Full wwPDB X-ray Structure Validation Report (i)

Jun 22, 2024 – 08:40 PM EDT

PDB ID	:	4ZHT
Title	:	Crystal structure of UDP-GlcNAc 2-epimerase
Authors	:	Chen, S.C.; Yang, C.S.; Ko, T.P.; Chen, Y.
Deposited on	:	2015-04-27
Resolution	:	2.69 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

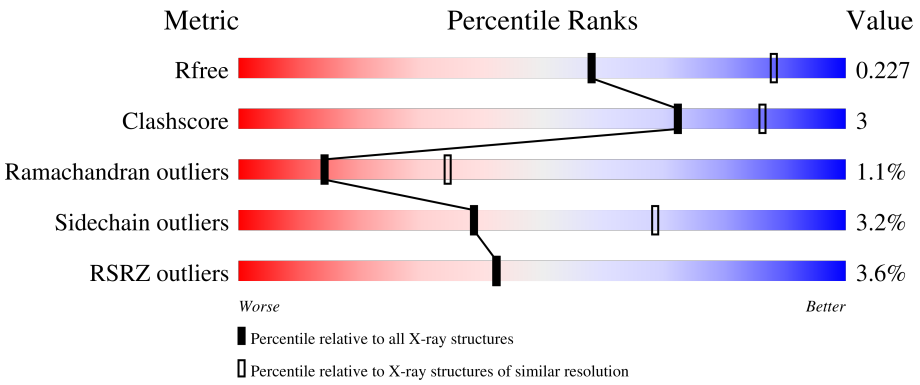
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.69 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	411	2%, 83%, 8%, • 7%
1	B	411	4%, 82%, 10%, • 7%
1	C	411	2%, 84%, 8%, • 7%
1	D	411	5%, 84%, 8%, • 7%

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
4	BM7	C	503	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12681 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 UDP-N-acetylglucosamine 2-epimerase/N-acetyl mannosamine kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	384	3026	1917	536	551	10	12	0	0	0
1	B	382	3011	1908	533	548	10	12	0	0	0
1	C	383	3016	1907	536	551	10	12	0	0	0
1	D	383	3018	1913	534	549	10	12	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

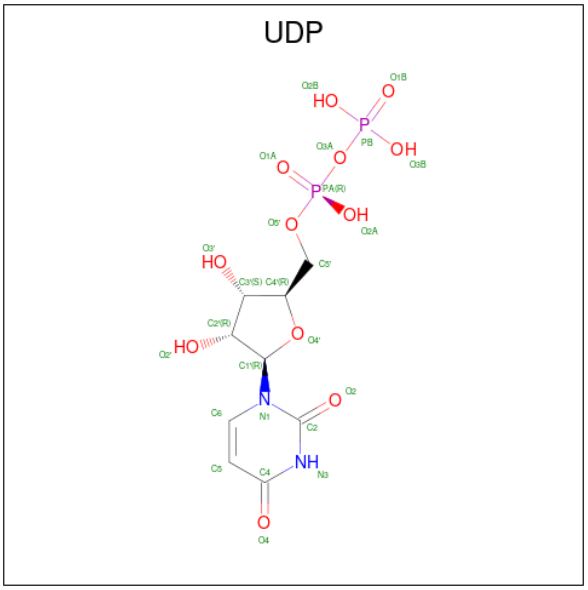
Chain	Residue	Modelled	Actual	Comment	Reference
A	406	HIS	-	expression tag	UNP Q9Y223
A	407	HIS	-	expression tag	UNP Q9Y223
A	408	HIS	-	expression tag	UNP Q9Y223
A	409	HIS	-	expression tag	UNP Q9Y223
A	410	HIS	-	expression tag	UNP Q9Y223
A	411	HIS	-	expression tag	UNP Q9Y223
B	406	HIS	-	expression tag	UNP Q9Y223
B	407	HIS	-	expression tag	UNP Q9Y223
B	408	HIS	-	expression tag	UNP Q9Y223
B	409	HIS	-	expression tag	UNP Q9Y223
B	410	HIS	-	expression tag	UNP Q9Y223
B	411	HIS	-	expression tag	UNP Q9Y223
C	406	HIS	-	expression tag	UNP Q9Y223
C	407	HIS	-	expression tag	UNP Q9Y223
C	408	HIS	-	expression tag	UNP Q9Y223
C	409	HIS	-	expression tag	UNP Q9Y223
C	410	HIS	-	expression tag	UNP Q9Y223
C	411	HIS	-	expression tag	UNP Q9Y223
D	406	HIS	-	expression tag	UNP Q9Y223
D	407	HIS	-	expression tag	UNP Q9Y223

Continued on next page...

Continued from previous page...

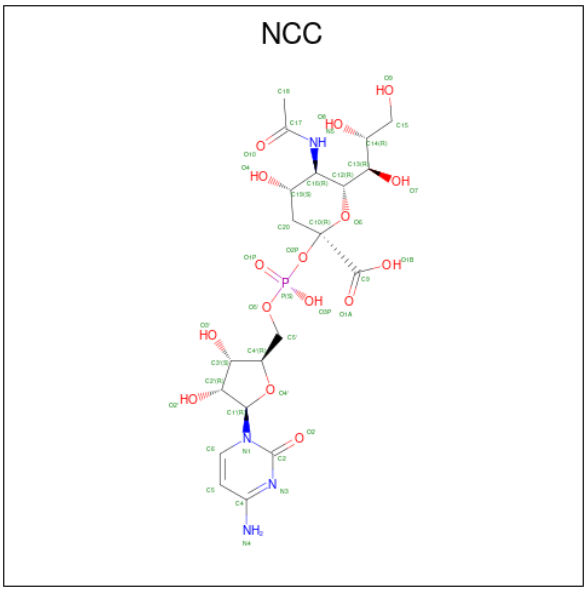
Chain	Residue	Modelled	Actual	Comment	Reference
D	408	HIS	-	expression tag	UNP Q9Y223
D	409	HIS	-	expression tag	UNP Q9Y223
D	410	HIS	-	expression tag	UNP Q9Y223
D	411	HIS	-	expression tag	UNP Q9Y223

- Molecule 2 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).



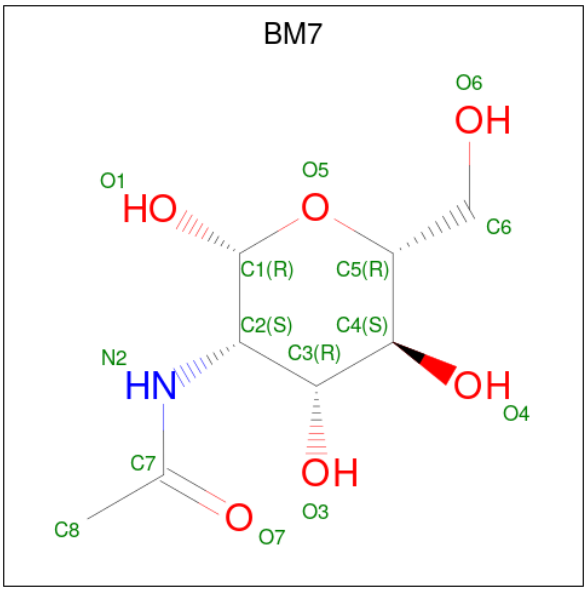
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	C	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	D	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 3 is CYTIDINE-5'-MONOPHOSPHATE-5-N-ACETYLNEURAMINIC ACID (three-letter code: NCC) (formula: $C_{20}H_{31}N_4O_{16}P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			41	20	4	16	1		
3	B	1	Total	C	N	O	P	0	0
			41	20	4	16	1		
3	C	1	Total	C	N	O	P	0	0
			41	20	4	16	1		
3	D	1	Total	C	N	O	P	0	0
			41	20	4	16	1		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-mannopyranose (three-letter code: BM7) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	C	1	15	8	1	6	0	0

- Molecule 5 is water.

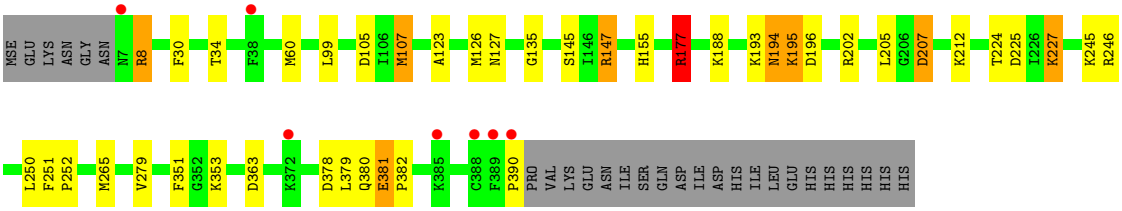
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	101	101	101	0	0
			Total	O		
5	B	78	78	78	0	0
			Total	O		
5	C	98	98	98	0	0
			Total	O		
5	D	54	54	54	0	0

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.

