



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 01:35 PM UTC

PDB ID : 4E80 / pdb\_00004e80  
Title : Structural Basis for the Activity of a Cytoplasmic RNA Terminal U-transferase  
Authors : Yates, L.A.; Fleurdepine, S.; Rissland, O.S.; DeColibus, L.; Harlos, K.; Norbury, C.J.; Gilbert, R.J.C.  
Deposited on : 2012-03-19  
Resolution : 3.02 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

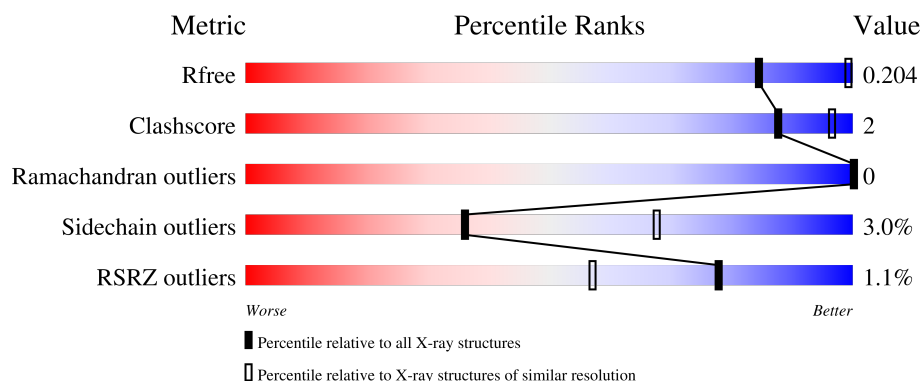
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.02 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))
$R_{free}$	180053	3131 (3.04-3.00)
Clashscore	190562	3444 (3.04-3.00)
Ramachandran outliers	187476	3319 (3.04-3.00)
Sidechain outliers	187428	3322 (3.04-3.00)
RSRZ outliers	180081	3130 (3.04-3.00)

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	405	<div> <div>%</div> <div> <div></div> <div>72%</div> <div>6%</div> <div>22%</div> </div> </div>
1	B	405	<div> <div>%</div> <div> <div></div> <div>70%</div> <div>6%</div> <div>23%</div> </div> </div>
1	C	405	<div> <div></div> <div> <div>74%</div> <div>6%</div> <div>19%</div> </div> </div>
1	D	405	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>6%</div> <div>20%</div> </div> </div>

## 2 Entry composition [i](#)

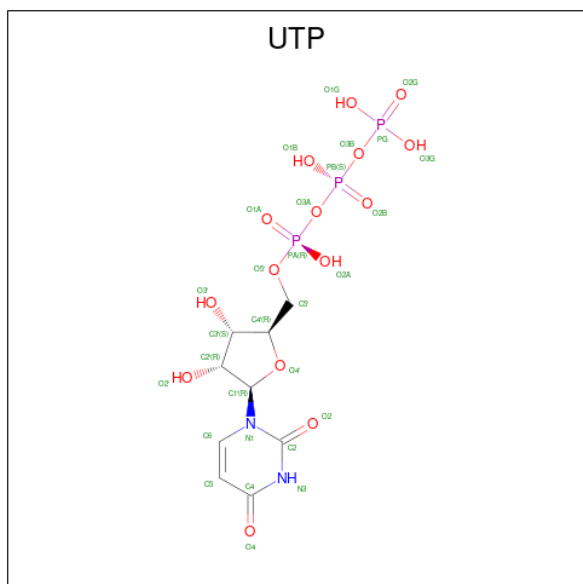
There are 3 unique types of molecules in this entry. The entry contains 10472 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Poly(A) RNA polymerase protein cid1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	317	Total	C	N	O	S	0	0	0
			2563	1667	425	462	9			
1	B	311	Total	C	N	O	S	0	0	0
			2520	1641	416	454	9			
1	C	330	Total	C	N	O	S	0	0	0
			2653	1719	439	486	9			
1	D	324	Total	C	N	O	S	0	0	0
			2603	1686	434	474	9			

- Molecule 2 is URIDINE 5'-TRIPHOSPHATE (CCD ID: UTP) (formula:  $C_9H_{15}N_2O_{15}P_3$ ).



