



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 30, 2020 – 12:16 PM JST

PDB ID : 6LHI  
Title : Quadruple mutant (N51I+C59R+S108N+I164L) plasmodium falciparum dihydrofolate reductase-thymidylate synthase (PfDHFR-TS) complexed with C466 (compound 42) and NADPH  
Authors : Vanichtanankul, J.; Vitsupakorn, D.  
Deposited on : 2019-12-09  
Resolution : 2.59 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.6  
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) : 2.14.6

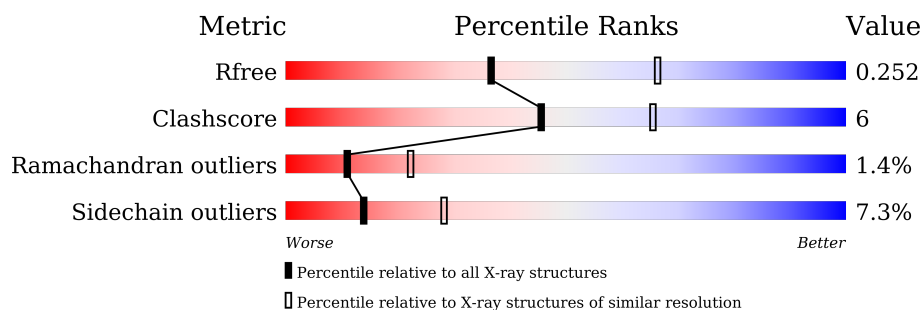
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



The reported resolution of this entry is 2.59 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 on 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	608	 71% 15% • 13%
1	B	608	 66% 18% • 13%

## 2 Entry composition [i](#)

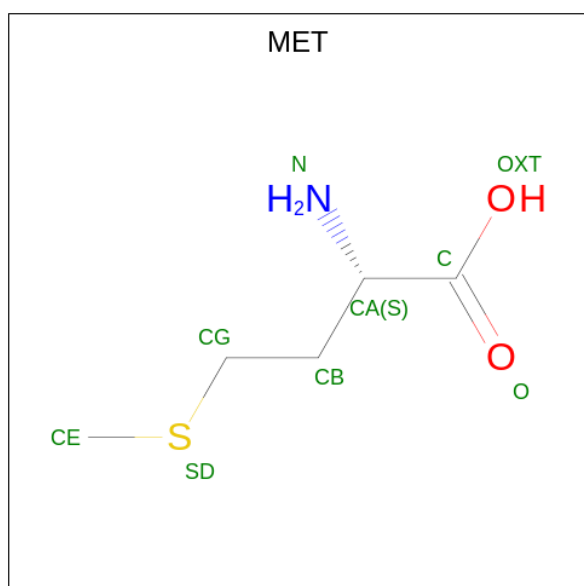
There are 6 unique types of molecules in this entry. The entry contains 18078 atoms, of which 8923 are hydrogens and 0 are deuteriums.

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 Bifunctional dihydrofolate reductase-thymidylate synthase.

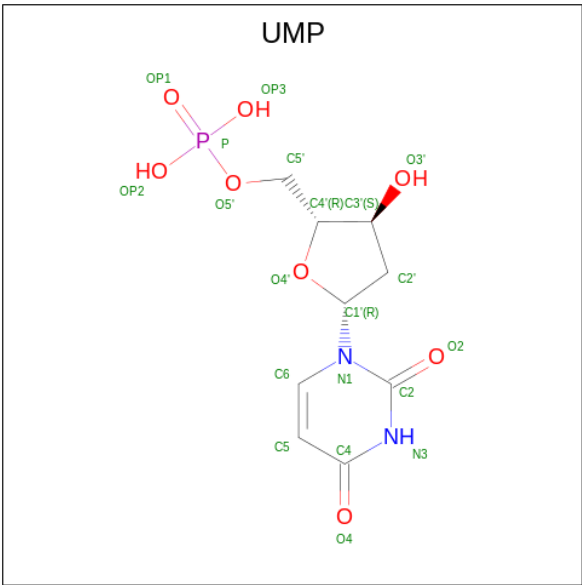
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	532	Total	C	H	N	O	S	256	0	0
			8832	2867	4399	733	809	24			
1	B	531	Total	C	H	N	O	S	254	0	0
			8813	2862	4389	730	807	25			

- Molecule 2 is METHIONINE (three-letter code: MET) (formula:  $C_5H_{11}NO_2S$ ).



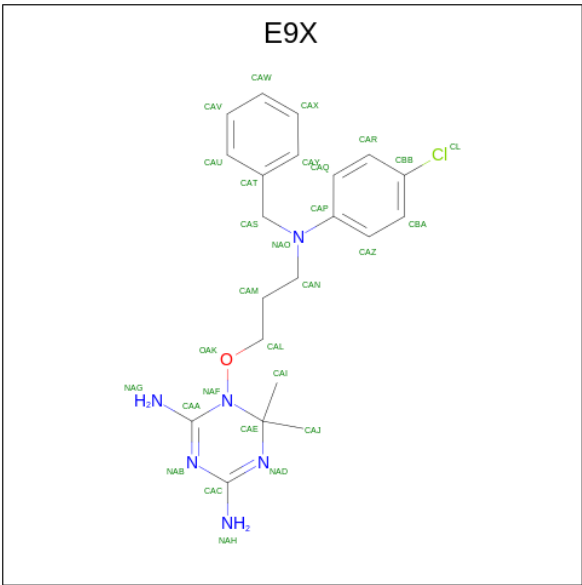
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			19	5	11	1	1	1		

