Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2024 – 12:38 PM EDT

PDB ID	:	4G5H
Title	:	Crystal structure of capsular polysaccharide synthesizing enzyme CapE from Staphylococcus aureus in complex with by-product
Authors	:	Miyafusa, T.; Caaveiro, J.M.M.; Tanaka, Y.; Tsumoto, K.
Deposited on	:	2012-07-18
Resolution	:	1.88 Å(reported)

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

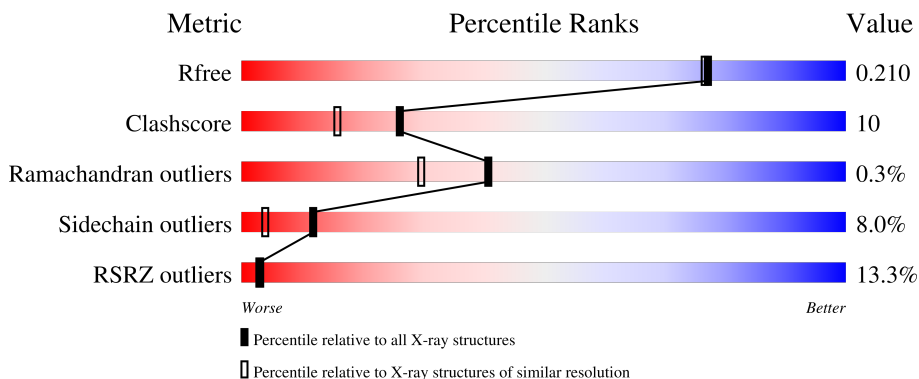|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.88 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		0.210
Clashscore		10
Ramachandran outliers		0.3%
Sidechain outliers		8.0%
RSRZ outliers		13.3%

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	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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	363	13% 72% 20% •• 5%

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	FMT	A	407	-	-	X	-

Continued on next page...

W O R L D W I D E

wwPDB

PROTEIN DATA BANK

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	FMT	A	410	-	-	X	-
4	FMT	A	415	-	-	X	-

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 3104 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 Capsular polysaccharide synthesis enzyme Cap8E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	346	2800	1769	471	544	16	0	10	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP Q7A2Y4
A	-19	ASN	-	expression tag	UNP Q7A2Y4
A	-18	HIS	-	expression tag	UNP Q7A2Y4
A	-17	LYS	-	expression tag	UNP Q7A2Y4
A	-16	HIS	-	expression tag	UNP Q7A2Y4
A	-15	HIS	-	expression tag	UNP Q7A2Y4
A	-14	HIS	-	expression tag	UNP Q7A2Y4
A	-13	HIS	-	expression tag	UNP Q7A2Y4
A	-12	HIS	-	expression tag	UNP Q7A2Y4
A	-11	HIS	-	expression tag	UNP Q7A2Y4
A	-10	SER	-	expression tag	UNP Q7A2Y4
A	-9	SER	-	expression tag	UNP Q7A2Y4
A	-8	GLY	-	expression tag	UNP Q7A2Y4
A	-7	LEU	-	expression tag	UNP Q7A2Y4
A	-6	VAL	-	expression tag	UNP Q7A2Y4
A	-5	PRO	-	expression tag	UNP Q7A2Y4
A	-4	ARG	-	expression tag	UNP Q7A2Y4
A	-3	GLY	-	expression tag	UNP Q7A2Y4
A	-2	SER	-	expression tag	UNP Q7A2Y4
A	-1	ALA	-	expression tag	UNP Q7A2Y4
A	0	MET	-	expression tag	UNP Q7A2Y4
A	1	GLY	-	expression tag	UNP Q7A2Y4

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$).

NAP

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 3 is [(2R,3R,4R,6R)-3-acetamido-6-methyl-4-oxidanyl-5-oxidanylidene-oxan-2 -yl] [[(2R,3S,4R,5R)-5-[2,4-bis(oxidanylidene)pyrimidin-1-yl]-3,4-bis(oxidanyl)oxolan-2 -yl]methoxy-oxidanyl-phosphoryl] hydrogen phosphate (three-letter code: UD7) (formula: $C_{17}H_{25}N_3O_{16}P_2$).
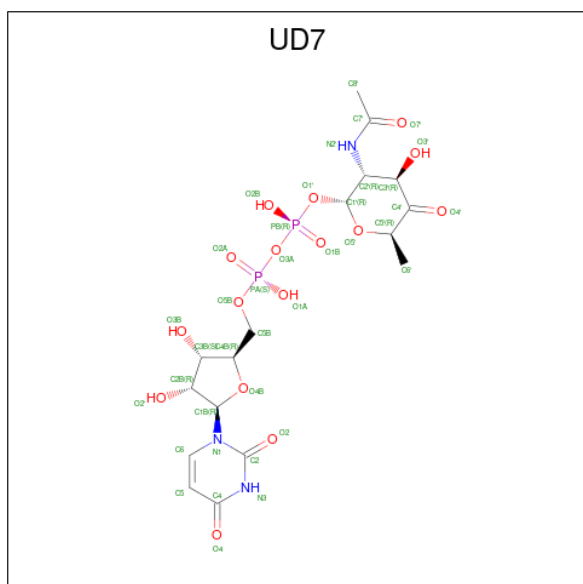
UD7
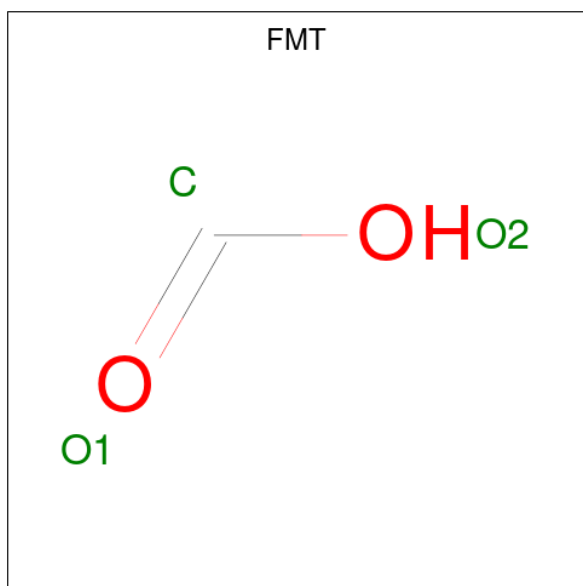
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			38	17	3	16	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 38	C 17	N 3	O 16	P 2	0	0