*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			29	9	2	15	3		
2	D	1	Total	C	N	O	P	0	0
			29	9	2	15	3		

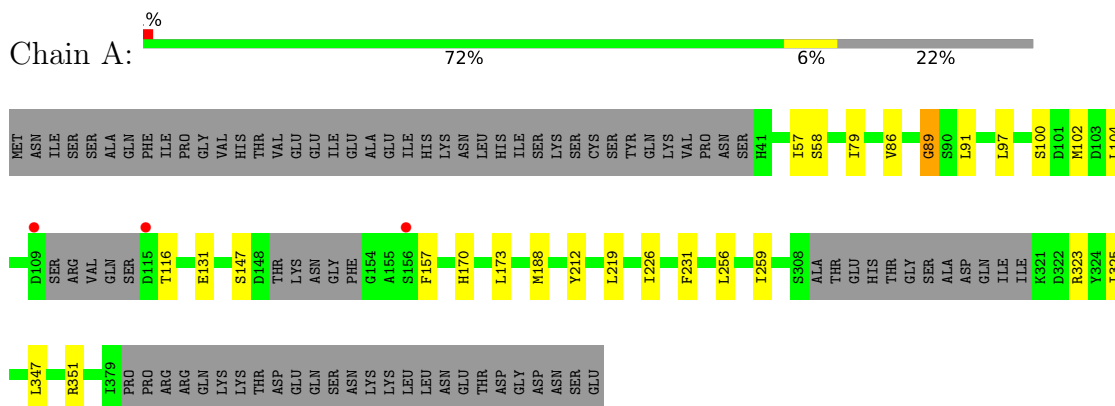
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	O	0	0
			3	3		
3	B	7	Total	O	0	0
			7	7		
3	C	2	Total	O	0	0
			2	2		
3	D	5	Total	O	0	0
			5	5		

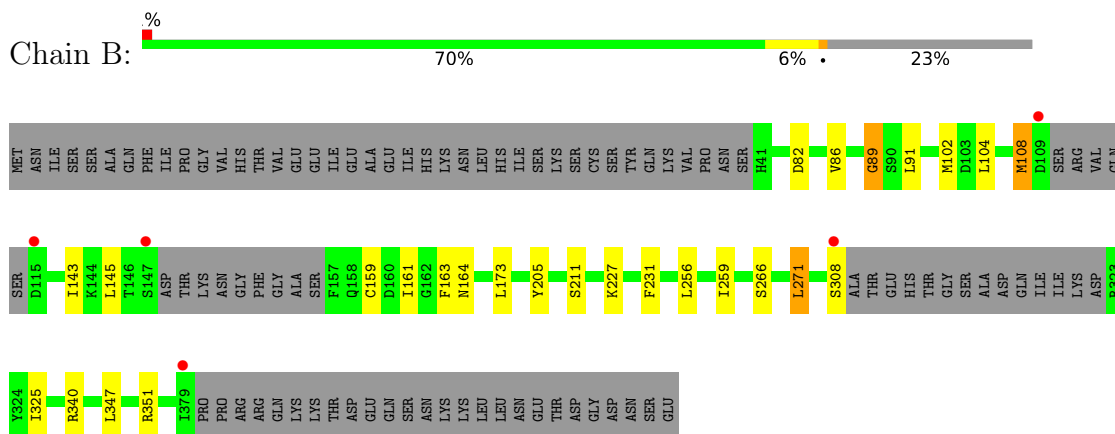
### 3 Residue-property plots [i](#)

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.

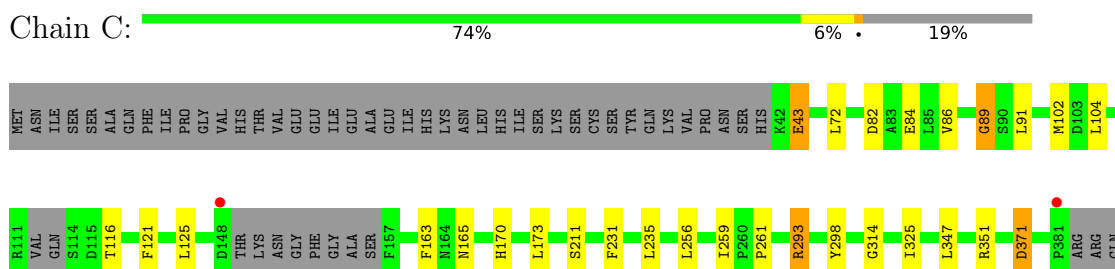
- Molecule 1: Poly(A) RNA polymerase protein cid1