- Molecule 3 is 2'-DEOXYURIDINE 5'-MONOPHOSPHATE (three-letter code: UMP) (formula:  $C_9H_{13}N_2O_8P$ ).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	P	1	0
			31	9	11	2	8	1		
3	B	1	Total	C	H	N	O	P	1	0
			31	9	11	2	8	1		

- Molecule 4 is 1-[3-[(4-chlorophenyl)-(phenylmethyl)amino]propoxy]-6,6-dimethyl-1,3,5-triazine-2,4-diamine (three-letter code: E9X) (formula: C<sub>21</sub>H<sub>27</sub>ClN<sub>6</sub>O) (labeled as "Ligand of Interest" by author).



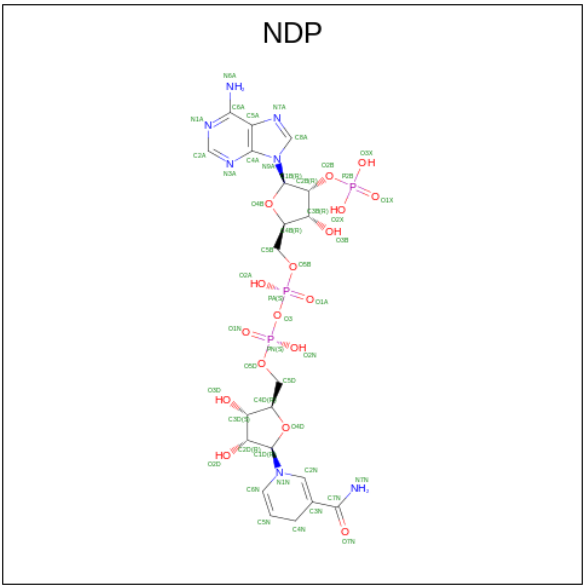
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	Cl	H	N	O	4	0
			56	21	1	27	6	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	B	1	Total	C	Cl	H	N	O	4	1
			56	21	1	27	6	1		

- Molecule 5 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C<sub>21</sub>H<sub>30</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	H	N	O	P		
5	A	1	72	21	24	7	17	3	3	0
5	B	1	72	21	24	7	17	3	3	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	47	Total O 47 47	0	0
6	B	49	Total O 49 49	0	0



- Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	56.31Å 154.49Å 163.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.41 – 2.59 46.41 – 2.59	Depositor EDS
% Data completeness (in resolution range)	90.1 (46.41-2.59) 84.3 (46.41-2.59)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 2.58Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.243 , 0.311 0.193 , 0.252	Depositor DCC
$R_{free}$ test set	2081 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.6	Xtriage
Anisotropy	0.448	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 39.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	18078	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: NDP, UMP, E9X

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	0.70	0/4537	0.91	2/6128 (0.0%)
1	B	0.72	0/4528	0.89	0/6116
All	All	0.71	0/9065	0.90	2/12244 (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	A	0	1
1	B	0	6
All	All	0	7

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	144	ASN	CB-CA-C	-6.47	97.46	110.40
1	A	525	TYR	CB-CA-C	5.04	120.48	110.40

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	137	ASP	Peptide
1	B	113	PRO	Peptide
1	B	117	LYS	Peptide
1	B	297	LYS	Peptide

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	B	302	LYS	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	4433	4399	4384	49	0
1	B	4424	4389	4374	69	0
2	A	8	11	8	1	0
3	A	20	11	11	0	0
3	B	20	11	11	2	0
4	A	29	27	0	2	0
4	B	29	27	0	0	0
5	A	48	24	26	2	0
5	B	48	24	26	4	0
6	A	47	0	0	1	1
6	B	49	0	0	0	1
All	All	9155	8923	8840	115	1

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 6.

The worst 5 of 115 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:605:ASP:C	2:A:701:MET:N	2.05	1.08
1:B:312:PHE:HA	1:B:565:ASN:HD21	1.21	1.00
1:A:312:PHE:HA	1:A:565:ASN:HD21	1.29	0.96
1:B:312:PHE:HA	1:B:565:ASN:ND2	1.94	0.81
1:B:310:ASN:ND2	1:B:310:ASN:H	1.83	0.75

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:846:HOH:O	6:B:946:HOH:O[1_455]	1.83	0.37

## 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	524/608 (86%)	492 (94%)	27 (5%)	5 (1%)	15	32
1	B	523/608 (86%)	470 (90%)	43 (8%)	10 (2%)	8	15
All	All	1047/1216 (86%)	962 (92%)	70 (7%)	15 (1%)	11	22

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	49	LYS
1	A	139	ASP
1	B	49	LYS
1	B	303	ASN
1	B	304	LYS

### 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	498/570 (87%)	470 (94%)	28 (6%)	21	42
1	B	497/570 (87%)	453 (91%)	44 (9%)	9	19
All	All	995/1140 (87%)	923 (93%)	72 (7%)	14	29

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	71	GLU
1	B	134	GLU
1	B	516	LEU
1	B	72	LYS
1	B	107	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	157	ASN
1	B	206	GLN
1	B	394	ASN
1	B	144	ASN
1	B	424	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

