYR:OH	2:D:301:EDO:C2	2.58	0.52
1:D:179:LEU:O	1:D:179:LEU:HD12	2.10	0.52
1:A:117:ARG:HD3	6:A:541:HOH:O	2.09	0.51
1:D:179:LEU:HD23	1:D:181:LEU:HD11	1.92	0.51
1:C:26[A]:GLU:OE1	1:C:29:ARG:CG	2.59	0.50
1:D:23:ARG:NH2	1:D:106:ASN:ND2	2.59	0.50
1:C:188:TRP:O	1:C:192:GLN:HG2	2.10	0.49
1:A:55:ALA:HA	1:A:162:LEU:HB2	1.94	0.49
1:C:64:ILE:HG22	1:C:67:GLY:HA3	1.95	0.49
1:A:96:TRP:HD1	1:A:151:LEU:HD22	1.78	0.48
1:A:68:GLN:OE1	1:A:70:LYS:HE3	2.14	0.48
1:A:196:ARG:NH1	6:A:500:HOH:O	2.45	0.48
1:B:59:ILE:O	1:B:143[B]:SER:HA	2.14	0.47
1:B:51:SER:HB2	1:B:121[B]:SER:OG	2.13	0.47
1:B:52:HIS:HA	6:B:508:HOH:O	2.14	0.47
1:B:160:GLU:OE2	1:B:160:GLU:HA	2.13	0.47
3:A:305:TYD:C5'	2:B:301:EDO:O1	2.63	0.47
1:C:76:ARG:HD2	1:D:76:ARG:HD2	1.98	0.45
1:C:158:ASP:OD1	1:C:158:ASP:N	2.47	0.45
1:B:51:SER:HB3	1:B:55:ALA:HB3	1.97	0.45
1:B:59:ILE:O	1:B:143[A]:SER:HA	2.17	0.45
1:A:1:MET:HG2	1:A:12:GLN:HE21	1.81	0.45
1:C:59:ILE:O	1:C:143:SER:HA	2.17	0.45
1:A:59:ILE:O	1:A:143[A]:SER:HA	2.17	0.44
1:D:66:PRO:CA	1:D:191:GLN:HE22	2.26	0.44
1:B:119:PHE:HE1	1:B:121[B]:SER:HG	1.66	0.44
1:C:63:GLU:HB2	1:C:140:ARG:HG3	2.00	0.43
1:D:179:LEU:HG	1:D:181:LEU:HG	2.00	0.43
1:C:83:VAL:O	1:C:96:TRP:HA	2.18	0.42
1:D:158:ASP:OD1	1:D:158:ASP:N	2.46	0.42
1:A:51:SER:HB3	1:A:55:ALA:HB3	2.02	0.41
1:A:59:ILE:O	1:A:143[B]:SER:HA	2.21	0.41
1:B:55:ALA:HA	1:B:162:LEU:HB2	2.03	0.41
1:B:119:PHE:HD1	6:B:463:HOH:O	2.03	0.41
1:B:140:ARG:NH1	6:B:509:HOH:O	2.46	0.41
1:C:66:PRO:CA	1:C:191:GLN:HE22	2.32	0.41
1:C:92:THR:HB	1:C:95:ARG:HG3	2.03	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:64:ILE:HG22	1:D:67:GLY:HA3	2.03	0.40
1:A:23:ARG:HA	2:A:302:EDO:H22	2.03	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	201/201 (100%)	200 (100%)	1 (0%)	0	100	100
1	B	202/201 (100%)	201 (100%)	1 (0%)	0	100	100
1	C	201/201 (100%)	197 (98%)	4 (2%)	0	100	100
1	D	199/201 (99%)	196 (98%)	3 (2%)	0	100	100
All	All	803/804 (100%)	794 (99%)	9 (1%)	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	167/165 (101%)	161 (96%)	6 (4%)	35	12
1	B	168/165 (102%)	164 (98%)	4 (2%)	49	24
1	C	167/165 (101%)	159 (95%)	8 (5%)	25	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	164/165 (99%)	159 (97%)	5 (3%)	41	16
All	All	666/660 (101%)	643 (96%)	23 (4%)	36	13

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	ARG
1	A	117	ARG
1	A	120	LEU
1	A	151	LEU
1	A	165	ASP
1	A	198	GLU
1	B	31	GLU
1	B	95	ARG
1	B	157	ASP
1	B	165	ASP
1	C	13	GLU
1	C	23	ARG
1	C	95	ARG
1	C	117	ARG
1	C	140	ARG
1	C	151	LEU
1	C	193	GLN
1	C	201	HIS
1	D	12	GLN
1	D	13	GLU
1	D	117	ARG
1	D	151	LEU
1	D	166	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	12	GLN
1	A	195	GLN
1	B	12	GLN
1	B	35	GLN
1	B	195	GLN
1	B	199	HIS
1	C	191	GLN
1	C	193	GLN