- Molecule 1: Poly(A) RNA polymerase protein cid1



- Molecule 1: Poly(A) RNA polymerase protein cid1



LYS  
LYS  
THR  
ASP  
GLU  
GLN  
SER  
ASN  
LYS  
LYS  
LEU  
LEU  
ASN  
GLU  
THR  
ASP  
GLY  
ASP  
ASN  
SER  
GLU

● Molecule 1: Poly(A) RNA polymerase protein cid1



MET  
ASN  
ILE  
SER  
SER  
ALA  
GLN  
PHE  
ILE  
PRO  
GLY  
VAL  
HIS  
THR  
VAL  
GLU  
GLU  
ILE  
GLU  
ALA  
GLU  
ILE  
HIS  
LYS  
ASN  
LEU  
HIS  
ILE  
SER  
LYS  
SER  
CYS  
SER  
TYR  
GLN  
LYS  
VAL  
PRO  
ASN  
SER  
H41  
L72  
V86  
G89  
S90  
L91  
M102  
D103  
L104  
M108  
D109  
SER  
ARG  
VAL  
GLN

SER  
D115  
T116  
L125  
R139  
I143  
S147  
ASP  
THR  
LYS  
ASN  
GLY  
PHE  
GLY  
ALA  
S156  
I161  
H170  
L173  
Y212  
F231  
L256  
I259  
E311  
S315  
A316  
D317  
GLN  
ILE  
ILE  
K321  
I325  
L347  
R351  
P380  
PRO  
ARG  
ARG  
GLN  
LYS  
THR  
ASP

GLU  
GLN  
SER  
ASN  
LYS  
LEU  
LEU  
THR  
GLY  
ASP  
ASN  
SER  
GLU

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	164.69Å 77.96Å 151.68Å 90.00° 109.29° 90.00°	Depositor
Resolution (Å)	28.72 – 3.02 28.72 – 3.02	Depositor EDS
% Data completeness (in resolution range)	99.1 (28.72-3.02) 99.1 (28.72-3.02)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.84 (at 3.01Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, $R_{free}$	0.176 , 0.201 0.182 , 0.204	Depositor DCC
$R_{free}$ test set	1784 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.1	Xtriage
Anisotropy	0.131	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 63.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10472	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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: UTP

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	0.79	0/2623	1.22	5/3538 (0.1%)
1	B	0.79	0/2579	1.21	5/3480 (0.1%)
1	C	0.80	0/2715	1.24	10/3669 (0.3%)
1	D	0.78	0/2662	1.23	6/3592 (0.2%)
All	All	0.79	0/10579	1.22	26/14279 (0.2%)

There are no bond length outliers.

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	371	ASP	CA-CB-CG	8.12	120.72	112.60
1	D	380	PRO	N-CA-CB	6.33	109.97	103.00
1	D	116	THR	CA-C-N	5.86	128.49	120.46
1	D	116	THR	C-N-CA	5.86	128.49	120.46
1	A	231	PHE	CA-CB-CG	5.62	119.42	113.80

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	2563	0	2573	12	0
1	B	2520	0	2533	10	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2653	0	2650	14	0
1	D	2603	0	2600	7	0
2	A	29	0	11	2	0
2	B	29	0	11	3	0
2	C	29	0	11	3	0
2	D	29	0	11	1	0
3	A	3	0	0	0	0
3	B	7	0	0	0	0
3	C	2	0	0	0	0
3	D	5	0	0	0	0
All	All	10472	0	10400	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:D:256:LEU:HA	1:D:259:ILE:HD12	1.75	0.69
1:A:58:SER:CB	1:C:261:PRO:HG3	2.23	0.69
1:C:256:LEU:HA	1:C:259:ILE:HD12	1.75	0.68
1:A:256:LEU:HA	1:A:259:ILE:HD12	1.75	0.67
1:B:256:LEU:HA	1:B:259:ILE:HD12	1.74	0.67

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	309/405 (76%)	299 (97%)	10 (3%)	0	100	100
1	B	303/405 (75%)	298 (98%)	5 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	324/405 (80%)	314 (97%)	10 (3%)	0	100	100
1	D	316/405 (78%)	309 (98%)	7 (2%)	0	100	100
All	All	1252/1620 (77%)	1220 (97%)	32 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	279/363 (77%)	270 (97%)	9 (3%)	34	66
1	B	275/363 (76%)	265 (96%)	10 (4%)	31	64
1	C	289/363 (80%)	280 (97%)	9 (3%)	35	67
1	D	282/363 (78%)	276 (98%)	6 (2%)	47	74
All	All	1125/1452 (78%)	1091 (97%)	34 (3%)	36	68