8 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
5	NDP	A	704	-	45,52,52	0.68	0	53,80,80	0.97	2 (3%)
3	UMP	A	702	-	18,21,21	1.06	1 (5%)	21,31,31	1.07	1 (4%)
2	MET	A	701	-	6,7,8	0.75	0	2,7,9	0.96	0
5	NDP	B	803	-	45,52,52	0.63	0	53,80,80	0.79	0
4	E9X	A	703	-	28,31,31	0.59	0	32,43,43	1.09	2 (6%)
3	UMP	B	801	-	18,21,21	1.04	1 (5%)	21,31,31	1.22	1 (4%)

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	NDP	A	704	-	-	4/30/77/77	0/5/5/5
3	UMP	A	702	-	-	2/7/22/22	0/2/2/2
2	MET	A	701	-	-	4/5/6/8	-
5	NDP	B	803	-	-	8/30/77/77	0/5/5/5
4	E9X	A	703	-	-	10/15/34/34	0/3/3/3
3	UMP	B	801	-	-	2/7/22/22	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	801	UMP	C4-N3	3.27	1.38	1.33
3	A	702	UMP	C4-N3	3.21	1.38	1.33

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	703	E9X	CAT-CAS-NAO	4.34	121.16	114.18
3	B	801	UMP	C5-C4-N3	-4.05	114.40	123.31
3	A	702	UMP	C5-C4-N3	-3.85	114.85	123.31
5	A	704	NDP	C5A-C6A-N6A	2.71	124.47	120.35
4	A	703	E9X	CAS-NAO-CAP	2.40	124.89	120.78

There are no chirality outliers.

5 of 30 torsion outliers are listed below:

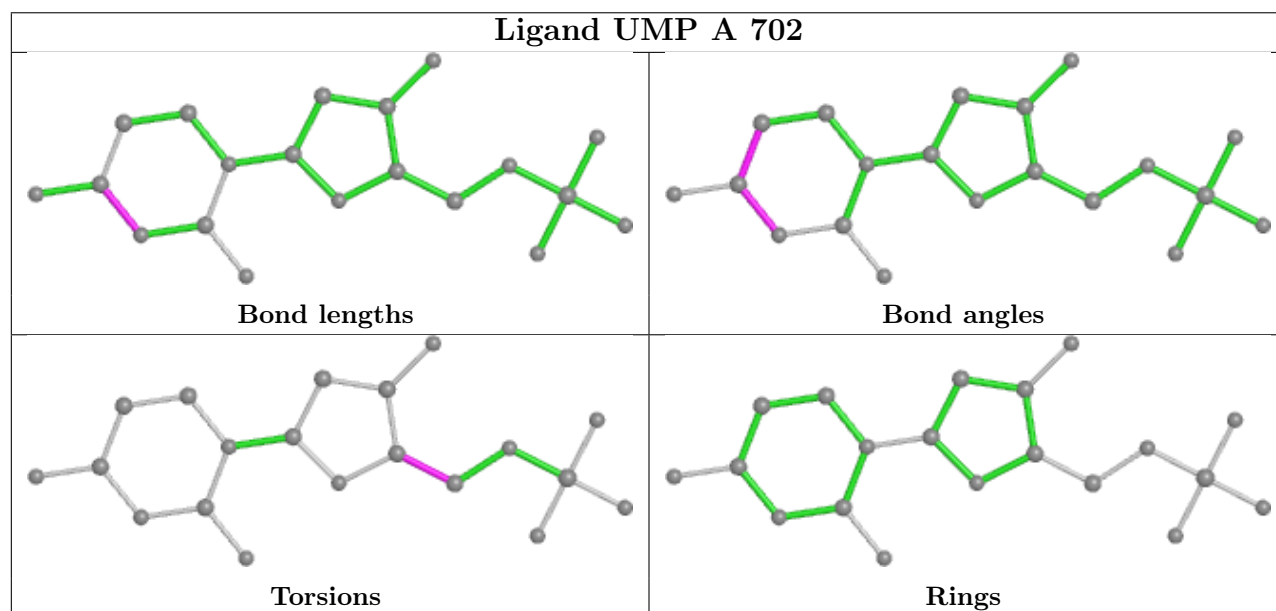
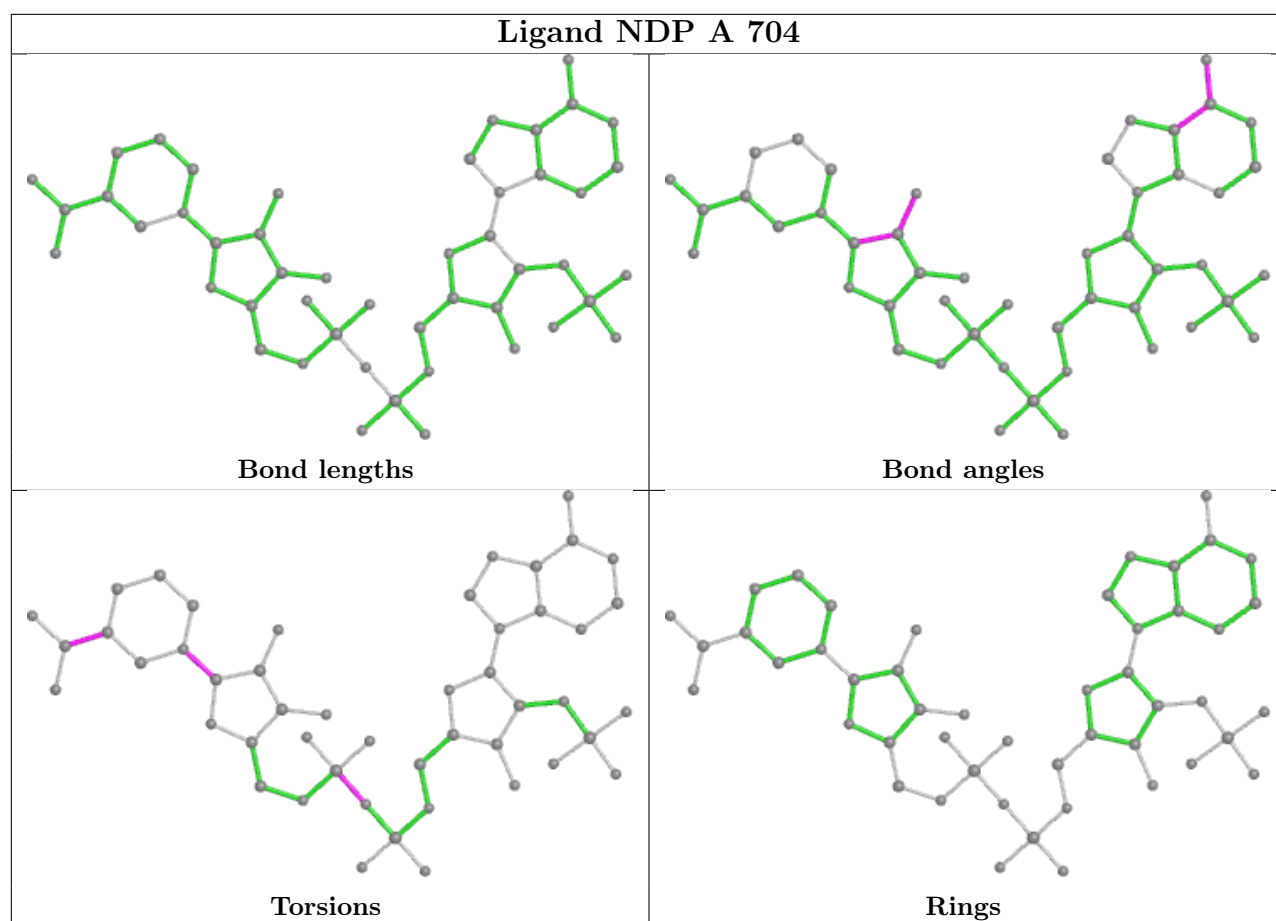
Mol	Chain	Res	Type	Atoms
2	A	701	MET	O-C-CA-CB
2	A	701	MET	N-CA-CB-CG
2	A	701	MET	C-CA-CB-CG
5	B	803	NDP	O4B-C4B-C5B-O5B
5	B	803	NDP	C3B-C4B-C5B-O5B