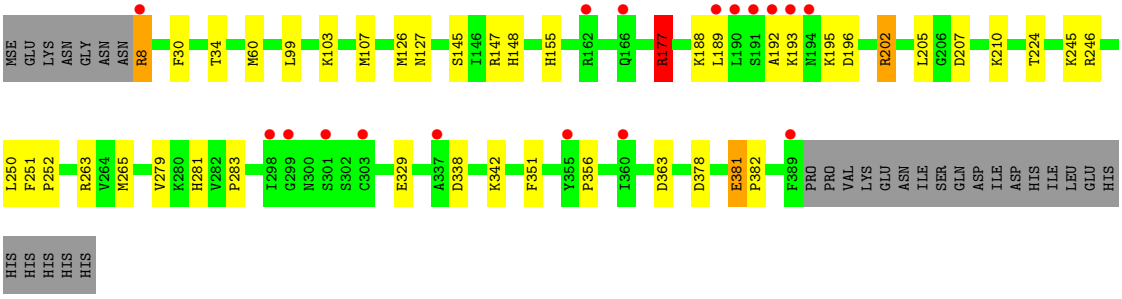
• Molecule 1: Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase

Chain A:



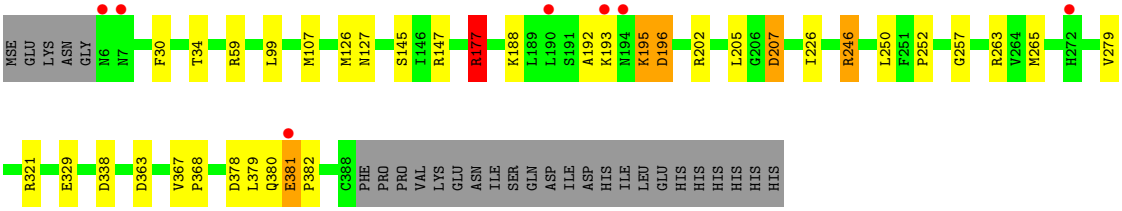
• Molecule 1: Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase

Chain B:



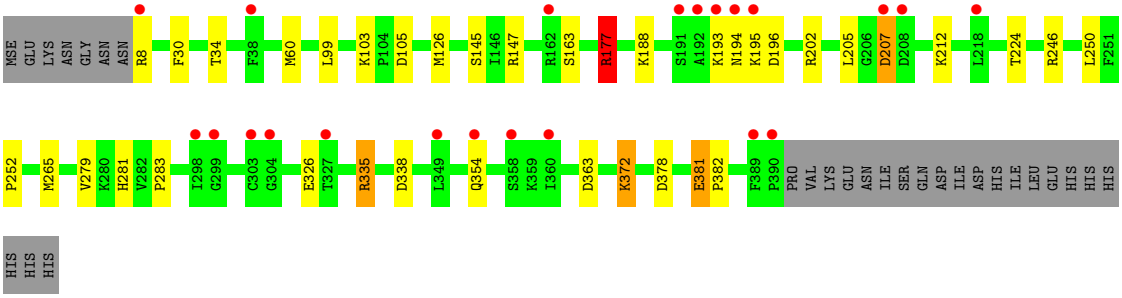
• Molecule 1: Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase

Chain C:



• Molecule 1: Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	106.04Å 98.10Å 154.72Å 90.00° 96.05° 90.00°	Depositor
Resolution (Å)	30.00 – 2.69 27.89 – 2.69	Depositor EDS
% Data completeness (in resolution range)	99.4 (30.00-2.69) 99.6 (27.89-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.51 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.189 , 0.225 0.194 , 0.227	Depositor DCC
R_{free} test set	2183 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	63.8	Xtriage
Anisotropy	0.297	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 30.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12681	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.13% 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: NCC, UDP, BM7

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	0.56	0/3071	0.87	10/4126 (0.2%)
1	B	0.54	0/3055	0.84	7/4103 (0.2%)
1	C	0.56	0/3059	0.83	7/4109 (0.2%)
1	D	0.52	0/3063	0.82	6/4115 (0.1%)
All	All	0.54	0/12248	0.84	30/16453 (0.2%)

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	B	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	8	ARG	NE-CZ-NH1	8.73	124.67	120.30
1	D	177	ARG	NE-CZ-NH1	8.36	124.48	120.30
1	C	177	ARG	NE-CZ-NH1	8.20	124.40	120.30
1	A	177	ARG	NE-CZ-NH1	8.02	124.31	120.30
1	B	177	ARG	NE-CZ-NH1	8.00	124.30	120.30
1	A	212	LYS	CD-CE-NZ	7.78	129.58	111.70
1	A	195	LYS	CD-CE-NZ	7.38	128.68	111.70
1	B	202	ARG	NE-CZ-NH1	-7.21	116.70	120.30
1	D	335	ARG	NE-CZ-NH1	7.20	123.90	120.30
1	B	263	ARG	NE-CZ-NH2	6.11	123.35	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	225	ASP	CB-CG-OD1	6.09	123.78	118.30
1	C	263	ARG	NE-CZ-NH2	6.02	123.31	120.30
1	A	147	ARG	NE-CZ-NH1	5.90	123.25	120.30
1	A	60	MSE	CG-SE-CE	5.73	111.50	98.90
1	D	147	ARG	NE-CZ-NH1	5.71	123.16	120.30
1	B	329	GLU	CA-CB-CG	5.66	125.85	113.40
1	B	8	ARG	NE-CZ-NH1	5.58	123.09	120.30
1	C	107	MSE	CA-CB-CG	-5.54	103.88	113.30
1	D	335	ARG	CG-CD-NE	5.49	123.33	111.80
1	D	177	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	C	147	ARG	NE-CZ-NH1	5.39	123.00	120.30
1	C	177	ARG	CG-CD-NE	5.38	123.09	111.80
1	D	177	ARG	CG-CD-NE	5.35	123.03	111.80
1	C	147	ARG	NE-CZ-NH2	-5.27	117.66	120.30
1	A	8	ARG	NE-CZ-NH2	-5.25	117.67	120.30
1	A	227	LYS	N-CA-CB	5.20	119.95	110.60
1	B	8	ARG	NE-CZ-NH2	-5.13	117.73	120.30
1	B	147	ARG	NE-CZ-NH1	5.13	122.87	120.30
1	A	147	ARG	NE-CZ-NH2	-5.07	117.76	120.30
1	C	321	ARG	NE-CZ-NH1	5.01	122.81	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	192	ALA	Peptide
1	C	192	ALA	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	3026	0	3066	24	0
1	B	3011	0	3053	24	0
1	C	3016	0	3056	19	0
1	D	3018	0	3060	20	0
2	A	25	0	11	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	25	0	11	0	0
2	C	25	0	11	0	0
2	D	25	0	11	0	0
3	A	41	0	29	0	0
3	B	41	0	29	0	0
3	C	41	0	29	0	0
3	D	41	0	29	0	0
4	C	15	0	0	0	0
5	A	101	0	0	4	0
5	B	78	0	0	5	0
5	C	98	0	0	3	0
5	D	54	0	0	2	0
All	All	12681	0	12395	84	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 3.

