WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Sep 25, 2023 – 05:48 AM EDT

PDB ID	:	5W0N
Title	:	Structure of human TUT7 catalytic module (CM) in complex with UMPNPP and U2 RNA
Authors	:	Faehnle, C.R.; Walleshauser, J.; Joshua-Tor, L.
Deposited on	:	2017-05-31
Resolution	:	2.50 Å(reported)

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

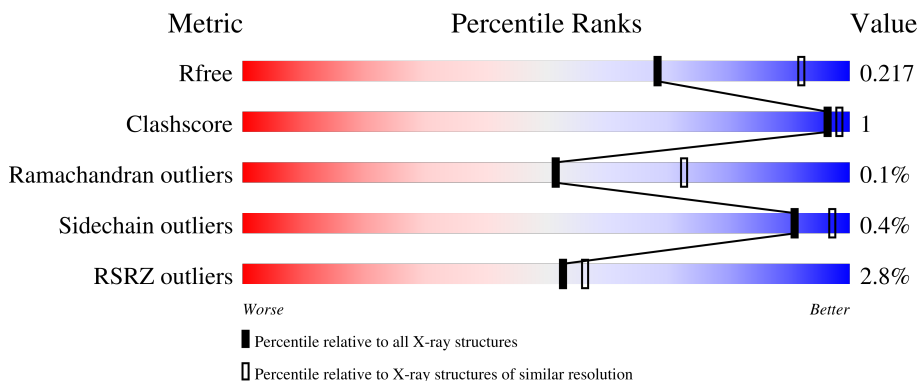|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.50 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	Percentile Ranks	Value
Rfree		0.217
Clashscore		1
Ramachandran outliers		0.1%
Sidechain outliers		0.4%
RSRZ outliers		2.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	403	4% 92% · 6%
1	B	403	2% 89% · 8%
1	C	403	% 84% · 13%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 18407 atoms, of which 9042 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 Terminal uridylyltransferase 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	379	Total 6163	C 1990	H 3085	N 515	O 553	S 20	0	0	0
1	B	372	Total 6025	C 1949	H 3010	N 503	O 543	S 20	0	0	0
1	C	350	Total 5688	C 1841	H 2845	N 470	O 518	S 14	0	0	0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Mg 1	0	0
2	B	1	Total 1	Mg 1	0	0
2	C	1	Total 1	Mg 1	0	0

- Molecule 3 is 5'-O-[(S)-hydroxy{[(S)-hydroxy(phosphonooxy)phosphoryl]amino}phosphoryl] uridine (three-letter code: 2KH) (formula: $C_9H_{16}N_3O_{14}P_3$).

2KH
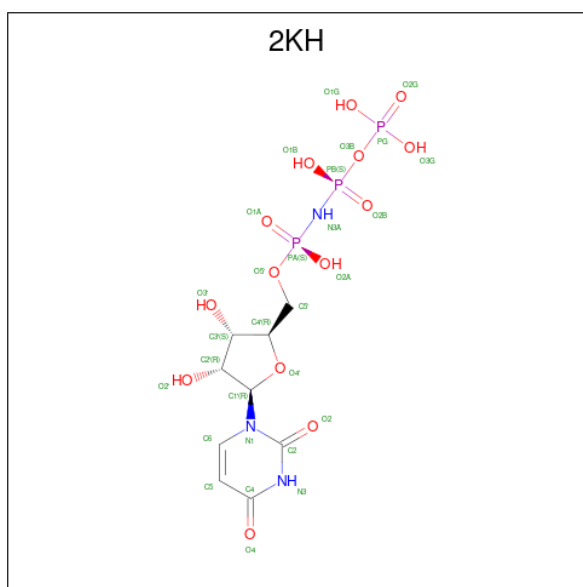
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	P	0	0
			41	9	12	3	14	3		
3	B	1	Total	C	H	N	O	P	0	0
			41	9	12	3	14	3		
3	C	1	Total	C	H	N	O	P	0	0
			41	9	12	3	14	3		

- Molecule 4 is [(2R,3S,4R,5R)-5-(2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL)-3,4 -DIHYDROXYTETRAHYDROFURAN-2-YL]METHYL (2R,3S,4R,5R)-5-(2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL)-4-HYDROXY-2-(HYDROXYMETHYL)TETRA HYDROFURAN-3-YL HYDROGEN (S)-PHOSPHATE (three-letter code: UPU) (formula: $C_{18}H_{23}N_4O_{14}P$).

UPU
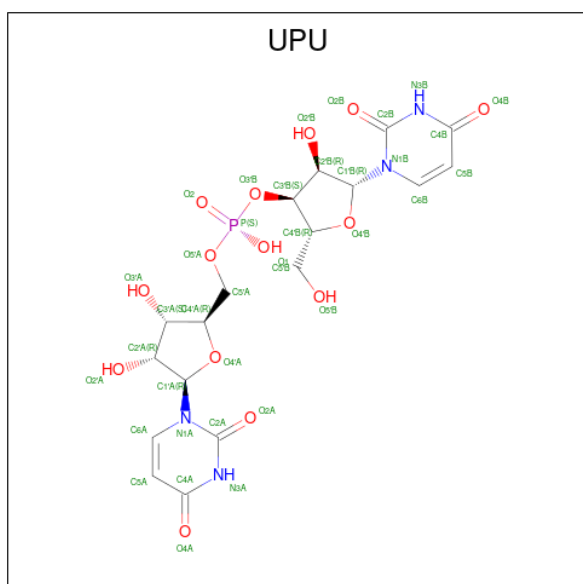
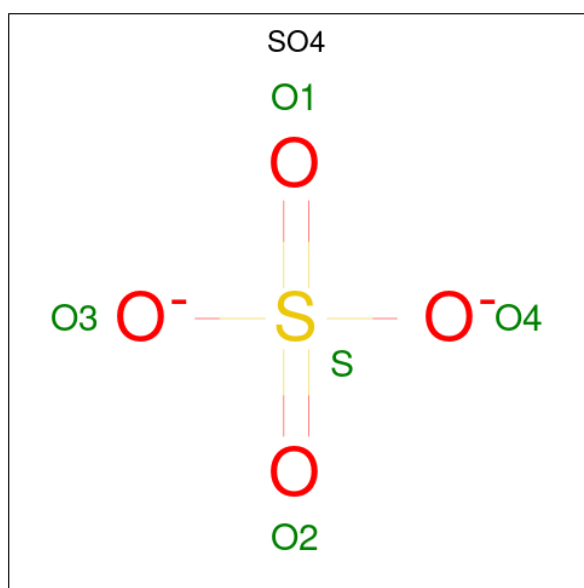| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	P	0	0
			59	18	22	4	14	1		
4	B	1	Total	C	H	N	O	P	0	0
			59	18	22	4	14	1		
4	C	1	Total	C	H	N	O	P	0	0
			59	18	22	4	14	1		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Zn	0	0
			1	1		
5	B	1	Total	Zn	0	0
			1	1		

- Molecule 6 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	I	0	0
			1	1		
6	B	1	Total	I	0	0
			1	1		
6	C	1	Total	I	0	0
			1	1		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 5	O 4	S 1	0	0

- Molecule 8 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	80	Total 80	O 80	0	0
8	B	88	Total 88	O 88	0	0
8	C	50	Total 50	O 50	0	0

