wwPDB X-ray Structure Validation Summary Report (i)

Feb 3, 2024 – 09:12 PM EST

PDB ID	:	1NB6
Title	:	HC-J4 RNA polymerase complexed with UTP
Authors	:	O'Farrell, D.J.; Trowbridge, R.; Rowlands, D.J.; Jaeger, J.
Deposited on	:	2002-12-02
Resolution	:	2.60 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 (1)) were used in the production of this report:

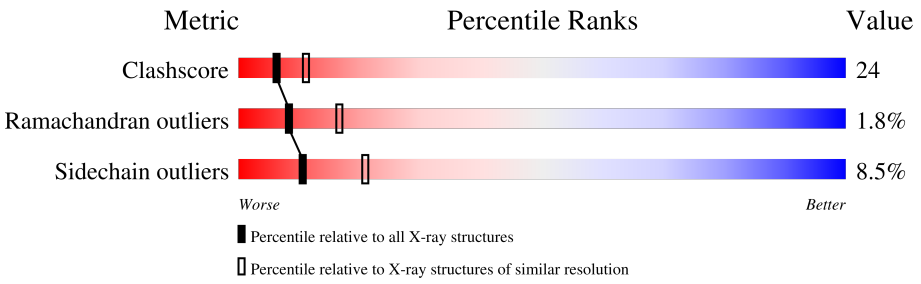
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.60 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))
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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	570	60% 35% • •
1	B	570	61% 32% 6% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9453 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.

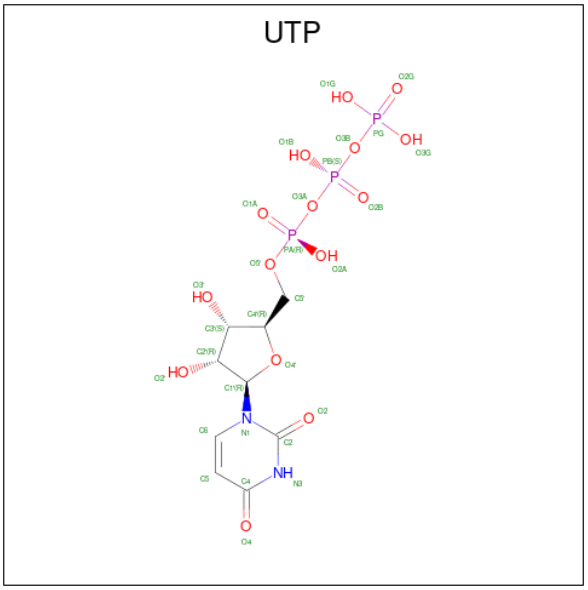
- Molecule 1 is a protein called polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	566	4388	2763	777	816	32	0	0	0
1	B	565	4381	2758	776	815	32	0	0	0

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mn		
2	A	2	2	2	0	0
2	B	2	2	2	0	0

- Molecule 3 is URIDINE 5'-TRIPHOSPHATE (three-letter code: UTP) (formula: $C_9H_{15}N_2O_{15}P_3$).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	270	270	270	0	0
			Total	O		
4	B	352	352	352	0	0

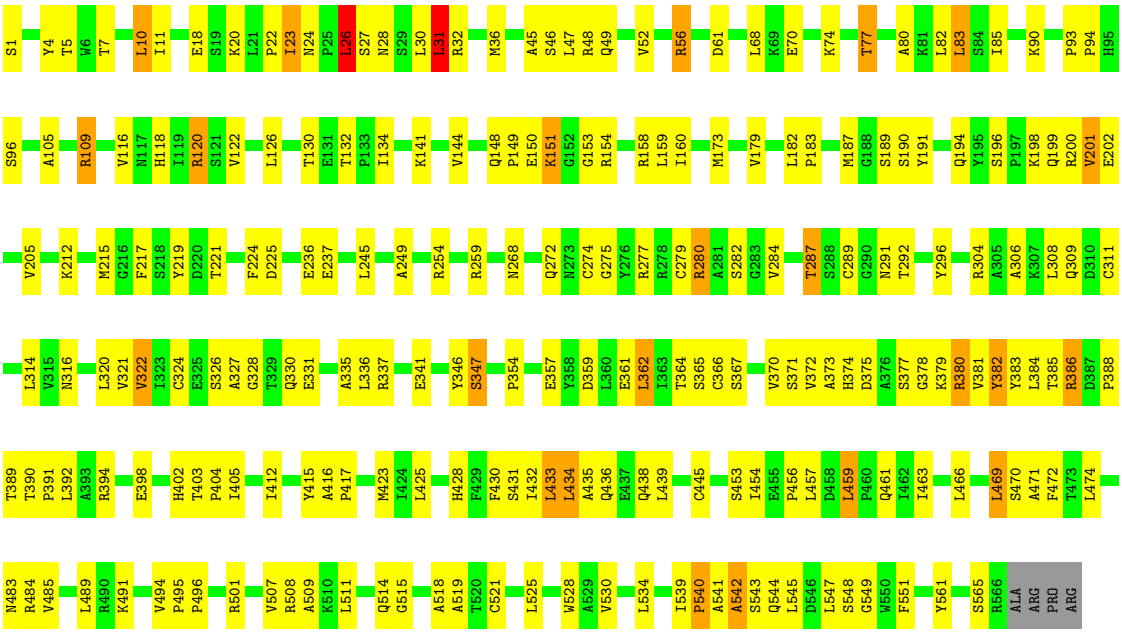
3 Residue-property plots (i)

