wwPDB X-ray Structure Validation Summary Report (i)

Nov 20, 2023 – 02:06 PM JST

PDB ID	:	7CTZ
Title	:	Wild-type plasmodium falciparum dihydrofolate reductase-thymidylate synthase (PfDHFR-TS) complexed with fragment 148, NADPH, and dUMP
Authors	:	Vitsupakorn, D.
Deposited on	:	2020-08-20
Resolution	:	2.65 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

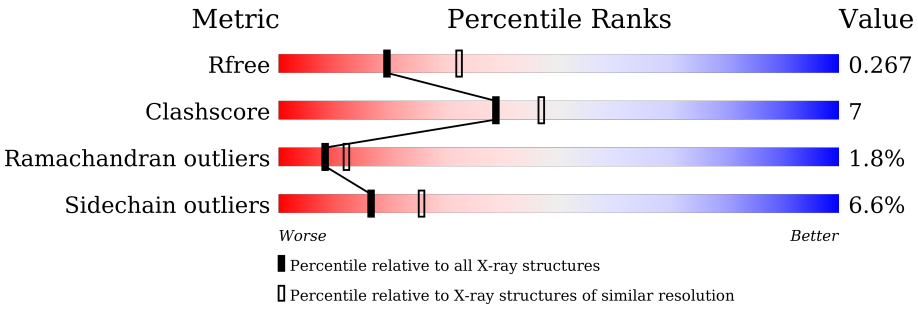
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	608	67% 19% • 12%
1	B	608	66% 20% • 12%

2 Entry composition

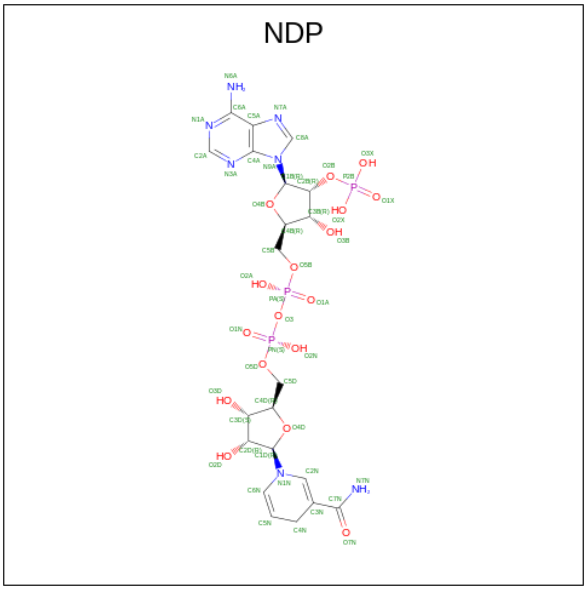
There are 5 unique types of molecules in this entry. The entry contains 17985 atoms, of which 8888 are hydrogens and 0 are deuteriums.

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.

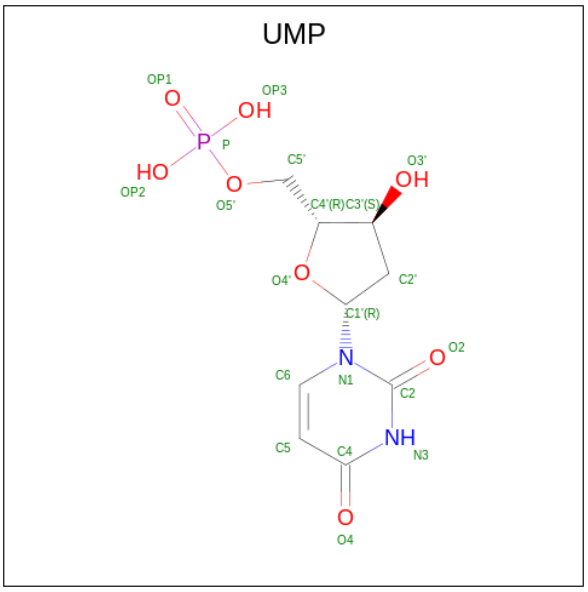
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	534	Total	C	H	N	O	S	253	0	0
			8858	2876	4411	731	813	27			
1	B	532	Total	C	H	N	O	S	257	0	0
			8817	2863	4389	730	809	26			

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: $C_{21}H_{30}N_7O_{17}P_3$).



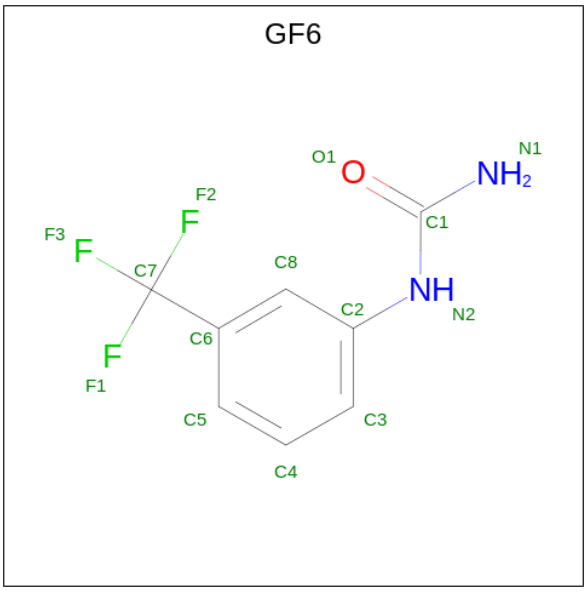
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	P	7	0
			74	21	26	7	17	3		
2	B	1	Total	C	H	N	O	P	7	0
			74	21	26	7	17	3		

- 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-(trifluoromethyl)phenyl]urea (three-letter code: GF6) (formula: $C_8H_7F_3N_2O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	F	H	N	O	4	1
			42	16	6	14	4	2		

- Molecule 5 is water.