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	86	VAL
1	D	91	LEU
1	D	311	GLU
1	B	108	MET
1	B	91	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	120	GLN
1	D	164	ASN
1	A	362	ASN
1	B	164	ASN
1	C	120	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
2	UTP	D	500	-	29,30,30	1.81	3 (10%)	43,47,47	1.68	9 (20%)
2	UTP	B	501	-	29,30,30	1.70	3 (10%)	43,47,47	1.77	10 (23%)
2	UTP	A	501	-	29,30,30	1.90	3 (10%)	43,47,47	1.96	12 (27%)
2	UTP	C	501	-	29,30,30	1.66	2 (6%)	43,47,47	1.70	11 (25%)

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	UTP	D	500	-	-	7/22/38/38	0/2/2/2
2	UTP	B	501	-	-	7/22/38/38	0/2/2/2
2	UTP	A	501	-	-	6/22/38/38	0/2/2/2
2	UTP	C	501	-	-	7/22/38/38	0/2/2/2

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	UTP	PB-O3A	6.54	1.66	1.59
2	D	500	UTP	PB-O3A	6.25	1.66	1.59
2	C	501	UTP	PB-O3A	6.03	1.66	1.59
2	A	501	UTP	PB-O3B	5.43	1.65	1.59
2	D	500	UTP	PB-O3B	5.42	1.65	1.59

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	UTP	C4-N3-C2	-5.38	119.94	126.61
2	A	501	UTP	C4-N3-C2	-5.20	120.16	126.61
2	C	501	UTP	C4-N3-C2	-4.90	120.53	126.61
2	D	500	UTP	C4-N3-C2	-4.87	120.57	126.61
2	A	501	UTP	C5'-C4'-C3'	-4.48	99.07	115.21

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	UTP	C5'-O5'-PA-O2A
2	A	501	UTP	PB-O3B-PG-O3G
2	B	501	UTP	PB-O3B-PG-O3G
2	C	501	UTP	PB-O3B-PG-O3G
2	D	500	UTP	PB-O3B-PG-O1G

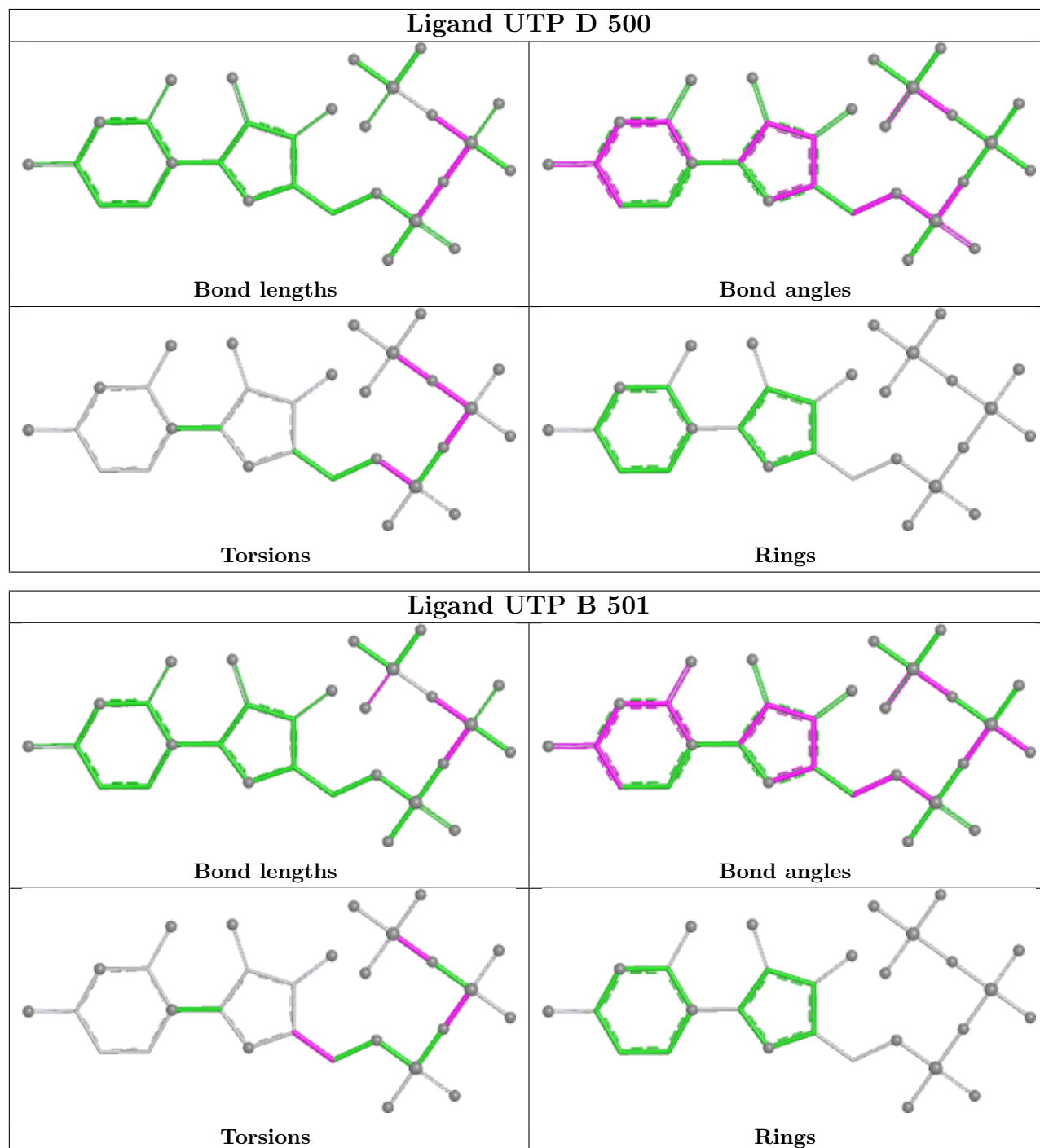
There are no ring outliers.

