



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 19, 2024 – 02:53 PM EDT

PDB ID : 4D0T  
Title : GalNAc-T2 crystal soaked with UDP-GalNAc, EA2 peptide and manganese  
Authors : Lira-Navarrete, E.; Iglesias-Fernandez, J.; Zandberg, W.F.; Companon, I.; Kong, Y.; Corzana, F.; Pinto, B.M.; Clausen, H.; Peregrina, J.M.; Vocadlo, D.; Rovira, C.; Hurtado-Guerrero, R.  
Deposited on : 2014-04-30  
Resolution : 2.45 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
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

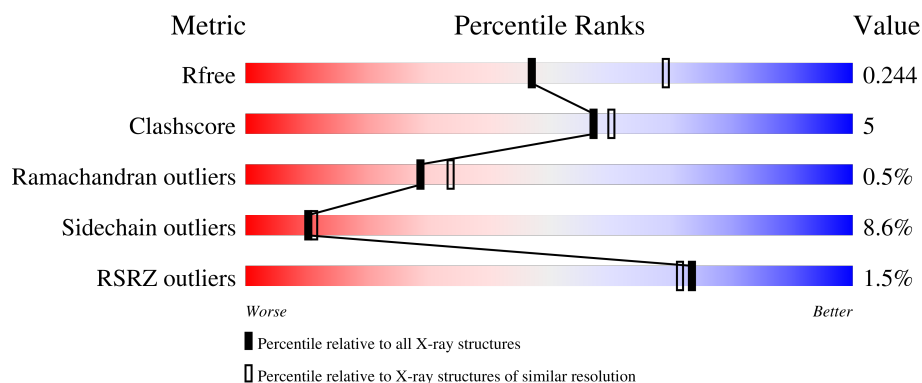
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*




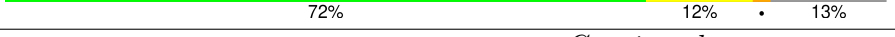
The reported resolution of this entry is 2.45 Å.

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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	571	
1	B	571	
1	C	571	
1	D	571	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	E	571	
1	F	571	
2	P	8	
2	Z	8	

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
7	NGA	P	1572	X	-	-	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 24500 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 POLYPEPTIDE N-ACETYL GALACTOSAMINYLTRANSFERASE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	496	Total	C	N	O	S	0	0	0
			3956	2490	718	724	24			
1	B	495	Total	C	N	O	S	0	0	0
			3947	2485	716	722	24			
1	C	495	Total	C	N	O	S	0	1	0
			3955	2490	719	722	24			
1	D	495	Total	C	N	O	S	0	3	0
			3967	2498	720	725	24			
1	E	495	Total	C	N	O	S	0	0	0
			3947	2485	716	722	24			
1	F	487	Total	C	N	O	S	0	0	0
			3892	2449	707	712	24			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	516	ASP	ASN	engineered mutation	UNP Q10471
B	516	ASP	ASN	engineered mutation	UNP Q10471
C	516	ASP	ASN	engineered mutation	UNP Q10471
D	516	ASP	ASN	engineered mutation	UNP Q10471
E	516	ASP	ASN	engineered mutation	UNP Q10471
F	516	ASP	ASN	engineered mutation	UNP Q10471

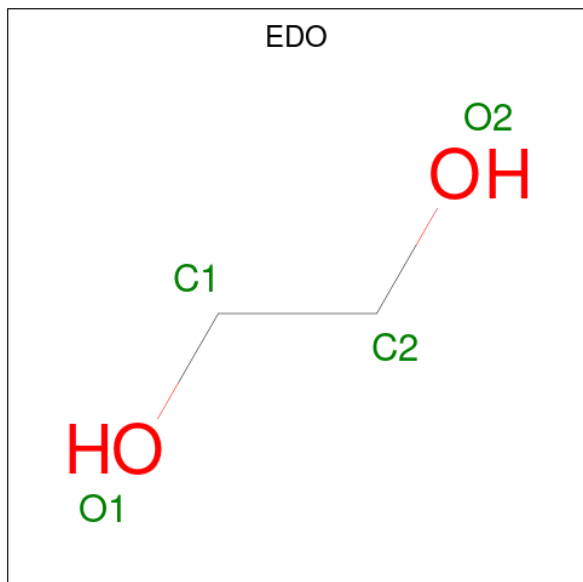
- Molecule 2 is a protein called PEPTIDE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	0	0	0
			47	28	7	12			
2	Z	6	Total	C	N	O	0	0	0
			40	25	6	9			

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mn 1 1	0	0
3	B	1	Total Mn 1 1	0	0
3	C	1	Total Mn 1 1	0	0
3	D	1	Total Mn 1 1	0	0
3	E	1	Total Mn 1 1	0	0
3	F	1	Total Mn 1 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



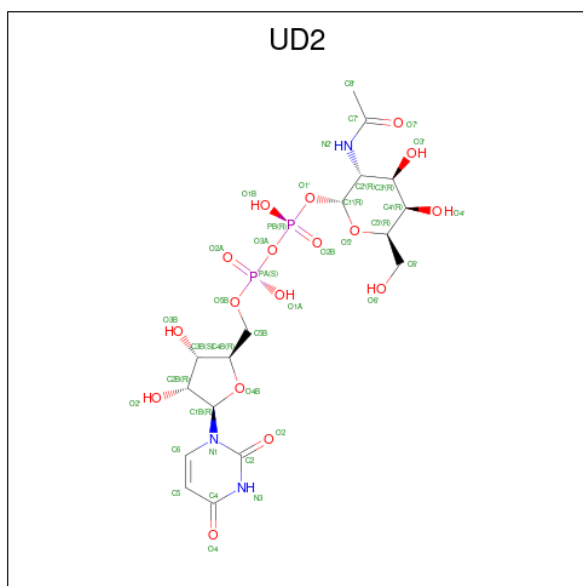
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0

*Continued on next page...*

Continued from previous page...

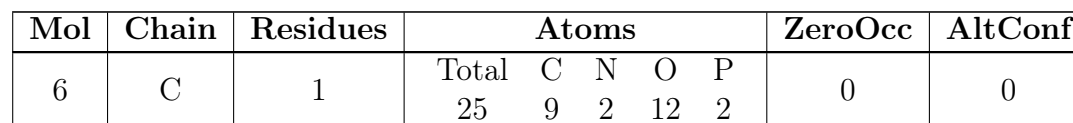
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			4	2	2		
4	D	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is URIDINE-DIPHOSPHATE-N-ACETYLGALACTOSAMINE (three-letter code: UD2) (formula:  $C_{17}H_{27}N_3O_{17}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			39	17	3	17	2		
5	B	1	Total	C	N	O	P	0	0
			39	17	3	17	2		
5	D	1	Total	C	N	O	P	0	0
			39	17	3	17	2		
5	E	1	Total	C	N	O	P	0	0
			39	17	3	17	2		
5	F	1	Total	C	N	O	P	0	0
			39	17	3	17	2		

- Molecule 6 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula:  $C_9H_{14}N_2O_{12}P_2$ ).