3 Residue-property plots ⓘ

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.
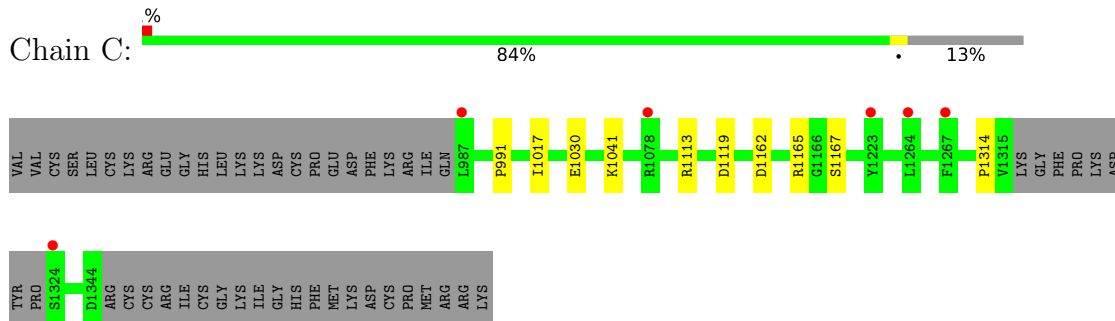
- Molecule 1: Terminal uridylyltransferase 7
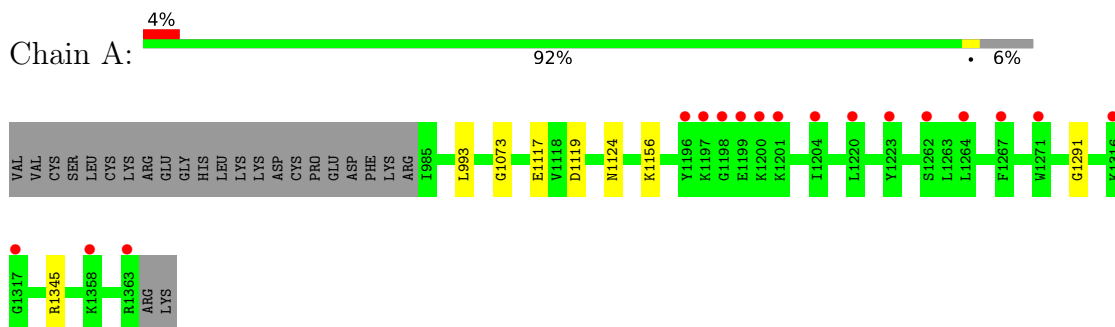
Chain A:

- Molecule 1: Terminal uridylyltransferase 7

Chain B:

- Molecule 1: Terminal uridylyltransferase 7

Chain C:
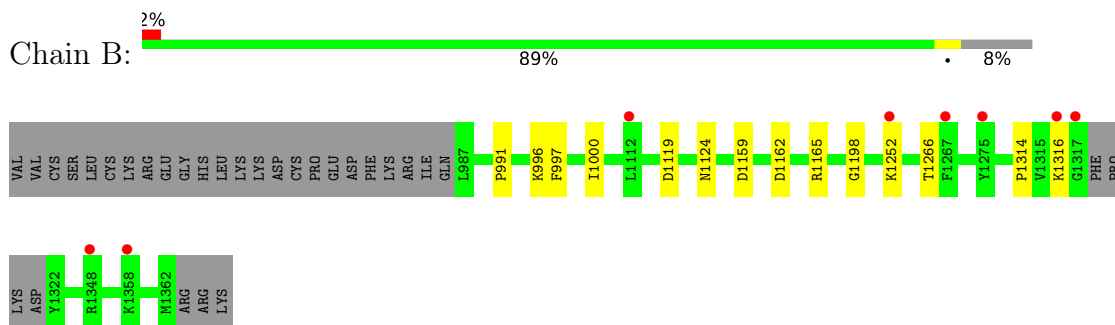
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	135.97Å 135.97Å 181.57Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	58.88 – 2.50 71.90 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (58.88-2.50) 99.6 (71.90-2.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.13 (at 2.48Å)	Xtriage
Refinement program	PHENIX 1.10_2155	Depositor
R, R_{free}	0.187 , 0.215 0.188 , 0.217	Depositor DCC
R_{free} test set	3252 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	48.7	Xtriage
Anisotropy	0.501	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 49.6	EDS
L-test for twinning [2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.036 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	18407	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% 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: UPU, ZN, IOD, SO4, 2KH, MG

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	0.26	0/3154	0.43	0/4262
1	B	0.26	0/3088	0.44	0/4172
1	C	0.26	0/2911	0.42	0/3938
All	All	0.26	0/9153	0.43	0/12372

There are no bond length outliers.

There are no bond angle outliers.

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	3078	3085	3091	4	1
1	B	3015	3010	3025	10	0
1	C	2843	2845	2851	5	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	29	12	13	0	0
3	B	29	12	15	0	0
3	C	29	12	14	0	0
4	A	37	22	21	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	37	22	21	2	0
4	C	37	22	21	1	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
7	A	5	0	0	0	0
8	A	80	0	0	0	0
8	B	88	0	0	0	0
8	C	50	0	0	0	0
All	All	9365	9042	9072	19	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (19) 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:C:1162:ASP:OD2	1:C:1165:ARG:NH1	2.21	0.73
1:B:996:LYS:HB3	1:B:1316:LYS:HE3	1.75	0.67
1:B:1124:ASN:ND2	4:B:1403:UPU:O2A	2.29	0.66
1:B:1162:ASP:OD2	1:B:1165:ARG:NH1	2.31	0.63
1:B:991:PRO:HG2	1:B:1314:PRO:HD2	1.85	0.58
1:A:1117:GLU:OE1	1:C:1113:ARG:NH2	2.36	0.52
1:B:1159:ASP:OD1	1:B:1165:ARG:NH2	2.43	0.52
1:B:1316:LYS:HG2	1:B:1316:LYS:O	2.12	0.50
1:C:991:PRO:HG2	1:C:1314:PRO:HD2	1.93	0.50
1:B:996:LYS:C	1:B:1316:LYS:HE3	2.34	0.48
1:B:1000:ILE:HD12	1:B:1316:LYS:HE2	1.97	0.46
1:B:997:PHE:CG	1:B:1314:PRO:HB3	2.51	0.46
1:C:1030:GLU:OE2	1:C:1041:LYS:HD2	2.17	0.45
4:B:1403:UPU:H6A	4:B:1403:UPU:O5'A	2.17	0.44
1:A:1124:ASN:ND2	4:A:1403:UPU:O2A	2.49	0.44
4:C:1403:UPU:H6A	4:C:1403:UPU:O5'A	2.18	0.44
1:B:1252:LYS:O	1:B:1266:THR:HG21	2.18	0.44
1:A:1291:GLY:O	1:A:1345:ARG:NH1	2.52	0.42
1:A:1156:LYS:HD2	1:C:1017:ILE:HD12	2.03	0.41

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:993:LEU:H	1:A:1073:GLY:O[5_454]	1.58	0.02

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	377/403 (94%)	368 (98%)	9 (2%)	0	100	100
1	B	368/403 (91%)	358 (97%)	9 (2%)	1 (0%)	41	61
1	C	346/403 (86%)	338 (98%)	8 (2%)	0	100	100
All	All	1091/1209 (90%)	1064 (98%)	26 (2%)	1 (0%)	51	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1198	GLY

5.3.2 Protein sidechains ⓘ

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	344/367 (94%)	343 (100%)	1 (0%)	92	97
1	B	337/367 (92%)	336 (100%)	1 (0%)	92	97
1	C	318/367 (87%)	316 (99%)	2 (1%)	86	95
All	All	999/1101 (91%)	995 (100%)	4 (0%)	91	97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1119	ASP
1	B	1119	ASP
1	C	1119	ASP
1	C	1167	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1129	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 15 ligands modelled in this entry, 8 are monoatomic - leaving 7 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	$\#|Z| > 2$
7	SO4	A	1406	-	4,4,4	0.14	0	6,6,6	0.05	0
4	UPU	B	1403	-	40,40,40	4.58	27 (67%)	59,60,60	1.67	11 (18%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
3	2KH	B	1402	2	29,30,30	4.04	21 (72%)	42,47,47	1.63	7 (16%)
4	UPU	A	1403	-	40,40,40	4.55	28 (70%)	59,60,60	1.65	11 (18%)
3	2KH	C	1402	2	29,30,30	4.01	19 (65%)	42,47,47	1.60	7 (16%)
4	UPU	C	1403	-	40,40,40	4.56	28 (70%)	59,60,60	1.62	10 (16%)
3	2KH	A	1402	2	29,30,30	4.05	20 (68%)	42,47,47	1.58	7 (16%)

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
4	UPU	B	1403	-	-	2/21/53/53	0/4/4/4
3	2KH	B	1402	2	-	3/19/38/38	0/2/2/2
4	UPU	A	1403	-	-	2/21/53/53	0/4/4/4
3	2KH	C	1402	2	-	6/19/38/38	0/2/2/2
4	UPU	C	1403	-	-	0/21/53/53	0/4/4/4
3	2KH	A	1402	2	-	6/19/38/38	0/2/2/2