- Molecule 4 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0
4	A	1	Total 3	C 1	O 2	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Na 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	136	Total 137	O 137	0	1

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: Capsular polysaccharide synthesis enzyme Cap8E

Chain A:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	124.03Å 124.03Å 103.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.80 – 1.88 35.80 – 1.88	Depositor EDS
% Data completeness (in resolution range)	99.8 (35.80-1.88) 99.8 (35.80-1.88)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 1.88Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.175 , 0.209 0.175 , 0.210	Depositor DCC
R_{free} test set	1546 reflections (4.02%)	wwPDB-VP
Wilson B-factor (Å^2)	25.6	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 52.3	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3104	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: UD7, FMT, NAP, NA

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	1.19	4/2859 (0.1%)	1.19	19/3852 (0.5%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	SER	CB-OG	-7.25	1.32	1.42
1	A	273	GLU	CD-OE1	6.25	1.32	1.25
1	A	268	GLU	CD-OE2	5.83	1.32	1.25
1	A	293	TYR	CE1-CZ	5.39	1.45	1.38

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	140[A]	MET	CG-SD-CE	-8.66	86.34	100.20
1	A	140[B]	MET	CG-SD-CE	-8.66	86.34	100.20
1	A	37	ARG	NE-CZ-NH2	8.22	124.41	120.30
1	A	23	ARG	NE-CZ-NH2	-8.12	116.24	120.30
1	A	37	ARG	NE-CZ-NH1	-7.35	116.63	120.30
1	A	33	ARG	NE-CZ-NH2	-7.27	116.67	120.30
1	A	254	ARG	NE-CZ-NH1	-6.63	116.98	120.30
1	A	274	ASP	CB-CG-OD2	6.23	123.91	118.30
1	A	312	THR	CB-CA-C	-6.06	95.23	111.60
1	A	191	THR	N-CA-CB	-6.02	98.86	110.30
1	A	191	THR	OG1-CB-CG2	5.87	123.50	110.00
1	A	257	GLU	CB-CA-C	-5.83	98.75	110.40
1	A	310	ASP	CB-CG-OD2	5.48	123.23	118.30
1	A	21	MET	CG-SD-CE	5.47	108.96	100.20
1	A	312	THR	OG1-CB-CG2	5.32	122.23	110.00
1	A	150	ARG	NE-CZ-NH2	5.24	122.92	120.30
1	A	53	LYS	CD-CE-NZ	-5.18	99.78	111.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	323	LYS	CD-CE-NZ	5.15	123.55	111.70
1	A	214	LYS	CB-CG-CD	-5.01	98.58	111.60

There are no chirality outliers.

There are no planarity outliers.

5.2　Too-close contacts ⓘ

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	2800	0	2808	56	0
2	A	48	0	22	6	0
3	A	76	0	48	1	0
4	A	42	0	18	11	0
5	A	1	0	0	0	0
6	A	137	0	0	9	0
All	All	3104	0	2896	61	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 10.

All (61) 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:16[A]:PHE:CZ	1:A:20:VAL:HG21	1.37	1.54
1:A:16[A]:PHE:CZ	1:A:20:VAL:CG2	2.23	1.20
1:A:16[A]:PHE:CE1	1:A:20:VAL:CG2	2.32	1.12
1:A:16[A]:PHE:CE1	1:A:20:VAL:HG21	1.93	1.03
1:A:226:ALA:O	1:A:312:THR:HG21	1.63	0.96
4:A:414:FMT:H	6:A:524:HOH:O	1.68	0.93
1:A:191:THR:HG21	3:A:402:UD7:O2'	1.76	0.86
1:A:16[A]:PHE:CE1	1:A:20:VAL:HG23	2.10	0.83
1:A:11:GLY:HA2	2:A:401:NAP:H1B	1.68	0.76
1:A:16[B]:PHE:CZ	1:A:205:ALA:HB1	2.21	0.75
1:A:155:GLU:HG3	6:A:588:HOH:O	1.85	0.75
1:A:20:VAL:HG11	1:A:78[A]:PHE:CD2	2.21	0.75