- 
- Chemical structure of NGA (N-Glycyl-L-alanine) showing a pyranose ring with substituents O1, O5, O6, O4, O3, and an amide group with N2 and C7.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	P	1	Total	C	N	O	0	0
			14	8	1	5		

- 

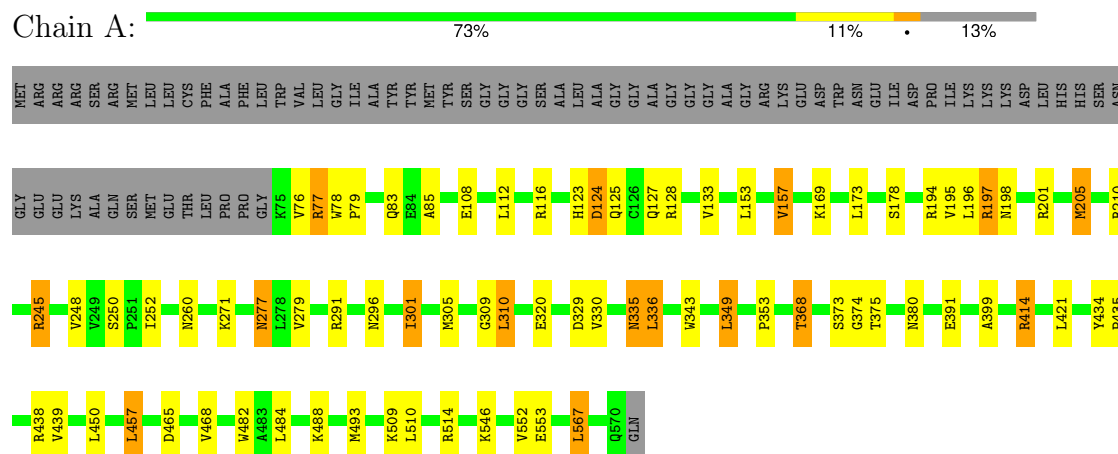
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	108	Total 108	O 108	0	0
8	B	99	Total 99	O 99	0	0
8	C	67	Total 67	O 67	0	0
8	D	55	Total 55	O 55	0	0
8	E	105	Total 105	O 105	0	0
8	F	34	Total 34	O 34	0	0
8	P	1	Total 1	O 1	0	0



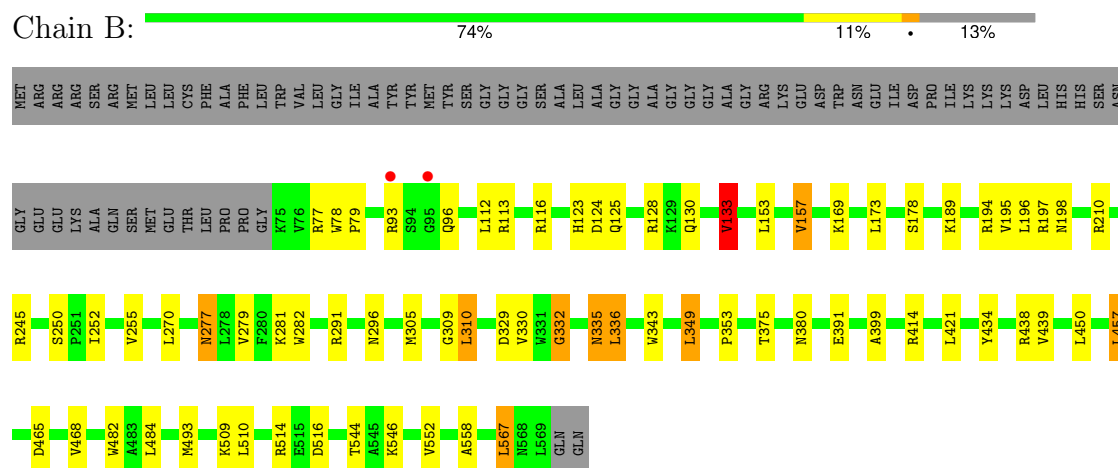
### 3 Residue-property plots

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.

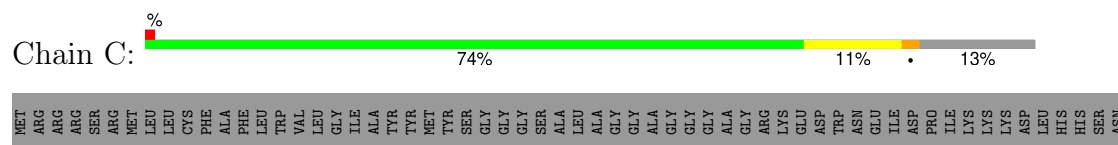
#### • Molecule 1: POLYPEPTIDE N-ACETYL GALACTOSAMINYLTRANSFERASE 2

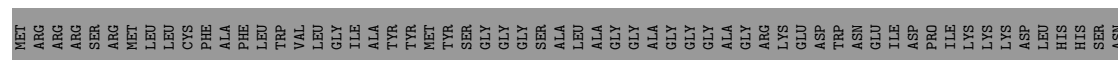


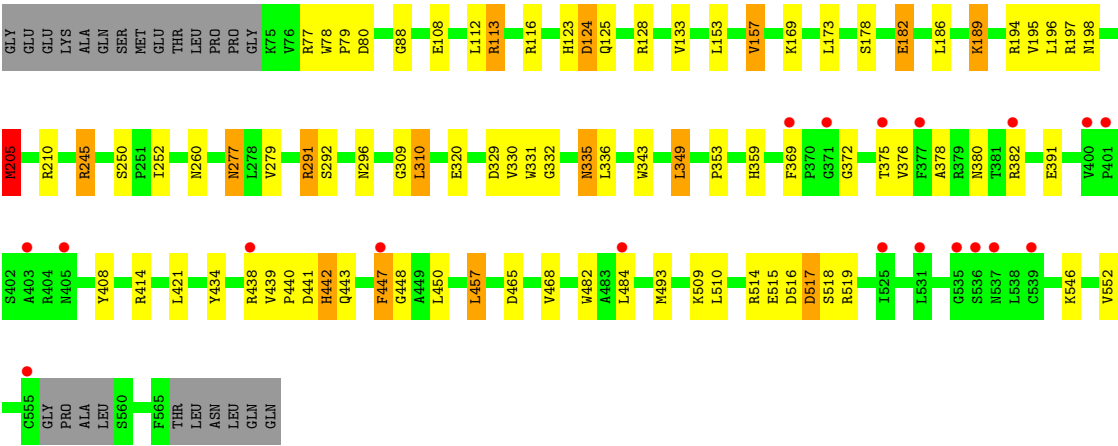
#### • Molecule 1: POLYPEPTIDE N-ACETYL GALACTOSAMINYLTRANSFERASE 2



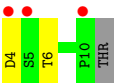
#### • Molecule 1: POLYPEPTIDE N-ACETYL GALACTOSAMINYLTRANSFERASE 2