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.

Note EDS was not executed.

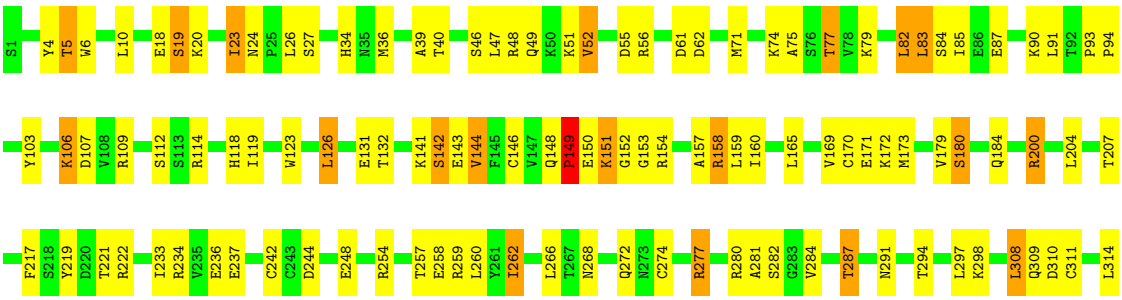
• Molecule 1: polyprotein

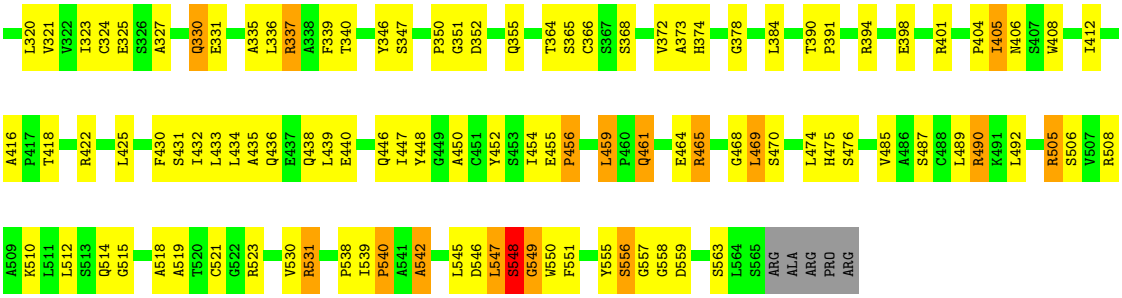
Chain A: 60% 35%



• Molecule 1: polyprotein

Chain B: 61% 32% 6%




L320 V321 V322 I323 C324 E325 S326 A327 Q330 E331 A335 L336 R337 A338 F339 T340 Y346 S347 P350 G351 D352 Q355 T364 S365 C366 S367 S368 V372 A373 H374 G378 L384 T390 P391 R394 E398 R401 P404 I405 N406 S407 W408 I412
A416 P417 T418 R422 L425 F430 S431 I432 L433 L434 A435 Q436 E437 Q438 L439 E440 Q446 I447 Y448 G449 A450 C451 Y452 S453 I454 E455 P456 L459 P460 Q461 E464 R465 G468 L469 S470 L474 H475 S476 V485 A486 S487 C488 L489 R490 K491 L492 R505 S506 V507 R508
A509 K510 L511 L512 S513 Q514 G515 A518 A519 T520 C521 G522 R523 V530 R531 P538 I539 P540 A541 A542 L545 D546 L547 S548 G549 W550 F551 Y555 S556 G557 G558 D559 S563 L564 S565 ARG ALA ARG PRO ARG

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.43Å 107.74Å 133.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.00 – 2.60	Depositor
% Data completeness (in resolution range)	95.1 (26.00-2.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.181 , 0.249	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9453	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

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, MN

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	0.64	1/4484 (0.0%)	0.83	5/6087 (0.1%)
1	B	0.69	2/4477 (0.0%)	0.85	3/6077 (0.0%)
All	All	0.66	3/8961 (0.0%)	0.84	8/12164 (0.1%)