All (84) 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:8:ARG:NH2	1:A:105:ASP:OD2	2.07	0.87
1:C:59:ARG:NH2	5:C:601:HOH:O	2.08	0.86
1:A:380:GLN:NE2	5:A:602:HOH:O	2.10	0.85
1:B:60:MSE:HE2	1:B:281:HIS:CD2	2.18	0.79
1:D:354:GLN:O	5:D:601:HOH:O	2.01	0.78
1:D:381:GLU:HB3	1:D:382:PRO:HA	1.68	0.76
1:B:381:GLU:HB3	1:B:382:PRO:HA	1.67	0.76
1:D:60:MSE:HE2	1:D:281:HIS:CD2	2.19	0.76
1:D:60:MSE:HE2	1:D:281:HIS:HD2	1.50	0.75
1:A:381:GLU:HB3	1:A:382:PRO:HA	1.68	0.75
1:C:381:GLU:HB3	1:C:382:PRO:HA	1.68	0.74
1:B:60:MSE:HE2	1:B:281:HIS:HD2	1.51	0.74
1:B:60:MSE:HE3	1:B:283:PRO:HA	1.69	0.74
1:D:60:MSE:HE3	1:D:283:PRO:HA	1.71	0.73
1:A:107:MSE:HE2	1:A:123:ALA:HB2	1.71	0.72
1:A:135:GLY:O	5:A:601:HOH:O	2.06	0.72
1:B:189:LEU:O	5:B:601:HOH:O	2.10	0.69
1:C:329:GLU:O	5:C:602:HOH:O	2.09	0.69
1:B:342:LYS:NZ	5:B:602:HOH:O	2.26	0.69
1:D:212:LYS:HD2	5:D:615:HOH:O	1.97	0.64
1:D:60:MSE:HE1	1:D:283:PRO:HG3	1.83	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:60:MSE:HE1	1:B:283:PRO:HG3	1.83	0.60
1:C:127:ASN:HD21	1:D:224:THR:HG22	1.67	0.59
1:A:194:ASN:HB3	5:A:646:HOH:O	2.04	0.57
1:C:202:ARG:HG2	1:C:207:ASP:HA	1.87	0.56
1:D:252:PRO:HD3	1:D:265:MSE:HE1	1.88	0.55
1:D:202:ARG:HG2	1:D:207:ASP:HA	1.90	0.54
1:B:8:ARG:N	5:B:606:HOH:O	2.41	0.54
1:C:378:ASP:OD1	1:C:379:LEU:N	2.41	0.53
1:A:99:LEU:HD12	1:A:126:MSE:HE1	1.90	0.53
1:A:202:ARG:HG2	1:A:207:ASP:HA	1.91	0.53
1:C:177:ARG:HG2	1:C:177:ARG:HH11	1.74	0.53
1:D:177:ARG:HH11	1:D:177:ARG:HG2	1.74	0.53
1:A:252:PRO:HD3	1:A:265:MSE:HE1	1.90	0.53
1:C:246:ARG:NH2	5:C:604:HOH:O	2.42	0.53
1:A:378:ASP:OD1	1:A:379:LEU:N	2.42	0.52
1:C:99:LEU:HD12	1:C:126:MSE:HE1	1.92	0.51
1:B:252:PRO:HD3	1:B:265:MSE:HE1	1.92	0.51
1:C:195:LYS:HG3	1:C:196:ASP:N	2.26	0.51
1:C:252:PRO:HD3	1:C:265:MSE:HE1	1.92	0.51
1:A:224:THR:HG22	1:B:127:ASN:HD21	1.76	0.51
1:C:226:ILE:HD13	1:C:257:GLY:HA3	1.92	0.50
1:A:250:LEU:HD13	1:A:279:VAL:HG23	1.93	0.49
1:C:250:LEU:HD13	1:C:279:VAL:HG23	1.93	0.49
1:B:155:HIS:O	1:B:177:ARG:HD3	2.13	0.48
1:B:250:LEU:HD13	1:B:279:VAL:HG23	1.94	0.48
1:D:60:MSE:HE1	1:D:283:PRO:CG	2.44	0.48
1:A:245:LYS:HD3	1:A:351:PHE:CZ	2.48	0.48
1:B:99:LEU:HD12	1:B:126:MSE:HE1	1.95	0.48
1:C:378:ASP:OD1	1:C:380:GLN:HG2	2.13	0.48
1:D:99:LEU:HD12	1:D:126:MSE:HE1	1.95	0.47
1:A:155:HIS:O	1:A:177:ARG:HD3	2.14	0.47
1:D:250:LEU:HD13	1:D:279:VAL:HG23	1.95	0.47
1:B:60:MSE:HE1	1:B:283:PRO:CG	2.44	0.47
1:B:245:LYS:HD3	1:B:351:PHE:CZ	2.50	0.47
1:A:8:ARG:HH22	1:A:105:ASP:CG	2.08	0.47
1:A:177:ARG:HH11	1:A:177:ARG:HG2	1.79	0.47
1:A:378:ASP:OD1	1:A:380:GLN:HG2	2.14	0.46
1:B:177:ARG:HG2	1:B:177:ARG:HH11	1.81	0.46
1:B:148:HIS:HD2	5:B:603:HOH:O	1.97	0.46
1:B:356:PRO:HG3	5:B:670:HOH:O	2.15	0.46
1:C:195:LYS:HG3	1:C:196:ASP:H	1.79	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:202:ARG:CG	1:C:207:ASP:HA	2.46	0.45
1:D:8:ARG:NH2	1:D:105:ASP:OD2	2.48	0.45
1:D:372:LYS:N	1:D:372:LYS:HD3	2.32	0.45
1:D:60:MSE:CE	1:D:281:HIS:CD2	2.96	0.45
1:D:30:PHE:O	1:D:34:THR:HG23	2.18	0.44
1:B:30:PHE:O	1:B:34:THR:HG23	2.17	0.44
1:D:381:GLU:CB	1:D:382:PRO:HA	2.43	0.44
1:D:163:SER:HB3	1:D:326:GLU:OE2	2.18	0.44
1:A:107:MSE:CE	1:A:123:ALA:HB2	2.45	0.44
1:C:30:PHE:O	1:C:34:THR:HG23	2.17	0.43
1:A:251:PHE:HA	1:A:265:MSE:HE1	2.00	0.43
1:B:381:GLU:CB	1:B:382:PRO:HA	2.43	0.43
1:A:30:PHE:O	1:A:34:THR:HG23	2.18	0.42
1:C:381:GLU:CB	1:C:382:PRO:HA	2.44	0.42
1:A:147:ARG:HG2	5:A:601:HOH:O	2.19	0.42
1:A:127:ASN:HD21	1:B:224:THR:HG22	1.84	0.41
1:A:381:GLU:CB	1:A:382:PRO:HA	2.44	0.41
1:B:60:MSE:HE3	1:B:283:PRO:CA	2.45	0.41
1:B:99:LEU:HD22	1:B:107:MSE:HE2	2.03	0.41
1:B:251:PHE:HA	1:B:265:MSE:HE1	2.03	0.41
1:C:367:VAL:HB	1:C:368:PRO:HD3	2.03	0.41
1:A:353:LYS:HE3	1:A:353:LYS:HB3	1.90	0.41

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	382/411 (93%)	365 (96%)	12 (3%)	5 (1%)	12 30
1	B	380/411 (92%)	364 (96%)	12 (3%)	4 (1%)	14 34
1	C	381/411 (93%)	363 (95%)	15 (4%)	3 (1%)	19 43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	381/411 (93%)	363 (95%)	13 (3%)	5 (1%)	12	30
All	All	1524/1644 (93%)	1455 (96%)	52 (3%)	17 (1%)	14	34

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	381	GLU
1	B	381	GLU
1	C	193	LYS
1	C	196	ASP
1	C	381	GLU
1	D	381	GLU
1	A	193	LYS
1	A	196	ASP
1	B	193	LYS
1	B	196	ASP
1	D	196	ASP
1	A	194	ASN
1	D	193	LYS
1	D	194	ASN
1	B	195	LYS
1	D	195	LYS
1	A	195	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	334/347 (96%)	324 (97%)	10 (3%)	41	70
1	B	332/347 (96%)	320 (96%)	12 (4%)	35	64
1	C	333/347 (96%)	324 (97%)	9 (3%)	44	74
1	D	333/347 (96%)	321 (96%)	12 (4%)	35	64
All	All	1332/1388 (96%)	1289 (97%)	43 (3%)	39	68

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	107	MSE
1	A	145	SER
1	A	177	ARG
1	A	188	LYS
1	A	205	LEU
1	A	207	ASP
1	A	227	LYS
1	A	246	ARG
1	A	363	ASP
1	A	390	PRO
1	B	103	LYS
1	B	145	SER
1	B	177	ARG
1	B	188	LYS
1	B	202	ARG
1	B	205	LEU
1	B	207	ASP
1	B	210	LYS
1	B	246	ARG
1	B	338	ASP
1	B	363	ASP
1	B	378	ASP
1	C	145	SER
1	C	177	ARG
1	C	188	LYS
1	C	195	LYS
1	C	205	LEU
1	C	207	ASP
1	C	246	ARG
1	C	338	ASP
1	C	363	ASP
1	D	103	LYS
1	D	145	SER
1	D	177	ARG
1	D	188	LYS
1	D	205	LEU
1	D	207	ASP
1	D	246	ARG
1	D	335	ARG
1	D	338	ASP
1	D	363	ASP
1	D	372	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	378	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	127	ASN
1	B	127	ASN
1	C	127	ASN
1	C	157	HIS

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)

9 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	UDP	D	501	-	24,26,26	1.12	3 (12%)	37,40,40	1.57	5 (13%)
3	NCC	A	502	-	40,43,43	0.78	0	49,65,65	1.43	7 (14%)
2	UDP	A	501	-	24,26,26	1.11	2 (8%)	37,40,40	1.69	5 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NCC	B	502	-	40,43,43	0.73	0	49,65,65	1.58	8 (16%)
3	NCC	C	502	-	40,43,43	0.74	0	49,65,65	1.68	11 (22%)
2	UDP	C	501	-	24,26,26	1.20	3 (12%)	37,40,40	1.72	6 (16%)
2	UDP	B	501	-	24,26,26	1.17	3 (12%)	37,40,40	1.55	6 (16%)
3	NCC	D	502	-	40,43,43	0.79	0	49,65,65	1.38	5 (10%)
4	BM7	C	503	-	15,15,15	1.15	1 (6%)	21,21,21	3.66	4 (19%)

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	UDP	D	501	-	-	7/16/32/32	0/2/2/2
3	NCC	A	502	-	-	3/30/70/70	0/3/3/3
2	UDP	A	501	-	-	1/16/32/32	0/2/2/2
3	NCC	B	502	-	-	3/30/70/70	0/3/3/3
3	NCC	C	502	-	-	5/30/70/70	0/3/3/3
2	UDP	C	501	-	-	5/16/32/32	0/2/2/2
2	UDP	B	501	-	-	0/16/32/32	0/2/2/2
3	NCC	D	502	-	-	3/30/70/70	0/3/3/3
4	BM7	C	503	-	-	5/6/26/26	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	503	BM7	C1-C2	3.27	1.56	1.52
2	C	501	UDP	C2-N1	2.80	1.42	1.38
2	B	501	UDP	C4-N3	-2.52	1.34	1.38
2	A	501	UDP	C2-N1	2.38	1.42	1.38
2	A	501	UDP	C5-C4	-2.27	1.38	1.43
2	D	501	UDP	C2-N1	2.27	1.42	1.38
2	C	501	UDP	C4-N3	-2.25	1.34	1.38
2	C	501	UDP	C2-N3	-2.22	1.34	1.38
2	B	501	UDP	C2-N1	2.21	1.42	1.38
2	D	501	UDP	C2-N3	-2.21	1.34	1.38
2	D	501	UDP	C4-N3	-2.14	1.34	1.38
2	B	501	UDP	C5-C4	-2.08	1.39	1.43