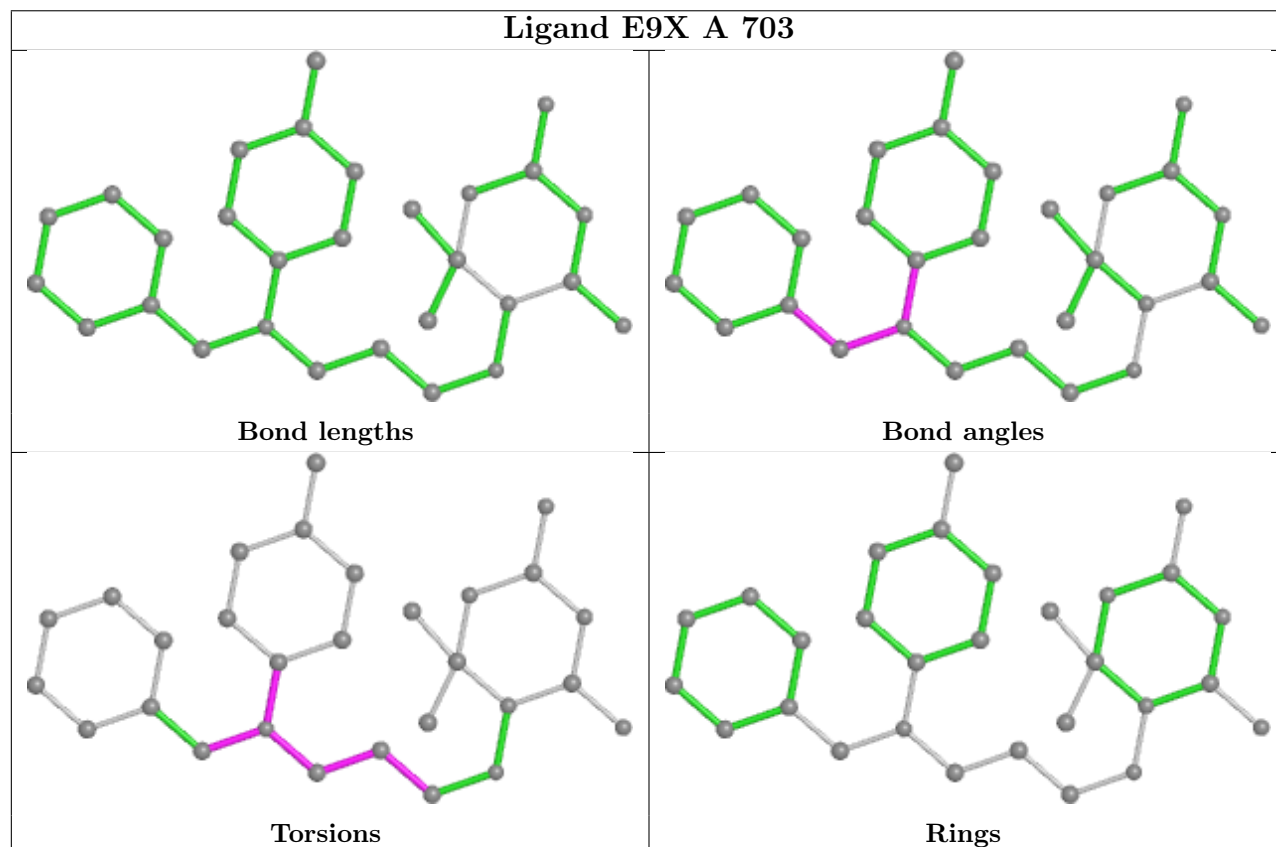
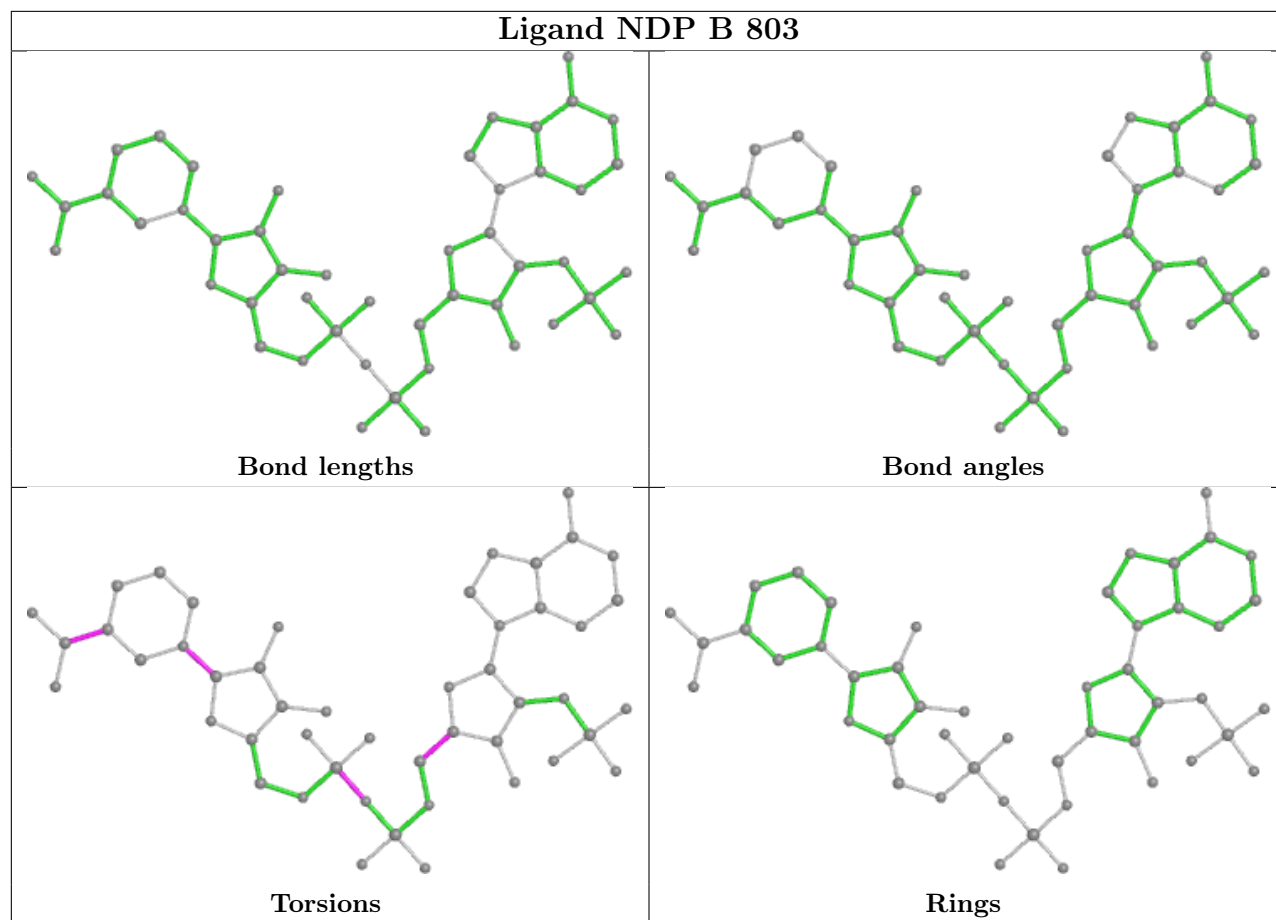
There are no ring outliers.