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

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	142	SER	CB-OG	8.42	1.53	1.42
1	A	274	CYS	CB-SG	-6.32	1.71	1.82
1	B	242	CYS	CB-SG	-5.57	1.72	1.81

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	280	ARG	NE-CZ-NH2	-6.89	116.85	120.30
1	B	200	ARG	NE-CZ-NH1	-6.53	117.03	120.30
1	A	11	ILE	N-CA-C	-6.20	94.26	111.00
1	A	31	LEU	CA-CB-CG	6.13	129.39	115.30
1	A	277	ARG	NE-CZ-NH1	6.10	123.35	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	382	TYR	Sidechain

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	4388	0	4383	224	0
1	B	4381	0	4374	204	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	29	0	11	2	0
3	B	29	0	11	2	0
4	A	270	0	0	28	0
4	B	352	0	0	38	0
All	All	9453	0	8779	428	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 24.

The worst 5 of 428 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:109:ARG:NH1	1:B:109:ARG:HB2	1.63	1.12
1:B:141:LYS:NZ	1:B:158:ARG:HH12	1.55	1.04
1:B:148:GLN:HE21	1:B:153:GLY:HA3	1.18	1.03
1:A:547:LEU:HD12	1:A:565:SER:H	1.25	1.01
1:B:277:ARG:HE	1:B:281:ALA:HB2	1.22	0.98

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	564/570 (99%)	517 (92%)	38 (7%)	9 (2%)	9	19
1	B	563/570 (99%)	515 (92%)	37 (7%)	11 (2%)	7	14
All	All	1127/1140 (99%)	1032 (92%)	75 (7%)	20 (2%)	8	16

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	151	LYS
1	A	540	PRO
1	B	469	LEU
1	B	542	ALA
1	B	549	GLY

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	475/485 (98%)	443 (93%)	32 (7%)	16	33
1	B	474/485 (98%)	425 (90%)	49 (10%)	7	13
All	All	949/970 (98%)	868 (92%)	81 (8%)	10	21

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	274	CYS
1	B	456	PRO
1	B	287	THR
1	B	355	GLN
1	B	487	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	34	HIS
1	B	148	GLN
1	B	514	GLN
1	B	49	GLN
1	B	206	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)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	UTP	A	1700	2	22,30,30	1.99	7 (31%)	27,47,47	2.71	13 (48%)
3	UTP	B	2700	2	22,30,30	1.89	6 (27%)	27,47,47	1.52	4 (14%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	UTP	A	1700	2	-	4/20/38/38	0/2/2/2
3	UTP	B	2700	2	-	5/20/38/38	0/2/2/2

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2700	UTP	C4-N3	4.69	1.41	1.33
3	A	1700	UTP	C4-N3	4.54	1.40	1.33
3	A	1700	UTP	PG-O3G	3.88	1.69	1.54
3	B	2700	UTP	O3'-C3'	3.57	1.51	1.43
3	A	1700	UTP	O4'-C1'	3.28	1.45	1.41

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1700	UTP	O3G-PG-O2G	-5.46	89.30	110.68
3	A	1700	UTP	O3G-PG-O3B	-4.71	88.83	104.64
3	A	1700	UTP	PB-O3B-PG	-4.30	118.06	132.83
3	A	1700	UTP	O1G-PG-O2G	4.24	127.27	110.68
3	A	1700	UTP	O3G-PG-O1G	-4.23	91.47	107.64

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	2700	UTP	O4'-C1'-N1-C6
3	B	2700	UTP	C2'-C1'-N1-C6
3	A	1700	UTP	O4'-C4'-C5'-O5'
3	A	1700	UTP	C5'-O5'-PA-O2A
3	B	2700	UTP	C5'-O5'-PA-O1A

There are no ring outliers.

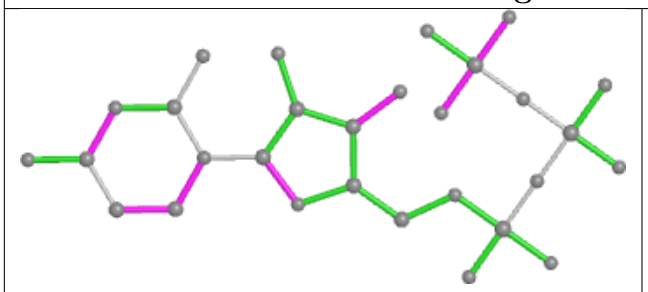
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1700	UTP	2	0
3	B	2700	UTP	2	0