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:20:VAL:HG11	1:A:78[A]:PHE:HD2	1.50	0.74
1:A:286:ARG:O	4:A:415:FMT:H	1.88	0.73
1:A:16[A]:PHE:HZ	1:A:20:VAL:HG21	1.46	0.71
1:A:16[A]:PHE:CE1	1:A:209:VAL:HG21	2.30	0.67
1:A:16[B]:PHE:HB2	2:A:401:NAP:O2N	1.96	0.66
1:A:16[A]:PHE:CE2	1:A:20:VAL:HG21	2.21	0.65
1:A:120:ILE:HD12	1:A:160:CYS:SG	2.39	0.62
1:A:4:ASP:OD1	1:A:4:ASP:N	2.29	0.61
1:A:22:LYS:HD2	1:A:49:TYR:OH	2.00	0.61
1:A:267:GLU:O	1:A:271:GLN:HG2	2.02	0.60
1:A:192:ASP:HB3	1:A:195:MET:HE3	1.84	0.59
2:A:401:NAP:N7N	6:A:620:HOH:O	2.31	0.59
1:A:118:LYS:NZ	6:A:522:HOH:O	2.35	0.59
1:A:51:ASN:HD22	1:A:53:LYS:H	1.51	0.57
1:A:84:LYS:HD2	2:A:401:NAP:H2D	1.86	0.57
1:A:286:ARG:O	4:A:415:FMT:C	2.54	0.54
1:A:16[A]:PHE:CD1	1:A:20:VAL:HG23	2.41	0.54
1:A:203:GLU:O	1:A:207:GLU:HG2	2.08	0.53
1:A:211:HIS:HD2	1:A:222:MET:HG2	1.73	0.53
4:A:415:FMT:O2	6:A:562:HOH:O	2.19	0.53
2:A:401:NAP:O7N	6:A:579:HOH:O	2.18	0.52
1:A:23:ARG:O	1:A:26:ASP:HB2	2.11	0.51
1:A:1:GLY:HA2	1:A:28:ASN:HD22	1.75	0.51
1:A:16[A]:PHE:C	1:A:16[A]:PHE:CD2	2.85	0.50
1:A:271:GLN:NE2	6:A:608:HOH:O	2.45	0.49
1:A:191:THR:CG2	1:A:192:ASP:N	2.76	0.49
1:A:0:MET:SD	1:A:2:PHE:CZ	3.06	0.49
1:A:191:THR:CG2	1:A:192:ASP:H	2.25	0.48
1:A:290:TYR:N	4:A:407:FMT:H	2.29	0.48
4:A:411:FMT:H	6:A:622:HOH:O	2.13	0.48
1:A:-5:PRO:HB3	1:A:215:HIS:HE1	1.80	0.47
1:A:138:LYS:HE3	2:A:401:NAP:O2D	2.15	0.47
1:A:191:THR:HG22	1:A:192:ASP:H	1.80	0.47
1:A:123:SER:O	1:A:164[B]:TYR:HD1	1.97	0.46
1:A:226:ALA:CB	4:A:410:FMT:H	2.45	0.46
1:A:2:PHE:HB3	1:A:29:ILE:HD12	1.98	0.46
1:A:120:ILE:HD13	1:A:120:ILE:N	2.30	0.45
1:A:68:GLU:HG3	4:A:413:FMT:O1	2.18	0.44
1:A:291:SER:H	4:A:407:FMT:C	2.30	0.44
1:A:48:LYS:HG3	1:A:49:TYR:CD1	2.54	0.42
1:A:126:LYS:HD2	1:A:134:MET:CE	2.49	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:164[A]:TYR:HD2	1:A:208:LEU:HD23	1.84	0.42
1:A:-6:VAL:HG11	1:A:-1:ALA:HA	2.02	0.42
1:A:242:PHE:CE1	1:A:334:LEU:HD13	2.55	0.42
1:A:15:SER:HB3	6:A:615:HOH:O	2.20	0.42
1:A:123:SER:O	1:A:164[B]:TYR:CD1	2.72	0.42
1:A:227:PRO:HD2	4:A:410:FMT:H	2.01	0.41
1:A:150:ARG:O	4:A:415:FMT:C	2.69	0.41
1:A:47[B]:LYS:HD2	1:A:47[B]:LYS:HA	1.99	0.40

There are no symmetry-related clashes.

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	354/363 (98%)	343 (97%)	10 (3%)	1 (0%)	41	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1	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	310/316 (98%)	286 (92%)	24 (8%)	13 4

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-7	LEU
1	A	-6	VAL
1	A	-4	ARG
1	A	0	MET
1	A	15	SER
1	A	21	MET
1	A	51	ASN
1	A	66	SER
1	A	72	ARG
1	A	84	LYS
1	A	145	PHE
1	A	156	GLN
1	A	170	SER
1	A	183	LYS
1	A	191	THR
1	A	231	VAL
1	A	239	LEU
1	A	245	ASP
1	A	248	ILE
1	A	303	GLN
1	A	312	THR
1	A	322	GLU
1	A	328	GLU
1	A	335	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	ASN
1	A	51	ASN
1	A	114	GLN
1	A	211	HIS
1	A	215	HIS
1	A	292	ASN
1	A	335	ASN

