wwPDB X-ray Structure Validation Summary Report (i)

Mar 9, 2024 – 05:14 PM EST

PDB ID	:	3UM6
Title	:	Double mutant (A16V+S108T) Plasmodium falciparum DHFR-TS (T9/94) complexed with cycloguanil, NADPH and dUMP
Authors	:	Vanichtanankul, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Yuthavong, Y.
Deposited on	:	2011-11-12
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:

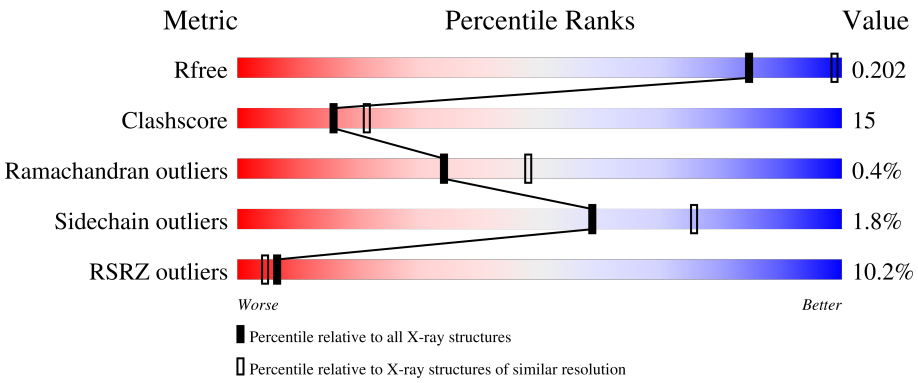
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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)
RSRZ outliers	127900	1318 (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% 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	608	6% 65% 23% • 10%
1	B	608	12% 62% 26% • 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	1CY	B	709	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9587 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.

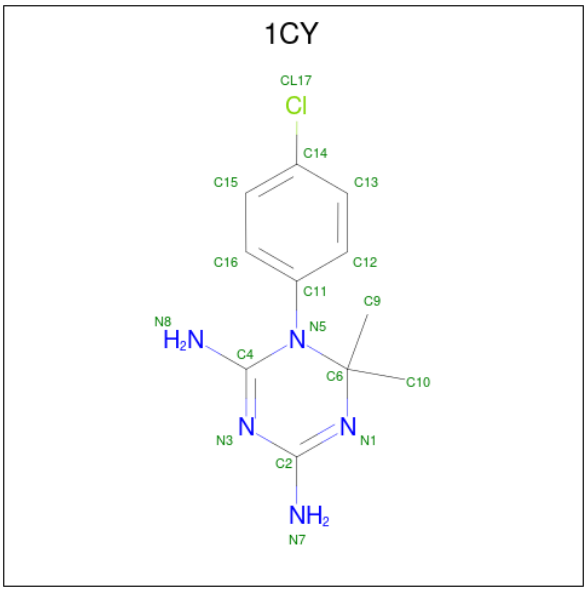
- Molecule 1 is a protein called Bifunctional dihydrofolate reductase-thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	546	4539	2930	749	832	28	0	0	0
1	B	544	4507	2910	745	826	26	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

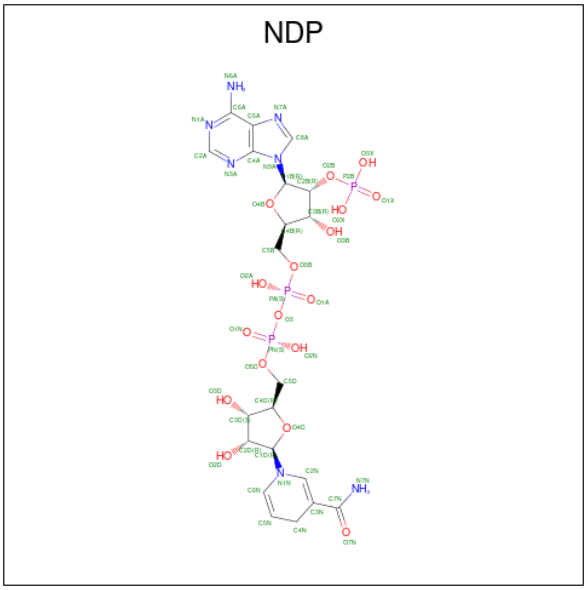
Chain	Residue	Modelled	Actual	Comment	Reference
A	16	VAL	ALA	engineered mutation	UNP A7UD81
A	108	THR	SER	engineered mutation	UNP A7UD81
B	16	VAL	ALA	engineered mutation	UNP A7UD81
B	108	THR	SER	engineered mutation	UNP A7UD81

- Molecule 2 is 1-(4-chlorophenyl)-6,6-dimethyl-1,6-dihydro-1,3,5-triazine-2,4-diamine (three-letter code: 1CY) (formula: $C_{11}H_{14}ClN_5$).



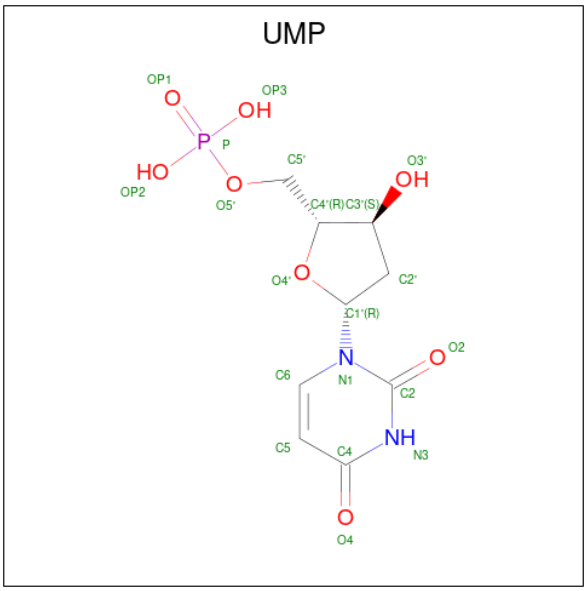
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	Cl	N		
2	A	1	17	11	1	5	0	0
			Total	C	Cl	N		
2	B	1	17	11	1	5	0	0

- Molecule 3 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
			Total	C	N	O	P		
3	A	1	48	21	7	17	3	0	0
			Total	C	N	O	P		
3	B	1	48	21	7	17	3	0	0

- Molecule 4 is 2'-DEOXYURIDINE 5'-MONOPHOSPHATE (three-letter code: UMP) (formula: $C_9H_{13}N_2O_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	20	9	2	8	1	0	0
			Total	C	N	O	P		
4	B	1	20	9	2	8	1	0	0

- Molecule 5 is water.