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type
1	D	12	GLN
1	D	106	ASN
1	D	191	GLN
1	D	193	GLN

### 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](#)

19 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	$\# Z  > 2$
2	EDO	A	302	-	3,3,3	0.43	0	2,2,2	1.30	0
2	EDO	D	302	-	3,3,3	0.63	0	2,2,2	0.39	0
2	EDO	B	301	-	3,3,3	0.44	0	2,2,2	0.19	0
3	TYD	C	304	-	21,26,26	0.67	0	27,40,40	1.72	3 (11%)
4	THM	B	305	-	18,18,18	1.53	3 (16%)	26,26,26	1.38	4 (15%)
2	EDO	D	301	-	3,3,3	0.43	0	2,2,2	0.73	0
2	EDO	B	303	-	3,3,3	0.49	0	2,2,2	0.28	0
2	EDO	C	302	-	3,3,3	0.55	0	2,2,2	0.22	0
2	EDO	C	301	-	3,3,3	0.63	0	2,2,2	0.35	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	THM	A	306	-	18,18,18	1.52	3 (16%)	26,26,26	1.42	5 (19%)
2	EDO	A	303	-	3,3,3	0.51	0	2,2,2	0.37	0
2	EDO	B	302	-	3,3,3	0.53	0	2,2,2	0.27	0
5	TDR	C	305	-	9,9,9	2.03	2 (22%)	12,12,12	2.52	5 (41%)
3	TYD	C	303	-	21,26,26	0.65	0	27,40,40	1.66	1 (3%)
3	TYD	B	304	-	21,26,26	0.67	0	27,40,40	1.62	4 (14%)
2	EDO	A	304	-	3,3,3	0.45	0	2,2,2	0.49	0
5	TDR	D	303	-	9,9,9	1.97	2 (22%)	12,12,12	3.41	7 (58%)
3	TYD	A	305	-	21,26,26	0.65	0	27,40,40	1.44	5 (18%)
2	EDO	A	301	-	3,3,3	0.52	0	2,2,2	0.68	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
2	EDO	A	302	-	-	1/1/1/1	-
2	EDO	D	302	-	-	0/1/1/1	-
2	EDO	B	301	-	-	1/1/1/1	-
3	TYD	C	304	-	-	3/13/28/28	0/2/2/2
4	THM	B	305	-	-	2/6/18/18	0/2/2/2
2	EDO	D	301	-	-	0/1/1/1	-
2	EDO	B	303	-	-	0/1/1/1	-
2	EDO	C	302	-	-	0/1/1/1	-
2	EDO	C	301	-	-	0/1/1/1	-
4	THM	A	306	-	-	2/6/18/18	0/2/2/2
2	EDO	A	303	-	-	0/1/1/1	-
2	EDO	B	302	-	-	1/1/1/1	-
5	TDR	C	305	-	-	-	0/1/1/1
3	TYD	C	303	-	-	2/13/28/28	0/2/2/2
3	TYD	B	304	-	-	7/13/28/28	0/2/2/2
2	EDO	A	304	-	-	0/1/1/1	-
5	TDR	D	303	-	-	-	0/1/1/1
3	TYD	A	305	-	-	2/13/28/28	0/2/2/2
2	EDO	A	301	-	-	0/1/1/1	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	305	TDR	C6-C5	4.44	1.40	1.34
5	D	303	TDR	C6-C5	4.33	1.40	1.34
4	A	306	THM	C6-C5	3.46	1.40	1.34
4	B	305	THM	C4-C5	-3.41	1.39	1.44
4	B	305	THM	C6-C5	3.29	1.40	1.34
5	D	303	TDR	C4-C5	-3.20	1.39	1.44
4	A	306	THM	C4-C5	-3.19	1.39	1.44
5	C	305	TDR	C4-C5	-3.19	1.39	1.44
4	B	305	THM	C2-N1	-2.52	1.34	1.38
4	A	306	THM	C2-N1	-2.45	1.34	1.38