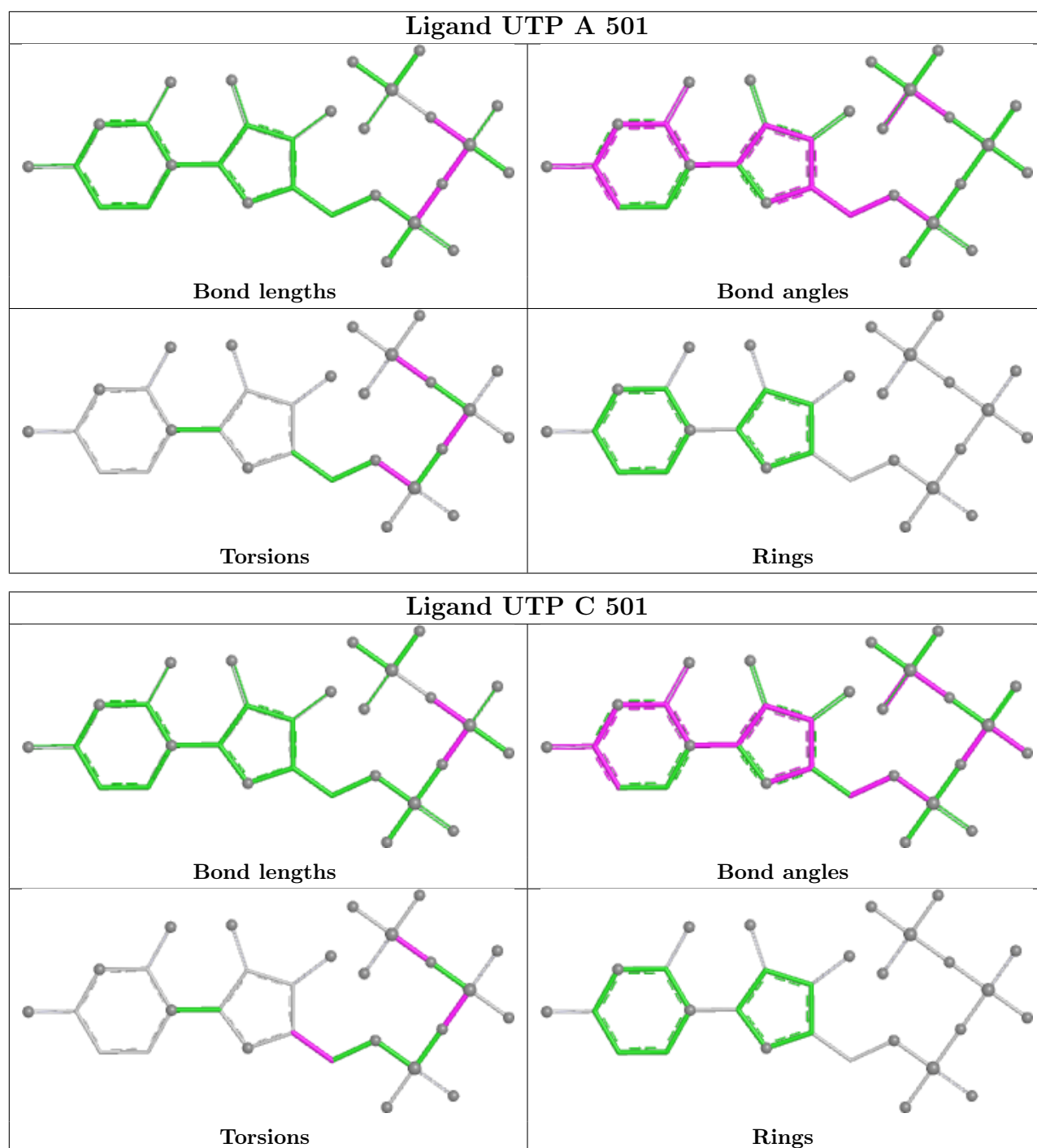
4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	500	UTP	1	0
2	B	501	UTP	3	0
2	A	501	UTP	2	0
2	C	501	UTP	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	317/405 (78%)	-0.21	3 (0%) 81 61	38, 66, 110, 140	0
1	B	311/405 (76%)	-0.53	5 (1%) 70 47	23, 41, 76, 108	0
1	C	330/405 (81%)	-0.29	2 (0%) 85 69	35, 56, 98, 122	0
1	D	324/405 (80%)	-0.29	4 (1%) 76 55	29, 58, 113, 146	0
All	All	1282/1620 (79%)	-0.33	14 (1%) 78 57	23, 56, 105, 146	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	148	ASP	5.0
1	B	109	ASP	3.7
1	A	115	ASP	3.6
1	D	115	ASP	3.6
1	A	109	ASP	3.3