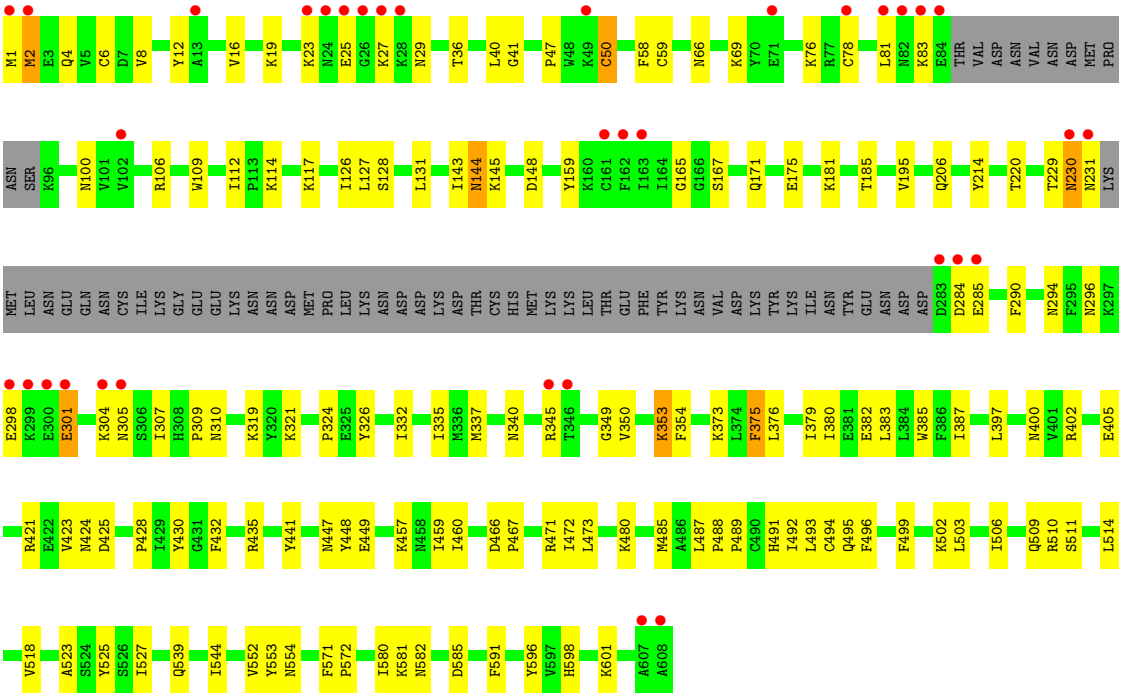
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	203	203	203	0	0
			Total	O		
5	B	168	168	168	0	0

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 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.

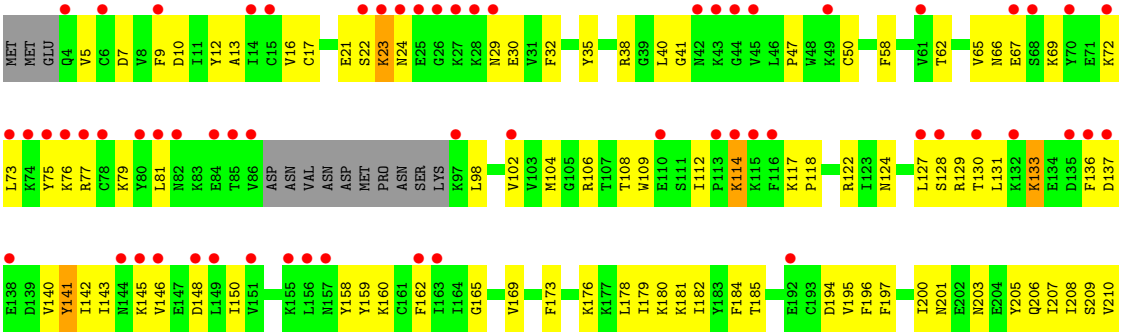
• Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

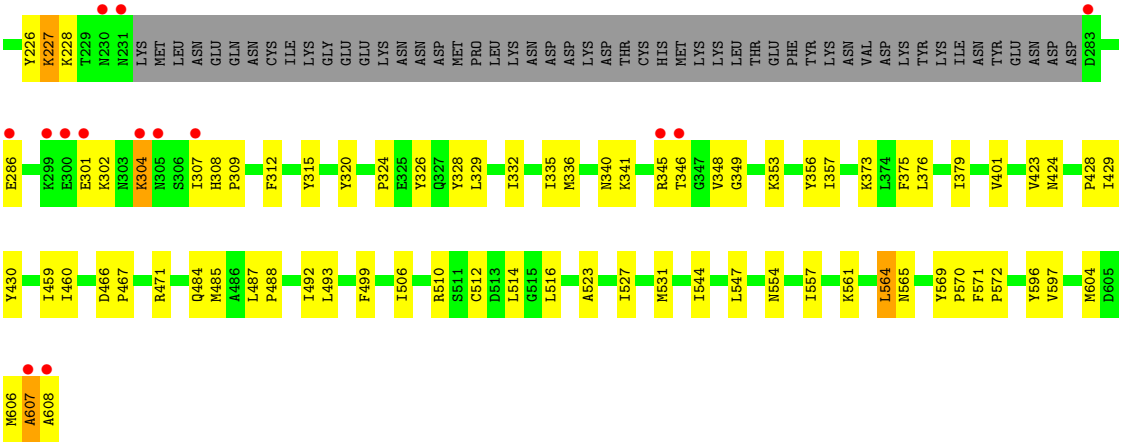
Chain A:



• Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

Chain B:




Y226 K227 K228 T229 N230 N231 LYS MET LEU ASN GLU GLN ASN CYS ILE LYS GLY GLU GLU LYS ASN ASN ASP MET PRO LEU LYS ASN ASP ASP LYS ASP THR CYS HIS MET LYS LYS LEU THR GLU PHE TYR LYS ASN VAL ASP LYS TYR LYS ILE ASN TYR GLU ASN ASP ASP D283
E286 K299 E300 E301 K302 N303 K304 N305 S306 I307 H308 P309 F312 Y315 Y320 P324 E325 Y326 Q327 Y328 L329 I332 I335 M336 N340 K341 R345 T346 G347 V348 G349 K353 Y356 I357 K373 L374 F375 L376 I379 V401 V423 N424 P428 I429
Y430 I459 I460 D466 P467 R471 Q484 M485 A486 L487 P488 I492 L493 F499 I506 R510 S511 C512 D513 L514 G515 L516 A523 I527 M531 I544 L547 N554 I557 K561 L564 N565 Y569 P570 F571 P572 Y596 V597 M604 D605
M606 A607 A608

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.70Å 155.55Å 164.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.11 – 2.65 29.11 – 2.64	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.11-2.65) 99.3 (29.11-2.64)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.86 (at 2.64Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.203 , 0.257 0.199 , 0.202	Depositor DCC
R_{free} test set	2179 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	42.4	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 53.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9587	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NDP, UMP, 1CY

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	0.37	0/4644	0.62	0/6267
1	B	0.37	0/4612	0.60	1/6229 (0.0%)
All	All	0.37	0/9256	0.61	1/12496 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	401	VAL	N-CA-C	-5.14	97.13	111.00

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	4539	0	4496	124	0
1	B	4507	0	4445	157	0
2	A	17	0	14	1	0
2	B	17	0	14	2	0
3	A	48	0	26	2	0
3	B	48	0	26	7	0
4	A	20	0	11	0	0
4	B	20	0	11	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	203	0	0	6	0
5	B	168	0	0	5	0
All	All	9587	0	9043	267	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 15.

The worst 5 of 267 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:345:ARG:HH11	1:B:345:ARG:HB3	1.16	1.08
1:A:27:LYS:HE2	1:A:29:ASN:HB2	1.22	1.07
1:A:304:LYS:HE3	1:A:309:PRO:HG2	1.41	1.01
1:B:345:ARG:HB3	1:B:345:ARG:NH1	1.80	0.94
1:B:429:ILE:HA	1:B:485:MET:HE2	1.57	0.84

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	540/608 (89%)	500 (93%)	38 (7%)	2 (0%)	34	48
1	B	538/608 (88%)	486 (90%)	50 (9%)	2 (0%)	34	48
All	All	1078/1216 (89%)	986 (92%)	88 (8%)	4 (0%)	34	48

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	607	ALA
1	A	430	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	301	GLU
1	B	430	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	510/571 (89%)	502 (98%)	8 (2%)	62	78
1	B	504/571 (88%)	494 (98%)	10 (2%)	55	73
All	All	1014/1142 (89%)	996 (98%)	18 (2%)	59	75

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	227	LYS
1	B	564	LEU
1	B	516	LEU
1	B	23	LYS
1	B	141	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	99	GLN
1	B	394	ASN
1	B	450	ASN
1	B	424	ASN
1	A	316	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	1CY	B	709	-	15,18,18	4.09	11 (73%)	15,27,27	1.89	3 (20%)
3	NDP	A	610	-	45,52,52	1.35	5 (11%)	53,80,80	1.74	8 (15%)
2	1CY	A	609	-	15,18,18	3.35	8 (53%)	15,27,27	1.84	7 (46%)
4	UMP	B	711	-	21,21,21	1.60	5 (23%)	31,31,31	1.89	11 (35%)
4	UMP	A	611	-	21,21,21	1.73	6 (28%)	31,31,31	1.87	10 (32%)
3	NDP	B	710	-	45,52,52	1.46	6 (13%)	53,80,80	1.46	7 (13%)

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	1CY	B	709	-	-	0/4/23/23	0/2/2/2
3	NDP	A	610	-	-	4/30/77/77	0/5/5/5
2	1CY	A	609	-	-	0/4/23/23	0/2/2/2
4	UMP	B	711	-	-	2/10/22/22	0/2/2/2
4	UMP	A	611	-	-	1/10/22/22	0/2/2/2
3	NDP	B	710	-	-	3/30/77/77	0/5/5/5

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	709	1CY	C13-C14	6.65	1.50	1.38
2	B	709	1CY	C16-C11	6.25	1.51	1.39
2	B	709	1CY	C12-C11	5.87	1.50	1.39
2	B	709	1CY	C15-C14	5.79	1.49	1.38
2	A	609	1CY	C13-C14	5.75	1.48	1.38

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	610	NDP	N3A-C2A-N1A	-5.72	119.74	128.68
3	A	610	NDP	C1D-N1N-C2N	-5.64	111.72	121.11
3	B	710	NDP	N3A-C2A-N1A	-5.57	119.98	128.68
2	B	709	1CY	C10-C6-C9	-5.07	103.43	110.86
4	B	711	UMP	O4'-C1'-N1	4.58	116.05	107.86

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	610	NDP	O4D-C1D-N1N-C2N
3	B	710	NDP	O4D-C1D-N1N-C2N
3	A	610	NDP	C2D-C1D-N1N-C2N
3	B	710	NDP	C2D-C1D-N1N-C2N
3	A	610	NDP	O4D-C1D-N1N-C6N

There are no ring outliers.