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 18 ligands modelled in this entry, 1 is monoatomic - leaving 17 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	$\#|Z| > 2$
4	FMT	A	413	-	2,2,2	0.60	0	1,1,1	0.81	0
4	FMT	A	408	-	2,2,2	1.17	0	1,1,1	0.55	0
3	UD7	A	402	-	36,40,40	1.27	2 (5%)	51,61,61	1.55	12 (23%)
4	FMT	A	409	-	2,2,2	1.36	0	1,1,1	0.45	0
4	FMT	A	410	-	2,2,2	1.11	0	1,1,1	0.16	0
4	FMT	A	407	-	2,2,2	1.14	0	1,1,1	0.03	0
4	FMT	A	405	-	2,2,2	1.69	1 (50%)	1,1,1	0.34	0
4	FMT	A	416	-	2,2,2	0.41	0	1,1,1	0.62	0
4	FMT	A	404	-	2,2,2	0.85	0	1,1,1	0.58	0
3	UD7	A	403	-	36,40,40	1.30	4 (11%)	51,61,61	1.80	11 (21%)
4	FMT	A	417	-	2,2,2	1.72	1 (50%)	1,1,1	0.84	0
4	FMT	A	412	-	2,2,2	1.30	0	1,1,1	0.43	0
4	FMT	A	415	-	2,2,2	0.66	0	1,1,1	0.26	0
2	NAP	A	401	-	45,52,52	1.65	6 (13%)	56,80,80	1.99	16 (28%)
4	FMT	A	414	-	2,2,2	1.06	0	1,1,1	0.19	0
4	FMT	A	406	-	2,2,2	0.97	0	1,1,1	0.39	0
4	FMT	A	411	-	2,2,2	1.23	0	1,1,1	0.51	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	UD7	A	402	-	-	5/24/61/61	0/3/3/3
3	UD7	A	403	-	-	7/24/61/61	0/3/3/3
2	NAP	A	401	-	-	8/31/67/67	0/5/5/5

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	NAP	C4A-N3A	5.61	1.43	1.35
2	A	401	NAP	C5A-C4A	4.10	1.51	1.40
2	A	401	NAP	P2B-O2B	4.06	1.67	1.59
3	A	402	UD7	C2-N1	3.86	1.44	1.38
2	A	401	NAP	C2A-N3A	3.61	1.37	1.32
3	A	403	UD7	C2-N1	3.55	1.44	1.38
2	A	401	NAP	O4D-C1D	3.48	1.45	1.41
3	A	403	UD7	C6-C5	2.85	1.41	1.35
3	A	402	UD7	C6-C5	2.83	1.41	1.35
3	A	403	UD7	C2-N3	-2.67	1.33	1.38
2	A	401	NAP	P2B-O1X	2.50	1.58	1.50
4	A	405	FMT	O2-C	2.18	1.39	1.28
4	A	417	FMT	O2-C	2.04	1.38	1.28
3	A	403	UD7	C2'-N2'	2.03	1.49	1.45

All (39) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	403	UD7	O5'-C1'-O1'	-5.32	104.41	111.36
2	A	401	NAP	N3A-C2A-N1A	-5.00	120.86	128.68
2	A	401	NAP	C5A-C6A-N6A	-4.91	112.89	120.35
2	A	401	NAP	N6A-C6A-N1A	4.75	128.43	118.57
3	A	403	UD7	O5'-C5'-C4'	-4.24	105.38	111.04
3	A	402	UD7	C5-C4-N3	4.12	121.00	114.84
3	A	403	UD7	N3-C2-N1	3.90	120.07	114.89
2	A	401	NAP	C2A-N1A-C6A	3.75	125.17	118.75
3	A	403	UD7	O4B-C1B-N1	3.61	116.61	108.36
2	A	401	NAP	C3N-C7N-N7N	3.60	122.07	117.75
3	A	402	UD7	O4-C4-C5	-3.58	118.86	125.16
2	A	401	NAP	C6N-N1N-C2N	-3.33	118.94	121.97
2	A	401	NAP	C3D-C2D-C1D	3.23	105.85	100.98