● Molecule 2: PEPTIDE



● Molecule 2: PEPTIDE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.22Å 122.37Å 249.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	249.72 – 2.45 70.07 – 2.45	Depositor EDS
% Data completeness (in resolution range)	99.7 (249.72-2.45) 99.7 (70.07-2.45)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.214 , 0.241 0.218 , 0.244	Depositor DCC
$R_{free}$ test set	3670 reflections (2.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.9	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 32.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	24500	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.86% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: EDO, UDP, MN, UD2, NGA

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	0.68	1/4048 (0.0%)	0.86	13/5478 (0.2%)
1	B	0.66	0/4039	0.86	10/5466 (0.2%)
1	C	0.65	1/4050 (0.0%)	0.89	12/5480 (0.2%)
1	D	0.62	1/4068 (0.0%)	0.85	6/5504 (0.1%)
1	E	0.67	2/4039 (0.0%)	0.85	12/5466 (0.2%)
1	F	0.61	2/3982 (0.1%)	0.89	11/5385 (0.2%)
2	P	0.91	0/48	1.22	0/67
2	Z	1.51	1/41 (2.4%)	1.57	1/58 (1.7%)
All	All	0.65	8/24315 (0.0%)	0.87	65/32904 (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	2
1	C	0	2
1	D	0	2
1	F	0	2
2	Z	0	1
All	All	0	9

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	292	SER	CB-OG	-6.18	1.34	1.42
1	A	553	GLU	CD-OE1	5.88	1.32	1.25
1	F	332	GLY	N-CA	5.86	1.54	1.46
1	E	319	GLU	CG-CD	5.55	1.60	1.51
2	Z	10	PRO	N-CD	5.18	1.55	1.47

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	194	ARG	NE-CZ-NH2	-18.29	111.15	120.30
1	F	194	ARG	NE-CZ-NH1	11.41	126.01	120.30
1	C	135	LEU	CA-CB-CG	9.32	136.74	115.30
1	B	113	ARG	NE-CZ-NH1	8.91	124.75	120.30
1	C	135	LEU	CB-CG-CD1	7.34	123.48	111.00

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	133	VAL	Peptide
1	B	332	GLY	Peptide
1	C	332	GLY	Peptide
1	C	370	PRO	Peptide
1	D	370	PRO	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	3956	0	3862	38	0
1	B	3947	0	3854	43	0
1	C	3955	0	3867	27	0
1	D	3967	0	3881	44	0
1	E	3947	0	3854	29	0
1	F	3892	0	3792	49	0
2	P	47	0	40	0	0
2	Z	40	0	39	28	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	8	0	12	2	0
4	B	8	0	12	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	8	0	12	0	0
4	D	8	0	12	1	0
4	E	8	0	12	3	0
5	A	39	0	25	5	0
5	B	39	0	25	3	0
5	D	39	0	25	5	0
5	E	39	0	25	2	0
5	F	39	0	25	3	0
6	C	25	0	11	0	0
7	P	14	0	13	0	0
8	A	108	0	0	2	0
8	B	99	0	0	4	0
8	C	67	0	0	2	0
8	D	55	0	0	1	0
8	E	105	0	0	1	0
8	F	34	0	0	1	0
8	P	1	0	0	0	0
All	All	24500	0	23398	232	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 5.

The worst 5 of 232 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:D:277:ASN:OD1	1:D:279:VAL:HB	1.56	1.06
1:B:270:LEU:HB2	2:Z:11:THR:HG21	1.37	1.06
1:B:270:LEU:HD12	2:Z:11:THR:CG2	1.87	1.04
1:F:442:HIS:HB3	1:F:443:GLN:HA	1.49	0.94
1:A:271:LYS:HG2	1:A:301:ILE:HD11	1.48	0.93

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	494/571 (86%)	482 (98%)	11 (2%)	1 (0%)	47	57
1	B	493/571 (86%)	481 (98%)	10 (2%)	2 (0%)	34	41
1	C	494/571 (86%)	478 (97%)	15 (3%)	1 (0%)	47	57
1	D	496/571 (87%)	480 (97%)	12 (2%)	4 (1%)	19	22
1	E	493/571 (86%)	481 (98%)	11 (2%)	1 (0%)	47	57
1	F	483/571 (85%)	466 (96%)	14 (3%)	3 (1%)	25	29
2	P	5/8 (62%)	4 (80%)	0	1 (20%)	0	0
2	Z	4/8 (50%)	1 (25%)	1 (25%)	2 (50%)	0	0
All	All	2962/3442 (86%)	2873 (97%)	74 (2%)	15 (0%)	29	34

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	94	SER
1	D	280	PHE
2	P	6	THR
2	Z	10	PRO
1	F	442	HIS

### 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	424/485 (87%)	388 (92%)	36 (8%)	10	12
1	B	423/485 (87%)	388 (92%)	35 (8%)	11	12
1	C	424/485 (87%)	385 (91%)	39 (9%)	9	9
1	D	426/485 (88%)	390 (92%)	36 (8%)	10	12
1	E	423/485 (87%)	386 (91%)	37 (9%)	10	11
1	F	417/485 (86%)	382 (92%)	35 (8%)	11	12

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	P	6/7 (86%)	5 (83%)	1 (17%)	2	1
2	Z	5/7 (71%)	5 (100%)	0	100	100
All	All	2548/2924 (87%)	2329 (91%)	219 (9%)	10	11

5 of 219 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	133	VAL
1	D	516	ASP
1	F	291	ARG
1	D	189	LYS
1	D	336	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 87 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	474	HIS
1	E	462	HIS
1	D	568	ASN
1	E	296	ASN
1	F	198	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 23 ligands modelled in this entry, 6 are 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	EDO	A	1573	-	3,3,3	0.64	0	2,2,2	0.07	0
4	EDO	B	1571	-	3,3,3	0.57	0	2,2,2	0.09	0
4	EDO	D	1573	-	3,3,3	0.63	0	2,2,2	0.16	0
5	UD2	D	1571	3	40,41,41	1.11	4 (10%)	59,62,62	1.91	18 (30%)
5	UD2	F	1567	3	40,41,41	0.89	2 (5%)	59,62,62	1.85	12 (20%)
4	EDO	E	1573	-	3,3,3	0.71	0	2,2,2	0.71	0
4	EDO	B	1572	-	3,3,3	0.66	0	2,2,2	0.09	0
4	EDO	C	1573	-	3,3,3	0.47	0	2,2,2	0.27	0
5	UD2	E	1571	3	40,41,41	0.96	1 (2%)	59,62,62	2.20	19 (32%)
4	EDO	E	1572	-	3,3,3	0.72	0	2,2,2	0.16	0
7	NGA	P	1572	2	14,14,15	1.05	1 (7%)	17,19,21	1.40	2 (11%)
6	UDP	C	1571	3	25,26,26	1.17	3 (12%)	38,40,40	1.52	8 (21%)
5	UD2	B	1573	3	40,41,41	1.06	3 (7%)	59,62,62	2.00	21 (35%)
5	UD2	A	1574	3	40,41,41	1.08	3 (7%)	59,62,62	1.86	17 (28%)
4	EDO	D	1572	-	3,3,3	0.75	0	2,2,2	0.26	0
4	EDO	C	1574	-	3,3,3	0.70	0	2,2,2	0.47	0
4	EDO	A	1572	-	3,3,3	0.57	0	2,2,2	0.27	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
4	EDO	A	1573	-	-	1/1/1/1	-
4	EDO	B	1571	-	-	1/1/1/1	-
4	EDO	D	1573	-	-	1/1/1/1	-
5	UD2	D	1571	3	-	7/26/63/63	0/3/3/3
5	UD2	F	1567	3	-	9/26/63/63	0/3/3/3
4	EDO	E	1573	-	-	1/1/1/1	-
4	EDO	B	1572	-	-	1/1/1/1	-
4	EDO	C	1573	-	-	0/1/1/1	-
5	UD2	E	1571	3	-	7/26/63/63	0/3/3/3

