



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 11, 2024 – 07:01 PM EDT

PDB ID : 6N0U  
Title : Crystal structure of a glucose-1-phosphate thymidyltransferase from Burkholderia phymatum bound to 2'-deoxy-thymidine-B-L-rhamnose  
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : 2018-11-07  
Resolution : 2.10 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 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	:	2.36.2
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)	:	2.36.2

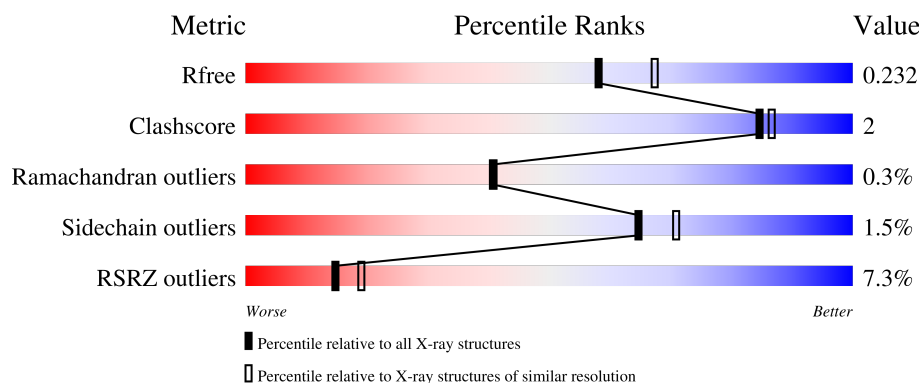
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	305	<div> <div>0%</div> <div>91%</div> <div>6%</div> <div>.</div> </div>
1	B	305	<div> <div>2%</div> <div>90%</div> <div>6%</div> <div>.</div> </div>
1	C	305	<div> <div>3%</div> <div>90%</div> <div>7%</div> <div>..</div> </div>
1	D	305	<div> <div>22%</div> <div>88%</div> <div>7%</div> <div>6%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9576 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 Glucose-1-phosphate thymidyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	295	Total	C	N	O	S	0	4	0
			2297	1467	390	431	9			
1	B	293	Total	C	N	O	S	0	0	0
			2238	1427	377	427	7			
1	C	295	Total	C	N	O	S	0	1	0
			2301	1470	392	431	8			
1	D	288	Total	C	N	O	S	0	0	0
			2096	1336	351	402	7			

There are 32 discrepancies between the modelled and reference sequences:

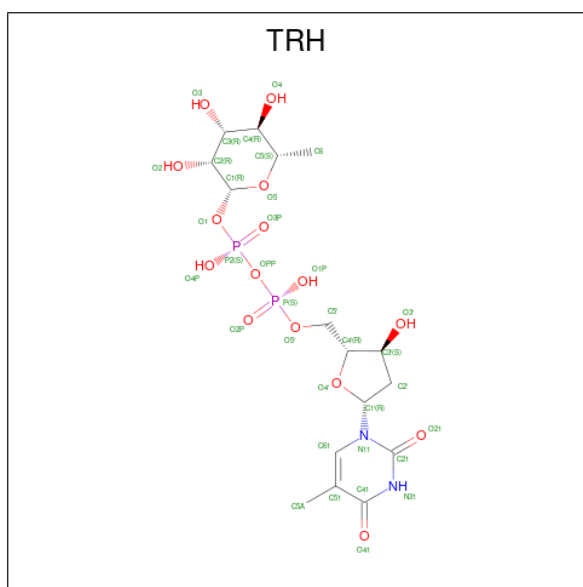
Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP B2JFC5
A	-6	ALA	-	expression tag	UNP B2JFC5
A	-5	HIS	-	expression tag	UNP B2JFC5
A	-4	HIS	-	expression tag	UNP B2JFC5
A	-3	HIS	-	expression tag	UNP B2JFC5
A	-2	HIS	-	expression tag	UNP B2JFC5
A	-1	HIS	-	expression tag	UNP B2JFC5
A	0	HIS	-	expression tag	UNP B2JFC5
B	-7	MET	-	initiating methionine	UNP B2JFC5
B	-6	ALA	-	expression tag	UNP B2JFC5
B	-5	HIS	-	expression tag	UNP B2JFC5
B	-4	HIS	-	expression tag	UNP B2JFC5
B	-3	HIS	-	expression tag	UNP B2JFC5
B	-2	HIS	-	expression tag	UNP B2JFC5
B	-1	HIS	-	expression tag	UNP B2JFC5
B	0	HIS	-	expression tag	UNP B2JFC5
C	-7	MET	-	initiating methionine	UNP B2JFC5
C	-6	ALA	-	expression tag	UNP B2JFC5
C	-5	HIS	-	expression tag	UNP B2JFC5
C	-4	HIS	-	expression tag	UNP B2JFC5
C	-3	HIS	-	expression tag	UNP B2JFC5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	HIS	-	expression tag	UNP B2JFC5
C	-1	HIS	-	expression tag	UNP B2JFC5
C	0	HIS	-	expression tag	UNP B2JFC5
D	-7	MET	-	initiating methionine	UNP B2JFC5
D	-6	ALA	-	expression tag	UNP B2JFC5
D	-5	HIS	-	expression tag	UNP B2JFC5
D	-4	HIS	-	expression tag	UNP B2JFC5
D	-3	HIS	-	expression tag	UNP B2JFC5
D	-2	HIS	-	expression tag	UNP B2JFC5
D	-1	HIS	-	expression tag	UNP B2JFC5
D	0	HIS	-	expression tag	UNP B2JFC5