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	UD7	C6-C5-C4	-3.15	115.20	119.52
3	A	403	UD7	C1B-N1-C6	3.14	127.69	120.84
3	A	403	UD7	C2'-N2'-C7'	3.10	130.72	123.18
2	A	401	NAP	O2A-PA-O1A	3.06	127.35	112.24
3	A	403	UD7	C5-C4-N3	3.02	119.36	114.84
2	A	401	NAP	PN-O3-PA	-2.96	122.66	132.83
3	A	402	UD7	O3'-C3'-C4'	2.92	118.16	109.94
3	A	403	UD7	C4-N3-C2	-2.82	122.86	126.58
2	A	401	NAP	O2N-PN-O1N	2.72	125.69	112.24
3	A	402	UD7	O5'-C5'-C4'	2.69	114.63	111.04
3	A	402	UD7	O4B-C4B-C5B	-2.68	100.57	109.37
2	A	401	NAP	O7N-C7N-C3N	-2.61	116.51	119.63
2	A	401	NAP	O4B-C1B-C2B	-2.58	102.11	106.59
3	A	402	UD7	O4'-C4'-C3'	-2.56	117.36	121.61
3	A	403	UD7	O2-C2-N1	-2.54	119.41	122.79
2	A	401	NAP	O2D-C2D-C3D	-2.36	104.17	111.82
2	A	401	NAP	O5B-PA-O1A	-2.36	99.84	109.07
3	A	402	UD7	O3B-C3B-C4B	-2.30	104.40	111.05
3	A	402	UD7	C2'-N2'-C7'	-2.27	117.66	123.18
3	A	402	UD7	O5'-C1'-O1'	2.27	114.33	111.36
3	A	402	UD7	O3'-C3'-C2'	-2.23	102.04	109.51
2	A	401	NAP	O3X-P2B-O2X	2.22	116.11	107.64
3	A	403	UD7	O5'-C1'-C2'	2.17	114.81	110.58
3	A	403	UD7	C1B-N1-C2	-2.15	113.68	117.57
3	A	402	UD7	O3A-PB-O1'	-2.05	98.34	102.48
2	A	401	NAP	C1B-N9A-C4A	-2.00	123.12	126.64

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	NAP	O4D-C1D-N1N-C2N
3	A	402	UD7	C5B-O5B-PA-O1A
3	A	402	UD7	C1'-O1'-PB-O3A
3	A	403	UD7	O5'-C1'-O1'-PB
3	A	403	UD7	C1'-C2'-N2'-C7'
3	A	403	UD7	O7'-C7'-N2'-C2'
3	A	403	UD7	C8'-C7'-N2'-C2'
2	A	401	NAP	C2N-C3N-C7N-O7N
2	A	401	NAP	C2N-C3N-C7N-N7N
2	A	401	NAP	C3D-C4D-C5D-O5D
2	A	401	NAP	O4D-C4D-C5D-O5D

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	401	NAP	C4N-C3N-C7N-N7N
2	A	401	NAP	C4N-C3N-C7N-O7N
3	A	402	UD7	PB-O3A-PA-O5B
3	A	403	UD7	PA-O3A-PB-O1'
3	A	403	UD7	C3'-C2'-N2'-C7'
2	A	401	NAP	O4B-C4B-C5B-O5B
3	A	402	UD7	C5B-O5B-PA-O3A
3	A	403	UD7	PA-O3A-PB-O1B
3	A	402	UD7	C5B-O5B-PA-O2A

There are no ring outliers.

8 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	413	FMT	1	0
3	A	402	UD7	1	0
4	A	410	FMT	2	0
4	A	407	FMT	2	0
4	A	415	FMT	4	0
2	A	401	NAP	6	0
4	A	414	FMT	1	0
4	A	411	FMT	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 UD7 A 402