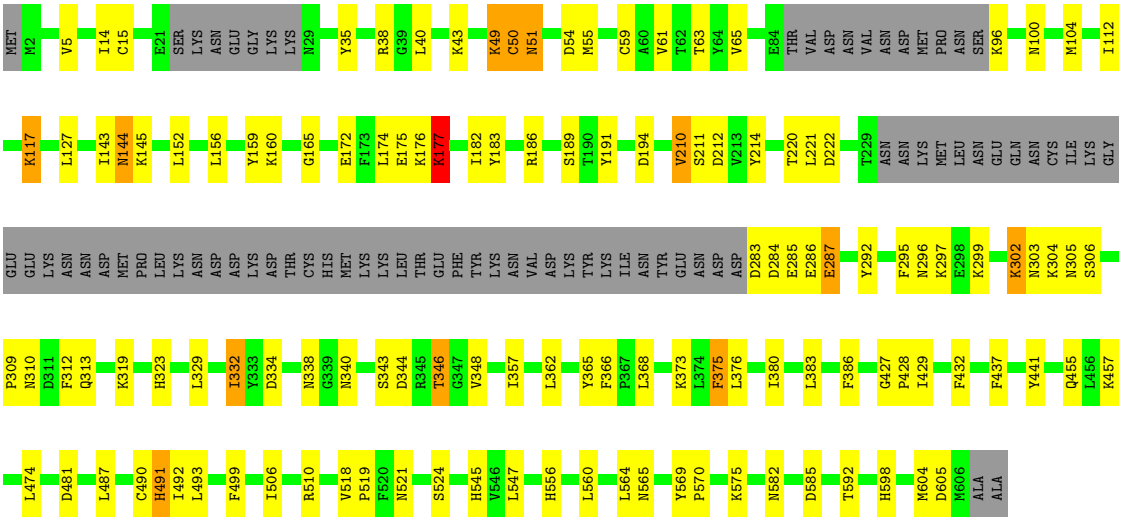
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	24	Total O 24 24	0	0
5	B	34	Total O 34 34	0	0

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. 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. 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.

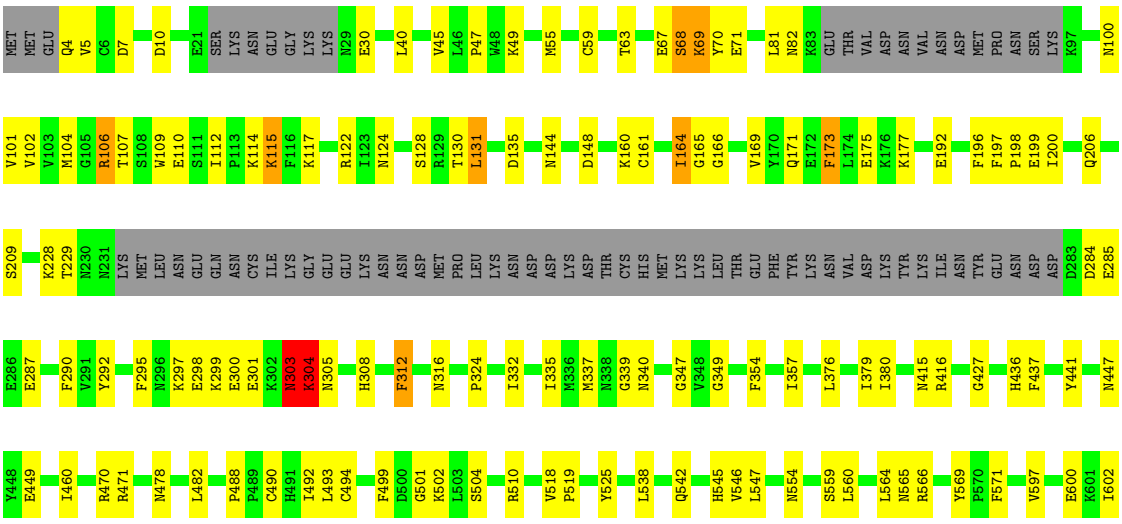
• Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

Chain A:



• Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

Chain B:




D605
M606
ALA
ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.64Å 155.64Å 164.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.83 – 2.65 24.83 – 2.65	Depositor EDS
% Data completeness (in resolution range)	97.8 (24.83-2.65) 92.3 (24.83-2.65)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.43 (at 2.64Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.255 , 0.325 0.203 , 0.267	Depositor DCC
R_{free} test set	2095 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	25.7	Xtriage
Anisotropy	0.142	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 45.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	17985	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GF6, UMP, NDP

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	0.75	0/4551	0.95	0/6144
1	B	0.75	0/4531	0.94	4/6118 (0.1%)
All	All	0.75	0/9082	0.95	4/12262 (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	2
1	B	0	4
All	All	0	6

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	470	ARG	NE-CZ-NH2	-6.41	117.10	120.30
1	B	415	ASN	CB-CA-C	5.82	122.04	110.40
1	B	416	ARG	NE-CZ-NH2	5.65	123.12	120.30
1	B	566	ARG	NE-CZ-NH2	-5.64	117.48	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	302	LYS	Peptide
1	A	306	SER	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	164	ILE	Peptide
1	B	165	GLY	Peptide
1	B	297	LYS	Peptide

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	4447	4411	4396	69	0
1	B	4428	4389	4373	70	0
2	A	48	26	26	2	0
2	B	48	26	26	2	0
3	A	20	11	11	1	0
3	B	20	11	11	1	0
4	A	28	14	0	1	0
5	A	24	0	0	0	0
5	B	34	0	0	1	0
All	All	9097	8888	8843	134	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 7.

The worst 5 of 134 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:312:PHE:HA	1:B:565:ASN:HD21	1.32	0.92
1:A:376:LEU:O	1:A:380:ILE:HG13	1.76	0.85
1:B:171:GLN:NE2	1:B:175:GLU:OE2	2.11	0.83
1:B:605:ASP:O	1:B:606:MET:N	2.12	0.82
1:A:334:ASP:OD1	1:A:338:ASN:ND2	2.22	0.72

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	526/608 (86%)	478 (91%)	39 (7%)	9 (2%)	9 13
1	B	523/608 (86%)	470 (90%)	43 (8%)	10 (2%)	8 11
All	All	1049/1216 (86%)	948 (90%)	82 (8%)	19 (2%)	8 12

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	304	LYS
1	A	310	ASN
1	B	68	SER
1	B	82	ASN
1	B	199	GLU

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	500/570 (88%)	470 (94%)	30 (6%)	19 30
1	B	498/570 (87%)	462 (93%)	36 (7%)	14 22
All	All	998/1140 (88%)	932 (93%)	66 (7%)	16 25

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	303	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	312	PHE
1	B	606	MET
1	A	487	LEU
1	A	457	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	424	ASN
1	B	394	ASN
1	A	424	ASN
1	B	303	ASN
1	A	394	ASN