- Molecule 2 is 2'-DEOXY-THYMIDINE-BETA-L-RHAMNOSE (three-letter code: TRH) (formula:  $C_{16}H_{26}N_2O_{15}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 35	C 16	N 2	O 15	P 2	0	0
2	B	1	Total 35	C 16	N 2	O 15	P 2	0	0
2	C	1	Total 35	C 16	N 2	O 15	P 2	0	0
2	D	1	Total 35	C 16	N 2	O 15	P 2	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	177	Total	O	0	2
			179	179		
4	B	160	Total	O	0	2
			162	162		
4	C	134	Total	O	0	0
			134	134		
4	D	20	Total	O	0	1
			21	21		



N264	A265	L268	L269	A270	L271	A272	R273	P274	L275	S276	K277	N278	A279	Y280	G281	Q282	Y283	L284	Q285	N286	L287	L288	T289	ASP	GLN	VAL	ALA	TRP	PRO	SER	VAL

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	152.32Å 80.18Å 131.04Å 90.00° 123.38° 90.00°	Depositor
Resolution (Å)	42.33 – 2.10 42.33 – 2.10	Depositor EDS
% Data completeness (in resolution range)	91.2 (42.33-2.10) 91.4 (42.33-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.33 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.14 _3260)	Depositor
R, $R_{free}$	0.182 , 0.233 0.182 , 0.232	Depositor DCC
$R_{free}$ test set	4271 reflections (6.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.8	Xtriage
Anisotropy	0.284	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 54.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.003 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9576	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, TRH

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	0.56	1/2359 (0.0%)	0.68	0/3210
1	B	0.51	0/2288	0.66	0/3123
1	C	0.44	0/2357	0.75	3/3209 (0.1%)
1	D	0.32	0/2144	0.53	0/2938
All	All	0.47	1/9148 (0.0%)	0.66	3/12480 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	37	TYR	CD1-CE1	5.30	1.47	1.39

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	294	TRP	C-N-CD	-19.48	77.74	120.60
1	C	294	TRP	C-N-CA	13.80	179.95	122.00
1	C	295	PRO	CA-N-CD	-6.63	102.21	111.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	294	TRP	Peptide

## 5.2 Too-close contacts

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	2297	0	2264	10	0
1	B	2238	0	2153	13	0
1	C	2301	0	2262	10	0
1	D	2096	0	1921	10	0
2	A	35	0	24	0	0
2	B	35	0	24	0	0
2	C	35	0	24	0	0
2	D	35	0	24	0	0
3	A	4	0	6	2	0
3	C	4	0	6	0	0
4	A	179	0	0	2	0
4	B	162	0	0	3	0
4	C	134	0	0	0	0
4	D	21	0	0	0	0
All	All	9576	0	8708	41	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 41 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:A:117:ASP:H	3:A:501:EDO:H11	1.41	0.86
1:D:82:GLN:NE2	1:D:84:SER:O	2.29	0.64
1:B:64:GLU:OE2	4:B:601:HOH:O	2.17	0.59
1:A:62[B]:ARG:HH21	1:A:62[B]:ARG:HB3	1.68	0.58
1:A:117:ASP:N	3:A:501:EDO:H11	2.16	0.58

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	297/305 (97%)	291 (98%)	5 (2%)	1 (0%)	41	41
1	B	291/305 (95%)	288 (99%)	2 (1%)	1 (0%)	41	41
1	C	294/305 (96%)	291 (99%)	2 (1%)	1 (0%)	41	41
1	D	286/305 (94%)	283 (99%)	2 (1%)	1 (0%)	41	41
All	All	1168/1220 (96%)	1153 (99%)	11 (1%)	4 (0%)	41	41

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	TYR
1	B	31	TYR
1	C	31	TYR
1	D	31	TYR

### 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	238/252 (94%)	234 (98%)	4 (2%)	60	67
1	B	227/252 (90%)	224 (99%)	3 (1%)	69	75
1	C	238/252 (94%)	236 (99%)	2 (1%)	81	86
1	D	196/252 (78%)	192 (98%)	4 (2%)	55	60
All	All	899/1008 (89%)	886 (99%)	13 (1%)	65	73