Bond lengths

Bond angles

Torsions

Rings

Ligand UD7 A 403

Bond lengths

Bond angles

Torsions

Rings

Ligand NAP A 401

Bond lengths

Bond angles

Torsions

Rings
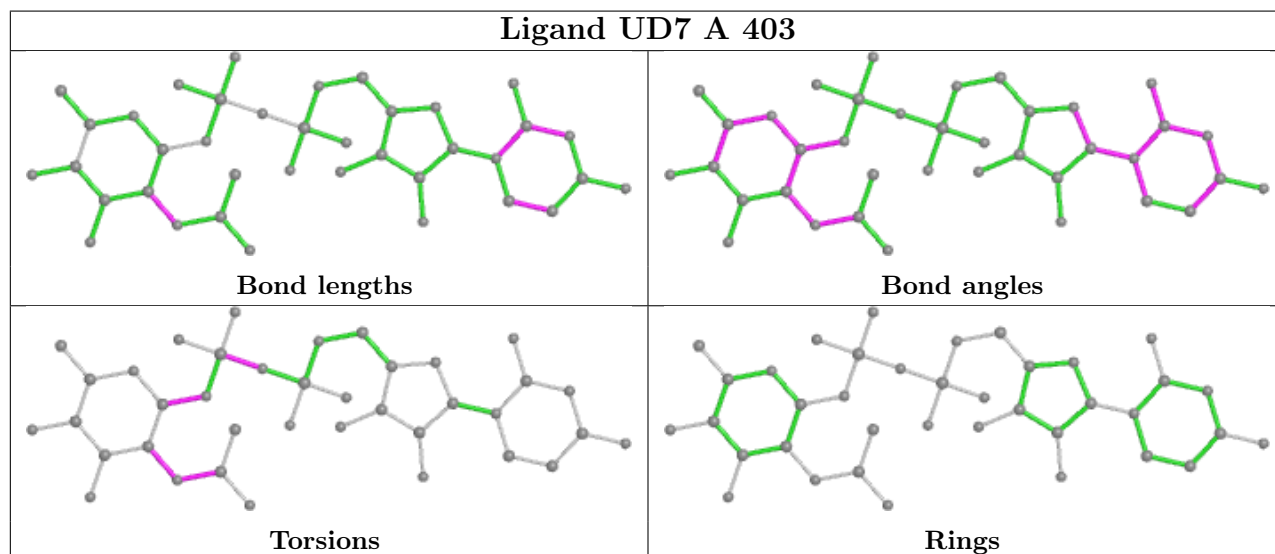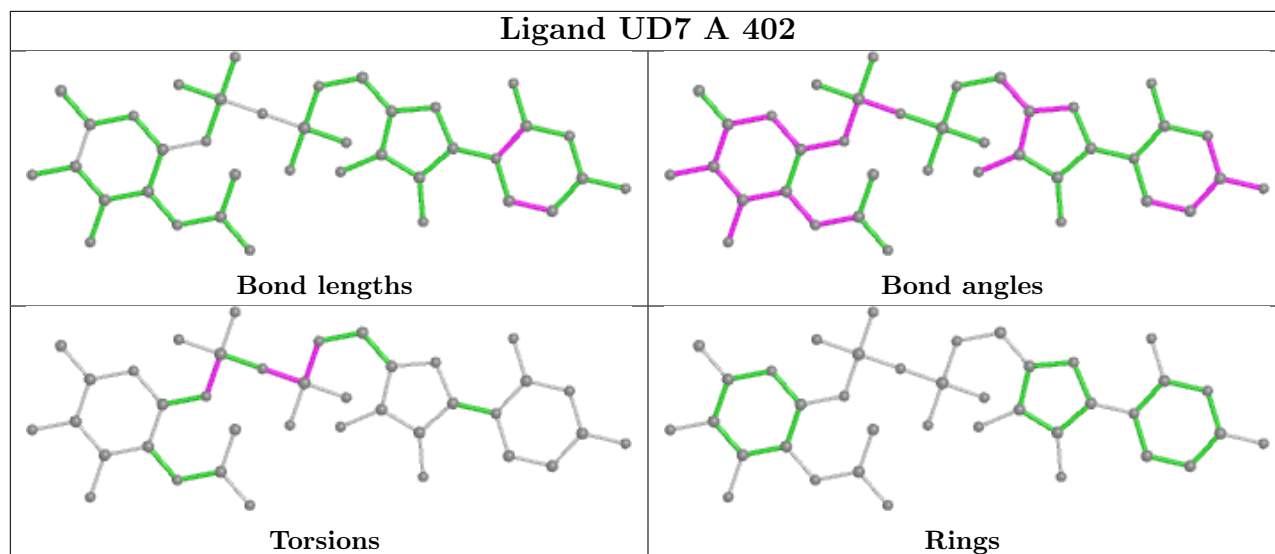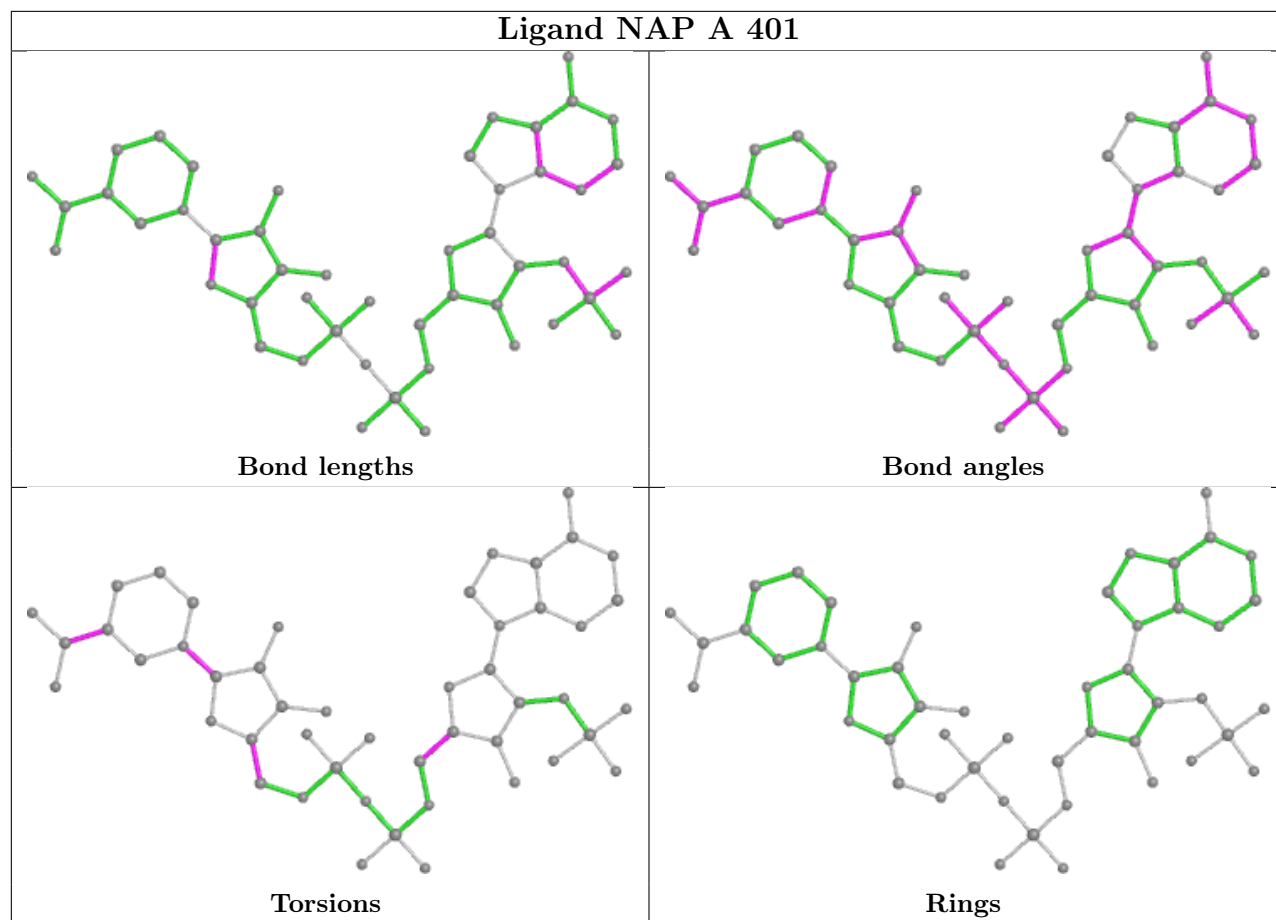
5.7 Other polymers ⓘ