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	303	TDR	O2-C2-N1	-7.26	114.79	122.79
3	C	303	TYD	C2-N3-C4	6.98	121.04	115.14
3	C	304	TYD	C2-N3-C4	6.70	120.80	115.14
3	B	304	TYD	C2-N3-C4	5.75	120.00	115.14
5	D	303	TDR	N1-C2-N3	5.47	121.33	115.13
5	C	305	TDR	N1-C2-N3	5.34	121.18	115.13
5	C	305	TDR	C6-N1-C2	-5.30	117.27	122.68
5	D	303	TDR	C6-N1-C2	-4.77	117.80	122.68
3	A	305	TYD	C2-N3-C4	3.89	118.43	115.14
3	C	304	TYD	O3B-PB-O3A	3.18	115.31	104.64
5	D	303	TDR	O4-C4-C5	-3.10	121.31	124.90
4	A	306	THM	N3-C2-N1	2.99	118.86	114.89
5	D	303	TDR	C4-N3-C2	-2.91	122.14	126.34
3	A	305	TYD	C5-C6-N1	-2.84	119.13	122.19
3	B	304	TYD	O5'-C5'-C4'	2.82	118.70	108.99
4	B	305	THM	O4-C4-C5	-2.66	121.82	124.90
3	B	304	TYD	C5M-C5-C4	2.63	126.08	121.37
3	A	305	TYD	O3B-PB-O2B	2.62	117.64	107.64
3	B	304	TYD	O3B-PB-O3A	2.62	113.41	104.64
4	B	305	THM	N3-C2-N1	2.53	118.25	114.89
5	D	303	TDR	C5-C4-N3	2.51	117.45	115.31
3	C	304	TYD	O5'-PA-O1A	-2.46	99.44	109.07
4	A	306	THM	C6-N1-C2	-2.31	118.95	121.30
5	C	305	TDR	C4-N3-C2	-2.28	123.05	126.34
5	D	303	TDR	CM5-C5-C6	-2.28	120.64	123.16
4	B	305	THM	O3'-C3'-C2'	-2.20	103.03	110.90
4	A	306	THM	O3'-C3'-C2'	-2.18	103.11	110.90
5	C	305	TDR	C5-C4-N3	2.18	117.17	115.31
3	A	305	TYD	O2B-PB-O3A	2.12	111.75	104.64

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	305	THM	O4'-C1'-N1	2.11	111.63	107.86
3	A	305	TYD	O2B-PB-O1B	-2.09	102.50	110.68
5	C	305	TDR	O2-C2-N3	-2.08	117.90	121.82
4	A	306	THM	C2'-C3'-C4'	2.05	107.03	102.76
4	A	306	THM	C2'-C1'-N1	-2.01	109.13	113.77

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	304	TYD	C5'-O5'-PA-O1A
3	B	304	TYD	C5'-O5'-PA-O2A
3	C	303	TYD	PA-O3A-PB-O3B
3	C	304	TYD	PA-O3A-PB-O3B
4	A	306	THM	O4'-C4'-C5'-O5'
4	A	306	THM	C3'-C4'-C5'-O5'
4	B	305	THM	O4'-C4'-C5'-O5'
4	B	305	THM	C3'-C4'-C5'-O5'
3	B	304	TYD	C3'-C4'-C5'-O5'
3	B	304	TYD	O4'-C4'-C5'-O5'
3	A	305	TYD	C5'-O5'-PA-O3A
3	B	304	TYD	PB-O3A-PA-O2A
3	B	304	TYD	PA-O3A-PB-O1B
3	A	305	TYD	PB-O3A-PA-O1A
2	B	301	EDO	O1-C1-C2-O2
3	C	303	TYD	PA-O3A-PB-O2B
3	C	304	TYD	PA-O3A-PB-O2B
3	B	304	TYD	C5'-O5'-PA-O3A
2	A	302	EDO	O1-C1-C2-O2
2	B	302	EDO	O1-C1-C2-O2
3	C	304	TYD	PA-O3A-PB-O1B

There are no ring outliers.

8 monomers are involved in 11 short contacts:

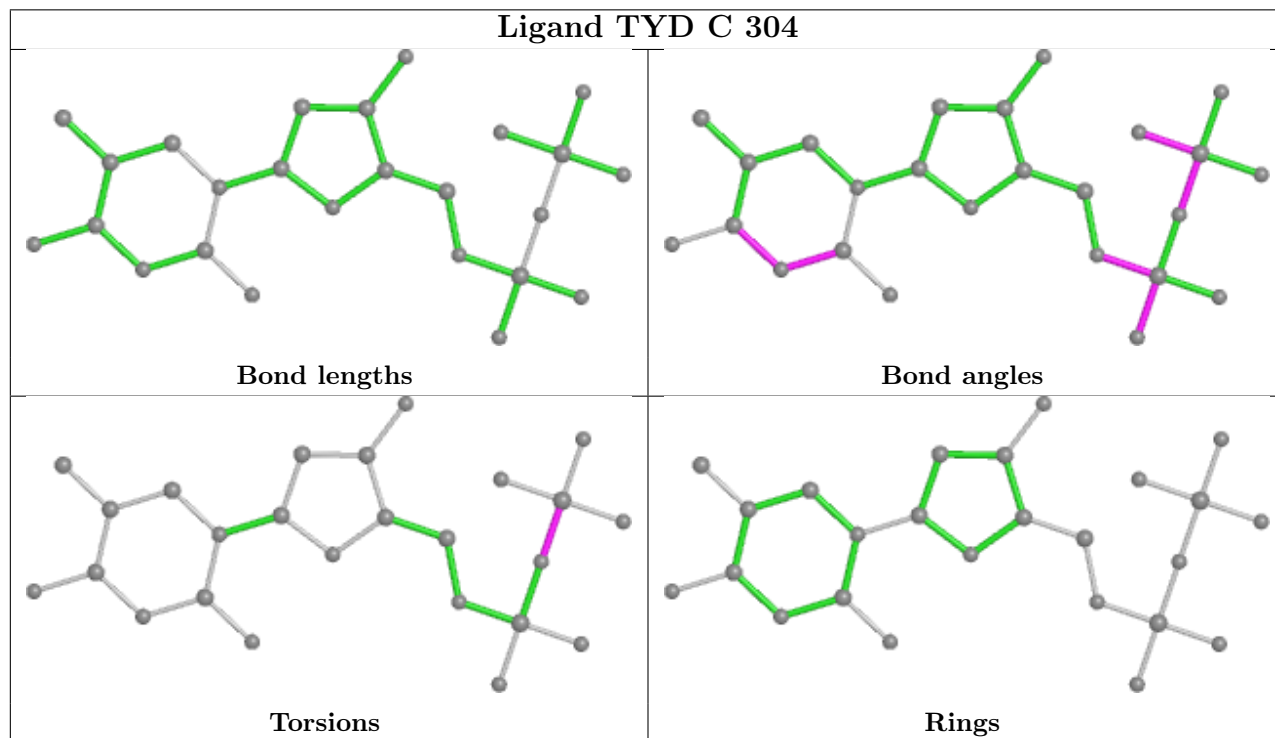
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	302	EDO	2	0
2	B	301	EDO	2	0
3	C	304	TYD	1	0
2	D	301	EDO	2	0
3	C	303	TYD	2	0