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1	MET
1	C	60	THR
1	D	213	ASP
1	D	111	ASN
1	D	129	THR

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

Mol	Chain	Res	Type
1	A	121	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](#)

6 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	EDO	A	501	-	3,3,3	0.64	0	2,2,2	0.19	0
2	TRH	D	500	-	34,37,37	0.33	0	52,57,57	1.29	2 (3%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	TRH	C	500	-	34,37,37	0.30	0	52,57,57	1.38	2 (3%)
2	TRH	B	500	-	34,37,37	0.41	0	52,57,57	1.71	4 (7%)
3	EDO	C	501	-	3,3,3	0.59	0	2,2,2	0.16	0
2	TRH	A	500	-	34,37,37	0.39	0	52,57,57	1.49	2 (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	EDO	A	501	-	-	0/1/1/1	-
2	TRH	D	500	-	-	3/21/53/53	0/3/3/3
2	TRH	C	500	-	-	3/21/53/53	0/3/3/3
2	TRH	B	500	-	-	3/21/53/53	0/3/3/3
3	EDO	C	501	-	-	1/1/1/1	-
2	TRH	A	500	-	-	4/21/53/53	0/3/3/3

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	500	TRH	O5-C1-O1	-10.87	97.16	111.36
2	C	500	TRH	O5-C1-O1	-8.99	99.61	111.36
2	A	500	TRH	O5-C1-O1	-8.63	100.09	111.36
2	D	500	TRH	O5-C1-O1	-7.50	101.56	111.36
2	A	500	TRH	OPP-P2-O1	5.33	113.23	102.48

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

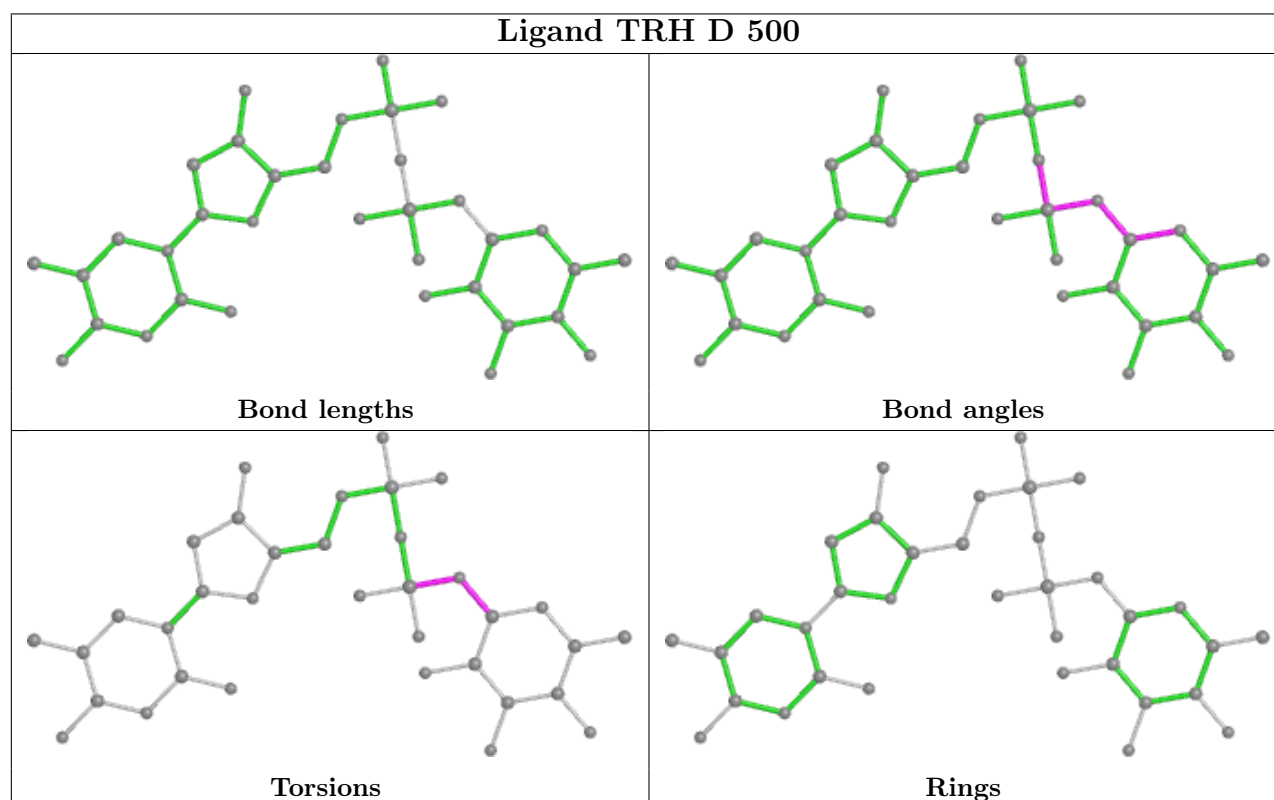
Mol	Chain	Res	Type	Atoms
2	A	500	TRH	O5-C1-O1-P2
2	B	500	TRH	C1-O1-P2-O4P
2	B	500	TRH	O5-C1-O1-P2
2	C	500	TRH	O5-C1-O1-P2
2	D	500	TRH	O5-C1-O1-P2