All (143) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1403	UPU	C2'B-C3'B	-13.52	1.22	1.52
4	B	1403	UPU	C2'B-C3'B	-13.45	1.23	1.52
4	C	1403	UPU	C2'B-C3'B	-13.37	1.23	1.52
3	A	1402	2KH	C3'-C4'	-9.34	1.29	1.53
3	B	1402	2KH	C3'-C4'	-9.34	1.29	1.53
3	C	1402	2KH	C3'-C4'	-9.27	1.29	1.53
4	B	1403	UPU	O4'A-C1'A	8.35	1.61	1.42
4	C	1403	UPU	O4'A-C1'A	8.25	1.61	1.42
4	A	1403	UPU	O4'A-C1'A	8.16	1.61	1.42
4	C	1403	UPU	C2B-N1B	7.69	1.50	1.38
4	B	1403	UPU	C2B-N1B	7.59	1.50	1.38
4	A	1403	UPU	C2B-N1B	7.42	1.50	1.38
3	C	1402	2KH	C2-N1	7.11	1.49	1.38
4	A	1403	UPU	C2A-N3A	6.98	1.50	1.38
3	A	1402	2KH	C2-N1	6.96	1.49	1.38
4	B	1403	UPU	C2'A-C1'A	-6.96	1.31	1.53
4	B	1403	UPU	C2A-N3A	6.95	1.50	1.38
4	C	1403	UPU	C2'A-C1'A	-6.92	1.31	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1403	UPU	C2'A-C1'A	-6.82	1.31	1.53
4	A	1403	UPU	C2A-N1A	6.80	1.49	1.38
4	B	1403	UPU	C2A-N1A	6.77	1.49	1.38
3	B	1402	2KH	C2-N1	6.76	1.49	1.38
4	C	1403	UPU	C2A-N3A	6.65	1.49	1.38
3	B	1402	2KH	C2'-C1'	-6.64	1.32	1.53
3	A	1402	2KH	C2'-C1'	-6.62	1.32	1.53
3	C	1402	2KH	C2-N3	6.58	1.49	1.38
3	C	1402	2KH	C2'-C1'	-6.58	1.32	1.53
4	C	1403	UPU	C2A-N1A	6.56	1.49	1.38
3	B	1402	2KH	C6-C5	6.46	1.50	1.35
3	C	1402	2KH	C6-C5	6.45	1.50	1.35
3	B	1402	2KH	C2-N3	6.44	1.49	1.38
4	B	1403	UPU	C6A-C5A	6.44	1.50	1.35
3	A	1402	2KH	C2-N3	6.43	1.49	1.38
3	A	1402	2KH	C6-C5	6.39	1.49	1.35
4	C	1403	UPU	C6A-C5A	6.35	1.49	1.35
4	C	1403	UPU	C6B-C5B	6.35	1.49	1.35
4	B	1403	UPU	C2B-N3B	6.34	1.49	1.38
4	A	1403	UPU	C6B-C5B	6.34	1.49	1.35
4	A	1403	UPU	C6A-C5A	6.33	1.49	1.35
4	B	1403	UPU	C6B-C5B	6.31	1.49	1.35
4	C	1403	UPU	C2B-N3B	6.30	1.49	1.38
4	B	1403	UPU	O4'A-C4'A	-6.27	1.31	1.45
4	A	1403	UPU	O4'A-C4'A	-6.19	1.31	1.45
4	C	1403	UPU	O4'A-C4'A	-6.13	1.31	1.45
4	A	1403	UPU	C2B-N3B	6.11	1.48	1.38
4	C	1403	UPU	O4'B-C1'B	-5.89	1.28	1.42
4	B	1403	UPU	O4'B-C1'B	-5.59	1.28	1.42
4	A	1403	UPU	O4'B-C1'B	-5.55	1.28	1.42
3	A	1402	2KH	O4'-C4'	5.20	1.56	1.45
3	B	1402	2KH	O4'-C4'	5.17	1.56	1.45
3	C	1402	2KH	O4'-C4'	5.14	1.56	1.45
3	A	1402	2KH	PB-O3B	5.10	1.65	1.59
4	B	1403	UPU	C4B-N3B	4.99	1.47	1.38
3	C	1402	2KH	PB-O3B	4.95	1.65	1.59
3	B	1402	2KH	PB-O3B	4.89	1.65	1.59
4	C	1403	UPU	C4B-N3B	4.89	1.47	1.38
4	A	1403	UPU	C4B-N3B	4.78	1.47	1.38
3	A	1402	2KH	PA-O1A	4.61	1.53	1.46
3	C	1402	2KH	O4'-C1'	4.46	1.52	1.42
4	A	1403	UPU	C4A-N3A	4.41	1.46	1.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1402	2KH	C4-N3	4.36	1.46	1.38
3	B	1402	2KH	O4'-C1'	4.33	1.52	1.42
3	A	1402	2KH	O4'-C1'	4.31	1.52	1.42
4	B	1403	UPU	C4A-N3A	4.30	1.46	1.38
4	C	1403	UPU	C4A-N3A	4.27	1.46	1.38
3	A	1402	2KH	C4-N3	4.18	1.46	1.38
3	B	1402	2KH	C4-N3	4.10	1.45	1.38
4	B	1403	UPU	C2'B-C1'B	4.08	1.66	1.53
4	A	1403	UPU	C2'B-C1'B	4.04	1.66	1.53
4	C	1403	UPU	C2'B-C1'B	4.04	1.66	1.53
4	A	1403	UPU	C6A-N1A	3.93	1.47	1.38
4	B	1403	UPU	C6A-N1A	3.89	1.47	1.38
4	B	1403	UPU	C6B-N1B	3.87	1.47	1.38
4	A	1403	UPU	C6B-N1B	3.86	1.47	1.38
4	C	1403	UPU	C6B-N1B	3.84	1.47	1.38
4	C	1403	UPU	C6A-N1A	3.82	1.47	1.38
3	A	1402	2KH	PA-N3A	3.82	1.73	1.63
3	C	1402	2KH	PA-N3A	3.80	1.73	1.63
4	A	1403	UPU	C5B-C4B	3.77	1.52	1.43
4	C	1403	UPU	C5B-C4B	3.75	1.51	1.43
4	B	1403	UPU	C5B-C4B	3.73	1.51	1.43
3	B	1402	2KH	PA-N3A	3.63	1.72	1.63
3	C	1402	2KH	PA-O1A	3.57	1.51	1.46
3	B	1402	2KH	PB-O2B	3.56	1.51	1.46
4	C	1403	UPU	C5'B-C4'B	-3.48	1.40	1.51
3	B	1402	2KH	PA-O1A	3.44	1.51	1.46
3	B	1402	2KH	O2-C2	-3.44	1.16	1.23
3	A	1402	2KH	O2-C2	-3.42	1.16	1.23
4	B	1403	UPU	C5'B-C4'B	-3.39	1.40	1.51
4	A	1403	UPU	C5'B-C4'B	-3.38	1.40	1.51
3	C	1402	2KH	C6-N1	3.31	1.46	1.38
3	C	1402	2KH	C5-C4	3.29	1.50	1.43
3	B	1402	2KH	C5-C4	3.28	1.50	1.43
3	A	1402	2KH	C6-N1	3.28	1.45	1.38
3	A	1402	2KH	C5-C4	3.24	1.50	1.43
3	B	1402	2KH	C6-N1	3.21	1.45	1.38
3	C	1402	2KH	O2-C2	-3.13	1.17	1.23
3	B	1402	2KH	C3'-C2'	3.09	1.61	1.53
3	C	1402	2KH	O4-C4	-3.09	1.18	1.24
3	A	1402	2KH	O4-C4	-3.07	1.18	1.24
3	C	1402	2KH	C3'-C2'	3.06	1.61	1.53
4	C	1403	UPU	O4A-C4A	-3.05	1.18	1.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1402	2KH	C3'-C2'	3.04	1.61	1.53
4	C	1403	UPU	O4'B-C4'B	3.03	1.51	1.45
3	B	1402	2KH	PB-O1B	-3.00	1.48	1.56
3	B	1402	2KH	O4-C4	-2.99	1.18	1.24
4	A	1403	UPU	O4'B-C4'B	2.98	1.51	1.45
4	B	1403	UPU	O4A-C4A	-2.96	1.18	1.24
4	B	1403	UPU	O4'B-C4'B	2.95	1.51	1.45
4	A	1403	UPU	O4A-C4A	-2.91	1.18	1.24
4	B	1403	UPU	C5A-C4A	2.90	1.50	1.43
4	B	1403	UPU	O3'A-C3'A	-2.90	1.36	1.43
4	A	1403	UPU	C5A-C4A	2.85	1.50	1.43
4	C	1403	UPU	O3'A-C3'A	-2.79	1.36	1.43
4	C	1403	UPU	C5A-C4A	2.78	1.49	1.43
4	C	1403	UPU	O2'A-C2'A	2.77	1.49	1.43
4	A	1403	UPU	O3'A-C3'A	-2.73	1.36	1.43
4	A	1403	UPU	O2'A-C2'A	2.71	1.49	1.43
4	B	1403	UPU	O2'A-C2'A	2.69	1.49	1.43
3	C	1402	2KH	PA-O5'	2.56	1.67	1.57
3	A	1402	2KH	PA-O5'	2.54	1.67	1.57
4	C	1403	UPU	O2'B-C2'B	2.53	1.48	1.43
3	B	1402	2KH	PA-O5'	2.52	1.67	1.57
4	A	1403	UPU	O4B-C4B	-2.51	1.19	1.24
4	A	1403	UPU	O2'B-C2'B	2.48	1.48	1.43
4	B	1403	UPU	O2'B-C2'B	2.48	1.48	1.43
4	B	1403	UPU	O4B-C4B	-2.45	1.19	1.24
3	B	1402	2KH	O3'-C3'	2.44	1.48	1.43
4	C	1403	UPU	O4B-C4B	-2.41	1.19	1.24
4	B	1403	UPU	P-O5'A	2.41	1.69	1.59
4	C	1403	UPU	P-O5'A	2.39	1.69	1.59
3	C	1402	2KH	O3'-C3'	2.36	1.48	1.43
4	A	1403	UPU	P-O5'A	2.31	1.68	1.59
4	C	1403	UPU	C3'B-C4'B	2.30	1.59	1.52
3	C	1402	2KH	PB-N3A	2.30	1.69	1.63
3	A	1402	2KH	O3'-C3'	2.29	1.48	1.43
3	A	1402	2KH	PB-N3A	2.27	1.69	1.63
3	B	1402	2KH	PB-N3A	2.19	1.69	1.63
4	A	1403	UPU	C3'B-C4'B	2.18	1.58	1.52
4	B	1403	UPU	P-O3'B	2.16	1.66	1.60
4	A	1403	UPU	O3'B-C3'B	2.11	1.51	1.44
3	A	1402	2KH	C5'-C4'	2.11	1.58	1.51
4	C	1403	UPU	C5'A-C4'A	2.05	1.58	1.51