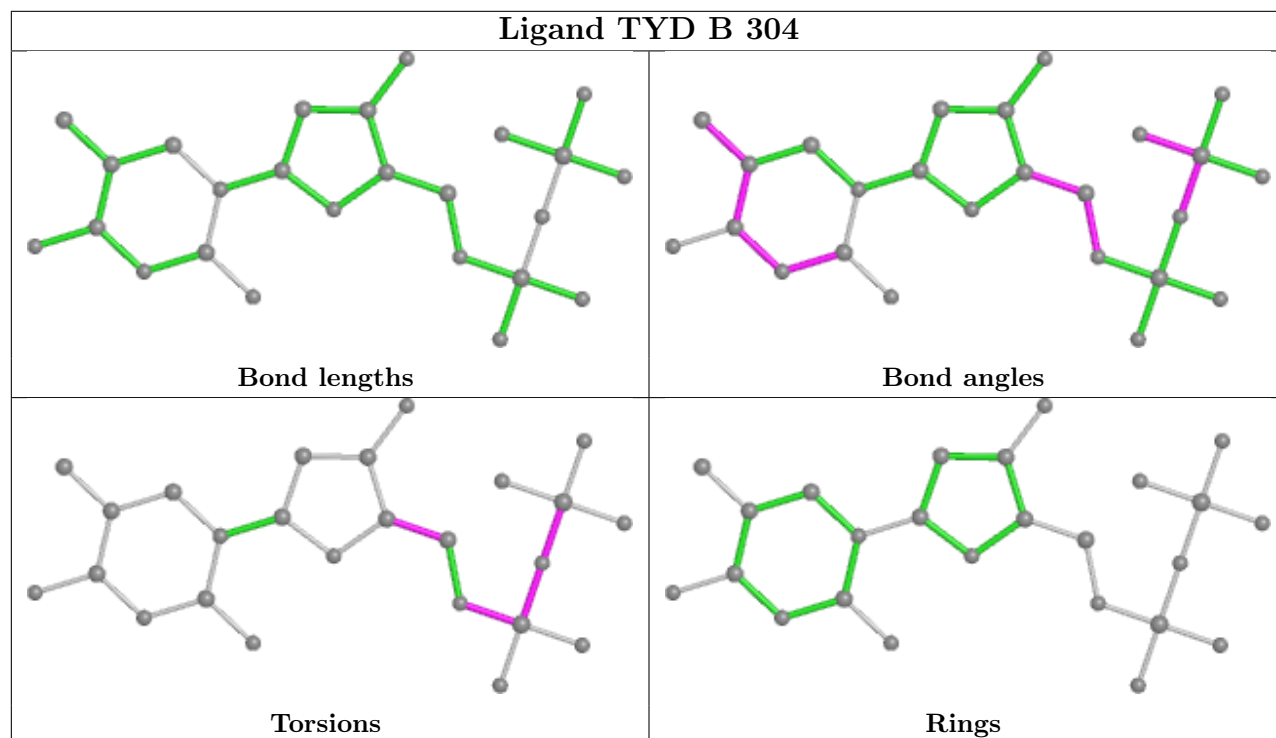
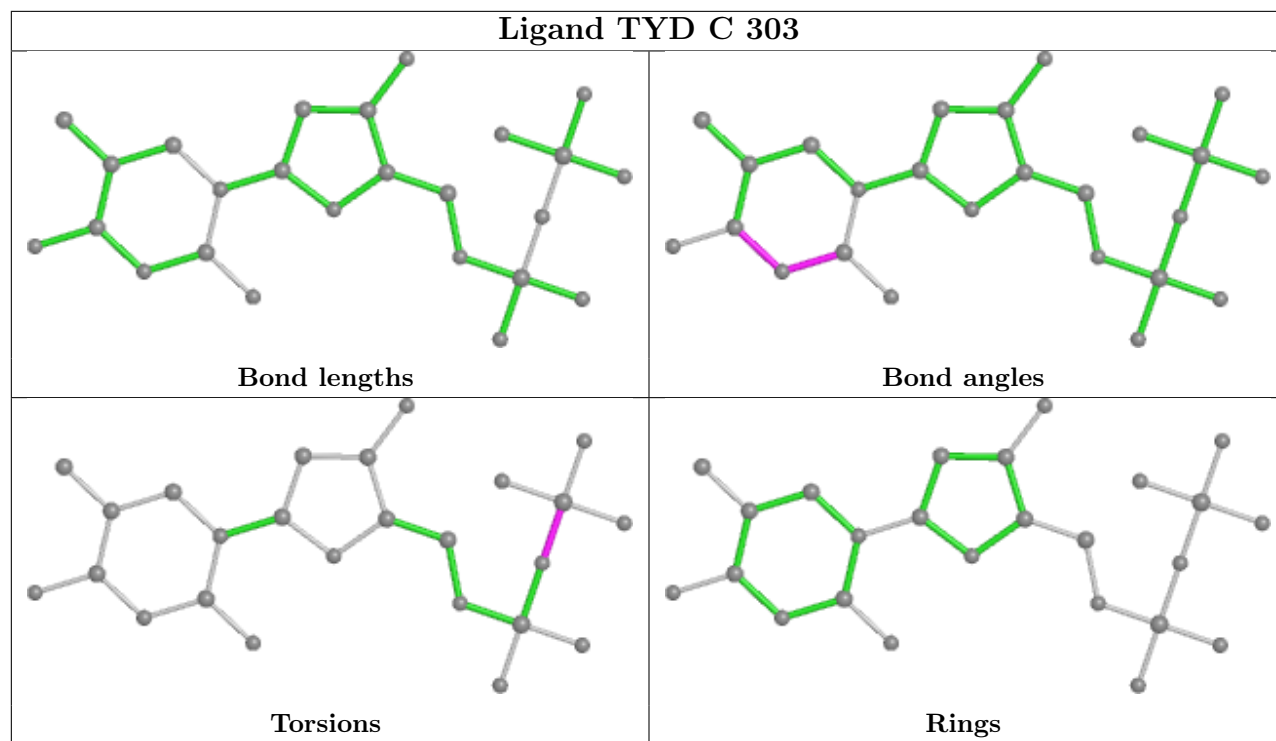
*Continued on next page...*