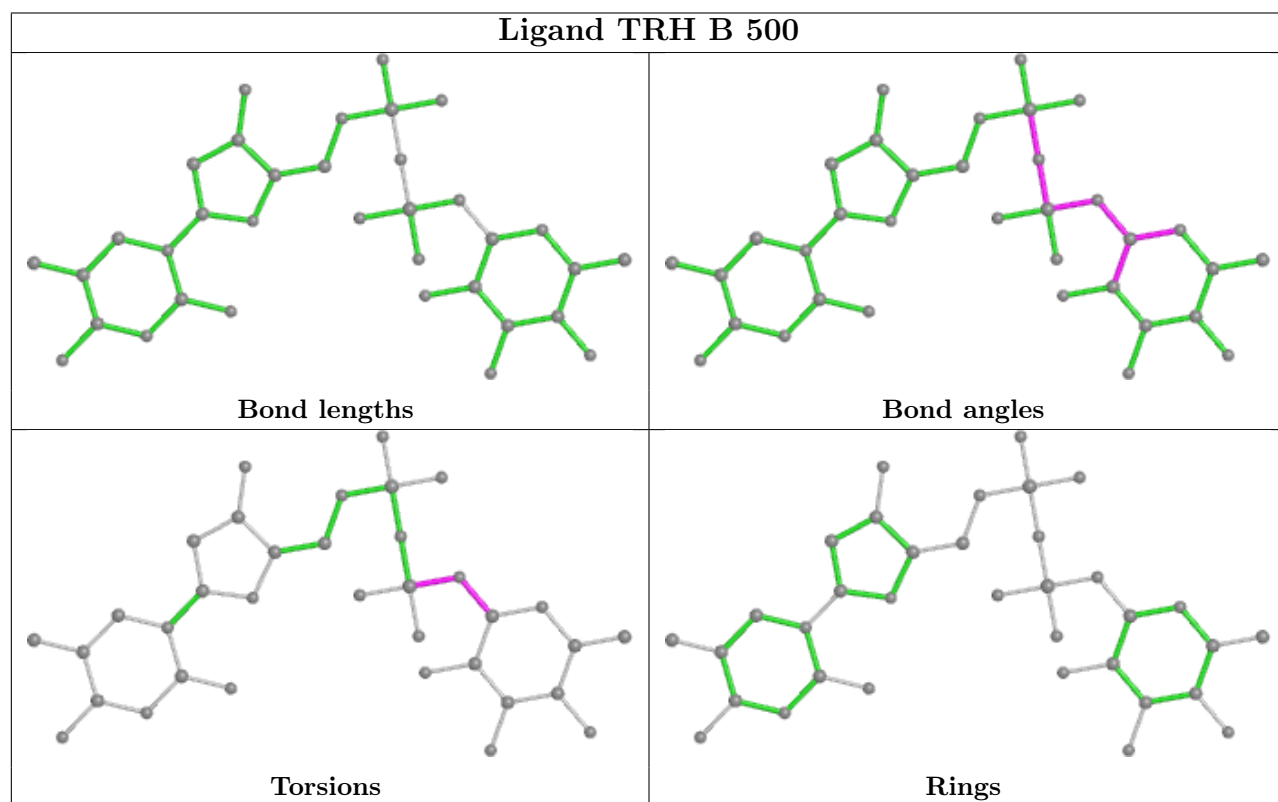
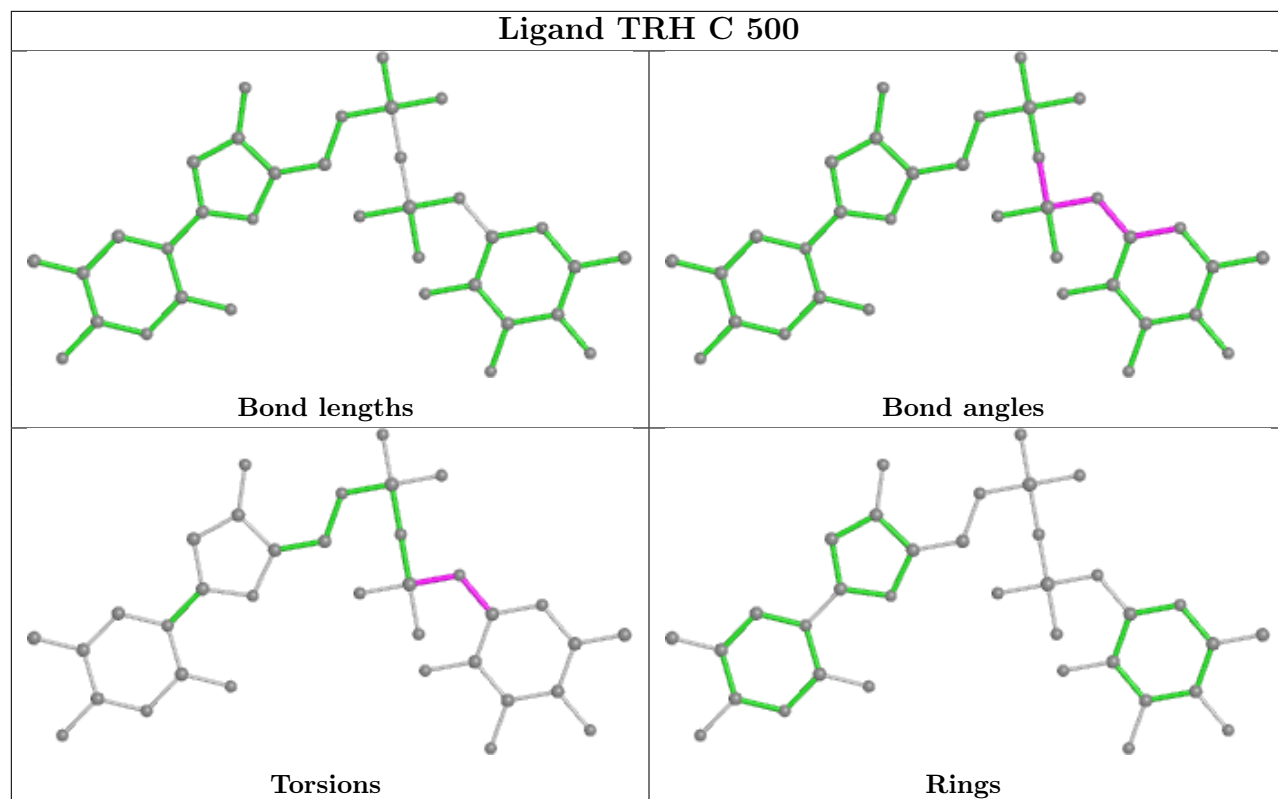
There are no ring outliers.