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	NGA	P	1572	2	1/1/5/7	1/6/23/26	0/1/1/1
4	EDO	E	1572	-	-	1/1/1/1	-
6	UDP	C	1571	3	-	1/16/32/32	0/2/2/2
5	UD2	B	1573	3	-	6/26/63/63	0/3/3/3
5	UD2	A	1574	3	-	8/26/63/63	0/3/3/3
4	EDO	D	1572	-	-	1/1/1/1	-
4	EDO	C	1574	-	-	0/1/1/1	-
4	EDO	A	1572	-	-	0/1/1/1	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	P	1572	NGA	C1-C2	3.21	1.56	1.52
5	A	1574	UD2	C2-N3	-2.90	1.32	1.38
5	B	1573	UD2	C2-N1	-2.81	1.34	1.38
5	D	1571	UD2	C4-N3	-2.54	1.34	1.38
5	A	1574	UD2	O2-C2	2.53	1.27	1.23

The worst 5 of 97 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	1571	UD2	O5'-C1'-C2'	-7.00	97.36	110.59
5	B	1573	UD2	O5'-C1'-C2'	-6.90	97.55	110.59
5	E	1571	UD2	N3-C2-N1	5.54	122.11	114.89
5	F	1567	UD2	O5'-C1'-C2'	-5.20	100.77	110.59
5	F	1567	UD2	N3-C2-N1	5.10	121.53	114.89

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	P	1572	NGA	C1

5 of 46 torsion outliers are listed below:

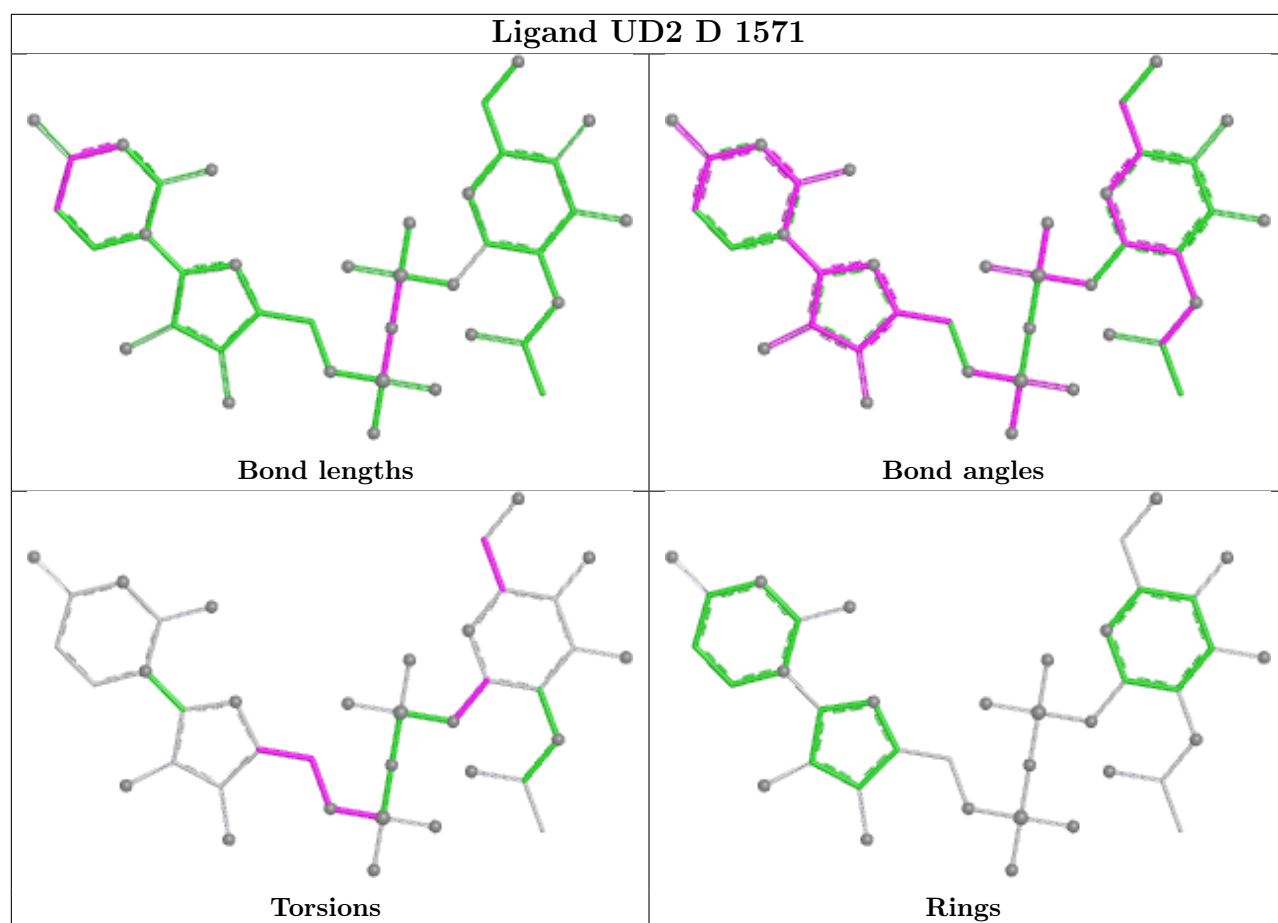
Mol	Chain	Res	Type	Atoms
5	A	1574	UD2	C5B-O5B-PA-O3A
5	B	1573	UD2	O5'-C1'-O1'-PB
5	B	1573	UD2	C5B-O5B-PA-O3A
5	D	1571	UD2	C5B-O5B-PA-O3A
5	E	1571	UD2	O5'-C1'-O1'-PB