All (53) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1403	UPU	C4A-N3A-C2A	-5.10	119.86	126.58
4	B	1403	UPU	C4B-N3B-C2B	-5.10	119.86	126.58
4	A	1403	UPU	C4B-N3B-C2B	-5.09	119.86	126.58
3	B	1402	2KH	C4-N3-C2	-5.01	119.97	126.58
3	C	1402	2KH	C4-N3-C2	-4.97	120.02	126.58
4	A	1403	UPU	C4A-N3A-C2A	-4.96	120.03	126.58
3	A	1402	2KH	C4-N3-C2	-4.91	120.10	126.58
4	C	1403	UPU	C4A-N3A-C2A	-4.91	120.11	126.58
4	C	1403	UPU	C4B-N3B-C2B	-4.89	120.13	126.58
4	B	1403	UPU	N3A-C2A-N1A	4.27	120.56	114.89
4	A	1403	UPU	N3B-C2B-N1B	4.17	120.42	114.89
4	B	1403	UPU	N3B-C2B-N1B	4.03	120.23	114.89
3	B	1402	2KH	N3-C2-N1	3.98	120.17	114.89
4	C	1403	UPU	N3B-C2B-N1B	3.97	120.16	114.89
4	C	1403	UPU	N3A-C2A-N1A	3.91	120.08	114.89
3	A	1402	2KH	N3-C2-N1	3.87	120.03	114.89
4	A	1403	UPU	N3A-C2A-N1A	3.75	119.86	114.89
3	C	1402	2KH	N3-C2-N1	3.71	119.81	114.89
4	A	1403	UPU	C5A-C4A-N3A	3.56	120.17	114.84
3	C	1402	2KH	C5-C4-N3	3.50	120.07	114.84
4	A	1403	UPU	C5B-C4B-N3B	3.45	120.00	114.84
4	C	1403	UPU	C5A-C4A-N3A	3.44	119.99	114.84
4	B	1403	UPU	C5B-C4B-N3B	3.43	119.98	114.84
4	B	1403	UPU	C5A-C4A-N3A	3.39	119.91	114.84
3	A	1402	2KH	C5-C4-N3	3.36	119.86	114.84
3	B	1402	2KH	C5-C4-N3	3.35	119.85	114.84
4	C	1403	UPU	C5B-C4B-N3B	3.34	119.84	114.84
4	B	1403	UPU	O4B-C4B-C5B	-3.14	119.63	125.16
3	C	1402	2KH	O4-C4-C5	-3.11	119.68	125.16
3	A	1402	2KH	O4-C4-C5	-2.95	119.97	125.16
3	B	1402	2KH	O2-C2-N1	-2.94	118.88	122.79
4	C	1403	UPU	O4B-C4B-C5B	-2.93	120.00	125.16
3	C	1402	2KH	C4'-O4'-C1'	-2.88	103.12	109.47
3	C	1402	2KH	PB-O3B-PG	-2.87	122.51	132.62
3	B	1402	2KH	O4-C4-C5	-2.86	120.12	125.16
3	A	1402	2KH	PB-O3B-PG	-2.78	122.81	132.62
4	A	1403	UPU	O4B-C4B-C5B	-2.76	120.31	125.16
3	B	1402	2KH	PB-O3B-PG	-2.70	123.12	132.62
4	A	1403	UPU	O4A-C4A-C5A	-2.69	120.43	125.16
4	B	1403	UPU	O4A-C4A-C5A	-2.63	120.53	125.16
4	B	1403	UPU	O2A-C2A-N1A	-2.61	119.31	122.79
3	B	1402	2KH	C4'-O4'-C1'	-2.61	103.72	109.47
4	C	1403	UPU	O4A-C4A-C5A	-2.60	120.58	125.16

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1402	2KH	C4'-O4'-C1'	-2.47	104.02	109.47
4	A	1403	UPU	O2B-C2B-N1B	-2.46	119.52	122.79
3	A	1402	2KH	O2-C2-N1	-2.42	119.57	122.79
4	A	1403	UPU	O2A-C2A-N1A	-2.39	119.60	122.79
4	A	1403	UPU	C3'A-C2'A-C1'A	2.35	105.88	101.43
3	C	1402	2KH	O2-C2-N1	-2.33	119.69	122.79
4	B	1403	UPU	O2B-C2B-N1B	-2.26	119.78	122.79
4	C	1403	UPU	O2A-C2A-N1A	-2.26	119.78	122.79
4	C	1403	UPU	O2B-C2B-N1B	-2.09	120.01	122.79
4	B	1403	UPU	C3'A-C2'A-C1'A	2.05	105.31	101.43

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1402	2KH	PB-N3A-PA-O5'
3	A	1402	2KH	PA-N3A-PB-O2B
3	A	1402	2KH	PB-O3B-PG-O1G
3	A	1402	2KH	PB-O3B-PG-O3G
3	B	1402	2KH	PB-N3A-PA-O5'
3	B	1402	2KH	PB-O3B-PG-O3G
3	C	1402	2KH	PB-N3A-PA-O1A
3	C	1402	2KH	PA-N3A-PB-O2B
3	C	1402	2KH	PB-O3B-PG-O1G
4	B	1403	UPU	O4'B-C4'B-C5'B-O5'B
4	B	1403	UPU	C3'B-C4'B-C5'B-O5'B
3	C	1402	2KH	C5'-O5'-PA-O2A
3	C	1402	2KH	PB-N3A-PA-O5'
3	C	1402	2KH	PB-O3B-PG-O2G
3	A	1402	2KH	PB-N3A-PA-O1A
4	A	1403	UPU	C3'B-C4'B-C5'B-O5'B
3	B	1402	2KH	PB-O3B-PG-O2G
3	A	1402	2KH	C5'-O5'-PA-O1A
4	A	1403	UPU	O4'B-C4'B-C5'B-O5'B

There are no ring outliers.