All (57) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	503	BM7	C1-C2-N2	11.91	124.53	110.73
4	C	503	BM7	O5-C1-C2	9.98	119.55	109.52
3	B	502	NCC	C18-C17-N5	4.79	124.21	116.10
2	C	501	UDP	C4-N3-C2	-4.73	120.35	126.58
3	C	502	NCC	C4-N3-C2	4.54	120.94	116.34
3	C	502	NCC	C19-C16-N5	4.54	119.35	110.38
2	C	501	UDP	N3-C2-N1	4.49	120.85	114.89
2	A	501	UDP	C4-N3-C2	-4.47	120.69	126.58
2	D	501	UDP	N3-C2-N1	4.42	120.75	114.89
2	D	501	UDP	C4-N3-C2	-4.36	120.83	126.58
3	A	502	NCC	C19-C16-N5	4.35	118.99	110.38
3	C	502	NCC	C18-C17-N5	4.29	123.36	116.10
3	B	502	NCC	C19-C16-N5	4.28	118.84	110.38
2	B	501	UDP	C4-N3-C2	-4.25	120.97	126.58
3	B	502	NCC	C4-N3-C2	4.23	120.63	116.34
3	D	502	NCC	C19-C16-N5	4.04	118.37	110.38
2	B	501	UDP	C5-C4-N3	3.90	120.67	114.84
2	C	501	UDP	C5-C4-N3	3.80	120.53	114.84
2	B	501	UDP	N3-C2-N1	3.76	119.88	114.89
2	A	501	UDP	C5-C4-N3	3.75	120.45	114.84
3	A	502	NCC	C18-C17-N5	3.73	122.41	116.10
2	A	501	UDP	O4-C4-C5	-3.68	118.69	125.16
2	A	501	UDP	N3-C2-N1	3.66	119.75	114.89
3	A	502	NCC	C4-N3-C2	3.64	120.03	116.34
3	D	502	NCC	C18-C17-N5	3.63	122.25	116.10
3	D	502	NCC	C4-N3-C2	3.61	120.00	116.34
2	D	501	UDP	C5-C4-N3	3.51	120.10	114.84
4	C	503	BM7	O5-C5-C6	3.47	115.06	106.44
2	C	501	UDP	O4-C4-C5	-3.14	119.63	125.16
3	A	502	NCC	C16-N5-C17	3.10	130.72	123.18
3	C	502	NCC	C16-N5-C17	3.09	130.68	123.18
2	B	501	UDP	O4-C4-C5	-3.06	119.78	125.16
3	B	502	NCC	C16-N5-C17	3.06	130.62	123.18
2	D	501	UDP	O4-C4-C5	-2.98	119.92	125.16
3	D	502	NCC	C16-N5-C17	2.97	130.40	123.18
4	C	503	BM7	C4-C3-C2	2.83	114.49	110.34
3	C	502	NCC	C3'-C2'-C1'	2.74	105.11	100.98
2	A	501	UDP	C2'-C1'-N1	2.72	120.93	113.22
2	D	501	UDP	O3B-PB-O2B	2.72	118.04	107.64
3	C	502	NCC	O6-C12-C16	-2.68	107.17	109.78
3	B	502	NCC	O10-C17-C18	-2.63	117.17	122.06
3	B	502	NCC	O6-C12-C16	-2.58	107.26	109.78

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	502	NCC	C10-O6-C12	2.48	119.97	114.20
3	D	502	NCC	C3'-C2'-C1'	2.44	104.66	100.98
3	A	502	NCC	O10-C17-C18	-2.37	117.65	122.06
3	C	502	NCC	O10-C17-N5	-2.31	117.70	121.95
3	A	502	NCC	O6-C12-C16	-2.29	107.55	109.78
2	B	501	UDP	PA-O3A-PB	-2.26	125.06	132.83
3	C	502	NCC	N4-C4-N3	2.20	119.96	116.49
3	B	502	NCC	C3'-C2'-C1'	2.16	104.23	100.98
2	C	501	UDP	C2'-C1'-N1	2.14	119.27	113.22
3	A	502	NCC	C10-O6-C12	2.09	119.06	114.20
2	B	501	UDP	O3B-PB-O2B	2.08	115.59	107.64
3	C	502	NCC	C10-O6-C12	2.05	118.98	114.20
3	C	502	NCC	O1A-C9-C10	-2.05	117.11	123.70
3	C	502	NCC	O7-C13-C14	-2.03	103.92	108.81
2	C	501	UDP	O2-C2-N3	-2.02	117.74	121.50

There are no chirality outliers.

All (32) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	501	UDP	C5'-O5'-PA-O2A
2	C	501	UDP	C5'-O5'-PA-O3A
2	D	501	UDP	O4'-C4'-C5'-O5'
2	D	501	UDP	C5'-O5'-PA-O3A
3	A	502	NCC	C19-C16-N5-C17
3	B	502	NCC	C19-C16-N5-C17
3	C	502	NCC	C19-C16-N5-C17
3	D	502	NCC	C19-C16-N5-C17
2	C	501	UDP	O4'-C4'-C5'-O5'
2	D	501	UDP	C3'-C4'-C5'-O5'
4	C	503	BM7	C8-C7-N2-C2
3	A	502	NCC	O10-C17-N5-C16
3	A	502	NCC	C18-C17-N5-C16
3	B	502	NCC	O10-C17-N5-C16
3	B	502	NCC	C18-C17-N5-C16
3	C	502	NCC	O10-C17-N5-C16
3	C	502	NCC	C18-C17-N5-C16
3	D	502	NCC	O10-C17-N5-C16
3	D	502	NCC	C18-C17-N5-C16
2	C	501	UDP	C3'-C4'-C5'-O5'
4	C	503	BM7	O5-C5-C6-O6
3	C	502	NCC	O8-C14-C15-O9

Continued on next page...

Continued from previous page...

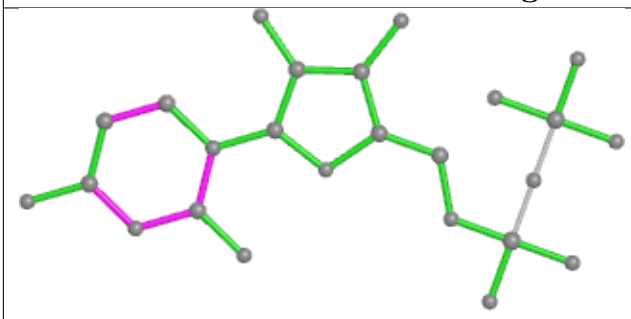
Mol	Chain	Res	Type	Atoms
2	D	501	UDP	PB-O3A-PA-O1A
4	C	503	BM7	O7-C7-N2-C2
2	A	501	UDP	C5'-O5'-PA-O3A
2	C	501	UDP	C5'-O5'-PA-O1A
2	D	501	UDP	C5'-O5'-PA-O1A
2	D	501	UDP	C5'-O5'-PA-O2A
3	C	502	NCC	C13-C14-C15-O9
4	C	503	BM7	C3-C2-N2-C7
2	D	501	UDP	PB-O3A-PA-O2A
4	C	503	BM7	C1-C2-N2-C7

There are no ring outliers.

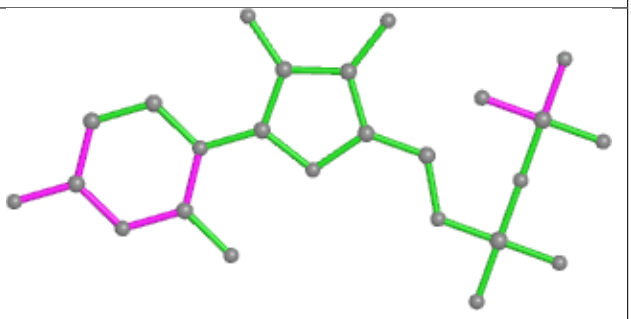
No monomer is involved in short contacts.

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.