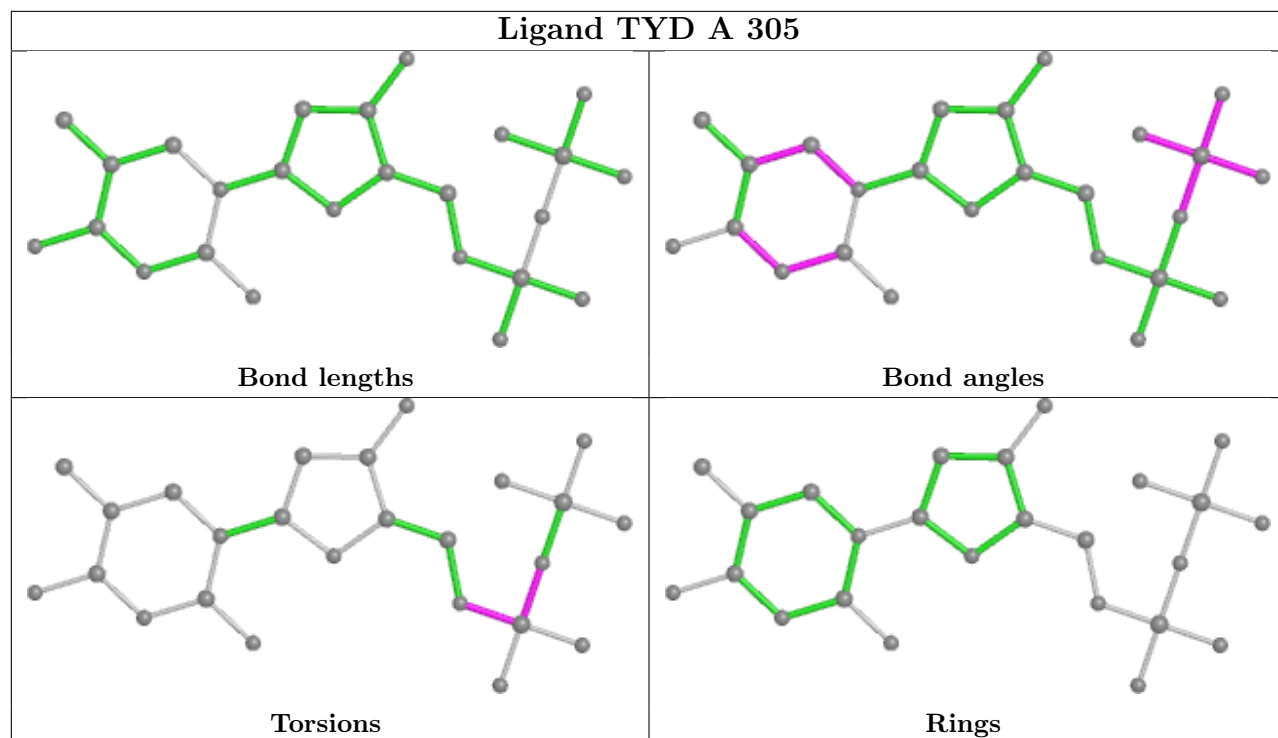
*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	304	TYD	1	0
2	A	304	EDO	1	0
3	A	305	TYD	3	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	199/201 (99%)	-0.14	2 (1%) 82 82	14, 24, 42, 65	0
1	B	199/201 (99%)	-0.11	3 (1%) 73 73	14, 24, 41, 67	0
1	C	201/201 (100%)	0.05	4 (1%) 65 64	17, 30, 48, 67	0
1	D	200/201 (99%)	0.06	3 (1%) 73 73	17, 30, 46, 56	0
All	All	799/804 (99%)	-0.03	12 (1%) 73 73	14, 26, 46, 67	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	198	GLU	4.2
1	C	200	HIS	3.8
1	C	201	HIS	3.6
1	D	179	LEU	3.4
1	B	197	LEU	3.0
1	D	65	PRO	2.8
1	A	197	LEU	2.7
1	B	199	HIS	2.7
1	C	88	ILE	2.3
1	C	157	ASP	2.2
1	A	198	GLU	2.1
1	D	150	ASP	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