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	$\#\|Z\|>2$
2	NDP	B	702	-	45,52,52	1.93	12 (26%)	53,80,80	1.81	12 (22%)
3	UMP	B	703	-	21,21,21	0.57	0	31,31,31	0.76	0
4	GF6	A	703[A]	-	14,14,14	0.50	0	20,20,20	1.18	1 (5%)
2	NDP	A	701	-	45,52,52	2.13	13 (28%)	53,80,80	1.36	8 (15%)
4	GF6	A	703[B]	-	14,14,14	0.29	0	20,20,20	0.79	0
3	UMP	A	702	-	21,21,21	0.56	0	31,31,31	0.72	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	NDP	B	702	-	-	15/30/77/77	0/5/5/5
3	UMP	B	703	-	-	4/10/22/22	0/2/2/2
4	GF6	A	703[A]	-	-	2/10/10/10	0/1/1/1
2	NDP	A	701	-	-	13/30/77/77	0/5/5/5
4	GF6	A	703[B]	-	-	4/10/10/10	0/1/1/1
3	UMP	A	702	-	-	1/10/22/22	0/2/2/2

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	702	NDP	O4B-C1B	5.77	1.49	1.41
2	A	701	NDP	C4N-C3N	-5.72	1.38	1.49
2	A	701	NDP	C2A-N3A	5.48	1.40	1.32
2	A	701	NDP	O4B-C1B	4.39	1.47	1.41
2	A	701	NDP	C4A-N3A	-4.15	1.29	1.35

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	702	NDP	N3A-C2A-N1A	-7.37	117.15	128.68
2	B	702	NDP	C1B-N9A-C4A	-4.24	119.19	126.64
2	B	702	NDP	C3B-C2B-C1B	-3.87	95.61	102.89
2	A	701	NDP	N3A-C2A-N1A	-3.53	123.16	128.68
2	A	701	NDP	C4A-C5A-N7A	-3.43	105.83	109.40

There are no chirality outliers.

5 of 39 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	NDP	C5B-O5B-PA-O1A
2	A	701	NDP	C5B-O5B-PA-O2A
2	A	701	NDP	C3B-C4B-C5B-O5B
2	A	701	NDP	C5D-O5D-PN-O1N
2	B	702	NDP	C5B-O5B-PA-O1A

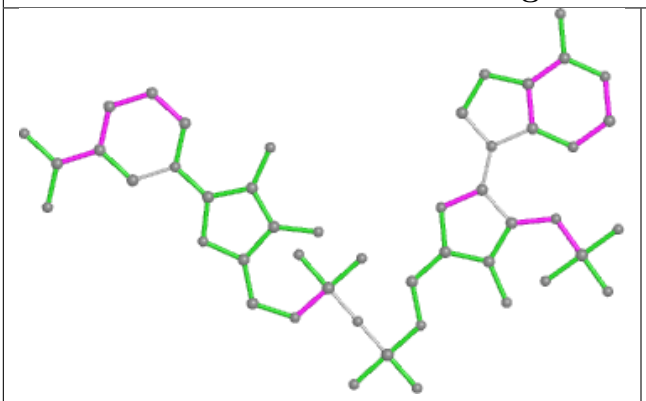
There are no ring outliers.

5 monomers are involved in 7 short contacts:

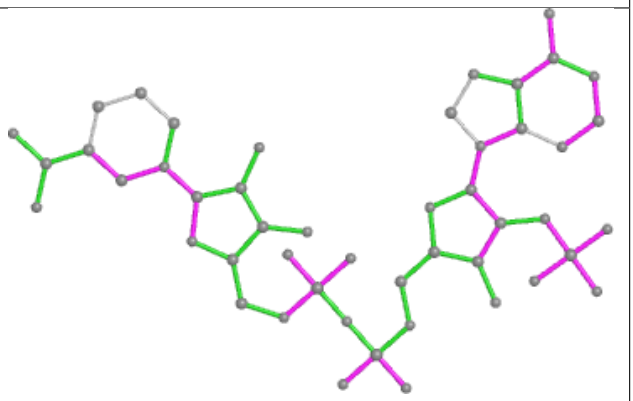
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	702	NDP	2	0
3	B	703	UMP	1	0
4	A	703[A]	GF6	1	0
2	A	701	NDP	2	0
3	A	702	UMP	1	0

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 then 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.