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	346/363 (95%)	0.55	46 (13%)	3	3	14, 34, 67, 98	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	0	MET	7.2
1	A	26	ASP	6.6
1	A	335	ASN	6.3
1	A	334	LEU	5.3
1	A	3[A]	ASP	5.2
1	A	332	ASN	4.7
1	A	336	ASP	4.7
1	A	329	TYR	4.4
1	A	49	TYR	4.2
1	A	-1	ALA	3.9
1	A	337	TYR	3.9
1	A	328	GLU	3.9
1	A	331	ARG	3.6
1	A	325	LEU	3.5
1	A	-6	VAL	3.4
1	A	28	ASN	3.2
1	A	184	ALA	3.1
1	A	322	GLU	3.0
1	A	136	ILE	2.9
1	A	171	ARG	2.9
1	A	139	ALA	2.7
1	A	-7	LEU	2.7
1	A	52	SER	2.7
1	A	2	PHE	2.7
1	A	1	GLY	2.7
1	A	122[A]	LEU	2.7
1	A	50	ASN	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	247	ALA	2.6
1	A	250	ILE	2.6
1	A	303	GLN	2.5
1	A	299	GLU	2.5
1	A	249	GLU	2.5
1	A	164[A]	TYR	2.4
1	A	327	LEU	2.3
1	A	16[A]	PHE	2.2
1	A	-8	GLY	2.2
1	A	186	GLU	2.2
1	A	128	ALA	2.2
1	A	330	VAL	2.2
1	A	48	LYS	2.1
1	A	4	ASP	2.1
1	A	243	GLU	2.1
1	A	333	GLU	2.1
1	A	242	PHE	2.0
1	A	183	LYS	2.0
1	A	130	PRO	2.0

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
4	FMT	A	405	3/3	0.69	0.18	39,39,48,49	0
4	FMT	A	409	3/3	0.78	0.29	46,46,51,60	0
4	FMT	A	414	3/3	0.79	0.28	53,53,59,63	0
4	FMT	A	410	3/3	0.81	0.26	59,59,61,68	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FMT	A	407	3/3	0.83	0.16	52,52,56,65	0
3	UD7	A	403	38/38	0.85	0.24	36,82,127,135	0
4	FMT	A	412	3/3	0.87	0.25	49,49,50,53	0
4	FMT	A	416	3/3	0.89	0.21	34,34,42,53	0
4	FMT	A	411	3/3	0.90	0.31	31,31,43,44	0
2	NAP	A	401	48/48	0.91	0.14	19,44,53,56	25
4	FMT	A	408	3/3	0.92	0.09	40,40,42,42	0
4	FMT	A	417	3/3	0.92	0.10	29,29,34,40	0
3	UD7	A	402	38/38	0.93	0.11	28,36,65,69	0
4	FMT	A	404	3/3	0.94	0.33	40,40,49,54	0
4	FMT	A	415	3/3	0.94	0.21	28,28,37,48	0
4	FMT	A	406	3/3	0.96	0.10	26,26,32,40	0
4	FMT	A	413	3/3	0.96	0.21	41,41,46,48	0
5	NA	A	418	1/1	0.99	0.06	21,21,21,21	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.