### 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](#)

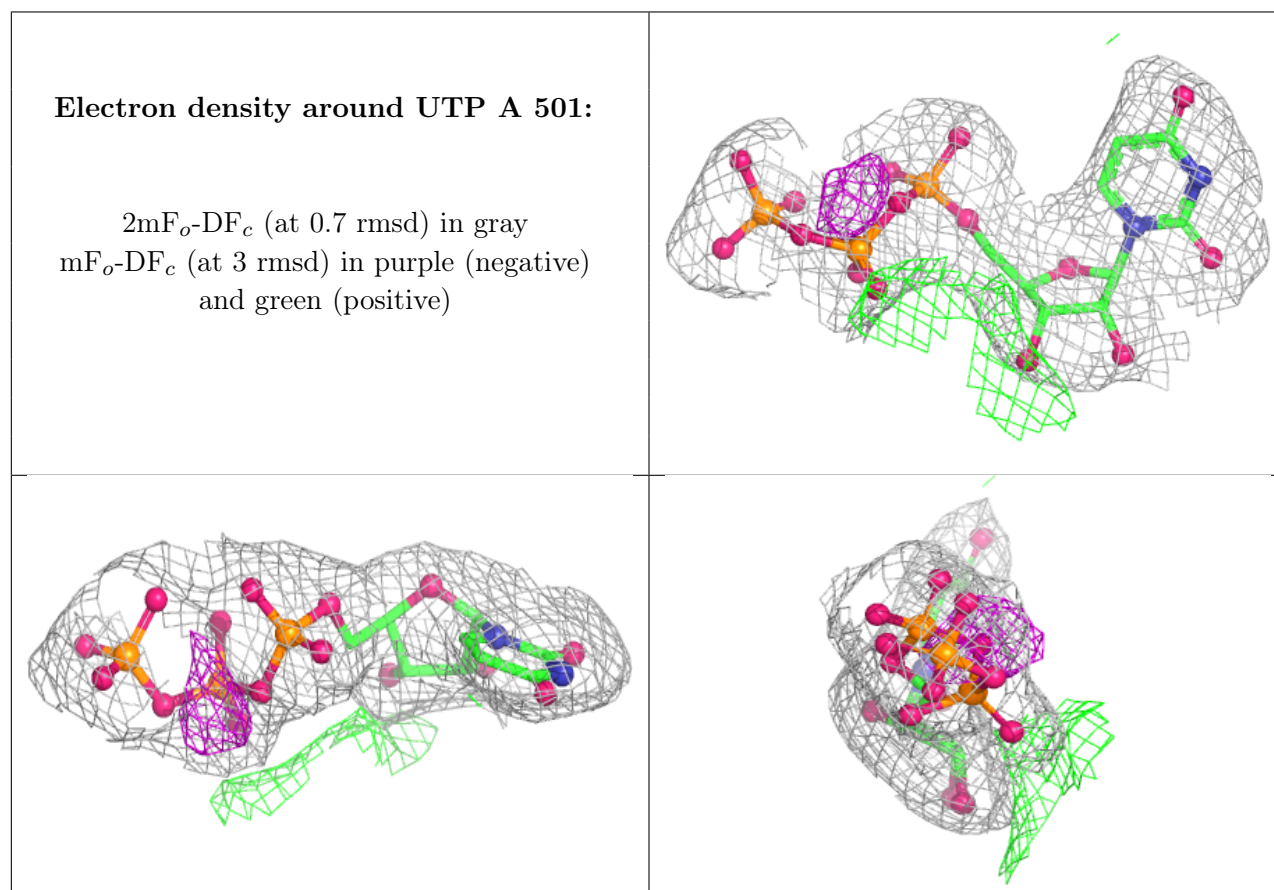
There are no oligosaccharides in this entry.