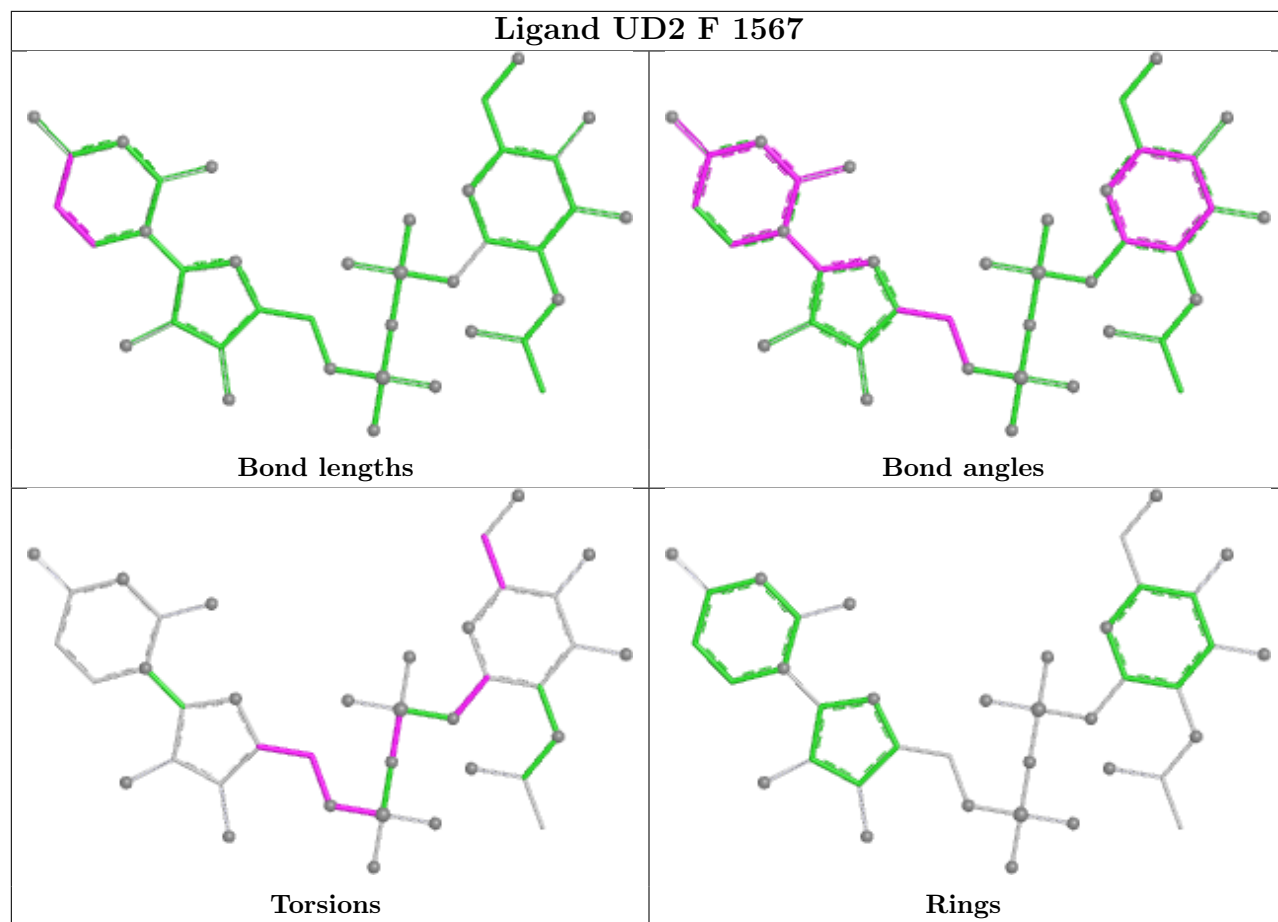
There are no ring outliers.

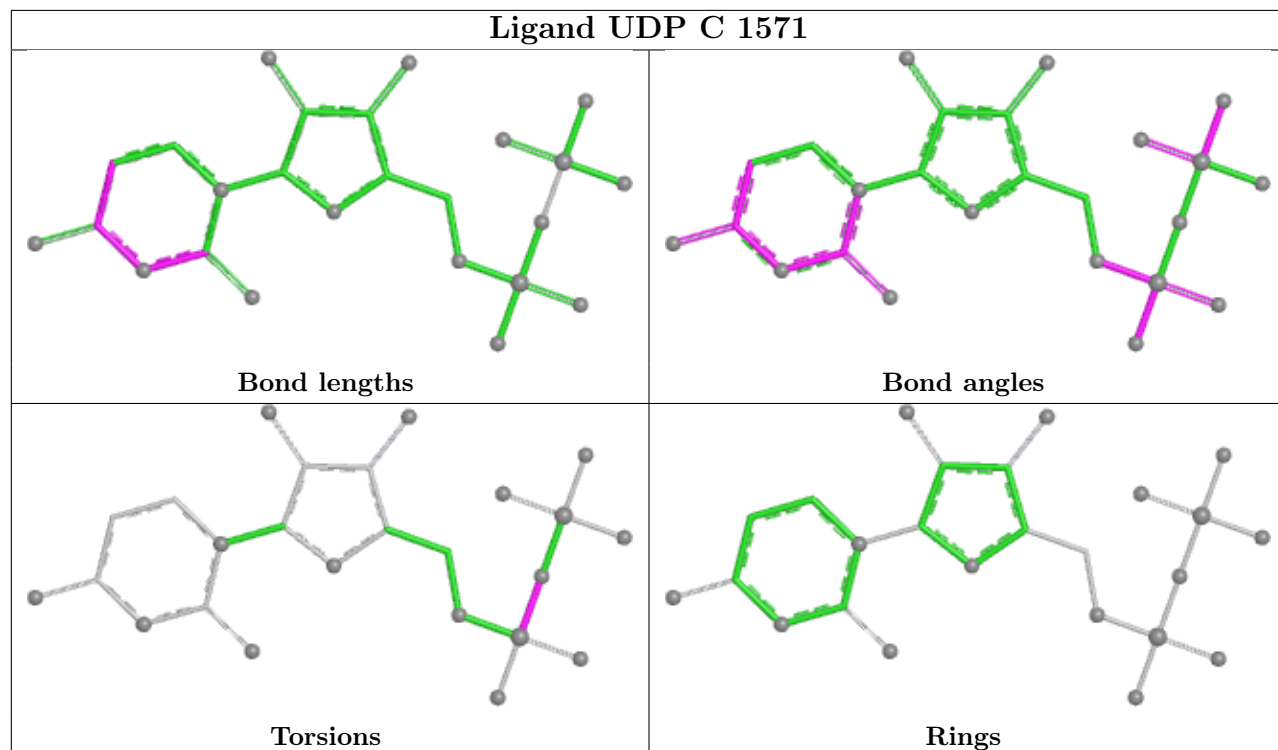
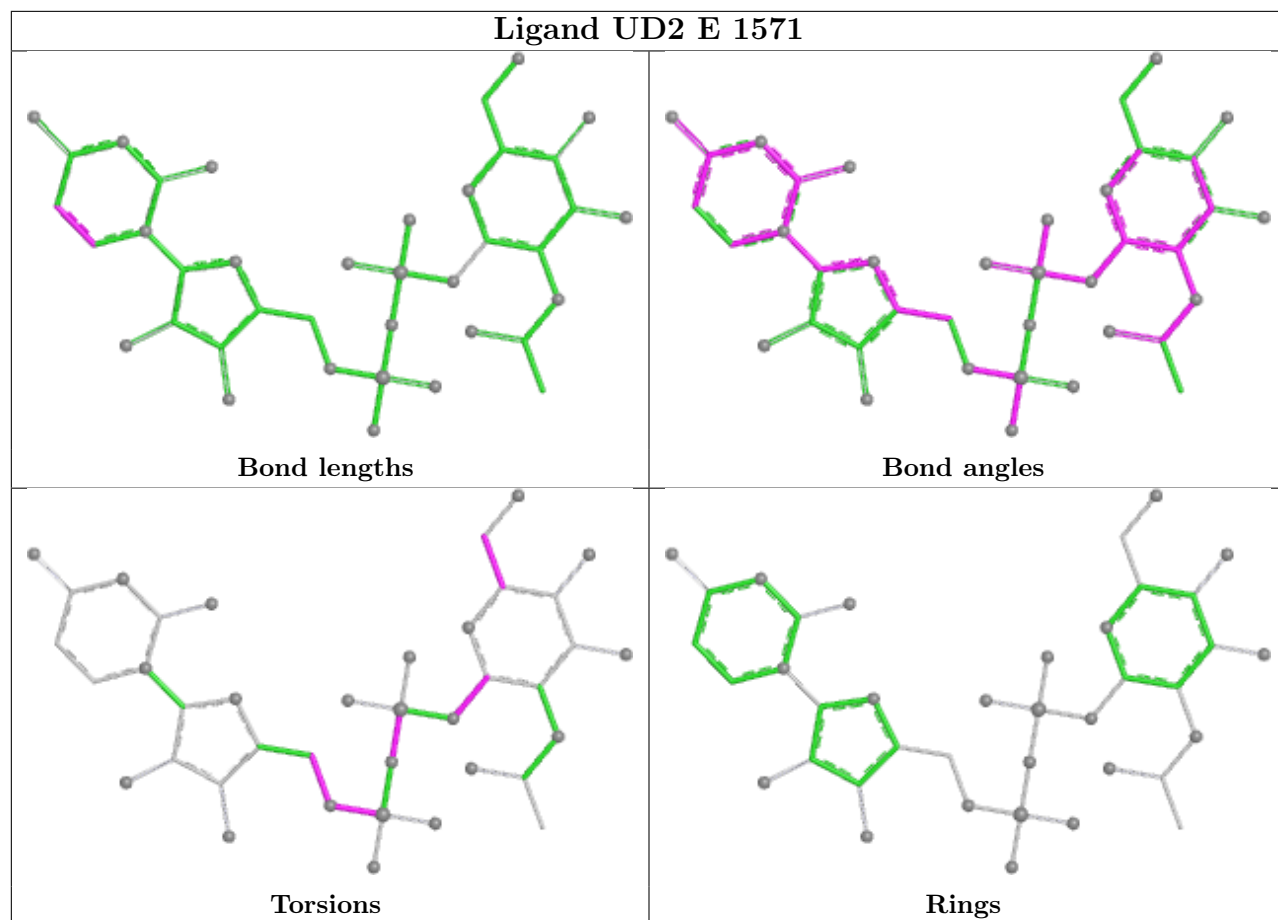
10 monomers are involved in 25 short contacts:

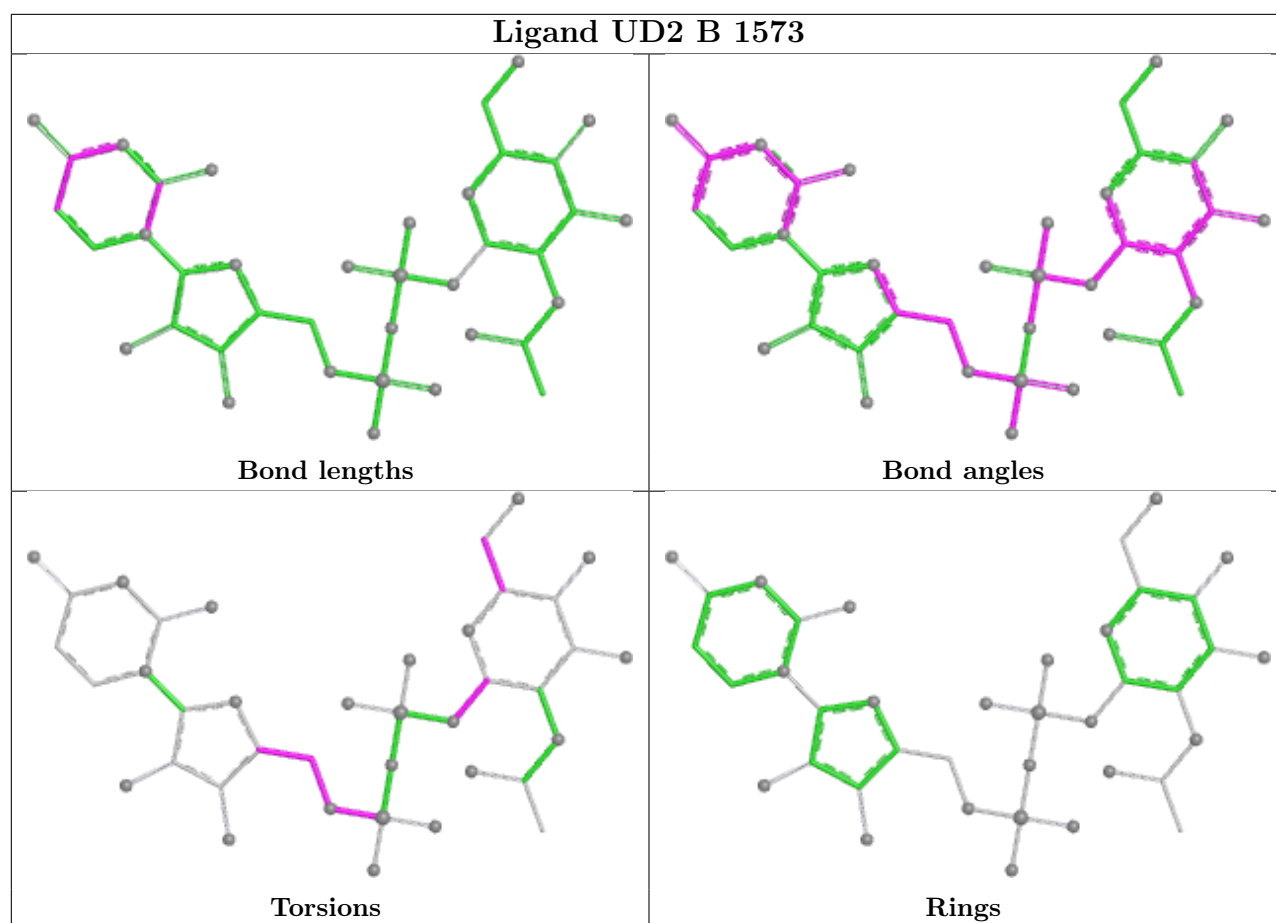
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1573	EDO	2	0
5	D	1571	UD2	5	0
5	F	1567	UD2	3	0
4	E	1573	EDO	1	0
4	B	1572	EDO	1	0
5	E	1571	UD2	2	0
4	E	1572	EDO	2	0
5	B	1573	UD2	3	0
5	A	1574	UD2	5	0
4	D	1572	EDO	1	0

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 than 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.