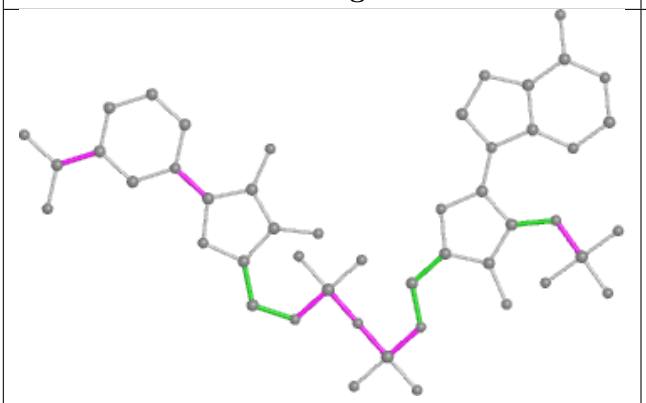
Ligand NDP B 702



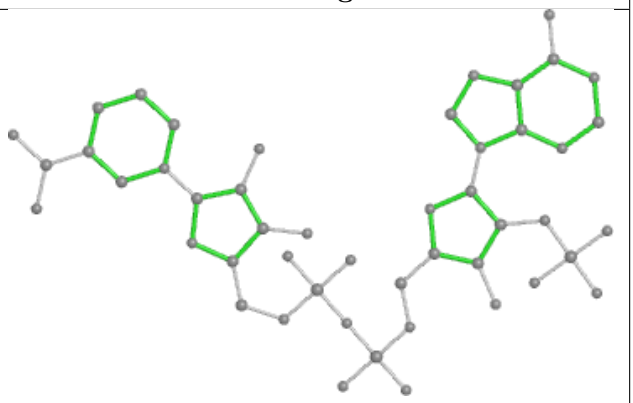
Bond lengths



Bond angles

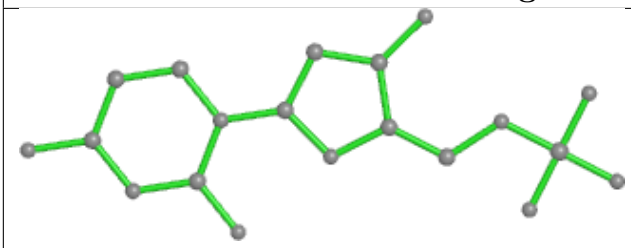


Torsions

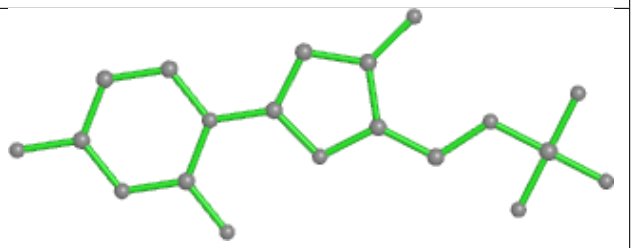


Rings

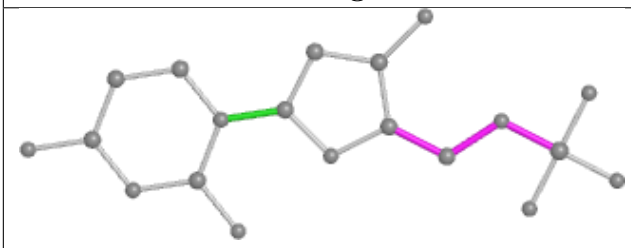
Ligand UMP B 703



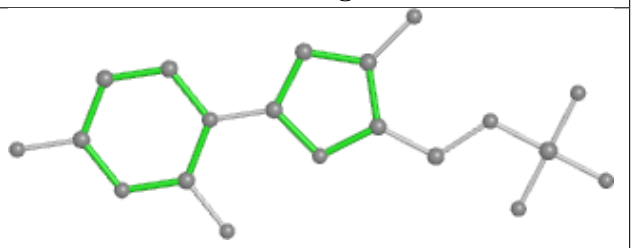
Bond lengths



Bond angles

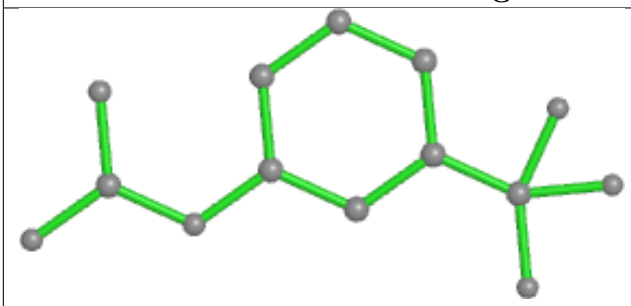


Torsions

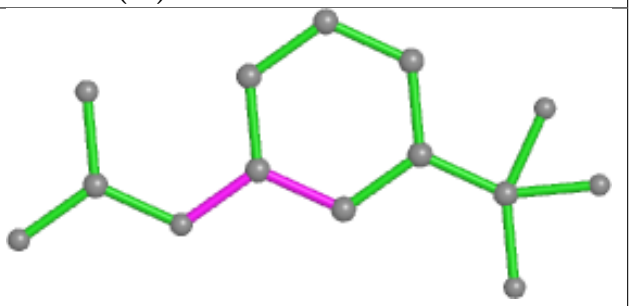


Rings

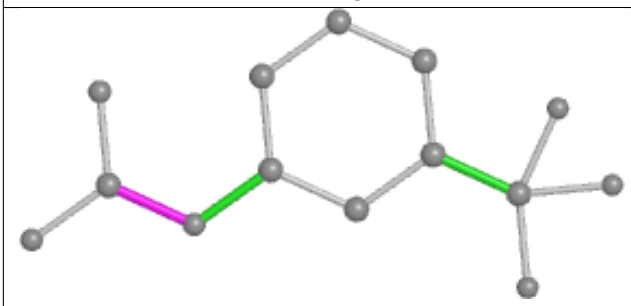
Ligand GF6 A 703 (A)