### 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
2	UTP	A	501	29/29	0.90	0.11	70,76,92,94	0
2	UTP	D	500	29/29	0.91	0.10	43,65,84,86	0
2	UTP	C	501	29/29	0.92	0.09	45,53,77,79	0
2	UTP	B	501	29/29	0.94	0.08	38,43,55,58	0

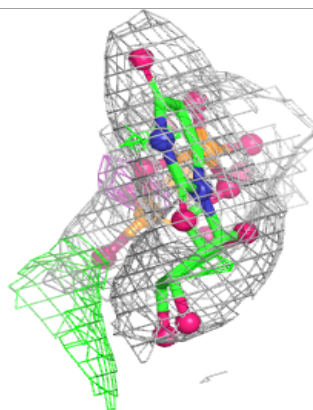
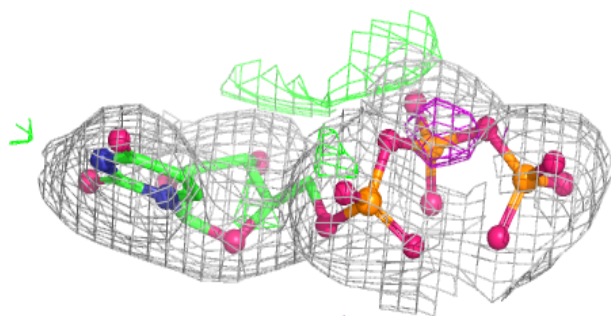
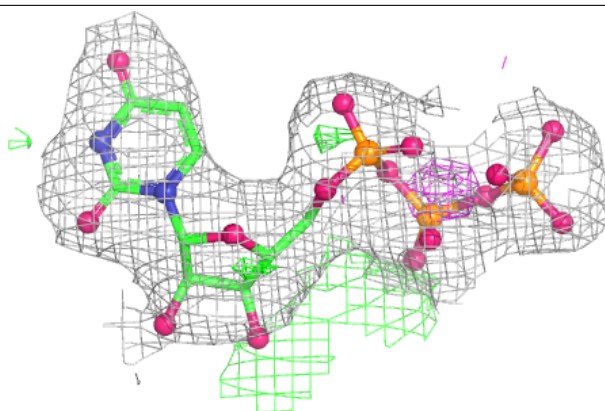
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



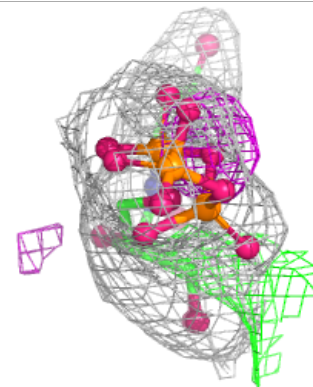
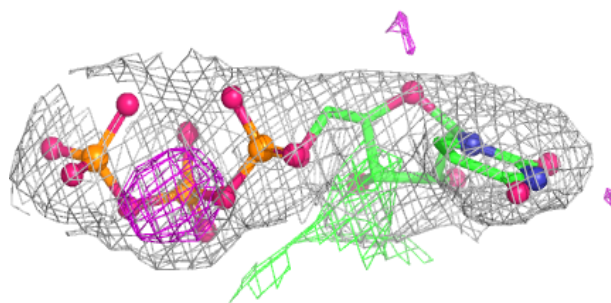
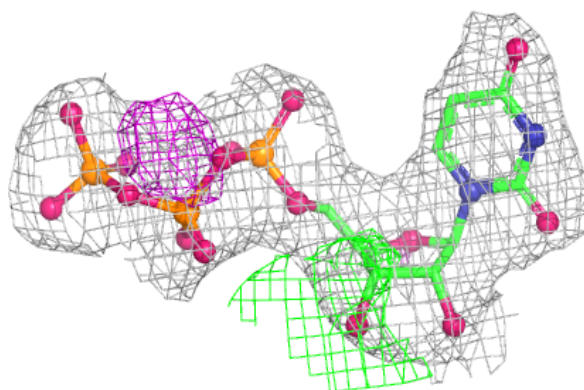