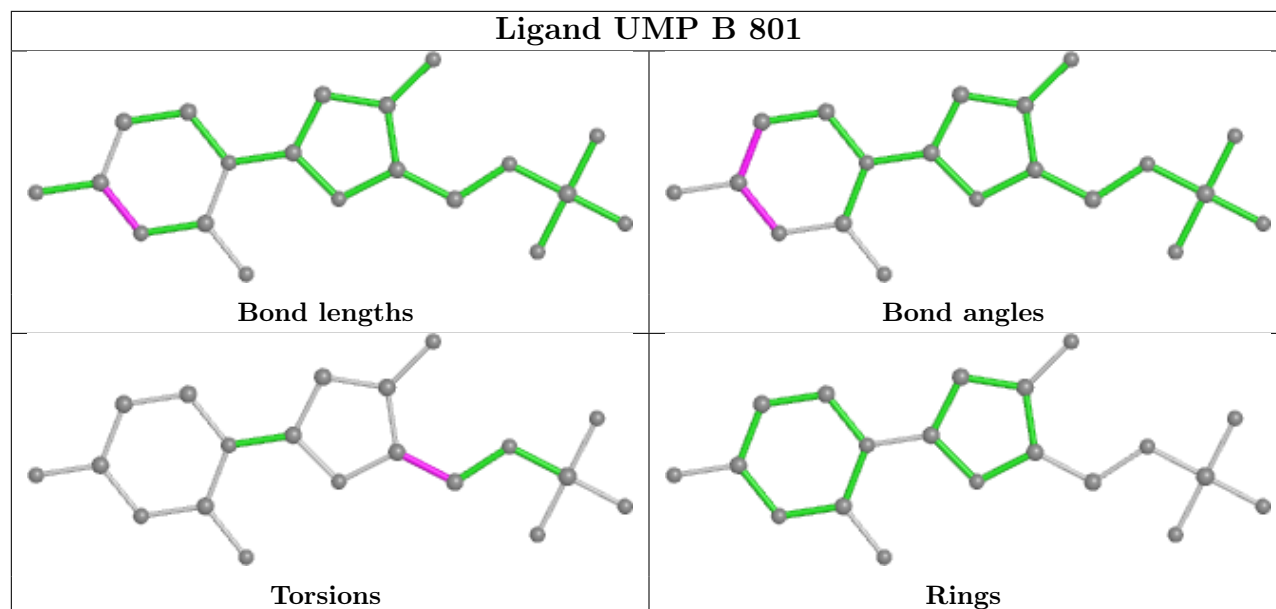
5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	704	NDP	2	0
2	A	701	MET	1	0
5	B	803	NDP	4	0
4	A	703	E9X	2	0
3	B	801	UMP	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