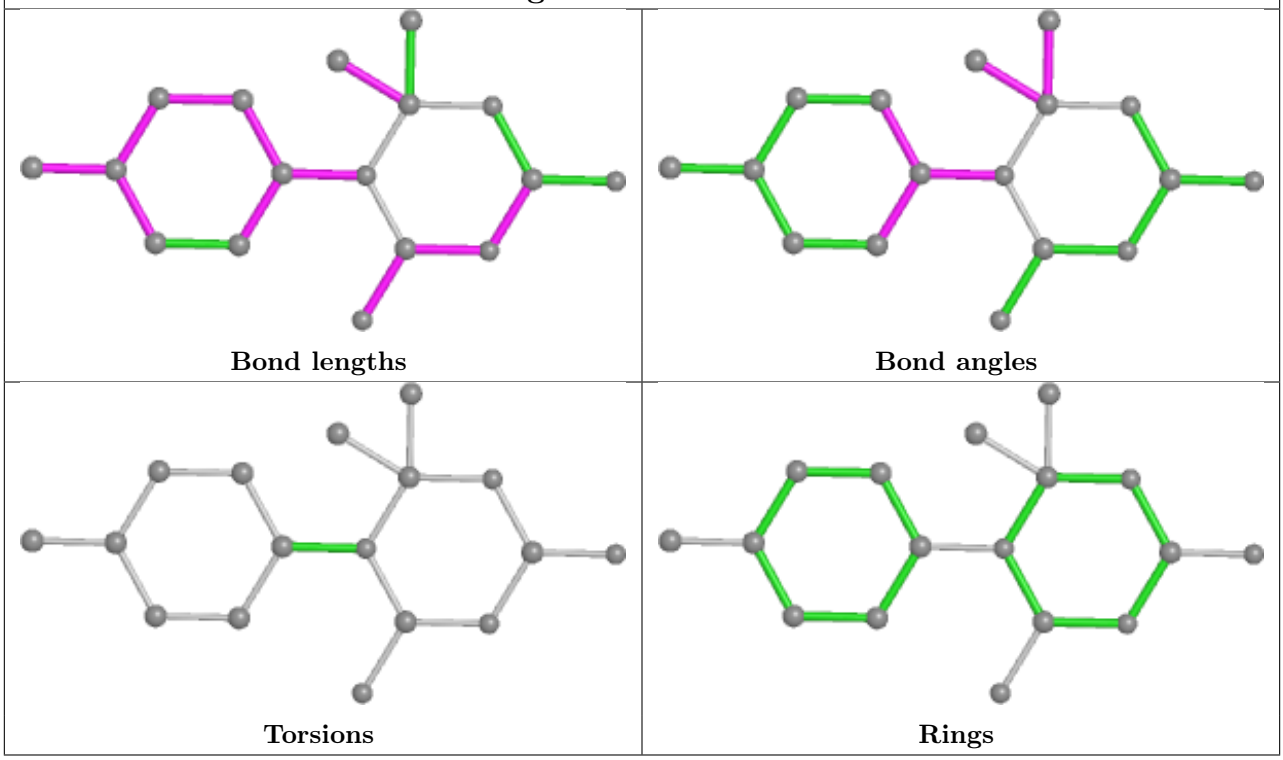
4 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	709	1CY	2	0
3	A	610	NDP	2	0
2	A	609	1CY	1	0
3	B	710	NDP	7	0

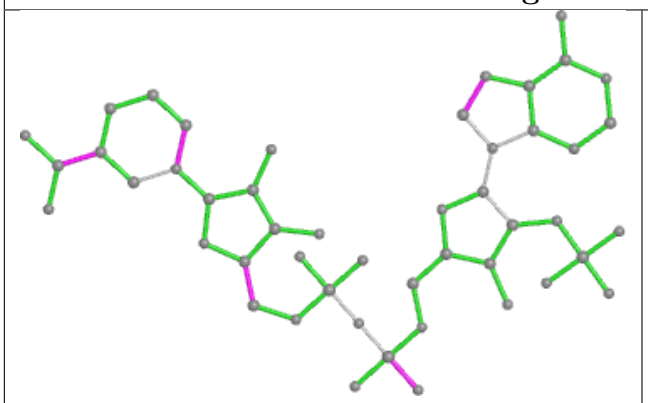
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