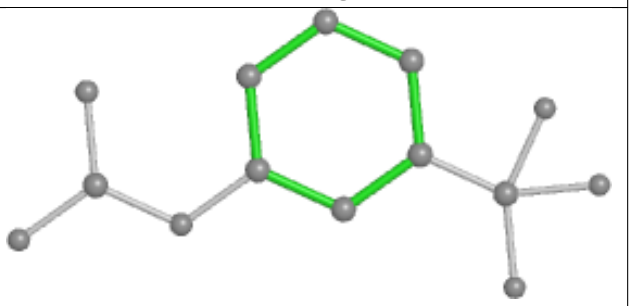
Bond lengths



Bond angles

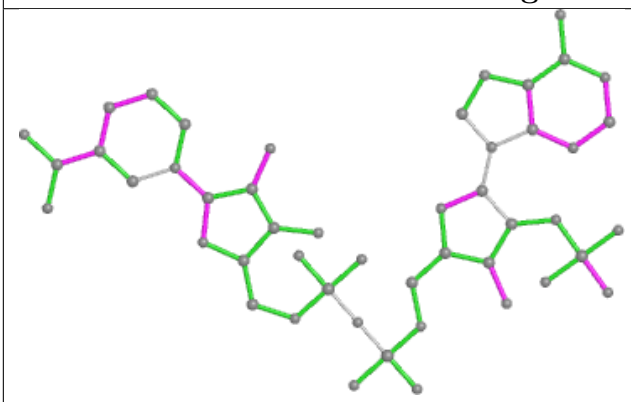


Torsions

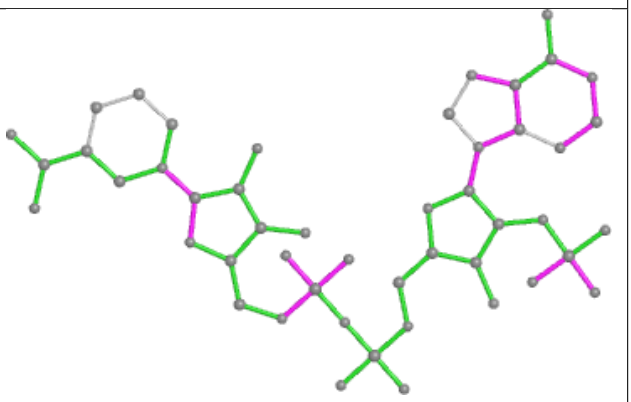


Rings

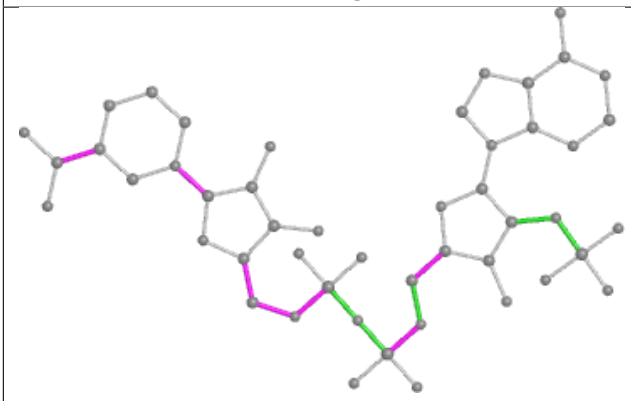
Ligand NDP A 701



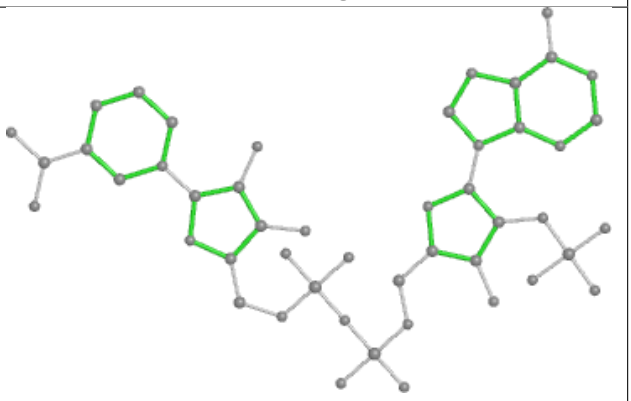
Bond lengths



Bond angles

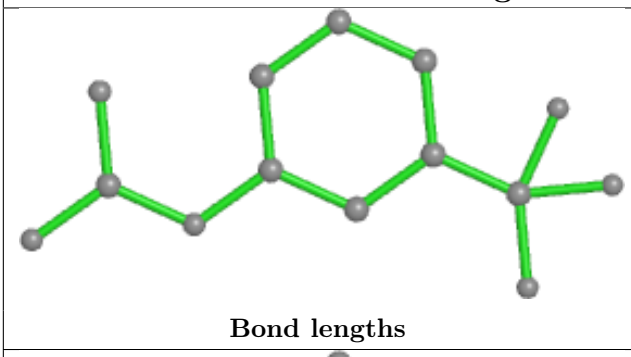


Torsions

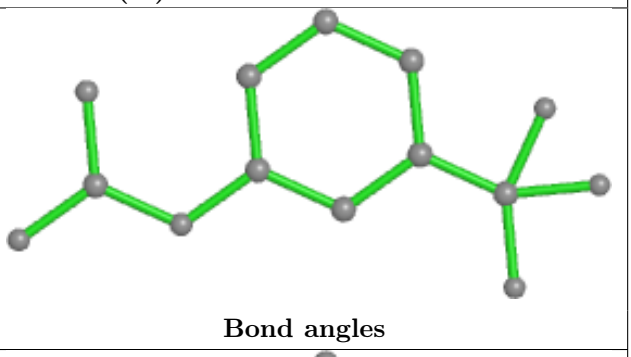


Rings

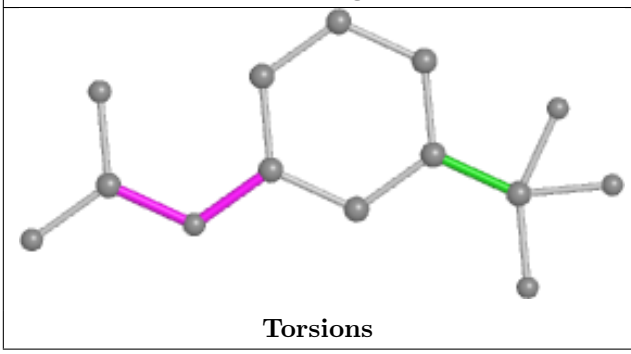
Ligand GF6 A 703 (B)



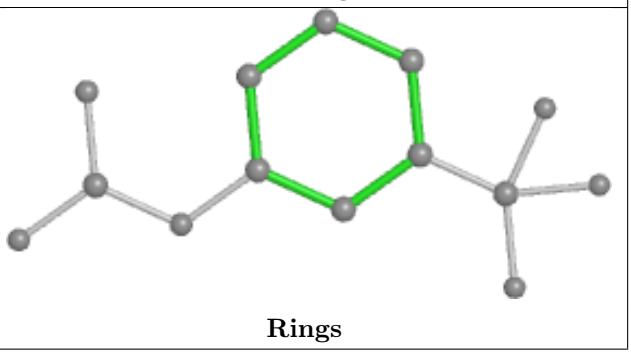
Bond lengths



Bond angles

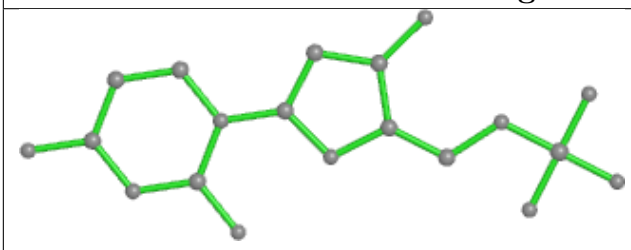


Torsions

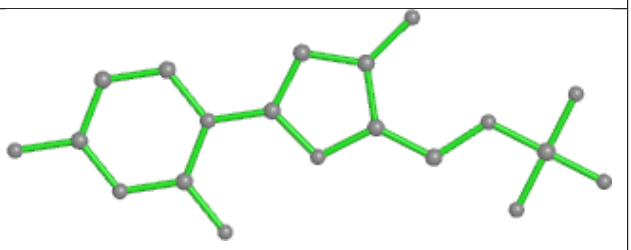


Rings

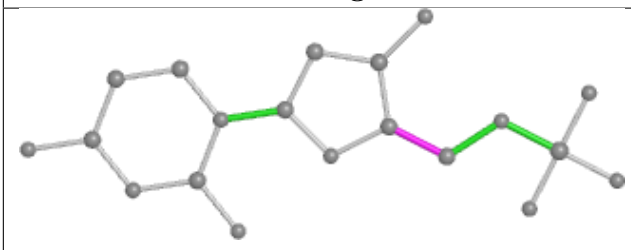
Ligand UMP A 702



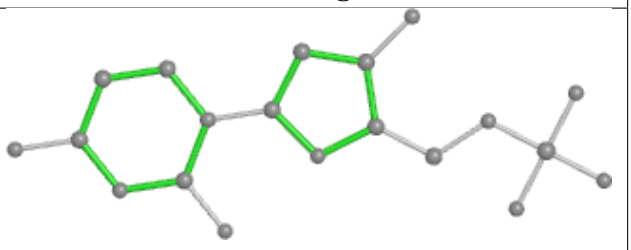
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	605:ASP	C	606:MET	N	3.30

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.