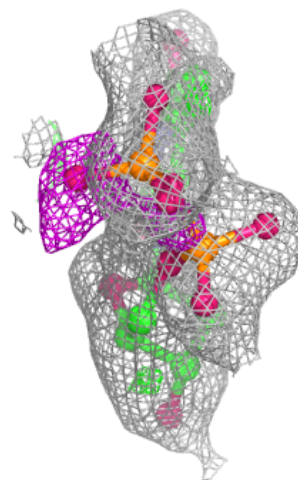
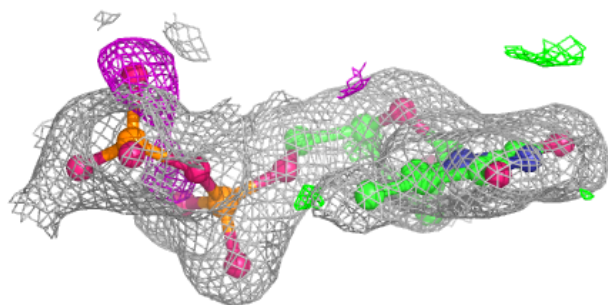
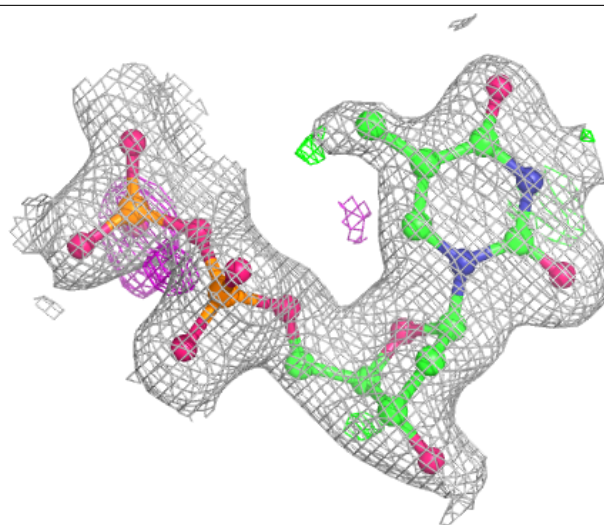
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	TDR	C	305	9/9	0.86	0.24	30,35,40,48	0
4	THM	B	305	17/17	0.87	0.17	27,42,55,61	0
2	EDO	D	301	4/4	0.87	0.11	35,37,37,39	0
5	TDR	D	303	9/9	0.87	0.17	30,35,41,42	0
4	THM	A	306	17/17	0.88	0.17	29,40,55,56	0
2	EDO	A	302	4/4	0.88	0.12	25,27,34,50	0
3	TYD	B	304	25/25	0.89	0.12	28,40,54,61	0
3	TYD	A	305	25/25	0.90	0.12	31,40,53,59	0
2	EDO	B	303	4/4	0.91	0.15	32,32,33,37	0
2	EDO	A	304	4/4	0.91	0.12	34,44,44,48	0
3	TYD	C	304	25/25	0.92	0.11	35,42,63,68	0
3	TYD	C	303	25/25	0.93	0.13	37,41,66,69	0
2	EDO	C	302	4/4	0.94	0.07	32,33,33,34	0
2	EDO	D	302	4/4	0.94	0.09	21,21,23,30	0
2	EDO	C	301	4/4	0.95	0.09	20,21,23,28	0
2	EDO	A	303	4/4	0.95	0.09	25,28,30,38	0
2	EDO	B	301	4/4	0.96	0.07	25,25,34,45	0
2	EDO	A	301	4/4	0.96	0.12	30,32,33,35	0
2	EDO	B	302	4/4	0.97	0.07	24,27,28,34	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.