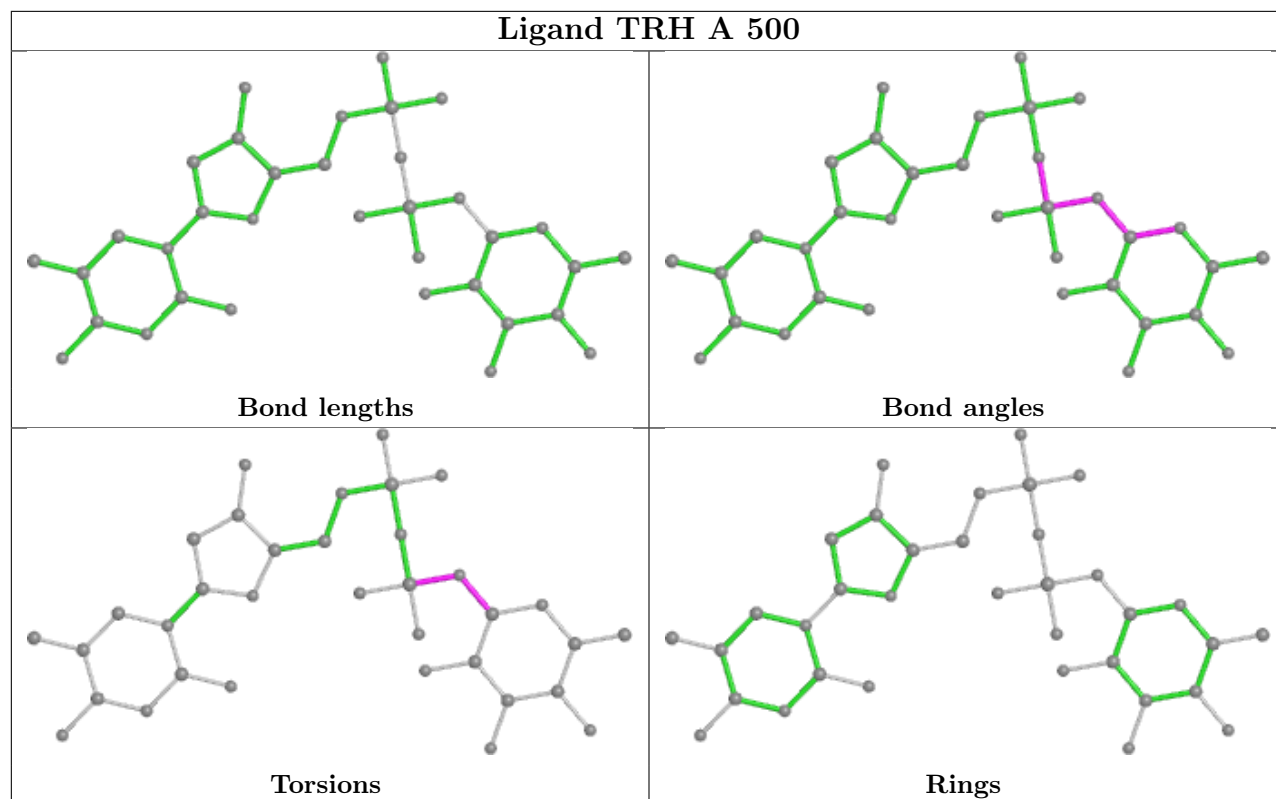
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	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	295/305 (96%)	-0.10	2 (0%) 87 89	17, 37, 63, 103	0
1	B	293/305 (96%)	-0.18	7 (2%) 59 64	19, 41, 64, 95	0
1	C	295/305 (96%)	-0.21	9 (3%) 49 55	34, 49, 69, 102	0
1	D	288/305 (94%)	1.09	68 (23%) 0 0	44, 86, 118, 139	0
All	All	1171/1220 (95%)	0.14	86 (7%) 15 19	17, 48, 104, 139	0