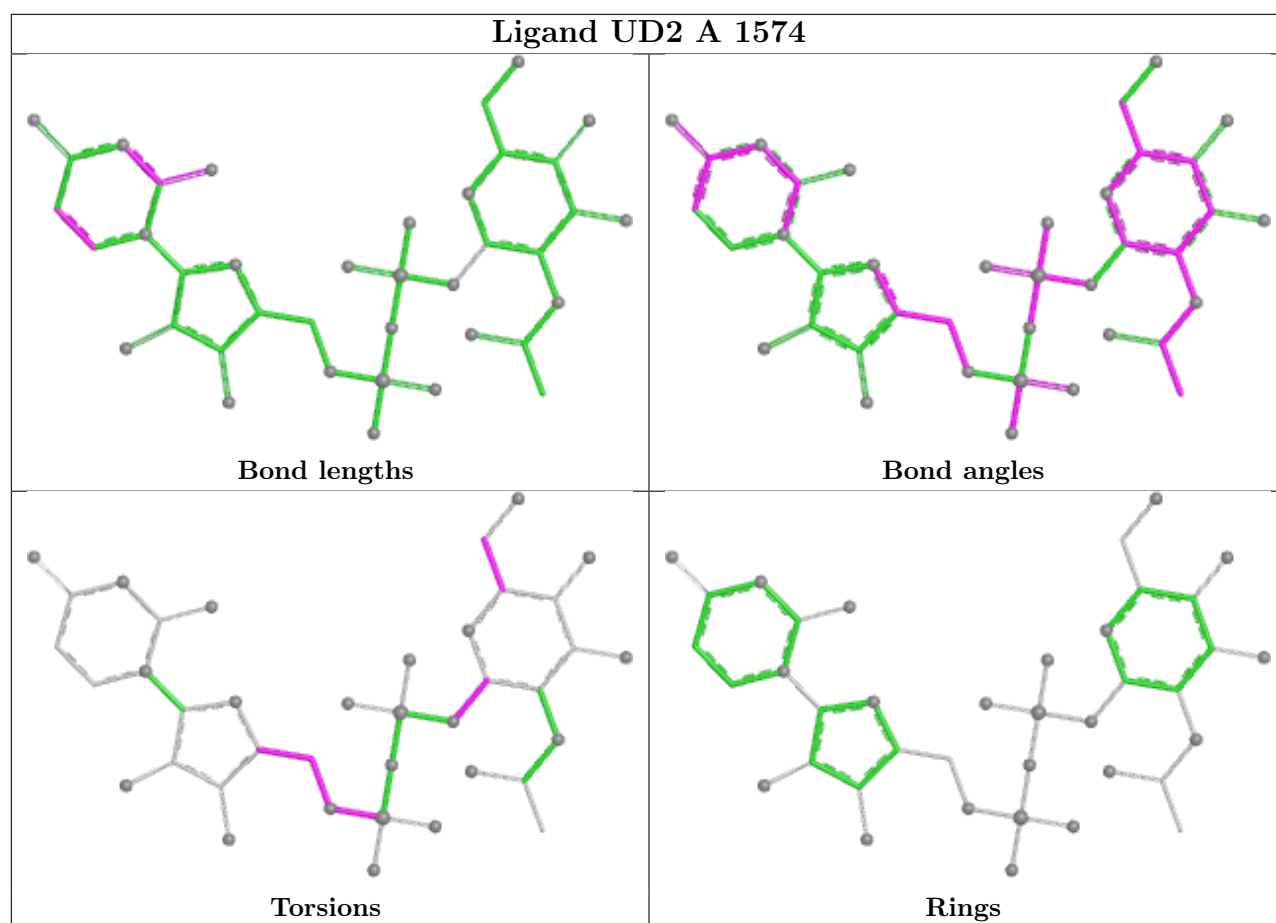












## 5.7 Other polymers [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	496/571 (86%)	-0.35	0 100 100	22, 36, 63, 96	0
1	B	495/571 (86%)	-0.32	2 (0%) 92 93	25, 39, 67, 119	0
1	C	495/571 (86%)	-0.28	8 (1%) 72 69	26, 43, 72, 119	0
1	D	495/571 (86%)	-0.11	7 (1%) 75 74	30, 50, 80, 111	0
1	E	495/571 (86%)	-0.35	1 (0%) 95 95	21, 37, 60, 93	0
1	F	487/571 (85%)	0.13	19 (3%) 39 36	23, 56, 97, 123	0
2	P	7/8 (87%)	1.92	3 (42%) 0 0	25, 73, 78, 78	0
2	Z	6/8 (75%)	3.92	4 (66%) 0 0	65, 76, 103, 103	0
All	All	2976/3442 (86%)	-0.20	44 (1%) 73 71	21, 42, 80, 123	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	Z	11	THR	8.2
2	Z	7	THR	5.6
2	Z	10	PRO	4.6
1	D	291	ARG	4.4
1	D	284	TYR	4.4