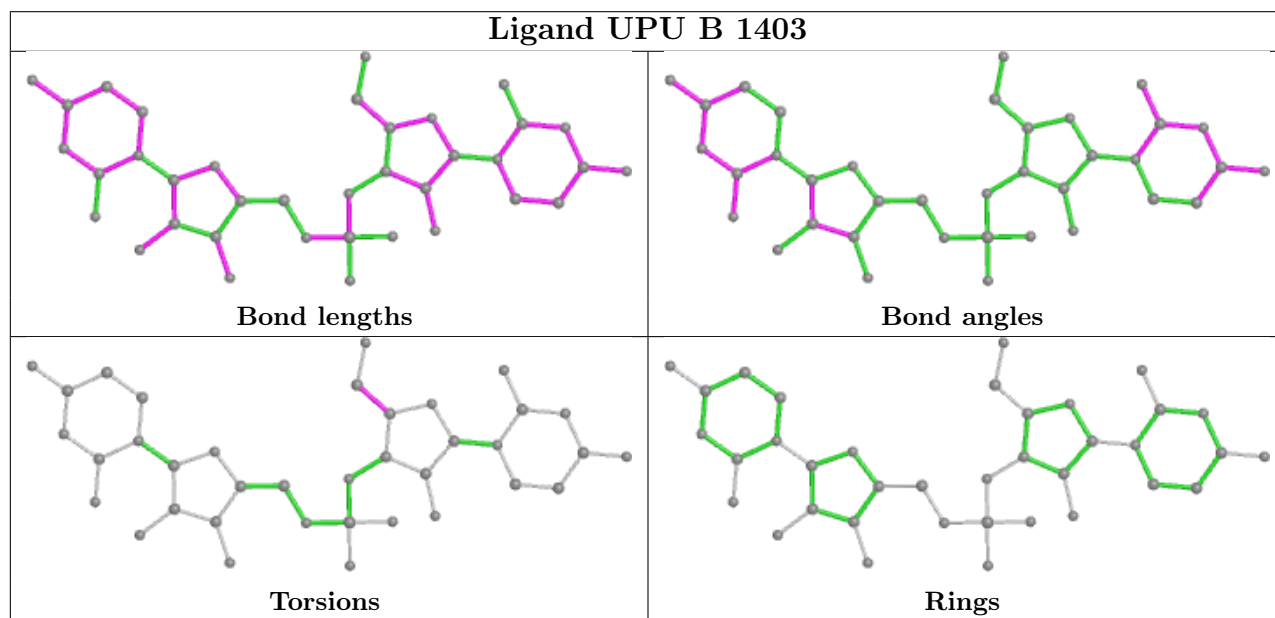3 monomers are involved in 4 short contacts:

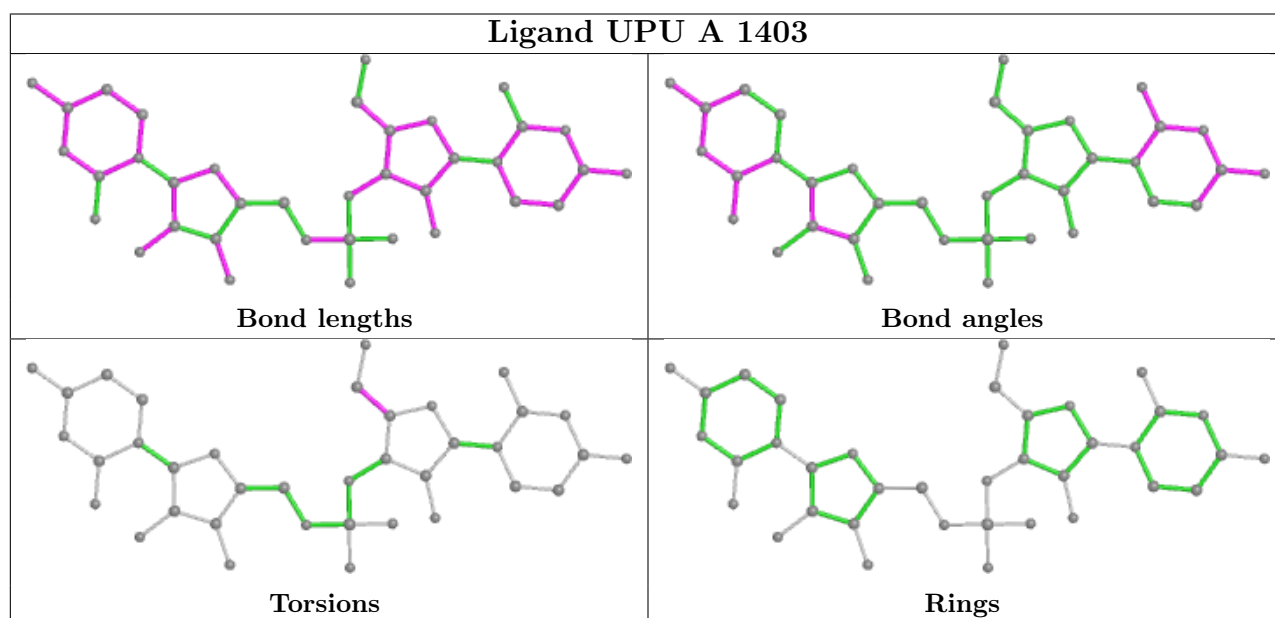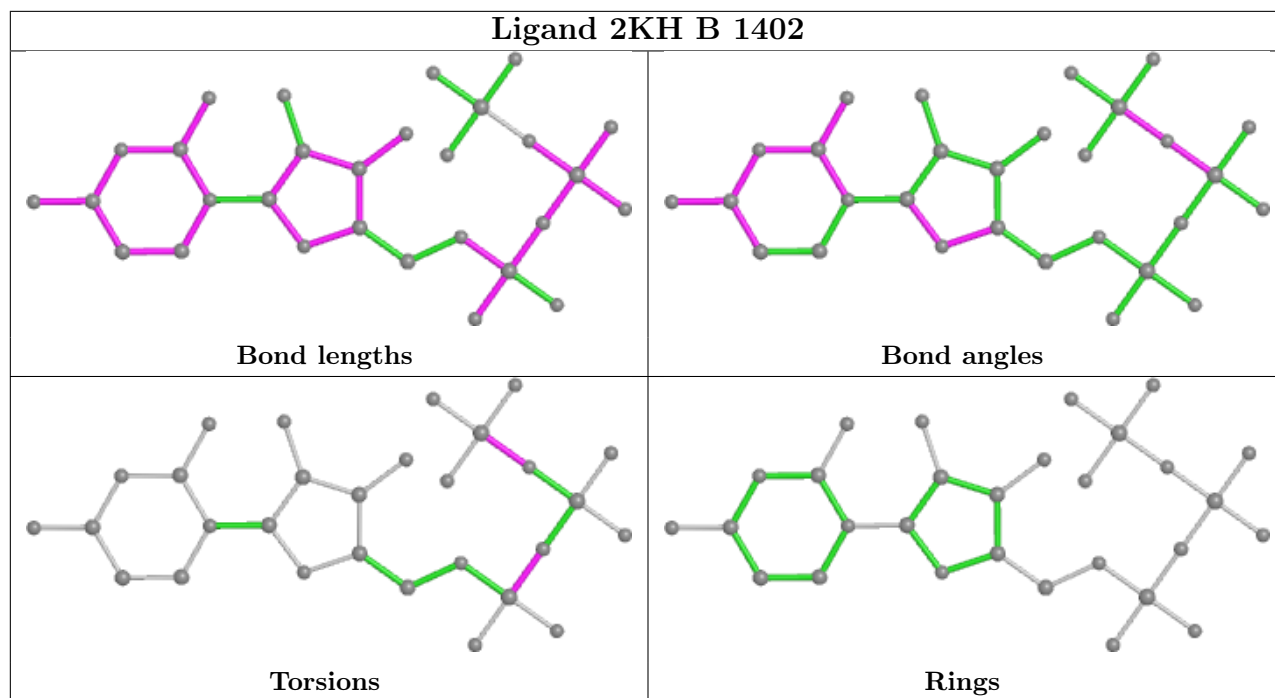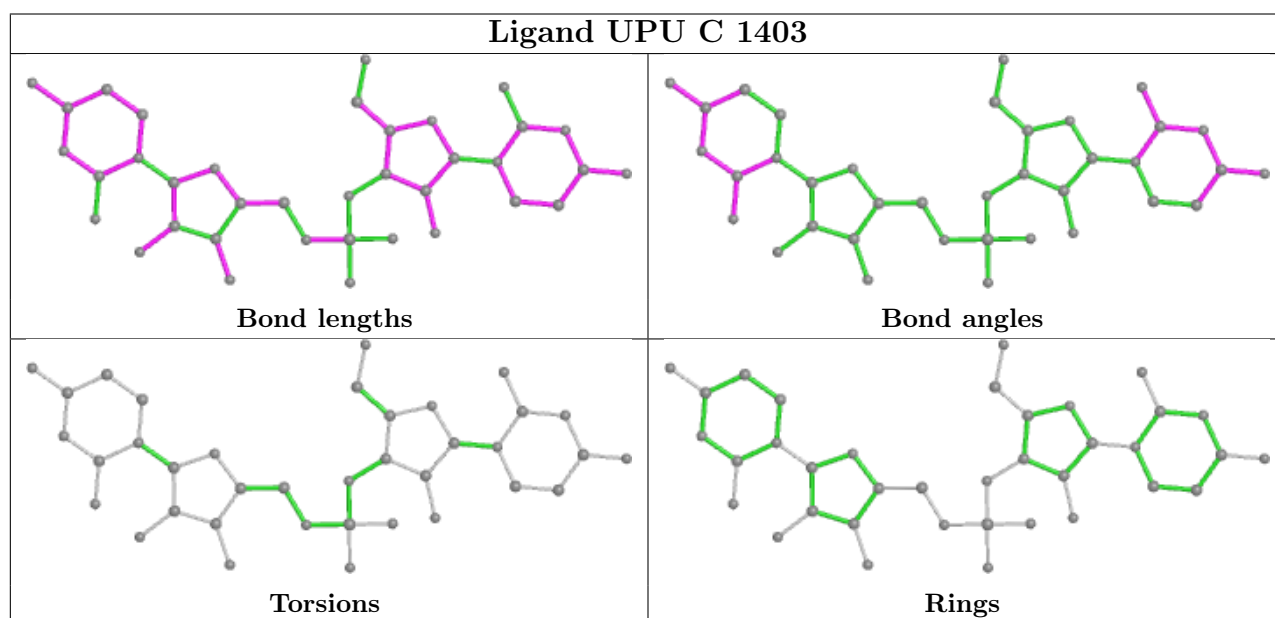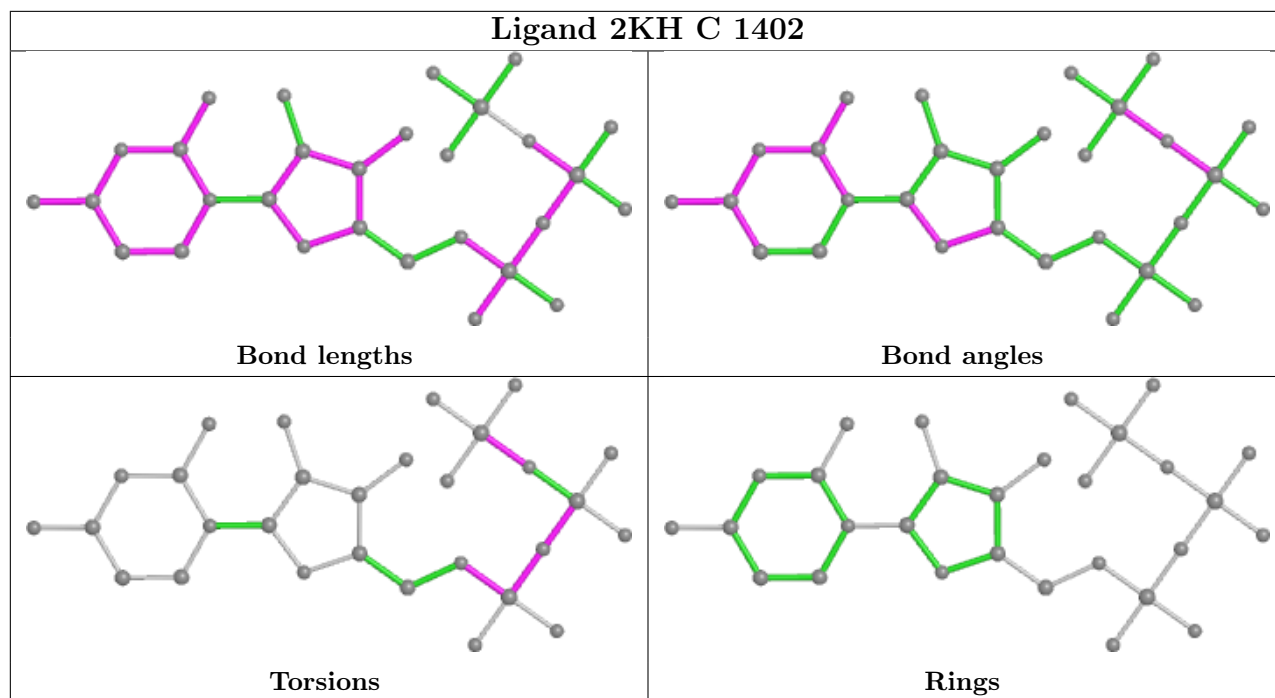| Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	1403	UPU	2	0
4	A	1403	UPU	1	0