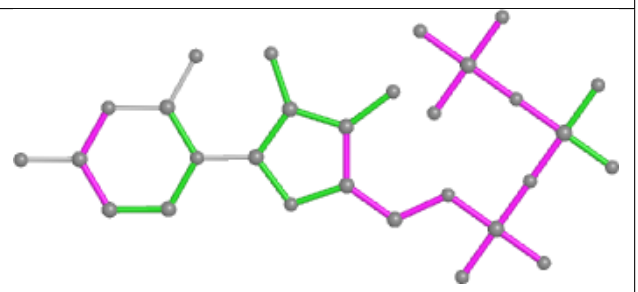
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

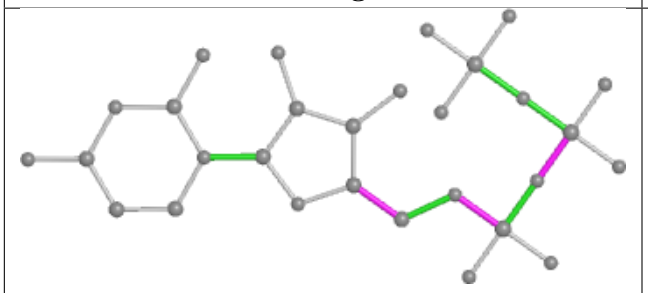
Ligand UTP A 1700



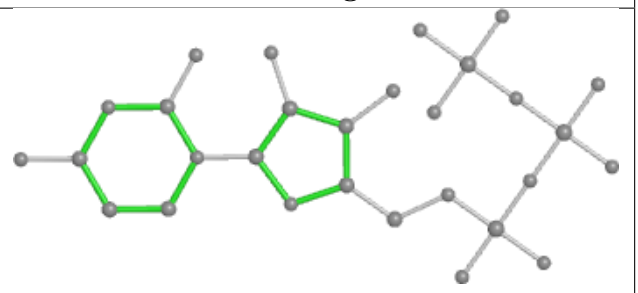
Bond lengths



Bond angles

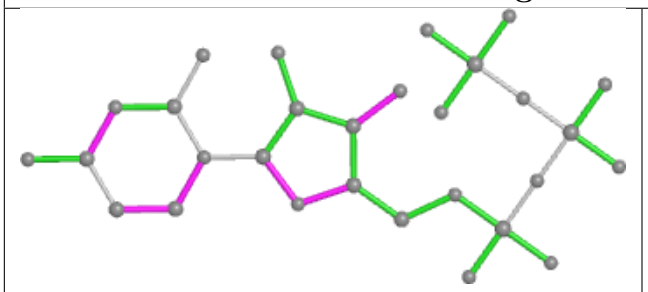


Torsions

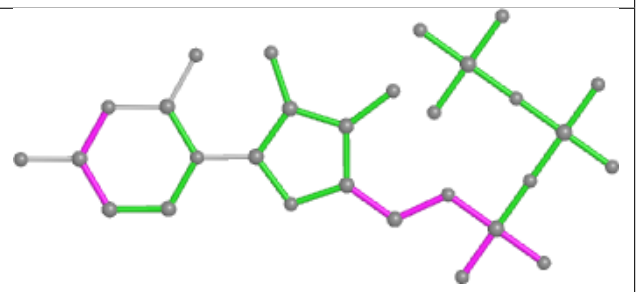


Rings

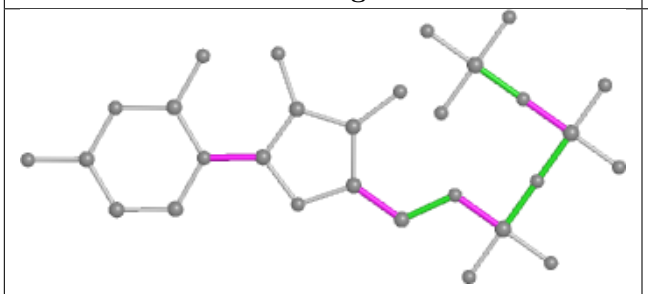
Ligand UTP B 2700



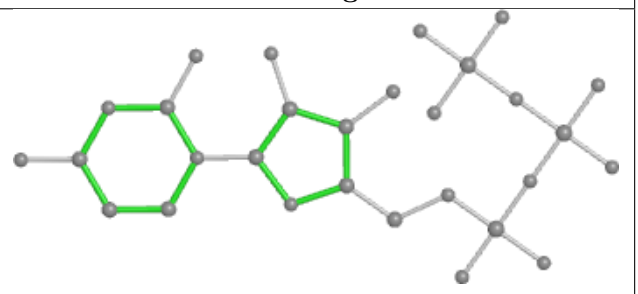
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.