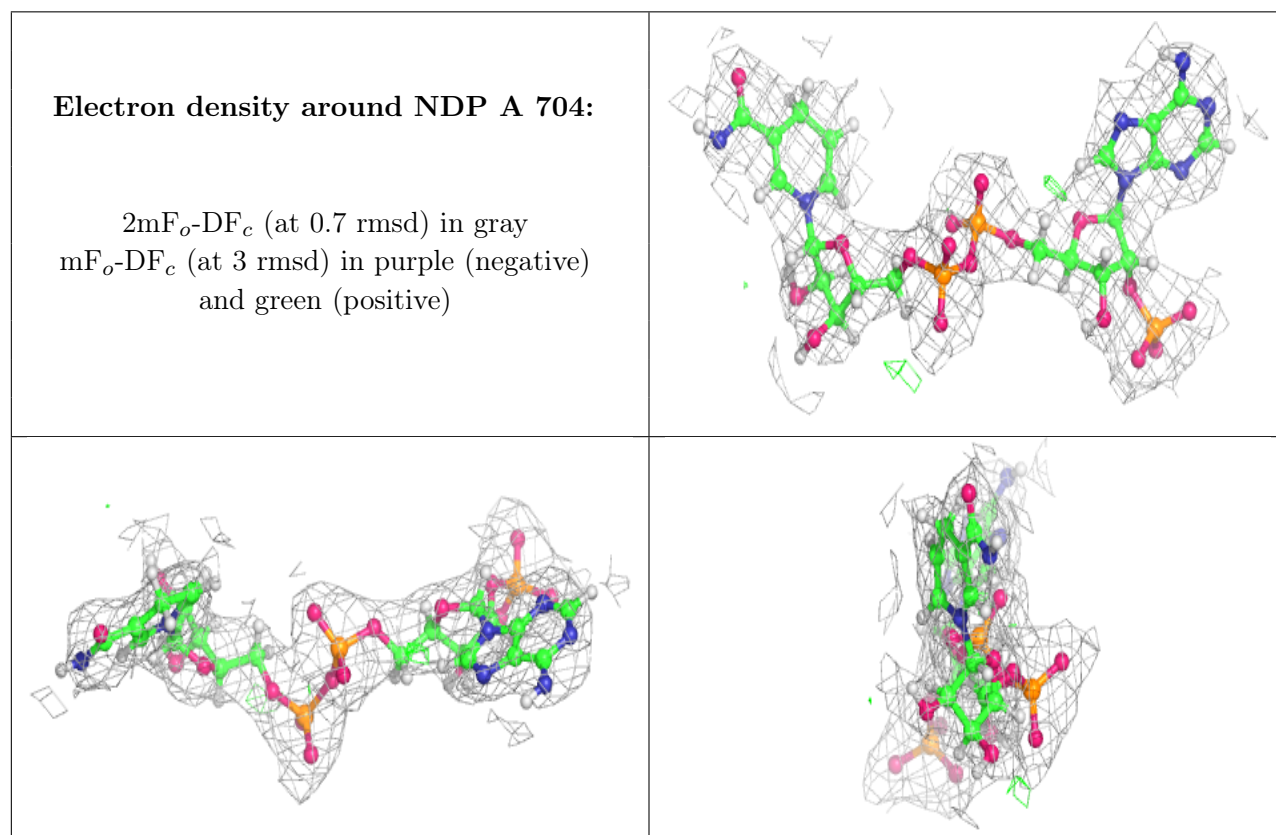
### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