Continued on next page...
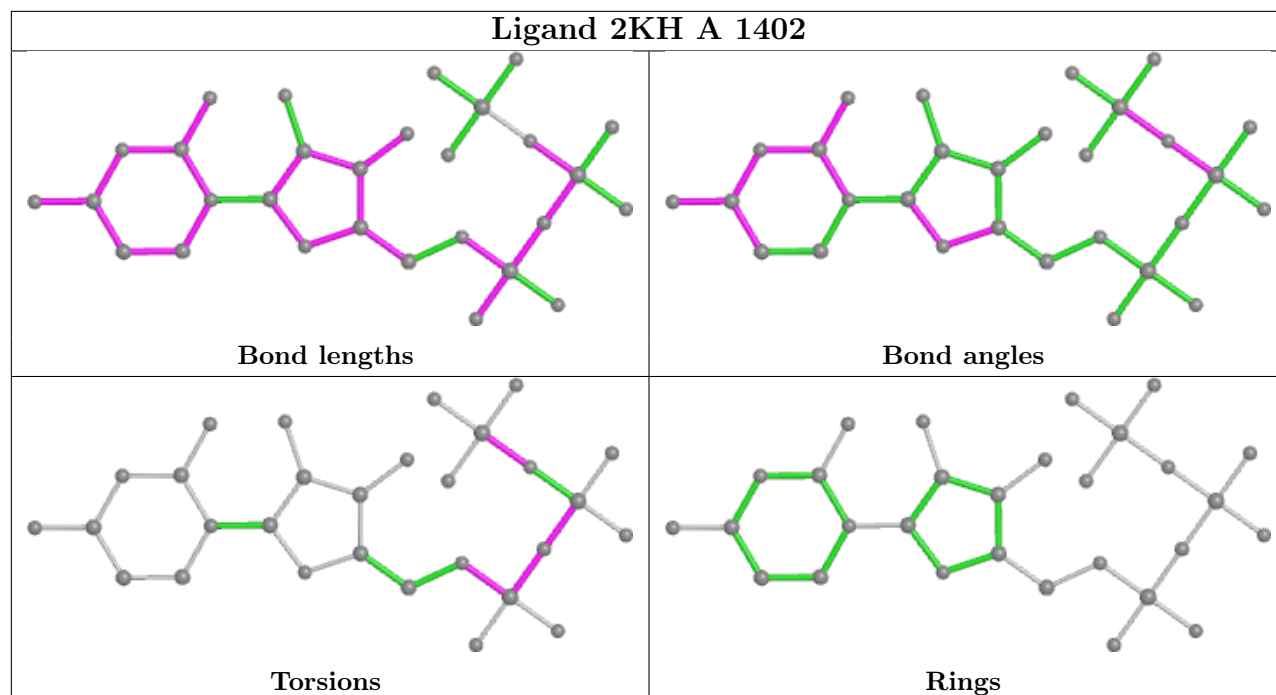
Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1403	UPU	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand UPU B 1403