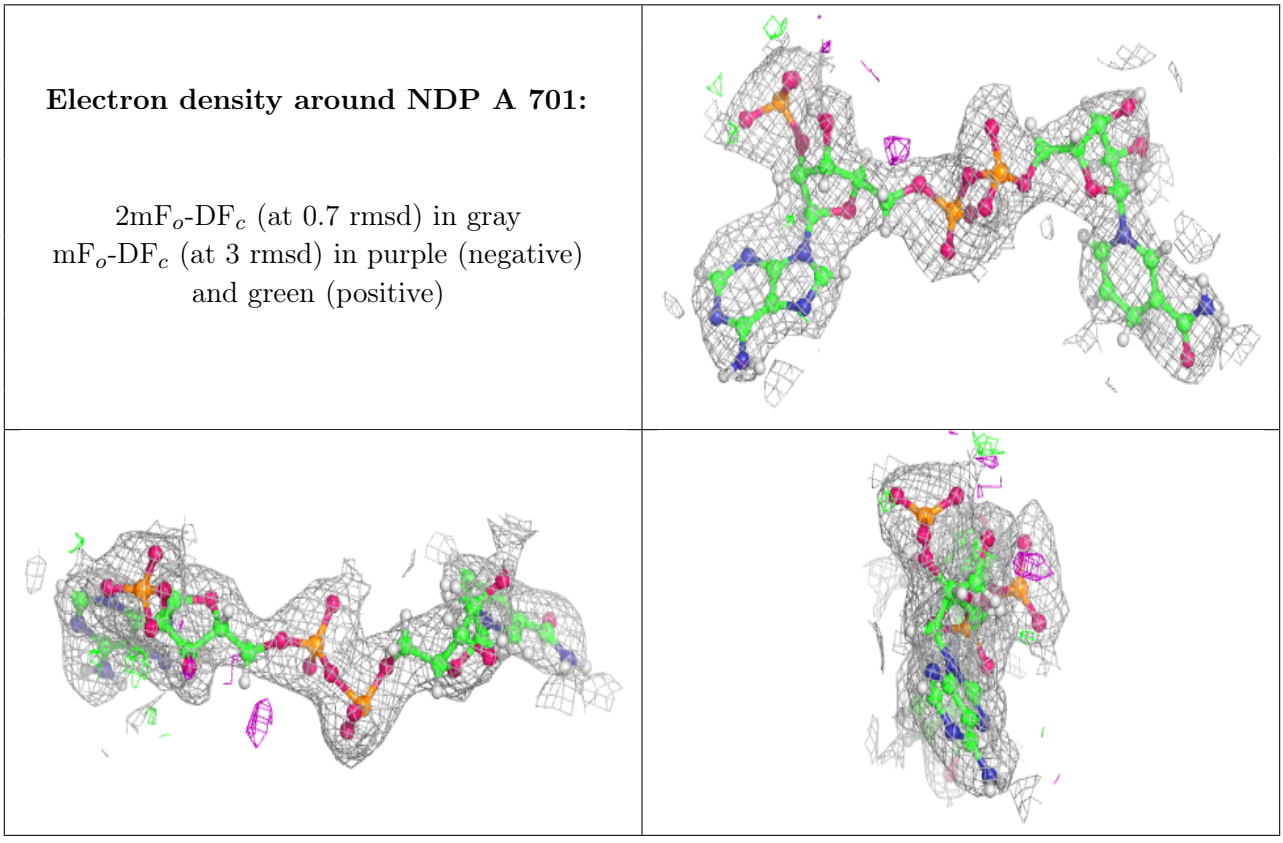
6.4 Ligands

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.

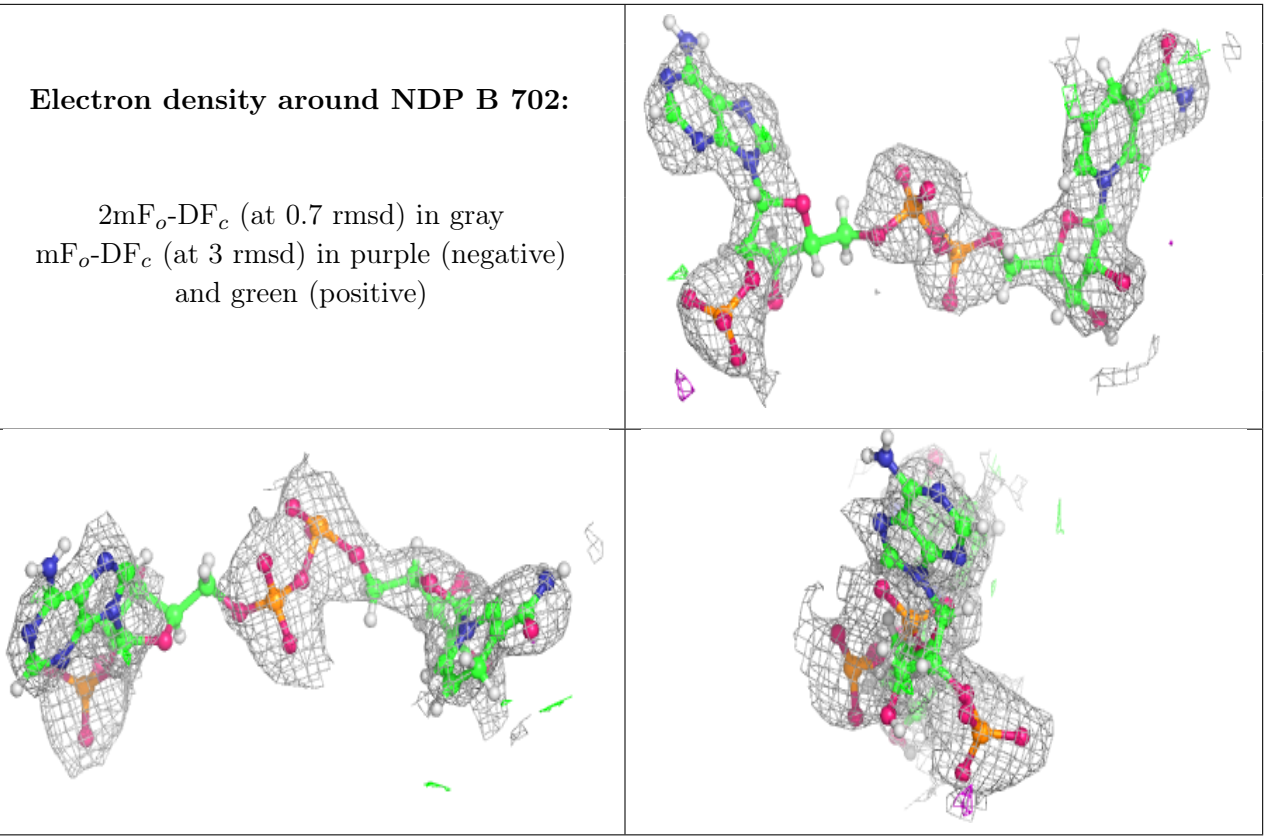
Electron density around NDP A 701:

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



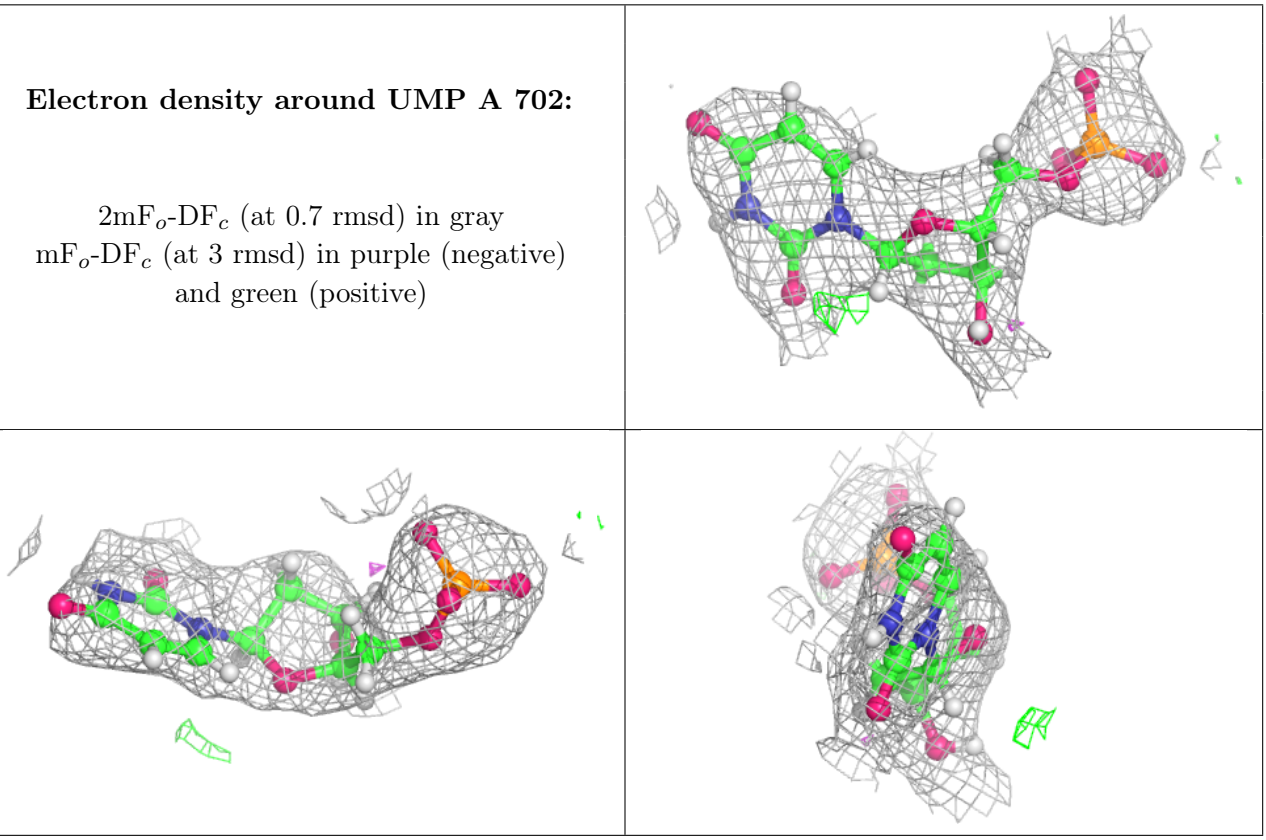
Electron density around NDP B 702:

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



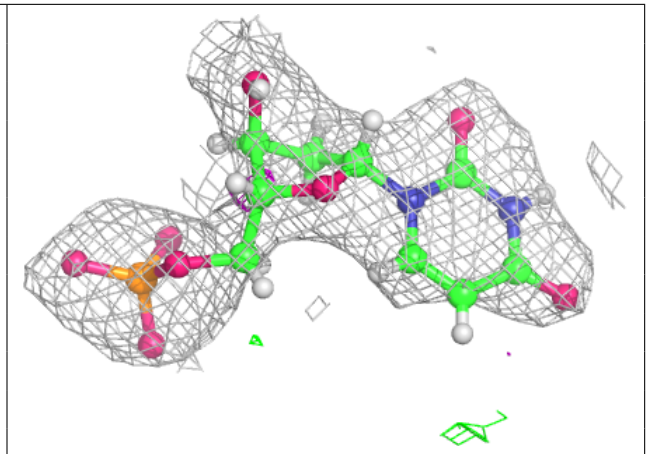
Electron density around UMP A 702:

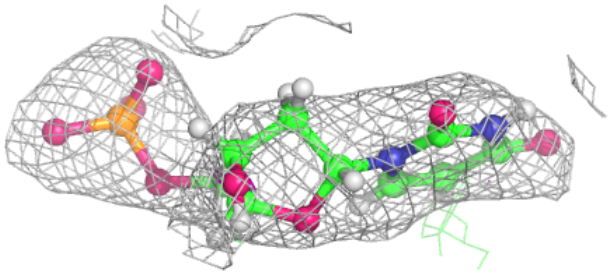
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