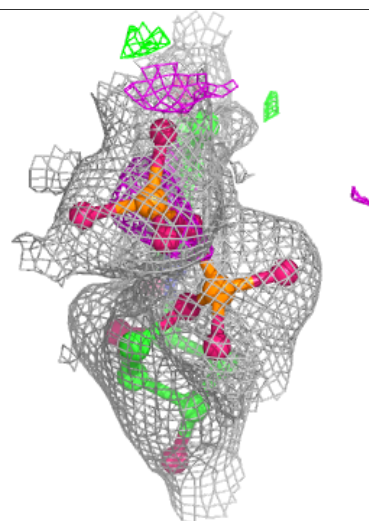
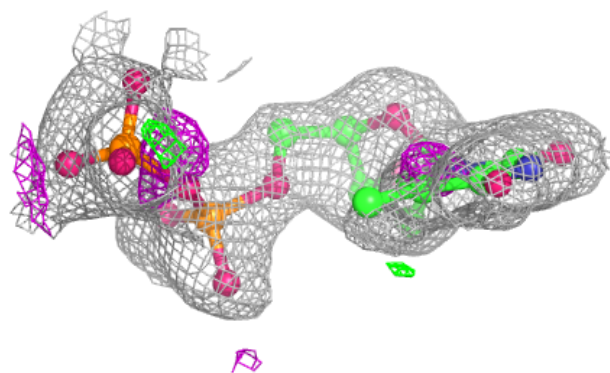
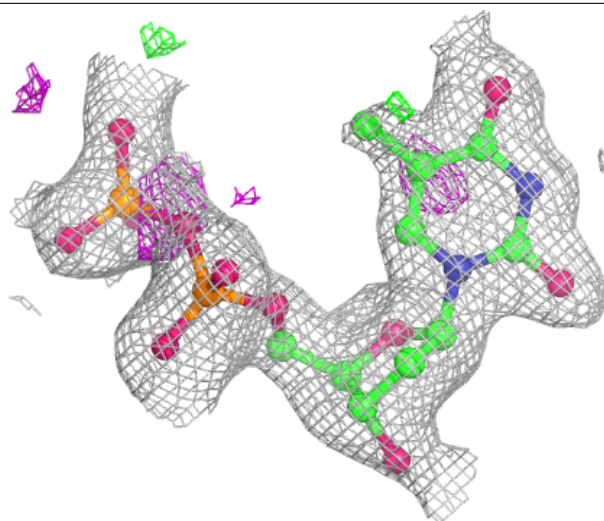
**Electron density around TYD B 304:**

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



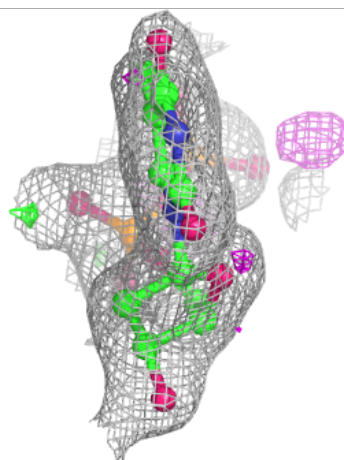
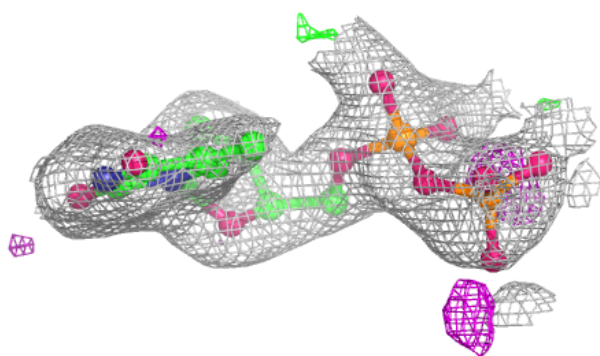
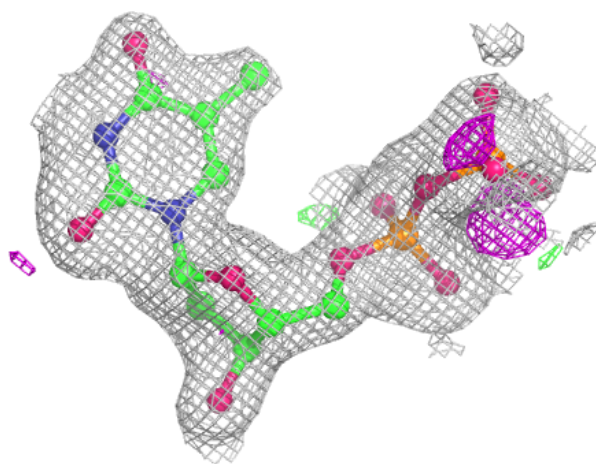
**Electron density around TYD A 305:**