Bond lengths

Bond angles

Torsions

Rings

Ligand 2KH B 1402

Bond lengths

Bond angles

Torsions

Rings

Ligand UPU A 1403

Bond lengths

Bond angles

Torsions

Rings

Ligand 2KH C 1402

Bond lengths

Bond angles

Torsions

Rings

Ligand UPU C 1403

Bond lengths

Bond angles

Torsions

Rings

Ligand 2KH A 1402

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	379/403 (94%)	0.45	17 (4%)	33	36	36, 56, 98, 141	0
1	B	372/403 (92%)	0.34	8 (2%)	62	65	36, 57, 95, 121	0
1	C	350/403 (86%)	0.35	6 (1%)	70	72	41, 62, 99, 136	0
All	All	1101/1209 (91%)	0.38	31 (2%)	53	56	36, 59, 97, 141	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1199	GLU	5.7
1	A	1223	TYR	4.0
1	A	1317	GLY	3.9
1	C	987	LEU	3.3
1	B	1252	LYS	3.2
1	B	1348	ARG	3.0
1	A	1196	TYR	3.0
1	A	1358	LYS	3.0
1	A	1200	LYS	3.0
1	A	1204	ILE	2.9
1	A	1363	ARG	2.8
1	A	1198	GLY	2.6
1	A	1201	LYS	2.5
1	B	1316	LYS	2.5
1	A	1267	PHE	2.5
1	B	1267	PHE	2.4
1	C	1078	ARG	2.3
1	A	1197	LYS	2.3
1	B	1358	LYS	2.2
1	C	1223	TYR	2.2
1	B	1317	GLY	2.2
1	C	1264	LEU	2.2
1	B	1275	TYR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	1267	PHE	2.1
1	A	1220	LEU	2.1
1	A	1264	LEU	2.1
1	A	1316	LYS	2.1
1	A	1271	TRP	2.1
1	B	1112	LEU	2.1
1	A	1262	SER	2.1
1	C	1324	SER	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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	MG	A	1401	1/1	0.91	0.24	78,78,78,78	0
3	2KH	A	1402	29/29	0.94	0.30	61,78,92,95	1
2	MG	B	1401	1/1	0.95	0.19	37,37,37,37	0
4	UPU	C	1403	37/37	0.95	0.17	46,59,72,81	0
4	UPU	B	1403	37/37	0.97	0.19	44,53,64,74	0
4	UPU	A	1403	37/37	0.97	0.20	40,51,62,67	0
5	ZN	A	1404	1/1	0.97	0.15	73,73,73,73	0
6	IOD	A	1405	1/1	0.97	0.12	96,96,96,96	0
6	IOD	C	1404	1/1	0.97	0.12	145,145,145,145	0
2	MG	C	1401	1/1	0.98	0.13	45,45,45,45	0
5	ZN	B	1404	1/1	0.98	0.23	64,64,64,64	0
3	2KH	B	1402	29/29	0.98	0.19	27,37,45,47	1
3	2KH	C	1402	29/29	0.98	0.17	38,51,61,66	1
7	SO4	A	1406	5/5	0.98	0.18	87,88,90,92	0
6	IOD	B	1405	1/1	0.99	0.14	65,65,65,65	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.
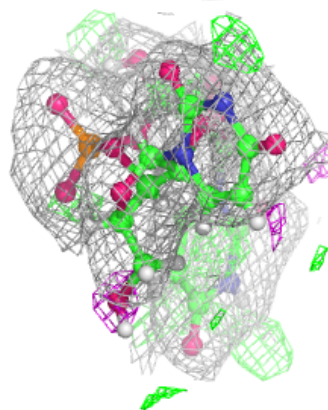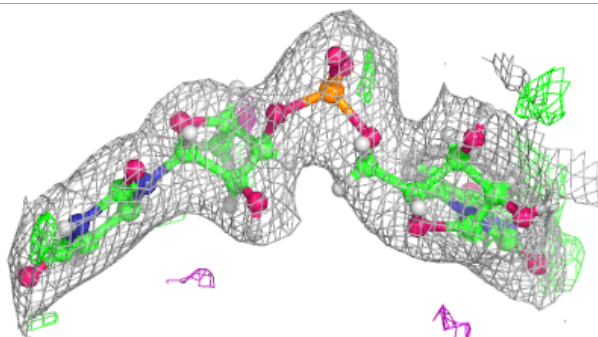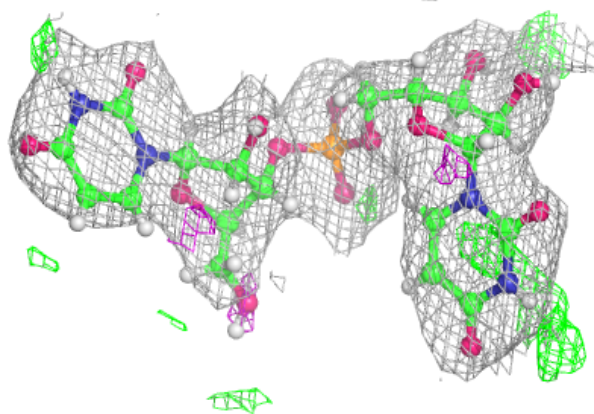
Electron density around 2KH A 1402:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)
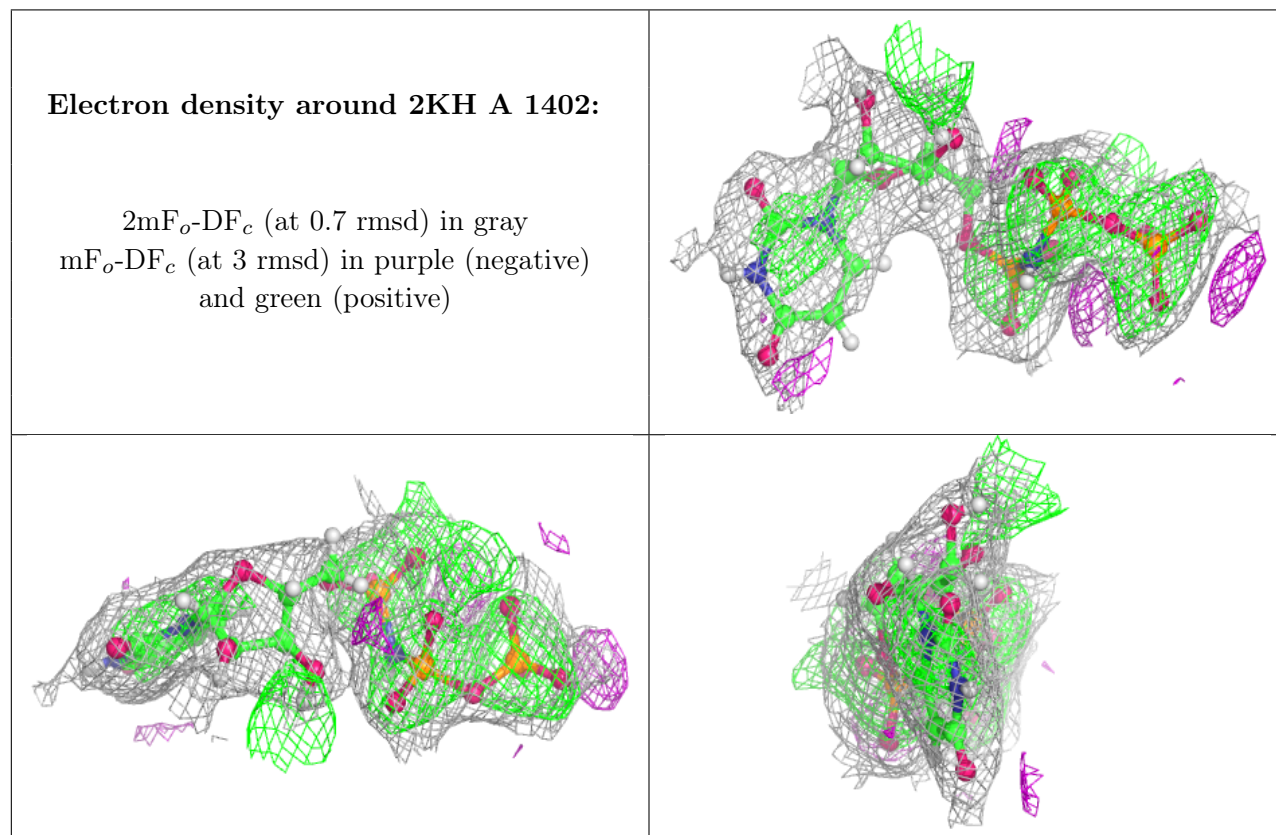
Electron density around UPU C 1403:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
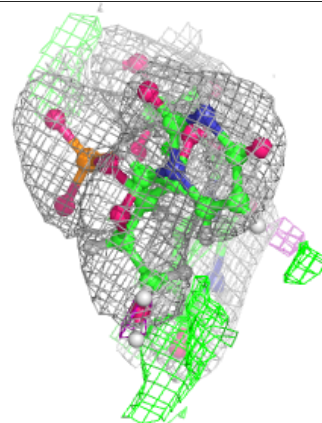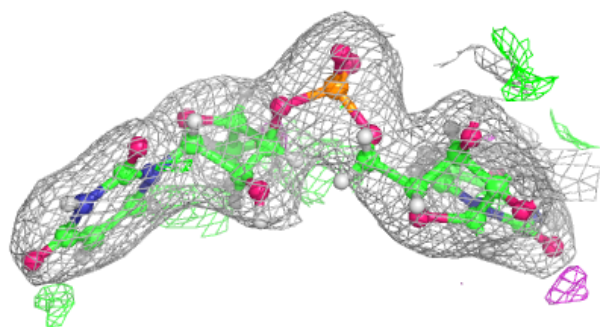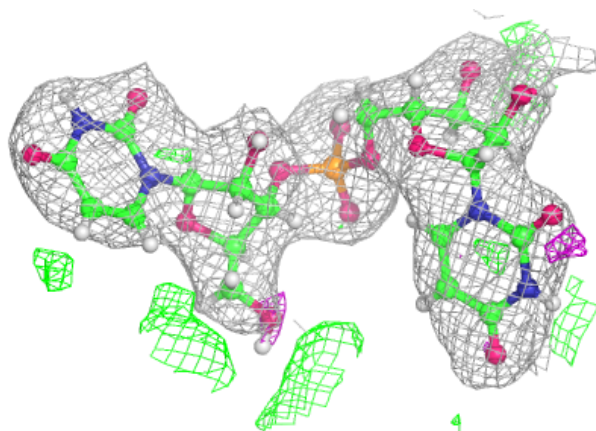and green (positive)
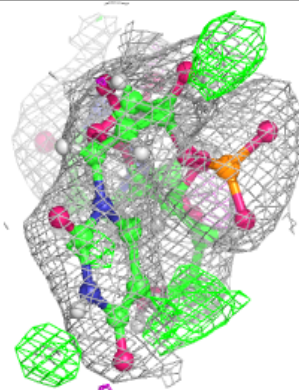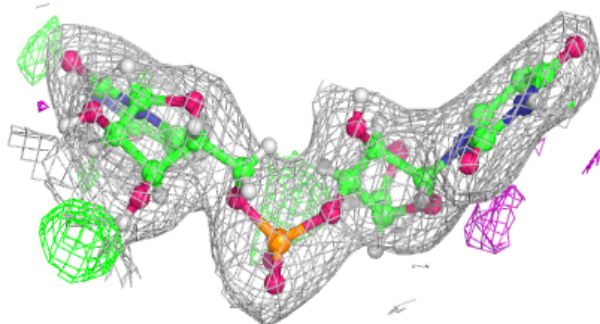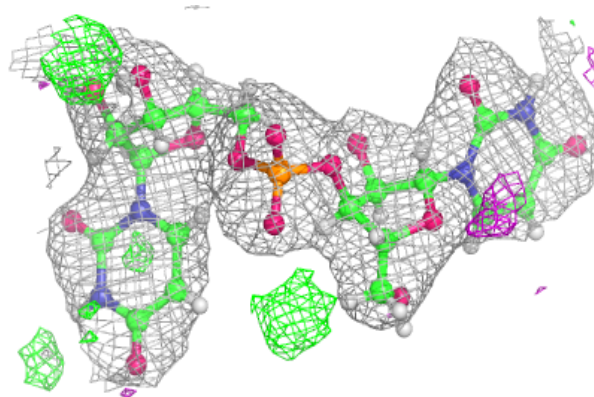
Electron density around UPU B 1403:

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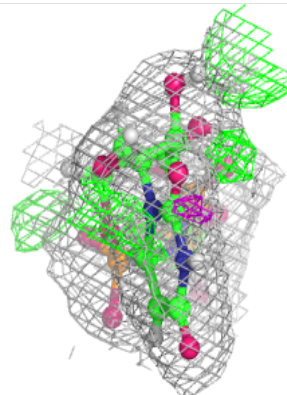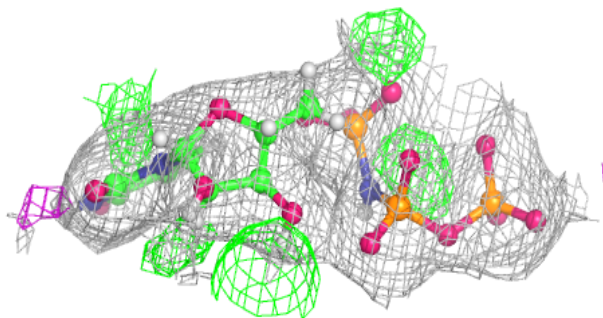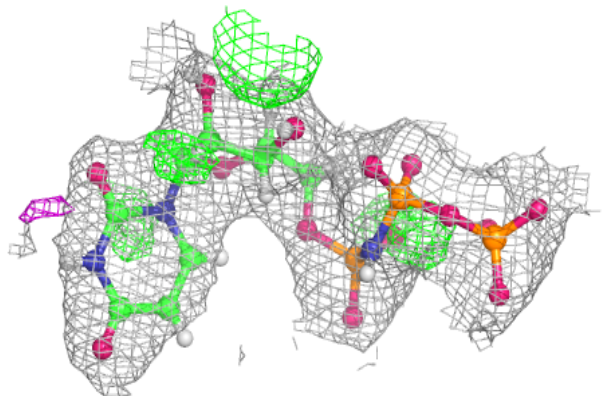$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
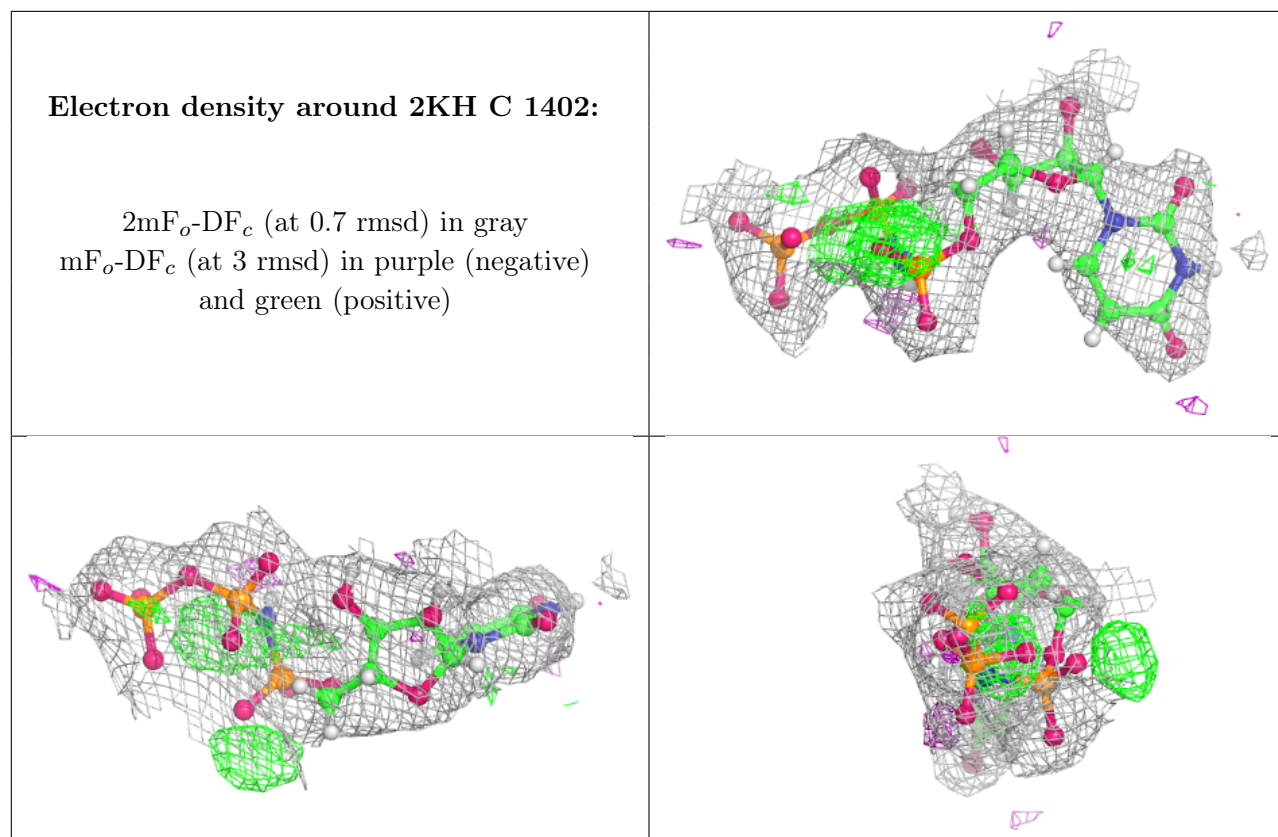and green (positive)

Electron density around UPU A 1403:

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

Electron density around 2KH B 1402:

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

Electron density around 2KH C 1402:

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 ⓘ

There are no such residues in this entry.