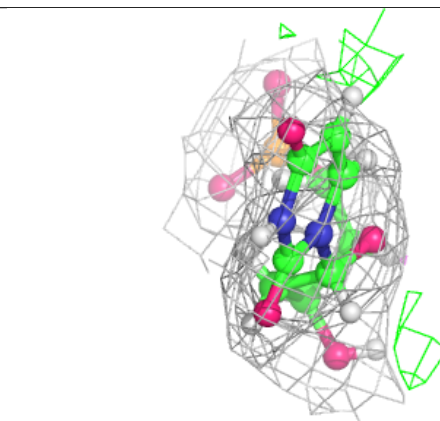
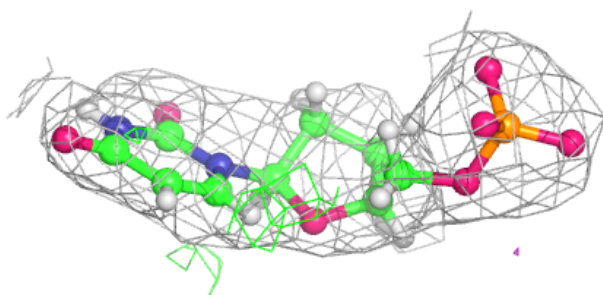
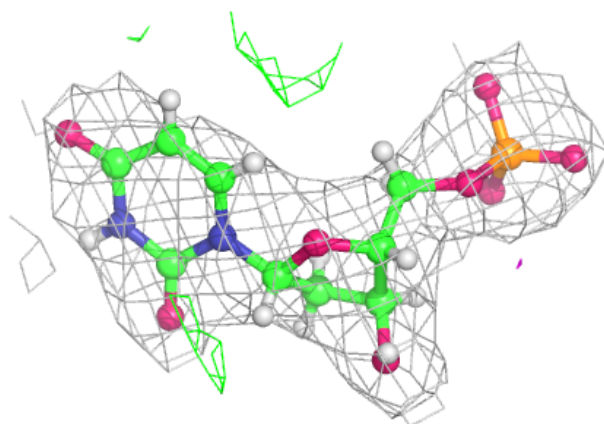
Unable to reproduce the depositors R factor - this section is therefore empty.

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.

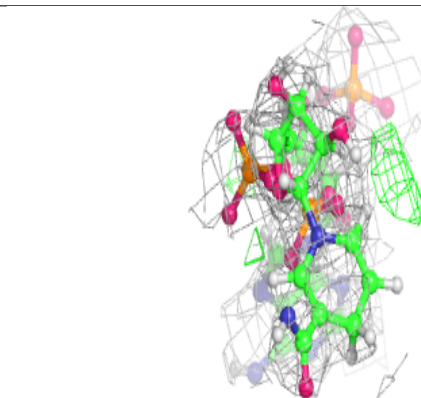
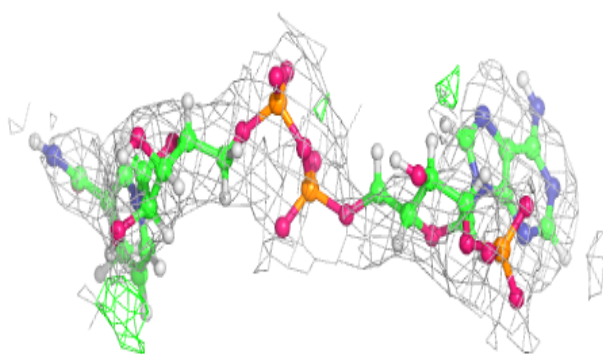
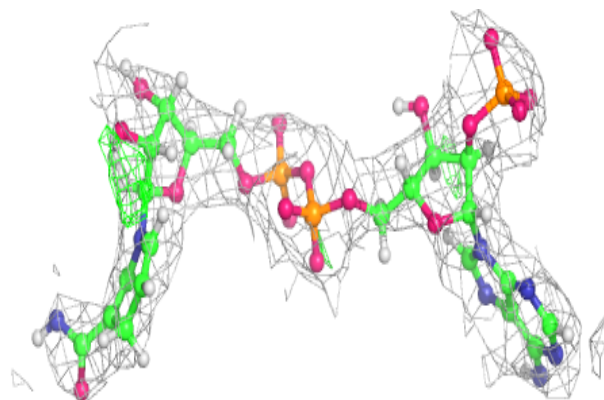