The worst 5 of 86 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	154	PHE	5.7
1	C	1	MET	5.0
1	D	153	ASP	4.9
1	D	202	VAL	4.5
1	D	268	LEU	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](#)

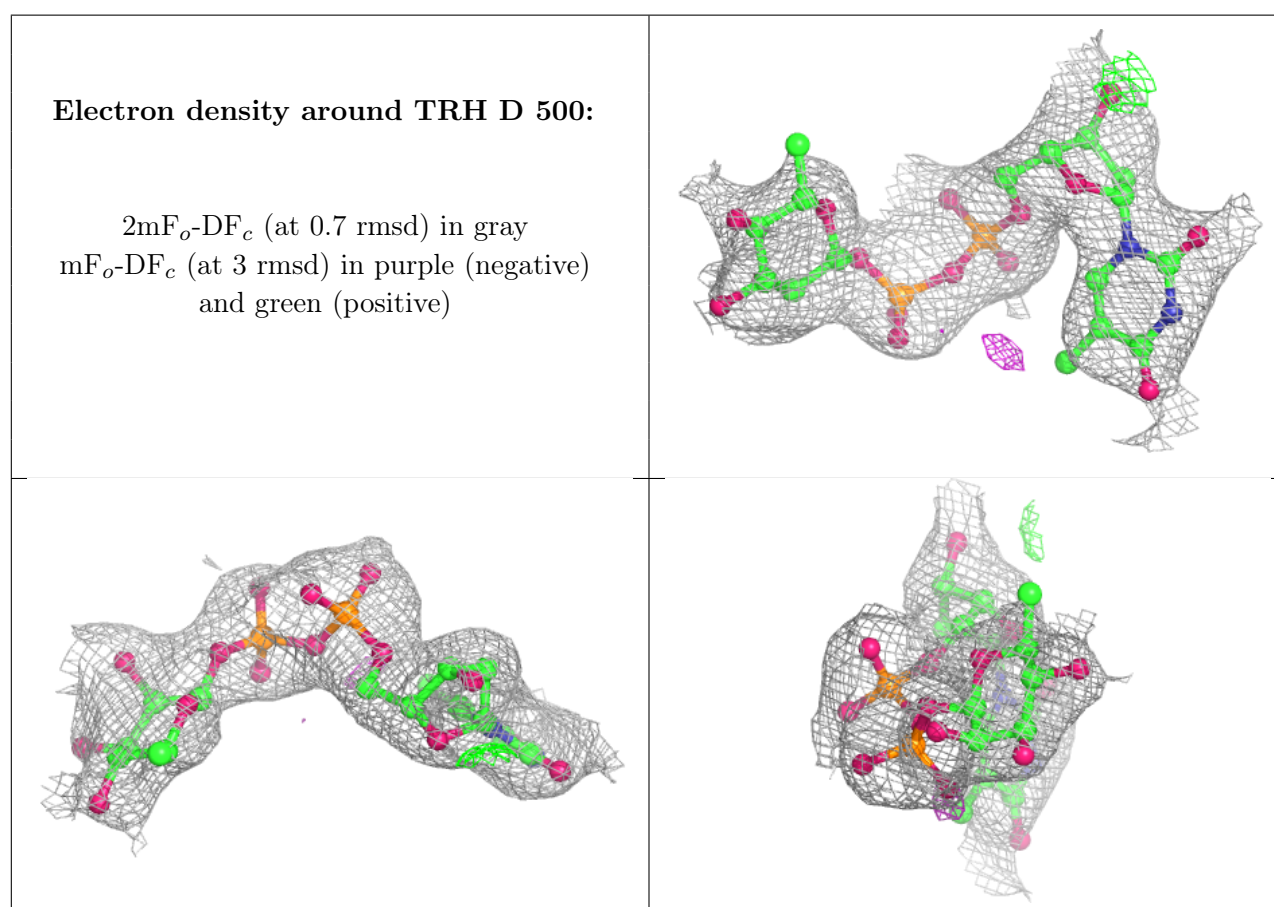
There are no monosaccharides in this entry.