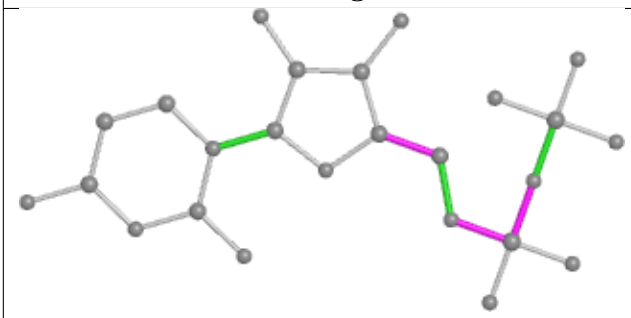
Ligand UDP D 501



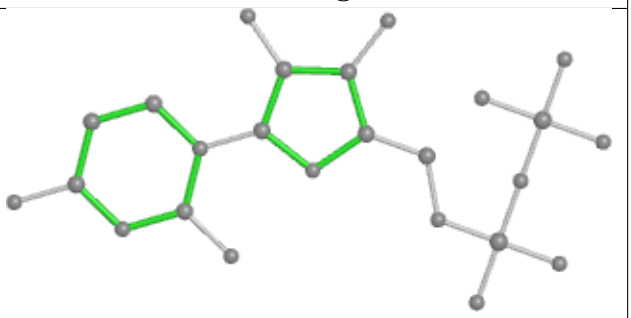
Bond lengths



Bond angles

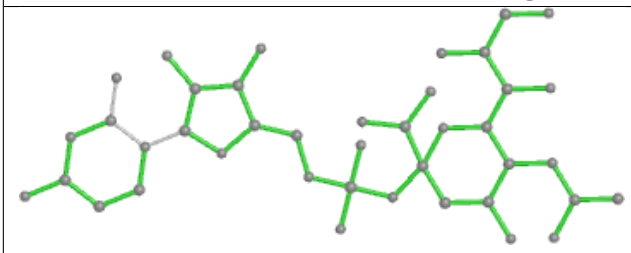


Torsions

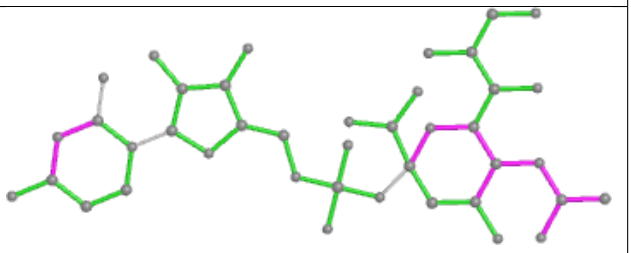


Rings

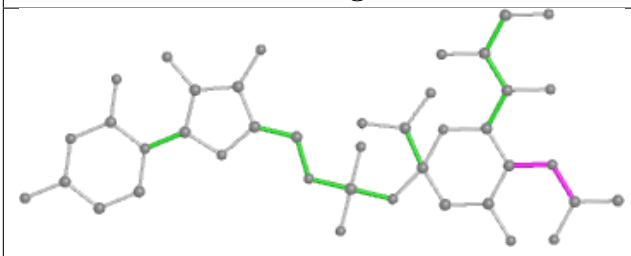
Ligand NCC A 502



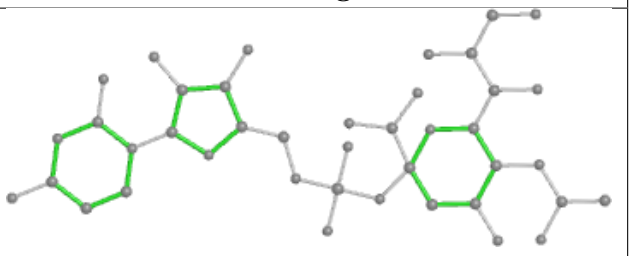
Bond lengths



Bond angles

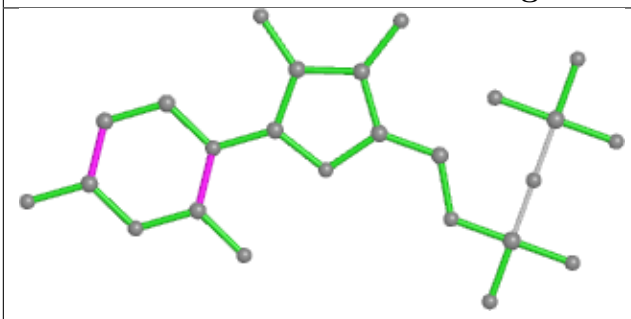


Torsions

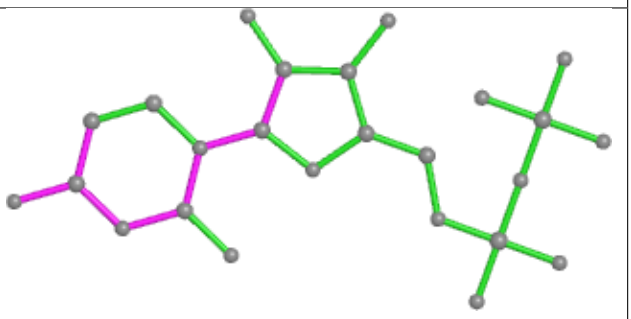


Rings

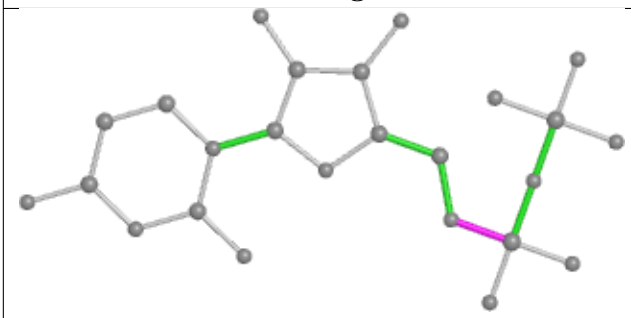
Ligand UDP A 501



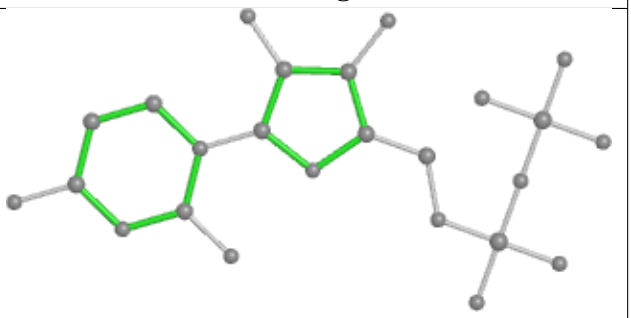
Bond lengths



Bond angles

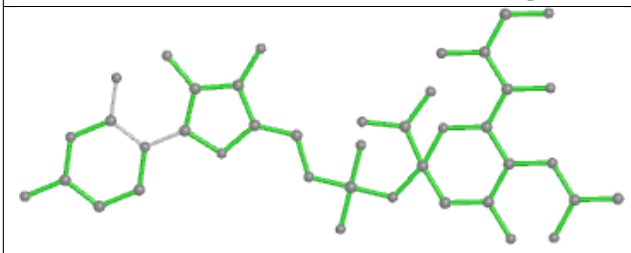


Torsions

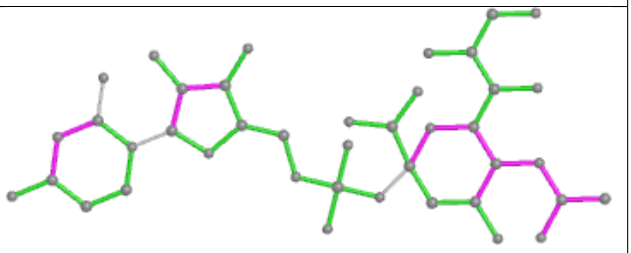


Rings

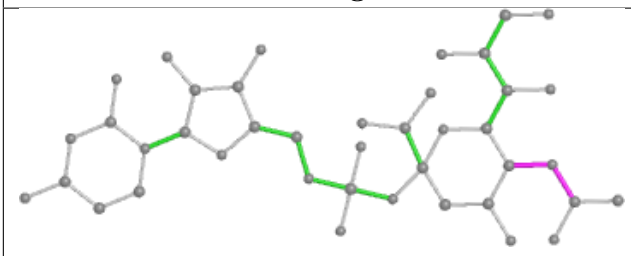