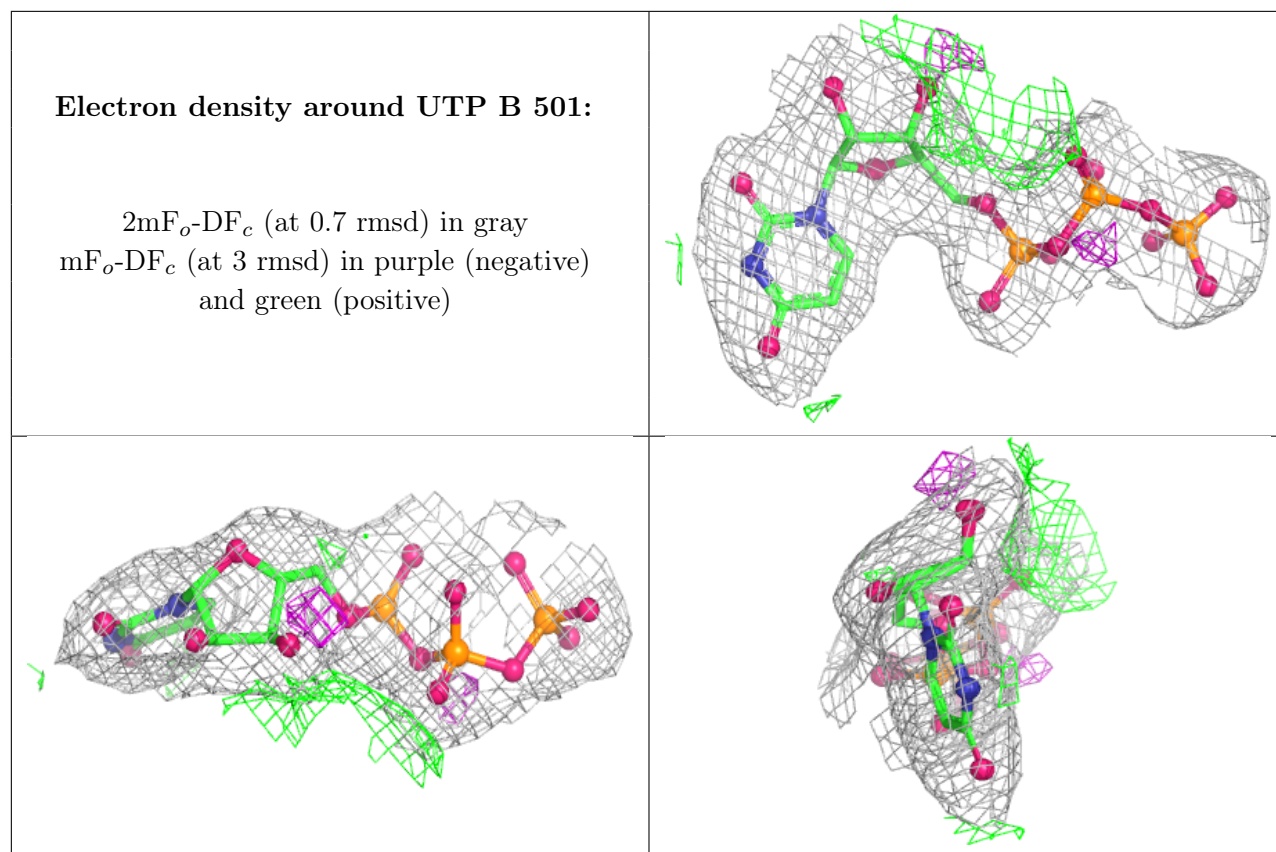
**Electron density around UTP D 500:**

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

**Electron density around UTP C 501:**

$2mF_o-DF_c$  (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.