### 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.

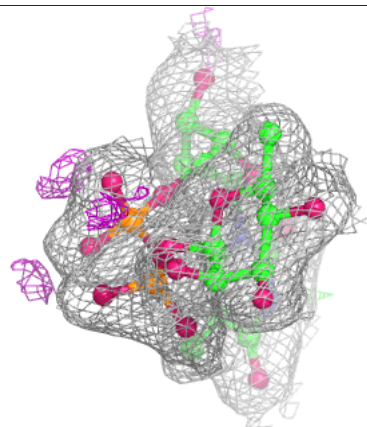
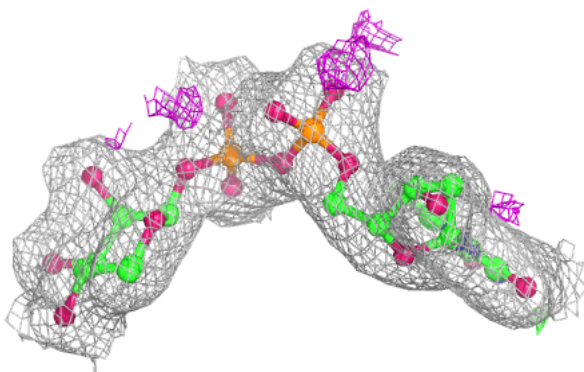
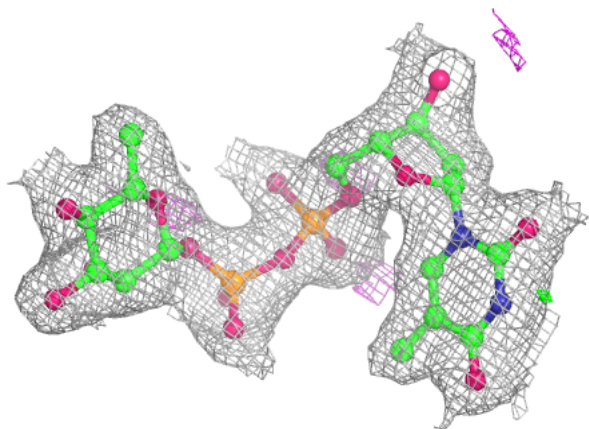
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	C	501	4/4	0.73	0.24	58,64,71,73	0
3	EDO	A	501	4/4	0.88	0.16	39,45,54,60	0
2	TRH	D	500	35/35	0.94	0.14	60,70,77,84	0
2	TRH	A	500	35/35	0.98	0.07	20,29,34,37	0
2	TRH	B	500	35/35	0.98	0.08	23,30,35,39	0
2	TRH	C	500	35/35	0.98	0.07	26,40,47,54	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.



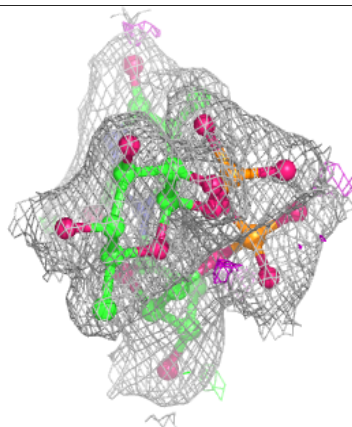
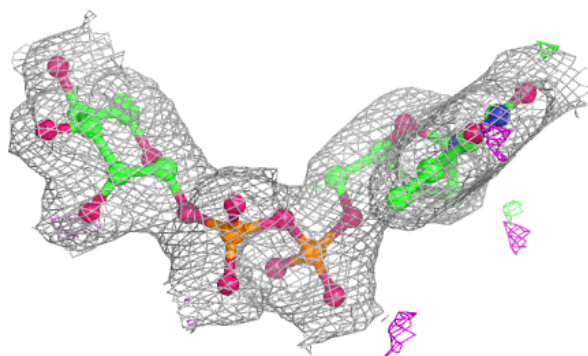
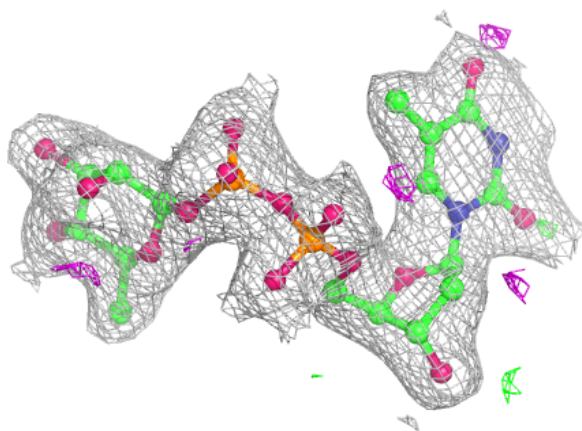
**Electron density around TRH A 500:**