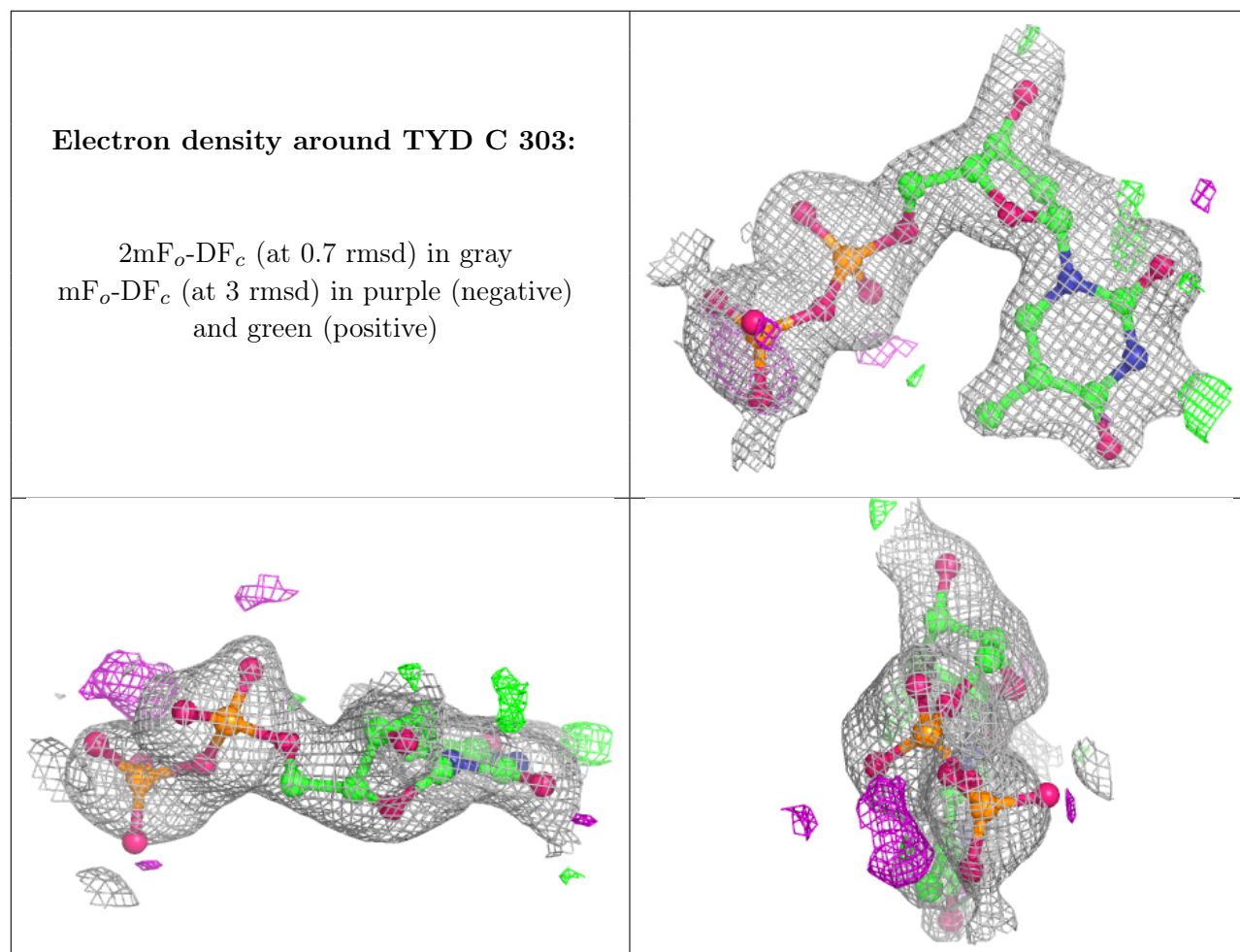
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around TYD C 304:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.