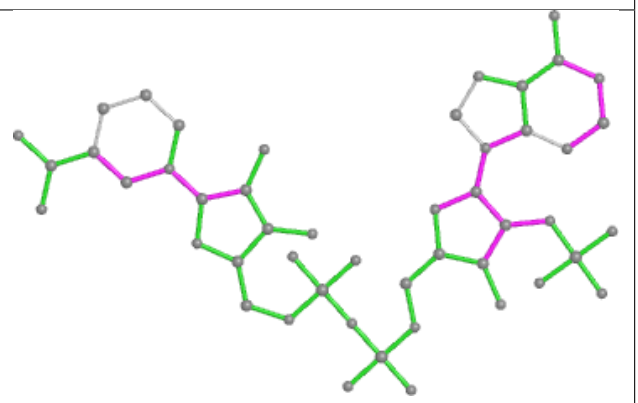
Ligand 1CY B 709



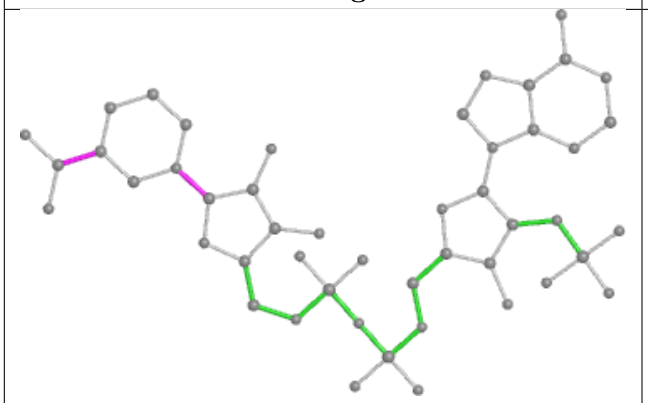
Ligand NDP A 610



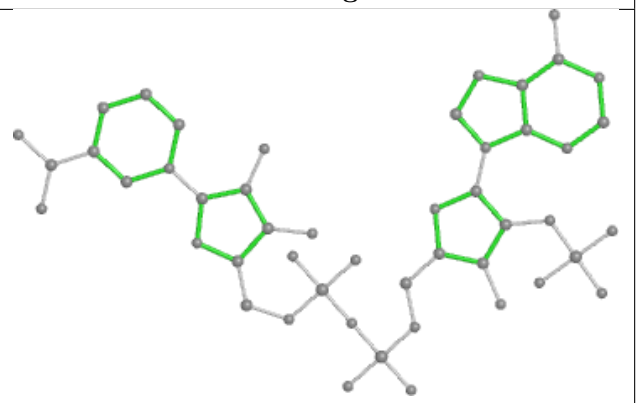
Bond lengths



Bond angles

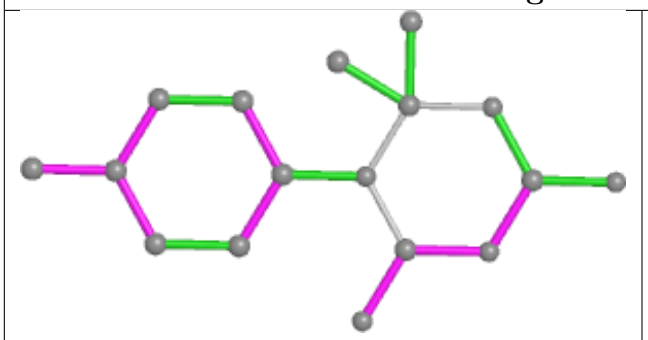


Torsions

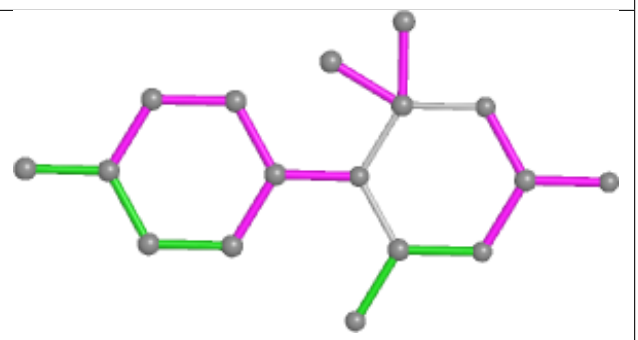


Rings

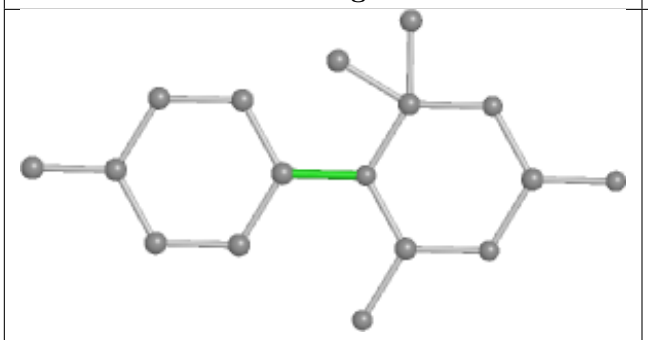
Ligand 1CY A 609



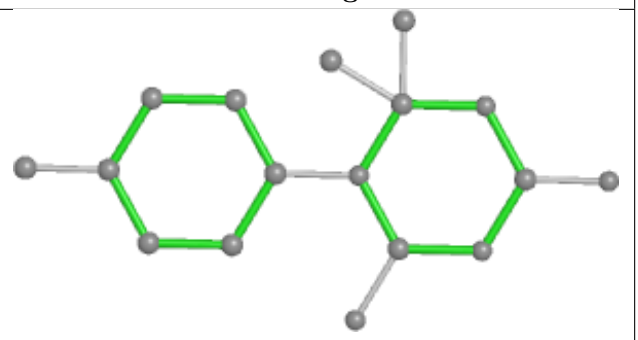
Bond lengths



Bond angles

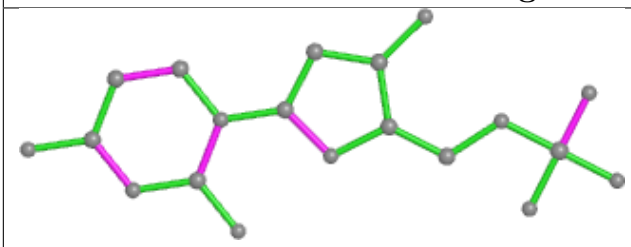


Torsions

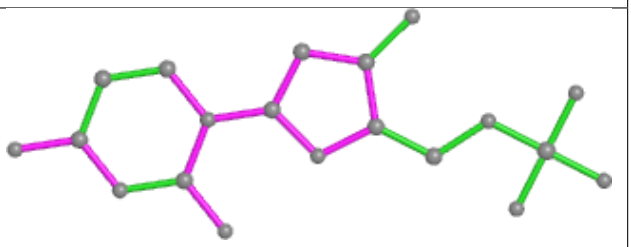


Rings

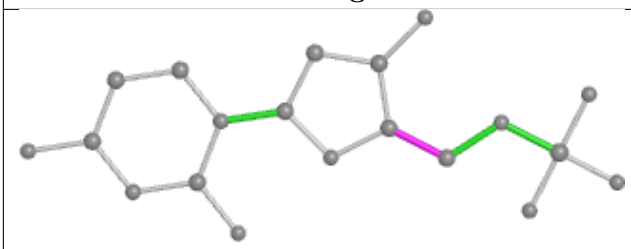
Ligand UMP B 711



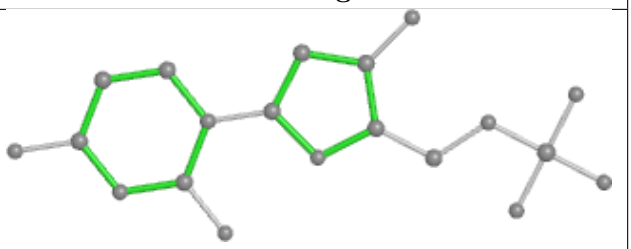