**Electron density around UMP A 702:**

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

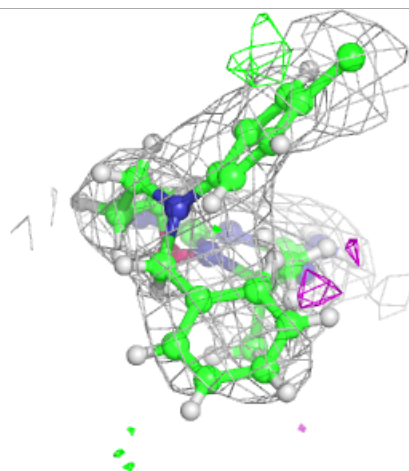
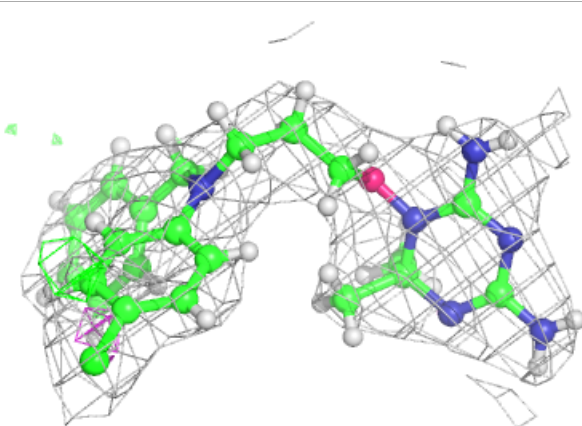
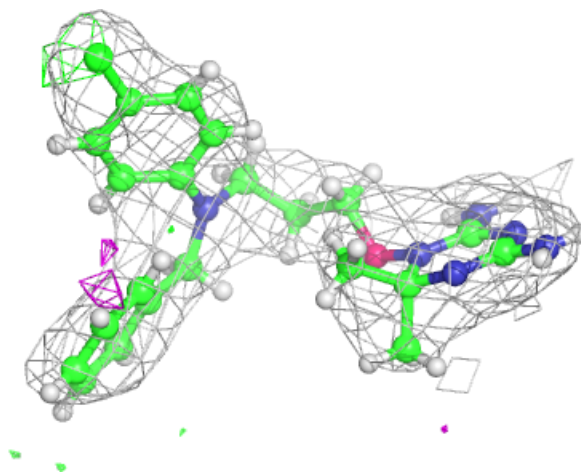
**Electron density around NDP B 803:**

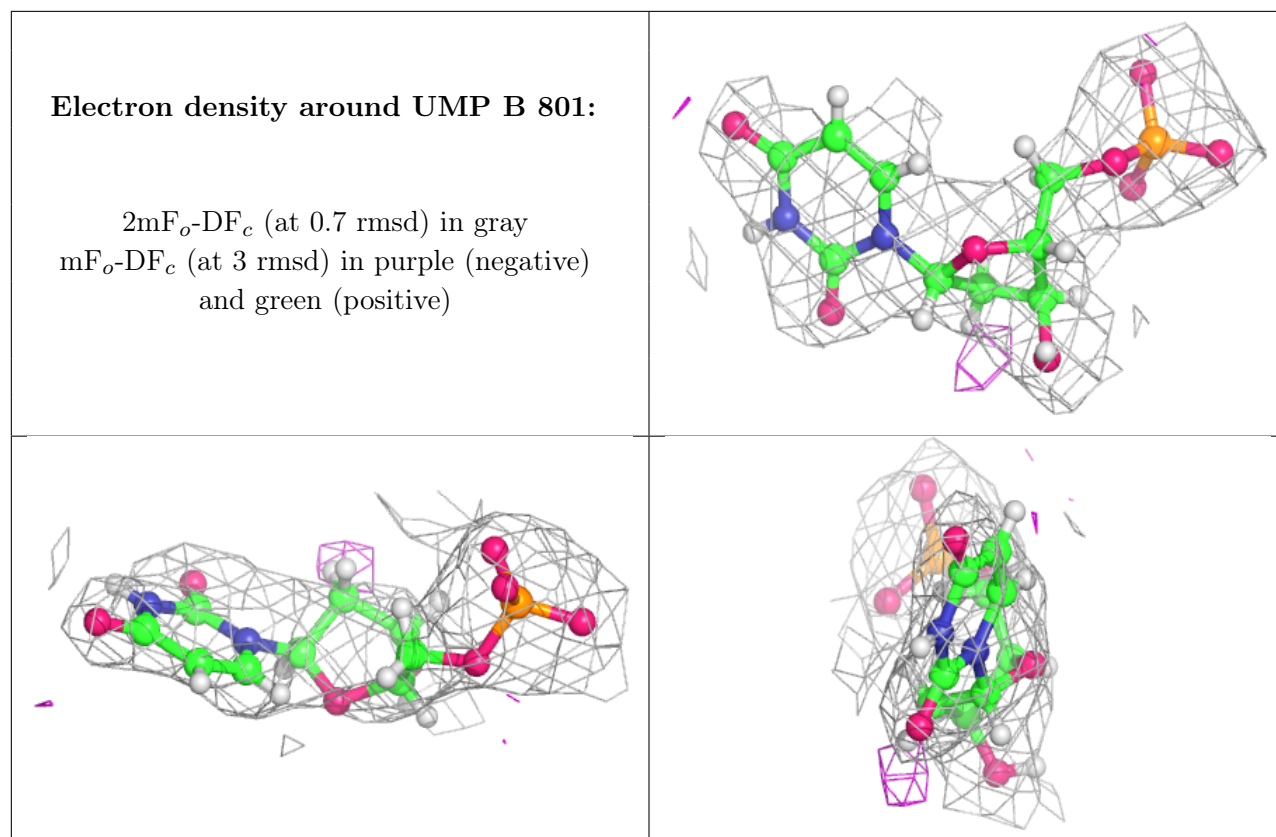
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around E9X A 703:**

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





## 6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.