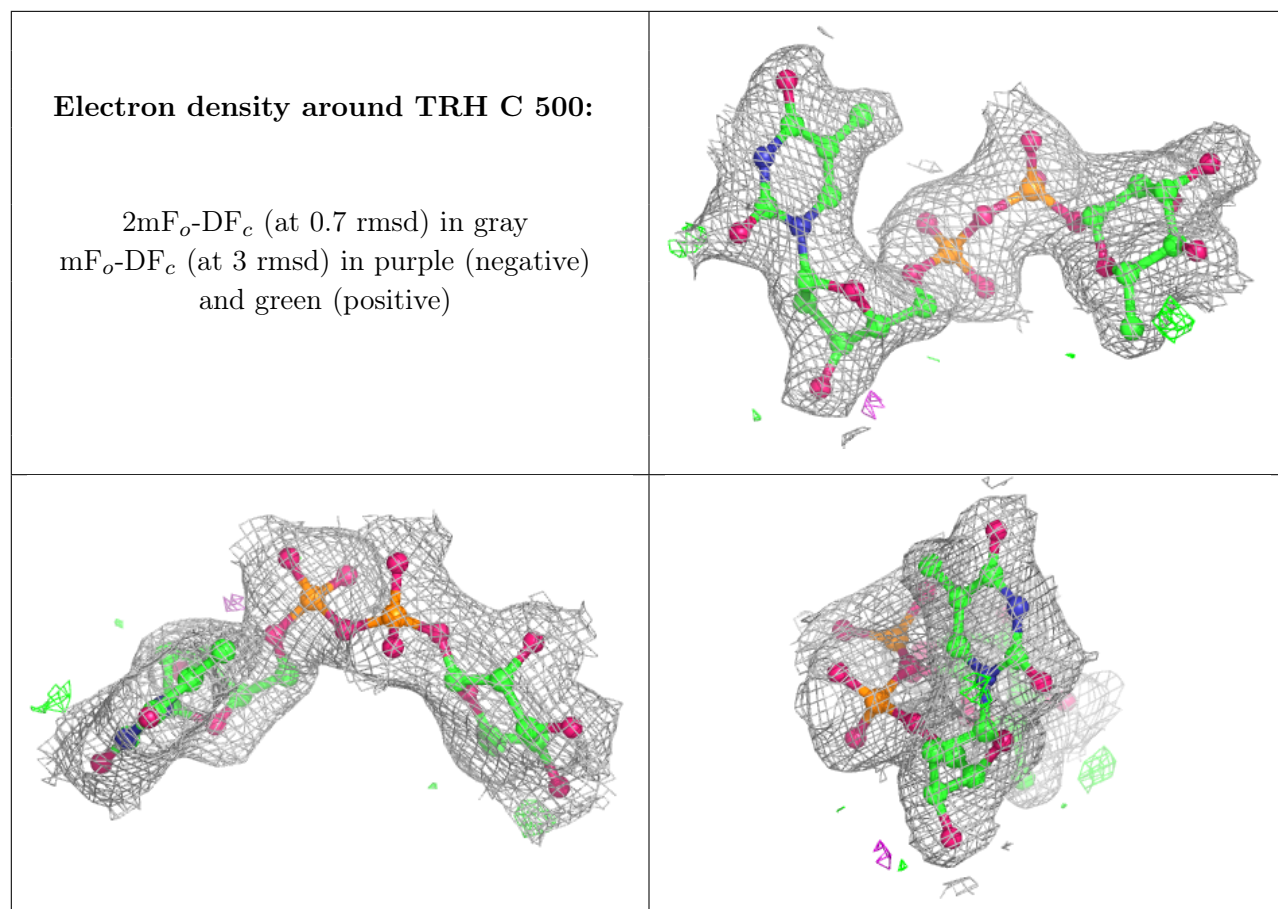
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around TRH B 500:**

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





## 6.5 Other polymers ⓘ

There are no such residues in this entry.