Bond lengths



Bond angles

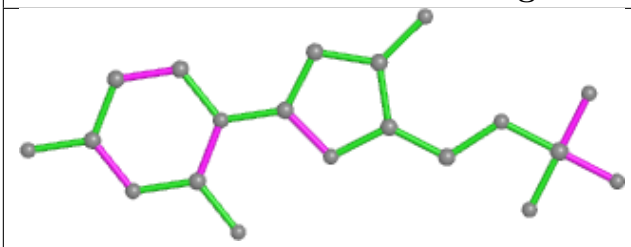


Torsions

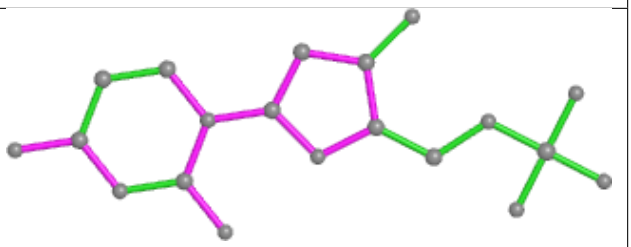


Rings

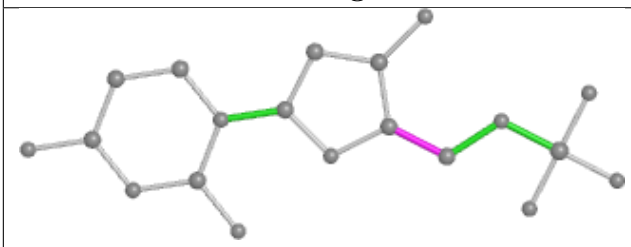
Ligand UMP A 611



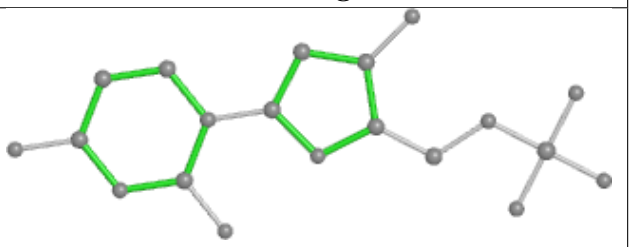
Bond lengths



Bond angles

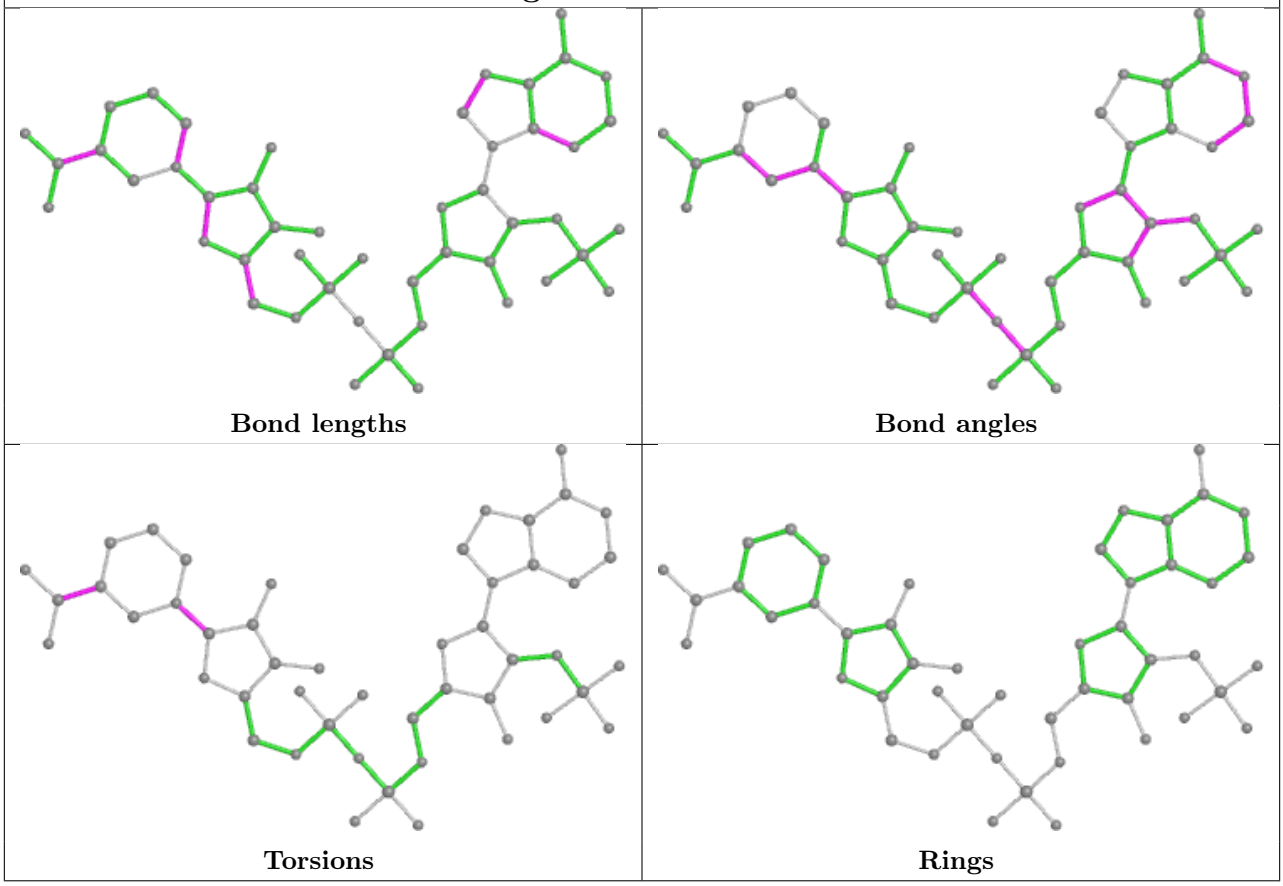


Torsions



Rings

Ligand NDP B 710



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, 95th 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^2)	Q<0.9
1	A	546/608 (89%)	-0.05	35 (6%) 19 16	20, 36, 86, 90	0
1	B	544/608 (89%)	0.32	76 (13%) 2 2	22, 45, 90, 90	0
All	All	1090/1216 (89%)	0.13	111 (10%) 6 4	20, 39, 90, 90	0