Ligand NCC B 502



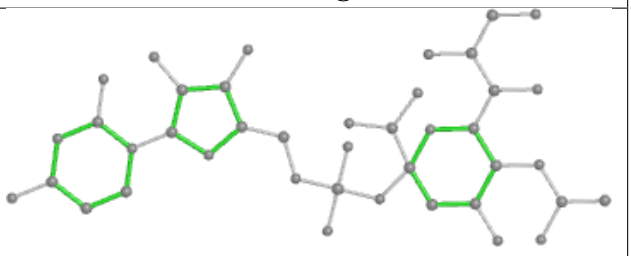
Bond lengths



Bond angles

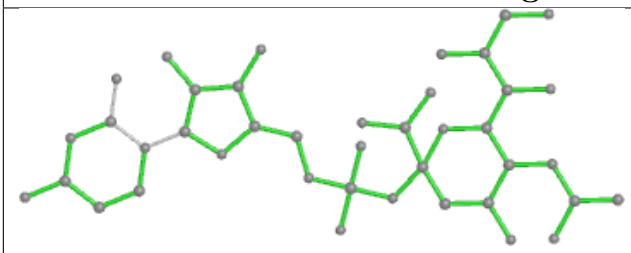


Torsions

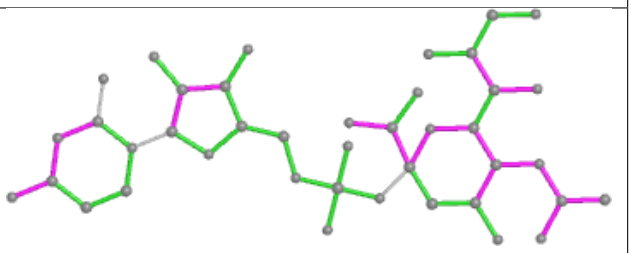


Rings

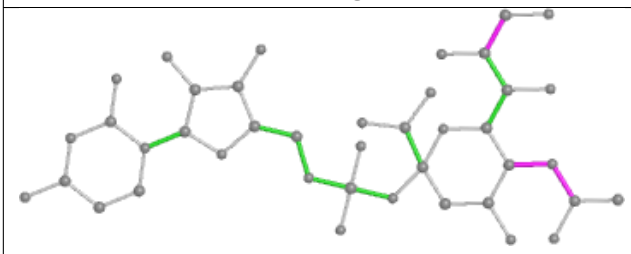
Ligand NCC C 502



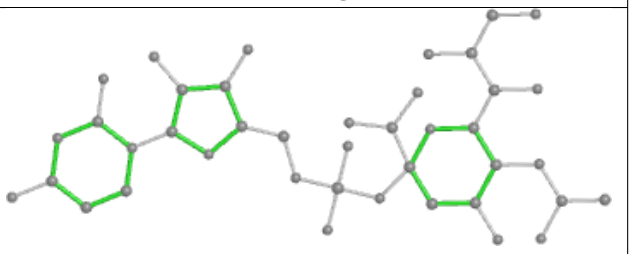
Bond lengths



Bond angles

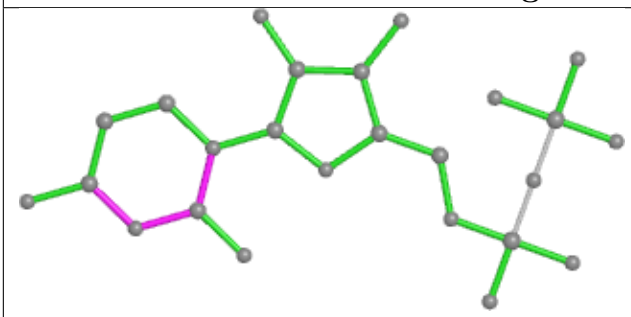


Torsions

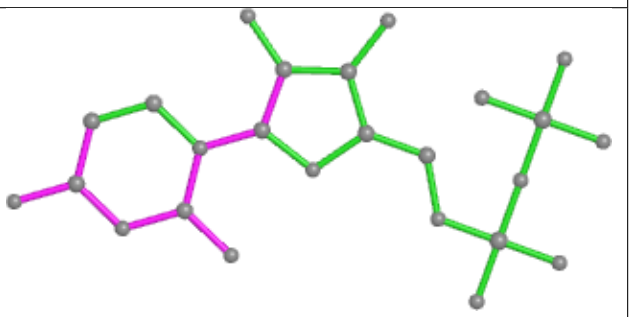


Rings

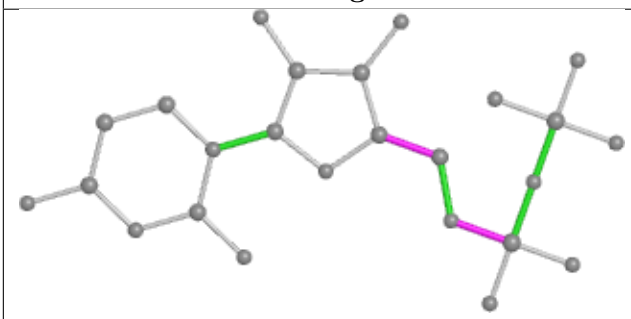
Ligand UDP C 501



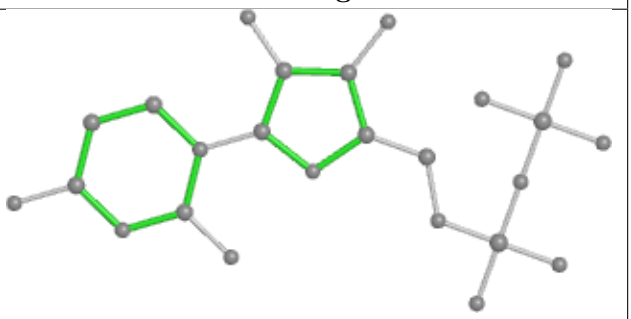
Bond lengths



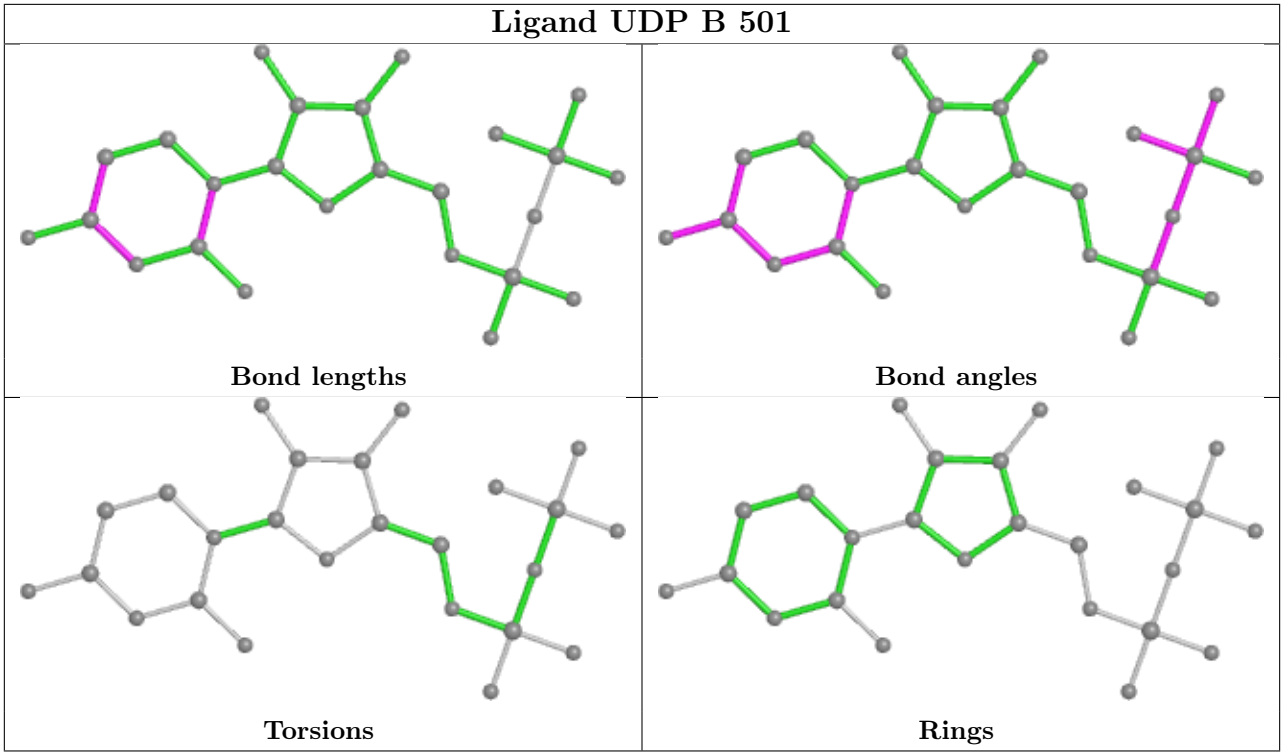
Bond angles

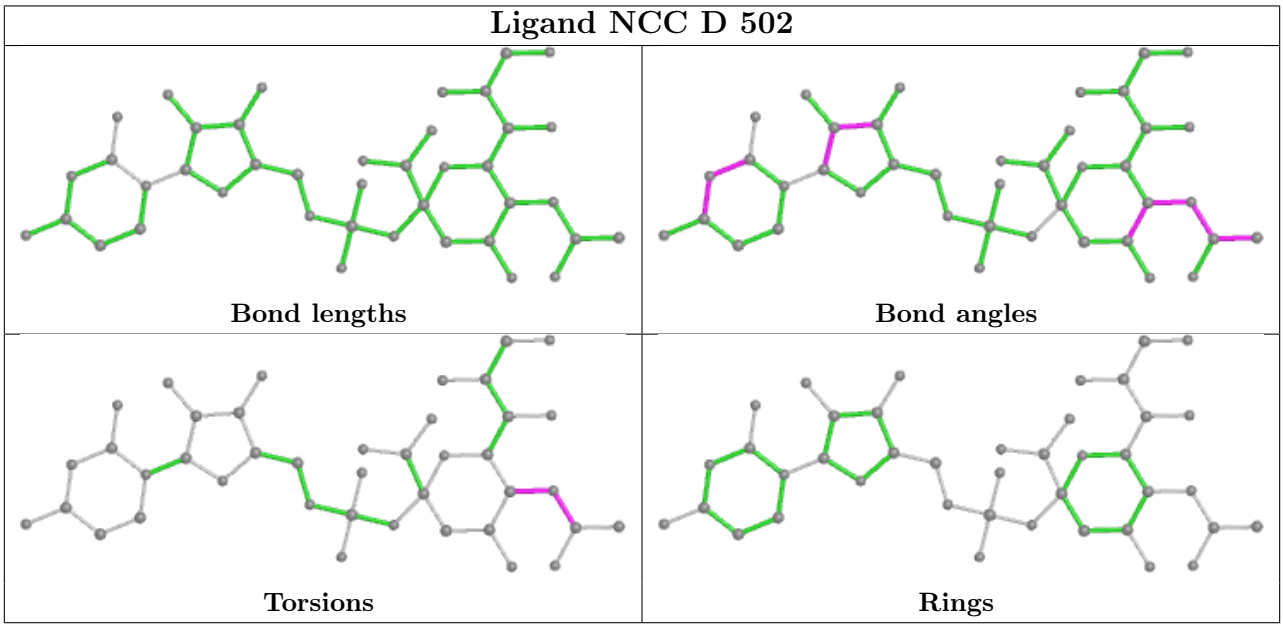


Torsions



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)