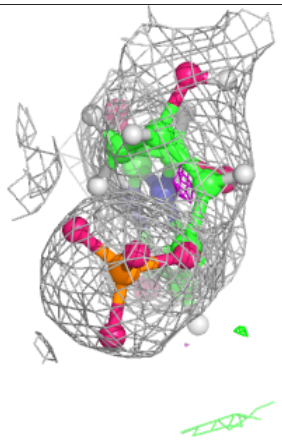


Electron density around UMP B 703:

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

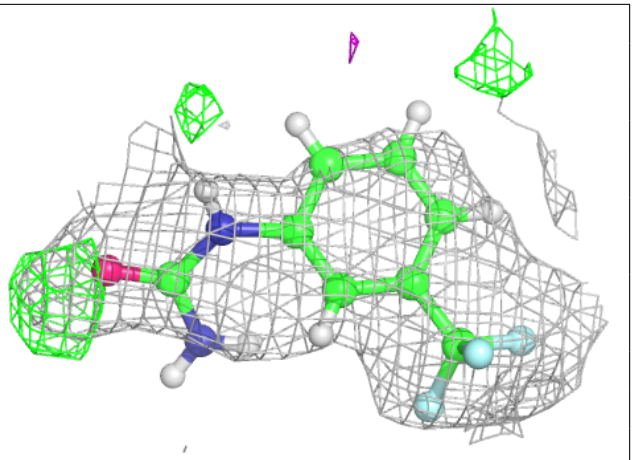


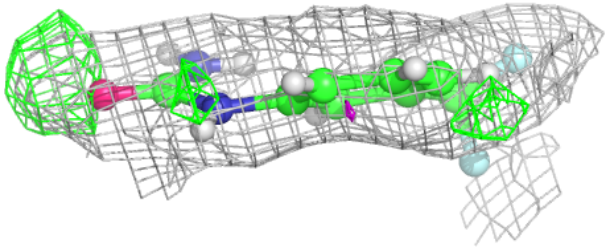


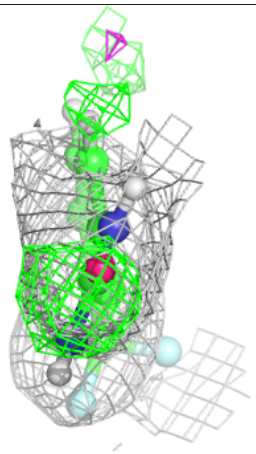


Electron density around GF6 A 703 (A):

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

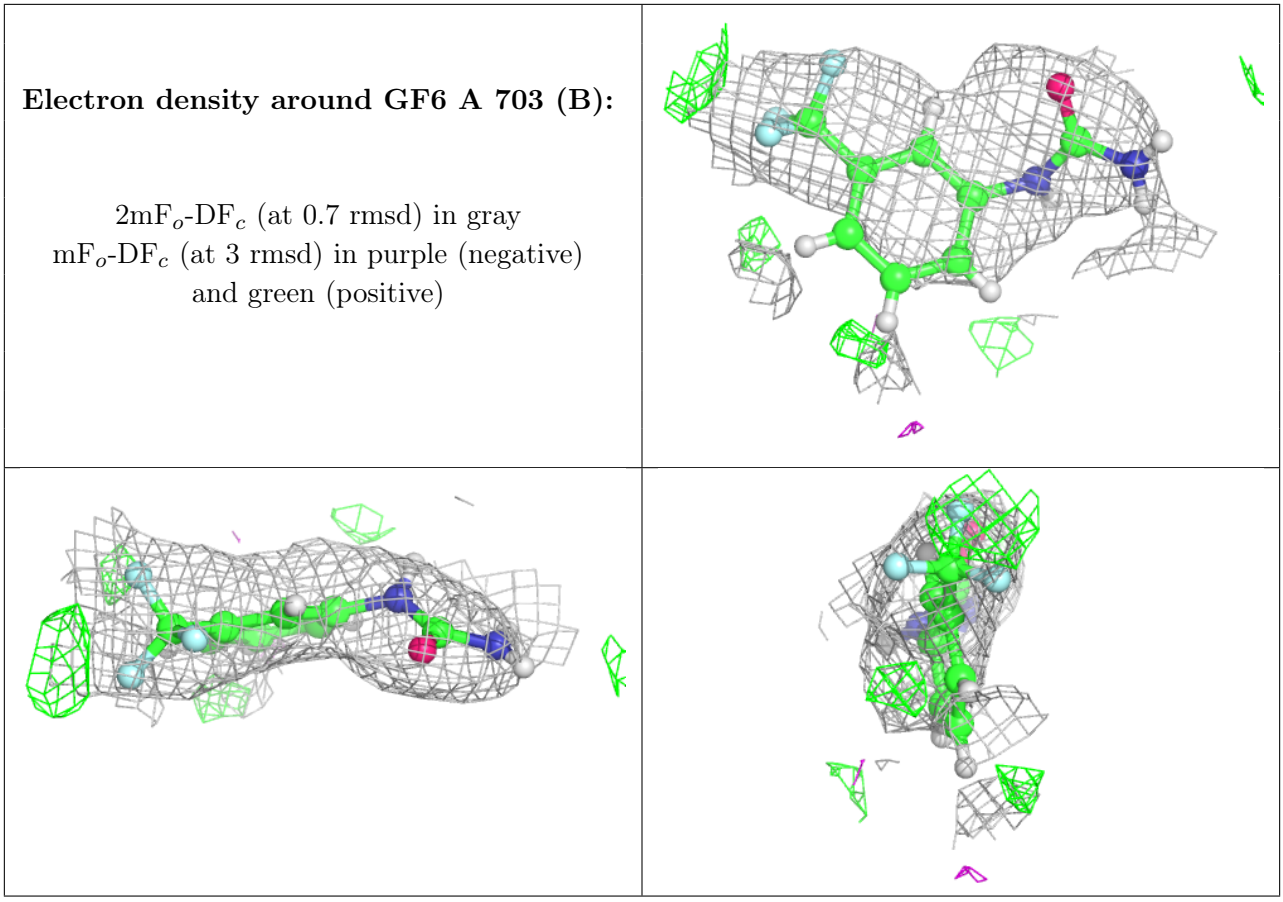






Electron density around GF6 A 703 (B):

$2mF_o$-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.