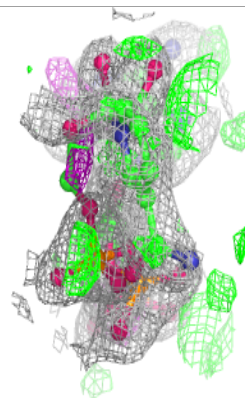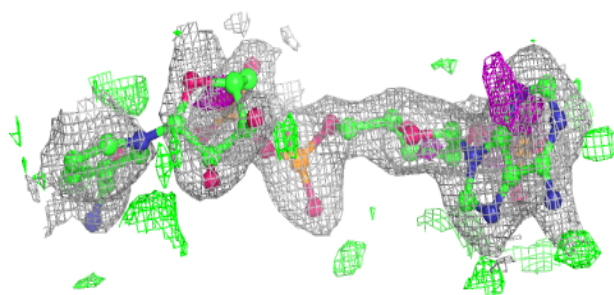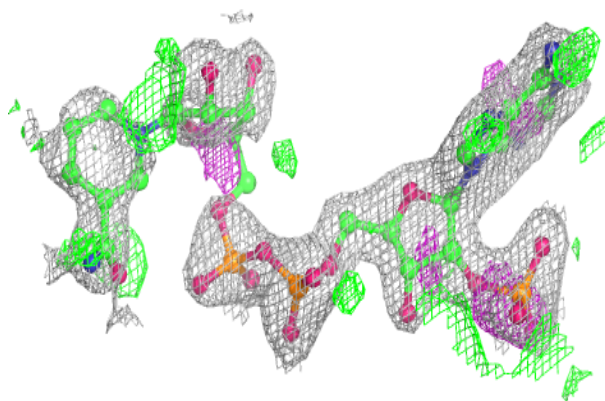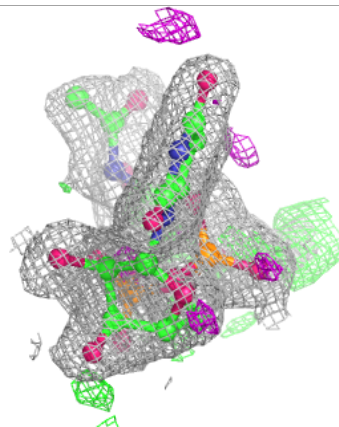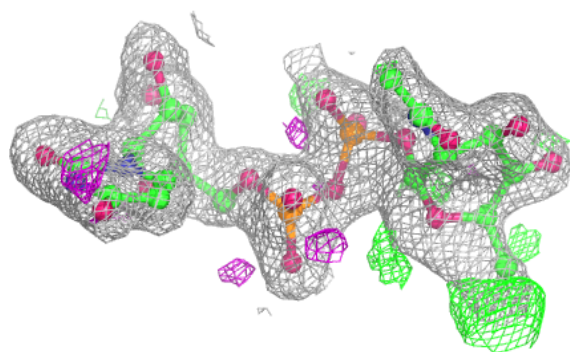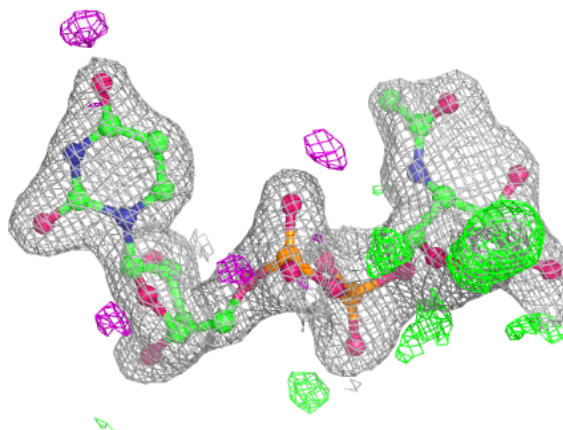**Electron density around UD7 A 403:**

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
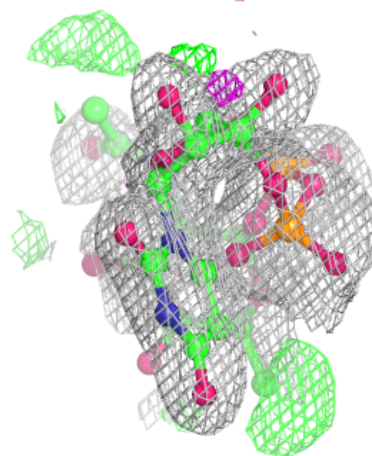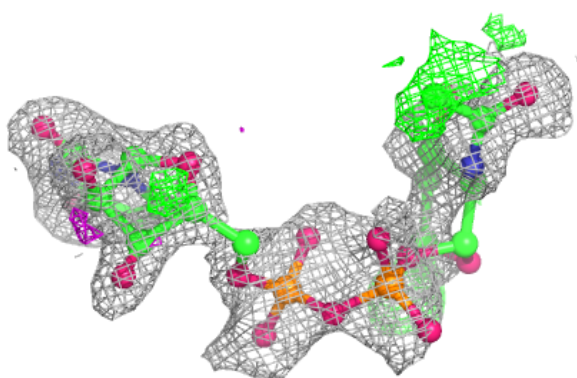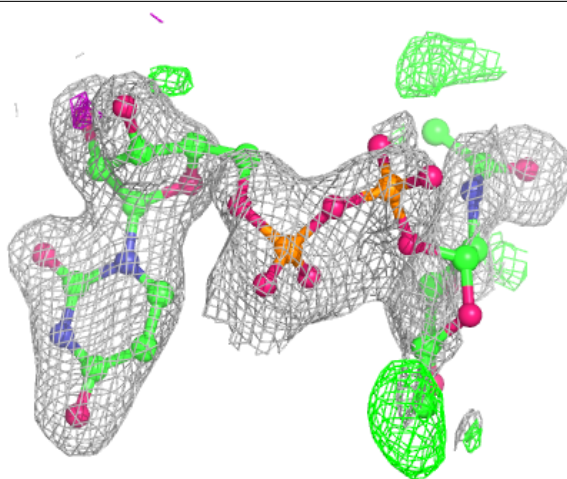
Electron density around NAP A 401:

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

Electron density around UD7 A 402:

$2mF_o$-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.