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	372/411 (90%)	-0.27	7 (1%)	66	69	47, 66, 106, 163	0
1	B	370/411 (90%)	0.02	17 (4%)	32	31	43, 78, 117, 165	0
1	C	371/411 (90%)	-0.30	7 (1%)	66	69	44, 63, 107, 150	0
1	D	371/411 (90%)	0.18	22 (5%)	22	21	50, 89, 135, 187	0
All	All	1484/1644 (90%)	-0.09	53 (3%)	42	42	43, 72, 122, 187	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	193	LYS	7.0
1	D	360	ILE	6.8
1	A	389	PHE	6.1
1	B	194	ASN	6.0
1	A	390	PRO	5.6
1	B	360	ILE	5.3
1	B	190	LEU	5.3
1	A	385	LYS	5.1
1	D	193	LYS	4.5
1	B	192	ALA	4.2
1	D	195	LYS	4.0
1	D	208	ASP	4.0
1	D	194	ASN	3.9
1	C	6	ASN	3.8
1	B	301	SER	3.8
1	A	388	CYS	3.7
1	D	299	GLY	3.7
1	C	193	LYS	3.5
1	A	7	ASN	3.5
1	C	272	HIS	3.5
1	B	298	ILE	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	7	ASN	3.3
1	D	191	SER	3.2
1	A	38	PHE	3.1
1	D	303	CYS	3.1
1	B	166	GLN	3.0
1	D	358	SER	2.9
1	D	8	ARG	2.9
1	B	299	GLY	2.9
1	B	303	CYS	2.9
1	D	327	THR	2.8
1	C	190	LEU	2.8
1	B	355	TYR	2.7
1	D	218	LEU	2.6
1	B	189	LEU	2.5
1	D	298	ILE	2.5
1	D	349	LEU	2.5
1	D	192	ALA	2.4
1	B	389	PHE	2.3
1	B	191	SER	2.3
1	D	38	PHE	2.3
1	B	337	ALA	2.3
1	B	8	ARG	2.3
1	A	372	LYS	2.3
1	D	162	ARG	2.2
1	C	381	GLU	2.2
1	D	390	PRO	2.2
1	D	207	ASP	2.2
1	C	194	ASN	2.2
1	D	354	GLN	2.1
1	B	162	ARG	2.1
1	D	389	PHE	2.0
1	D	304	GLY	2.0

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)

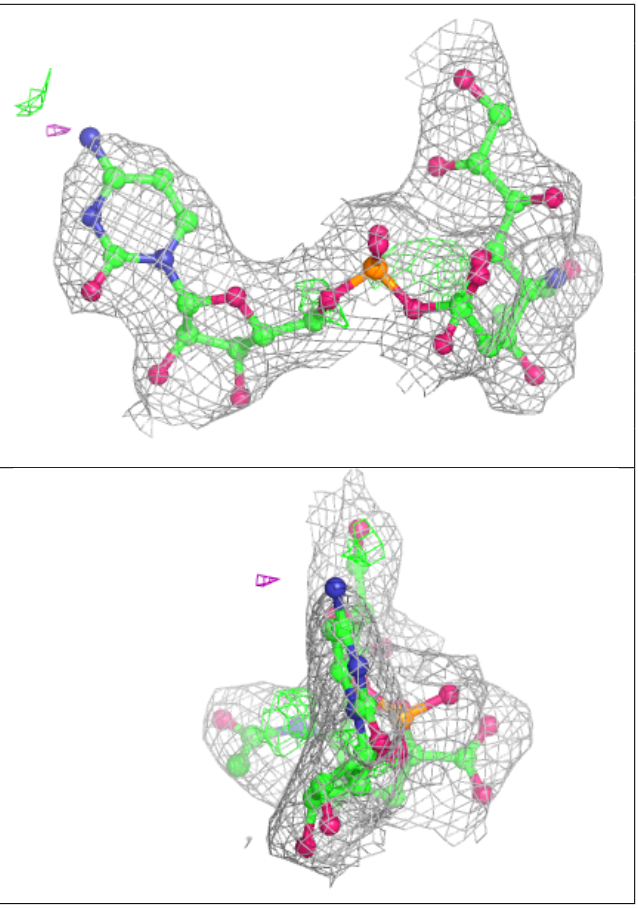
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	BM7	C	503	15/15	0.64	0.56	74,87,108,113	0
3	NCC	D	502	41/41	0.95	0.11	52,66,77,85	0
2	UDP	D	501	25/25	0.96	0.19	49,64,74,78	0
3	NCC	B	502	41/41	0.96	0.12	58,64,75,78	0
3	NCC	C	502	41/41	0.97	0.12	38,46,53,70	0
3	NCC	A	502	41/41	0.97	0.11	41,53,69,75	0
2	UDP	B	501	25/25	0.97	0.21	51,63,68,72	0
2	UDP	A	501	25/25	0.98	0.14	39,52,60,65	0
2	UDP	C	501	25/25	0.98	0.16	41,47,52,61	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.

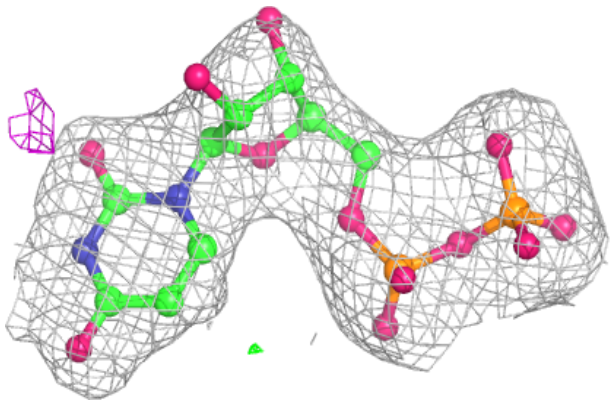
Electron density around NCC D 502:

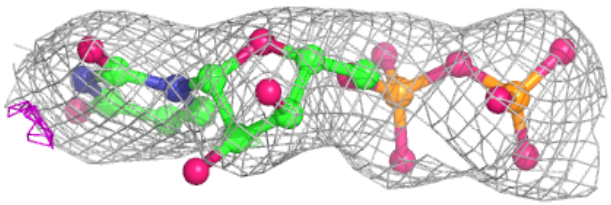
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

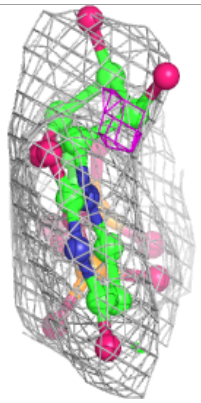


Electron density around UDP D 501:

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

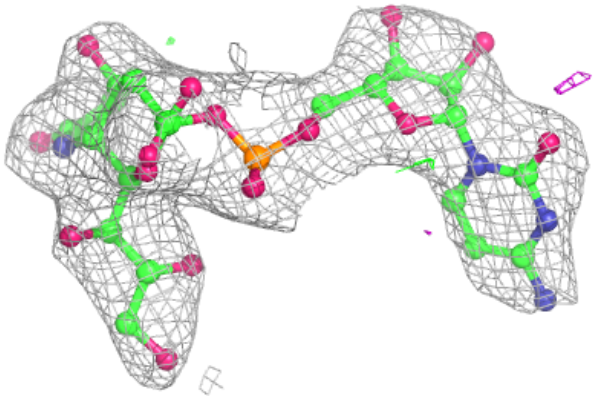


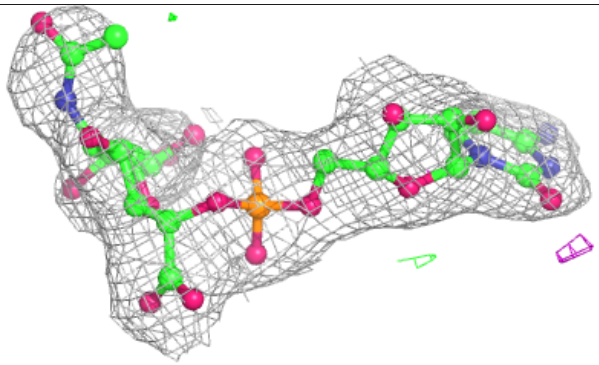


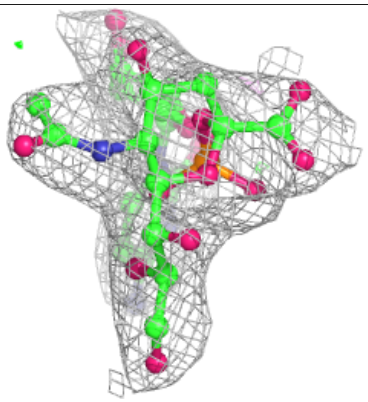


Electron density around NCC B 502:

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

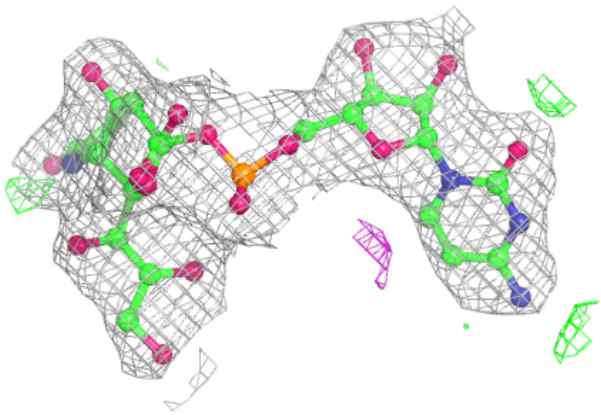


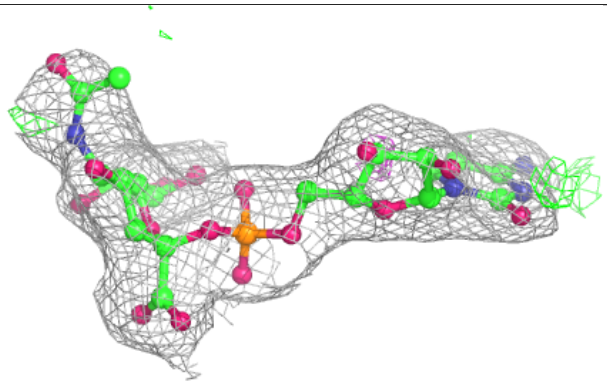


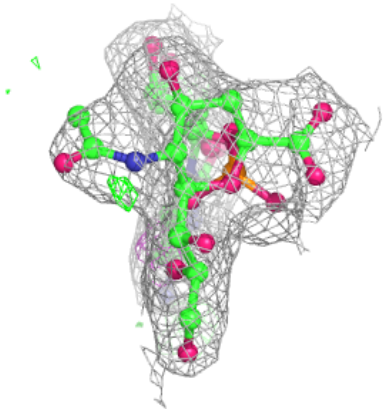


Electron density around NCC C 502:

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

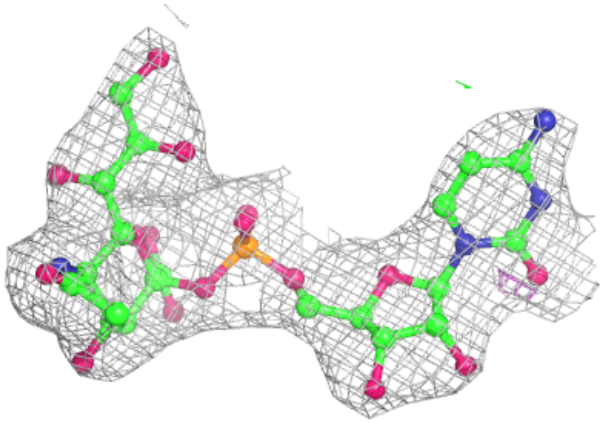


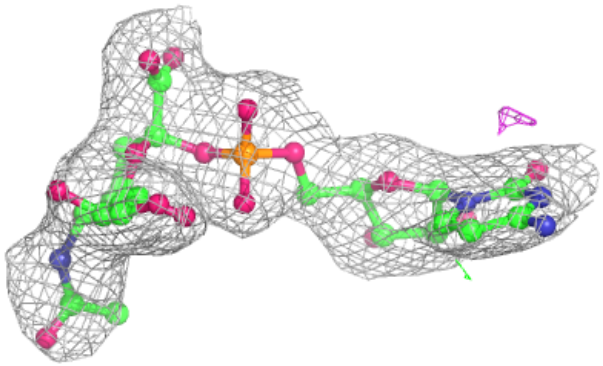


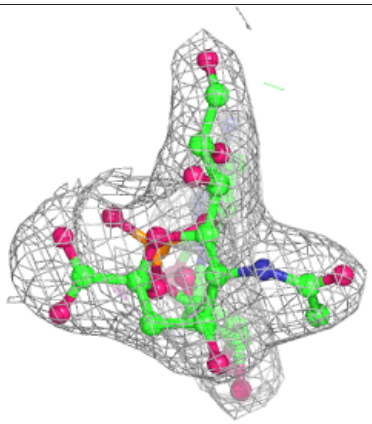


Electron density around NCC A 502:

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

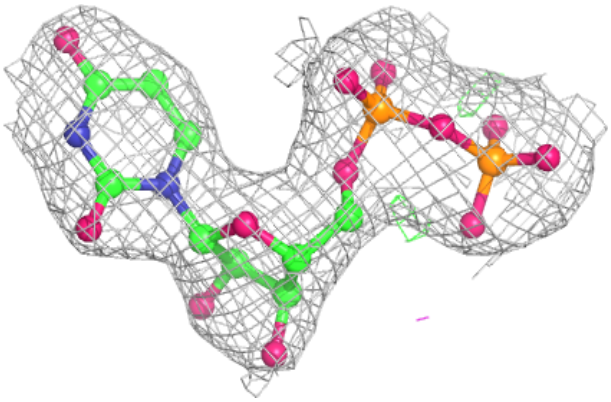


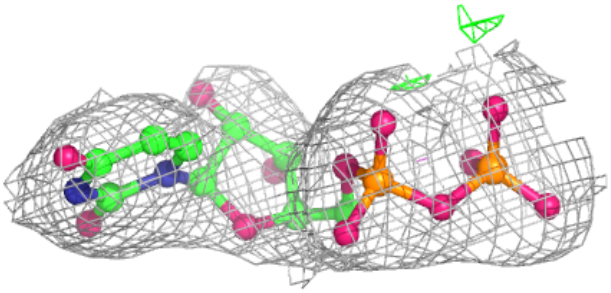


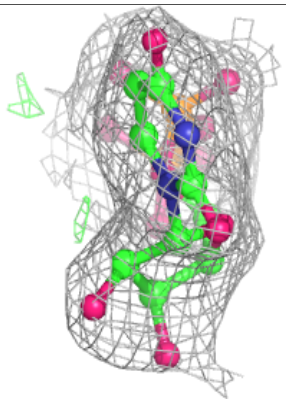


Electron density around UDP B 501:

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

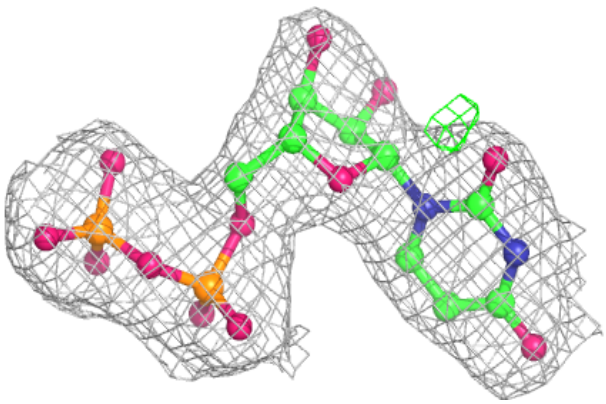


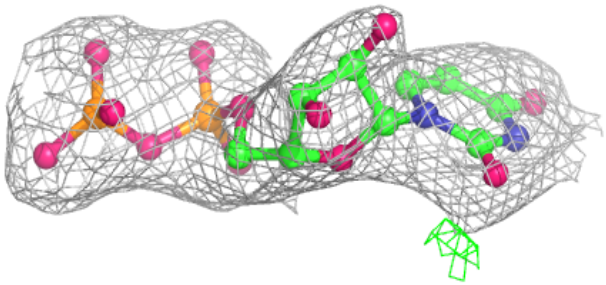


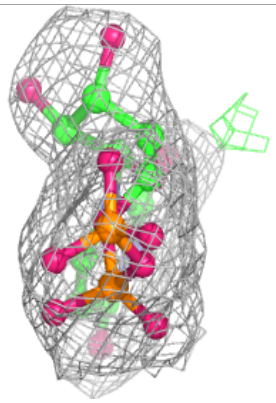


Electron density around UDP A 501:

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

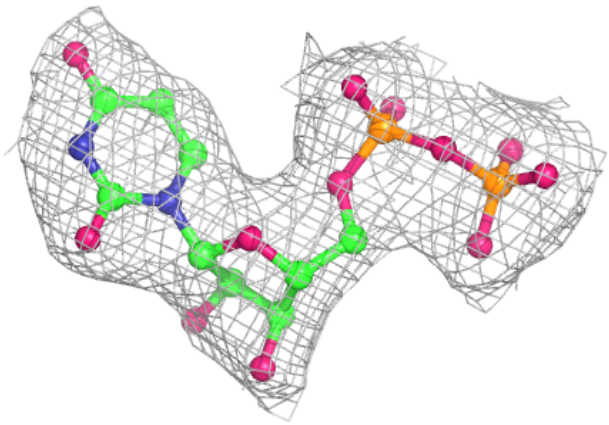


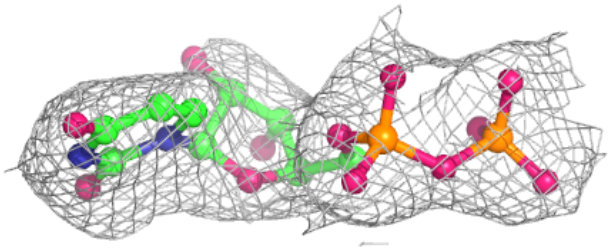


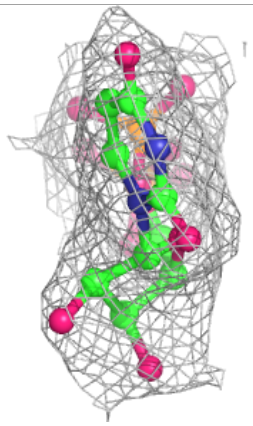


Electron density around UDP C 501:

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







6.5 Other polymers (i)

There are no such residues in this entry.