The worst 5 of 111 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	9.5
1	A	346	THR	7.2
1	B	26	GLY	6.4
1	B	27	LYS	6.2
1	A	299	LYS	5.9

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)

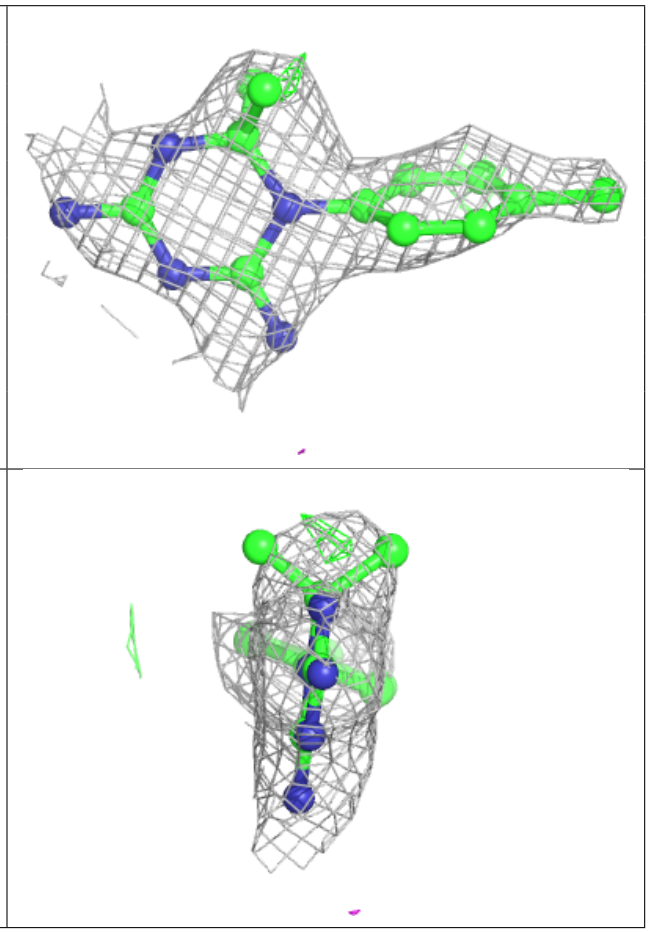
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	1CY	B	709	17/17	0.78	0.41	85,86,89,90	0
3	NDP	B	710	48/48	0.85	0.21	82,88,90,90	0
2	1CY	A	609	17/17	0.92	0.23	28,31,39,47	0
4	UMP	A	611	20/20	0.97	0.14	35,46,55,56	0
4	UMP	B	711	20/20	0.97	0.13	39,41,43,44	0
3	NDP	A	610	48/48	0.98	0.11	30,37,45,45	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.

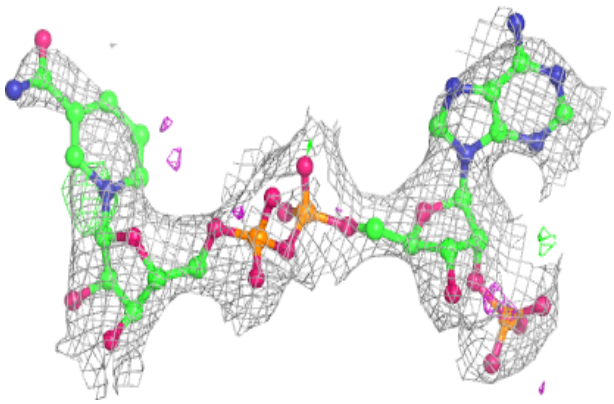
Electron density around 1CY B 709:

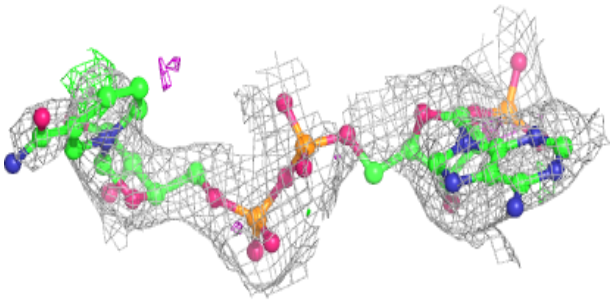
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

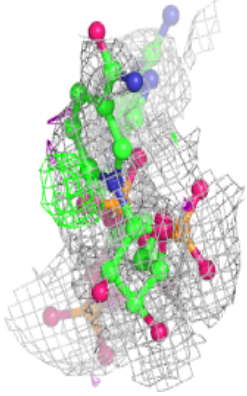


Electron density around NDP B 710:

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

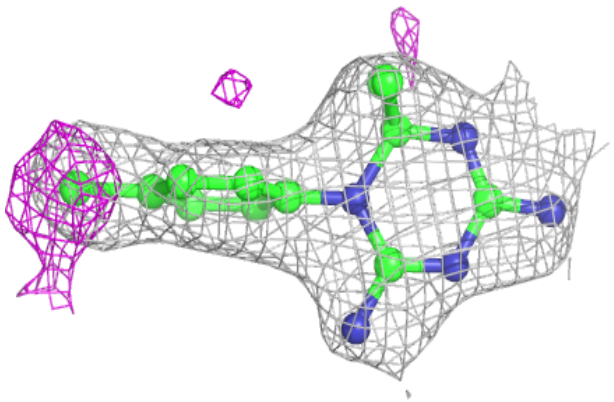


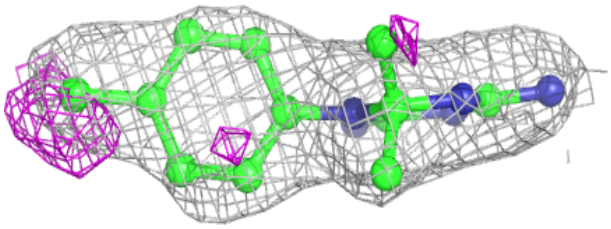


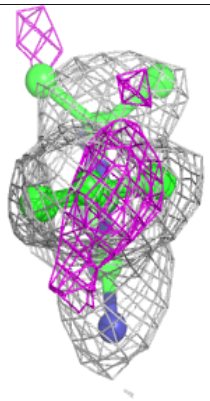


Electron density around 1CY A 609:

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

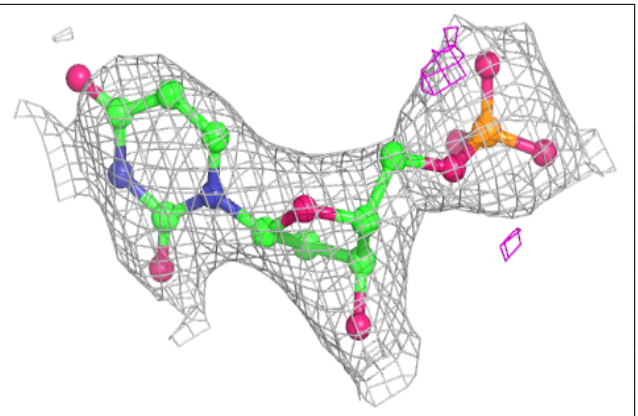


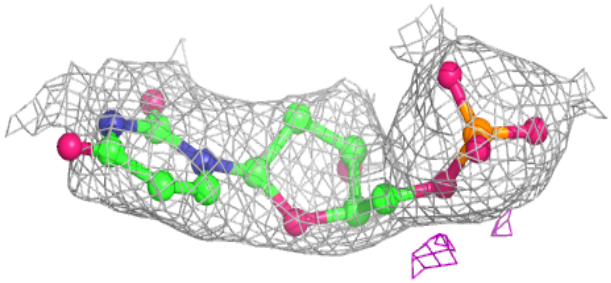


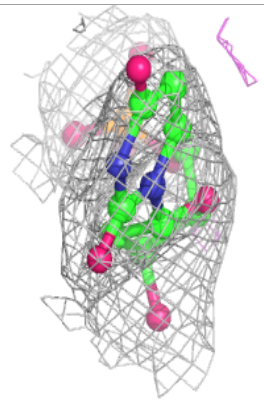


Electron density around UMP A 611:

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

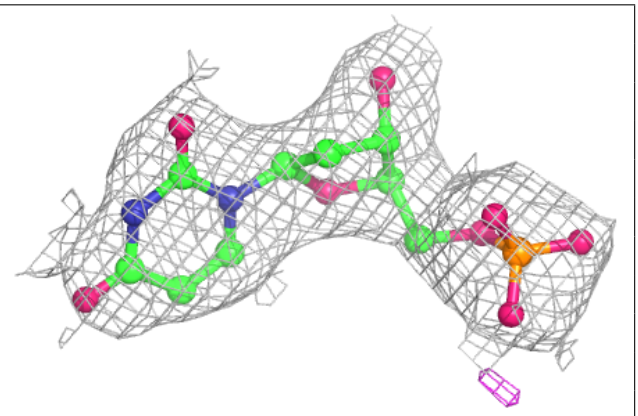


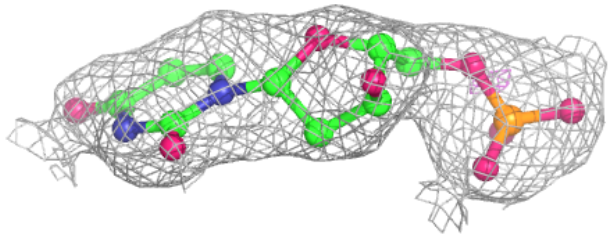


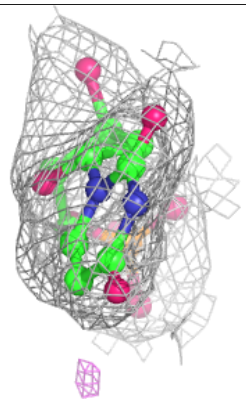


Electron density around UMP B 711:

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

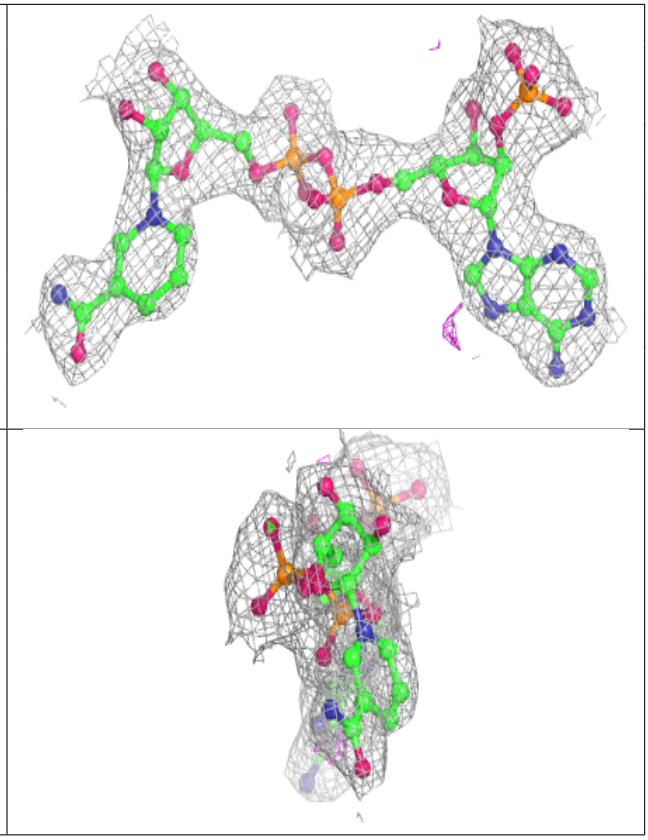






Electron density around NDP A 610:

$2mF_o$-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.