### 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 ⓘ

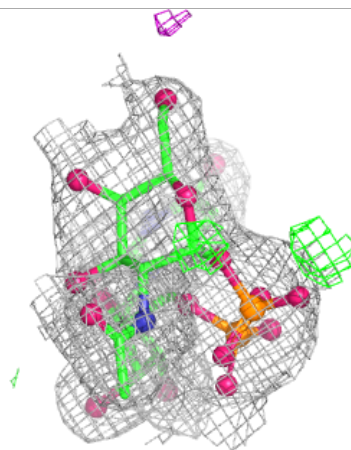
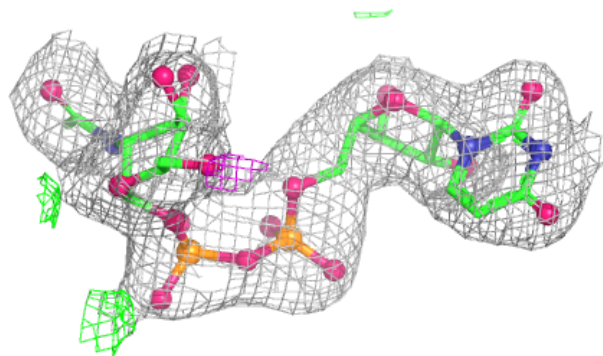
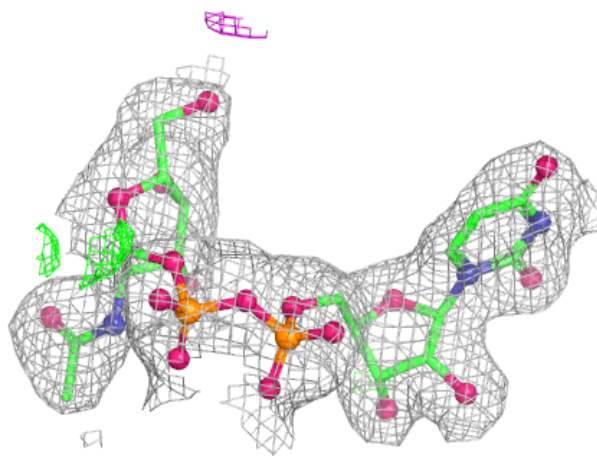
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	EDO	E	1572	4/4	0.76	0.25	54,58,60,62	0
4	EDO	D	1572	4/4	0.78	0.34	48,48,55,62	0
4	EDO	A	1572	4/4	0.82	0.21	51,54,55,56	0
4	EDO	D	1573	4/4	0.85	0.30	47,48,49,53	0
4	EDO	B	1572	4/4	0.85	0.19	51,55,56,59	0
4	EDO	C	1574	4/4	0.86	0.22	51,53,54,54	0
4	EDO	A	1573	4/4	0.86	0.35	48,48,49,51	0
4	EDO	C	1573	4/4	0.89	0.16	44,46,50,51	0
4	EDO	B	1571	4/4	0.90	0.18	55,55,57,58	0
4	EDO	E	1573	4/4	0.92	0.15	39,41,43,44	0
5	UD2	F	1567	39/39	0.95	0.12	32,50,65,68	0
7	NGA	P	1572	14/15	0.95	0.13	31,41,47,50	0
5	UD2	D	1571	39/39	0.96	0.12	29,37,44,47	0
3	MN	B	1570	1/1	0.96	0.13	34,34,34,34	0
5	UD2	B	1573	39/39	0.96	0.13	34,38,49,51	0
5	UD2	E	1571	39/39	0.97	0.13	23,30,38,42	0
5	UD2	A	1574	39/39	0.98	0.11	20,27,40,41	0
6	UDP	C	1571	25/25	0.98	0.11	31,34,36,38	0
3	MN	F	1566	1/1	0.98	0.10	40,40,40,40	0
3	MN	C	1570	1/1	0.99	0.11	29,29,29,29	0
3	MN	D	1570	1/1	0.99	0.14	32,32,32,32	0
3	MN	E	1570	1/1	0.99	0.12	29,29,29,29	0
3	MN	A	1571	1/1	1.00	0.13	25,25,25,25	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.

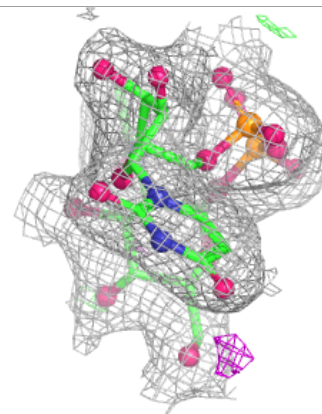
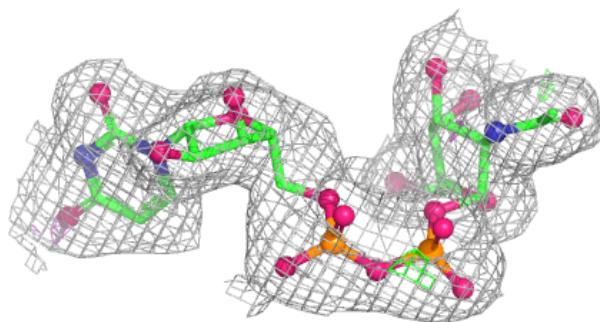
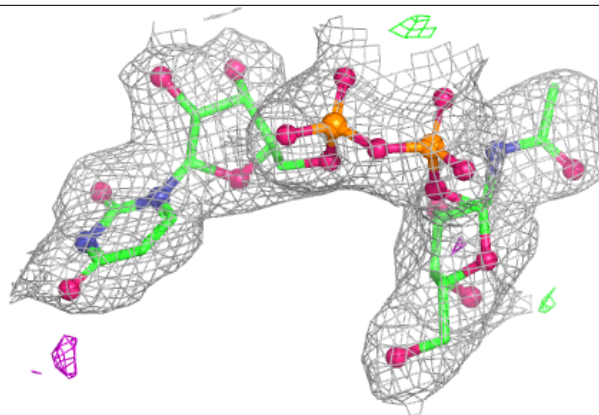
**Electron density around UD2 F 1567:**

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

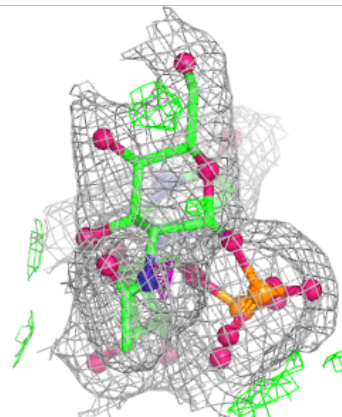
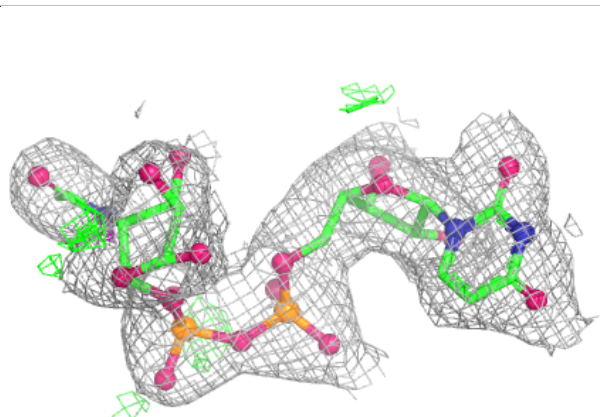
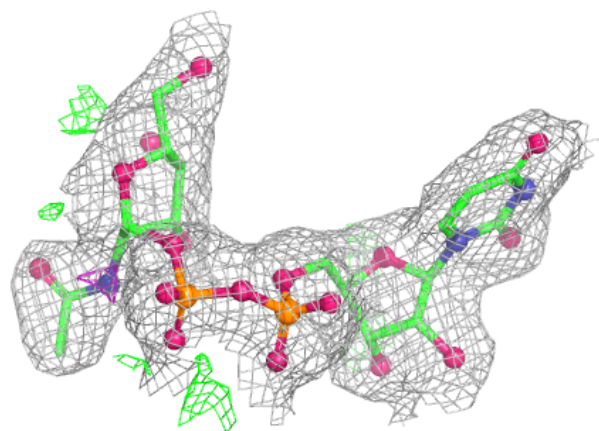


**Electron density around UD2 D 1571:**

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

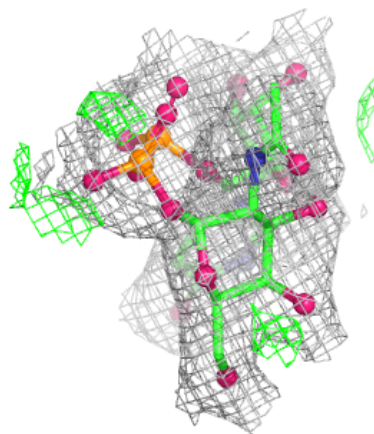
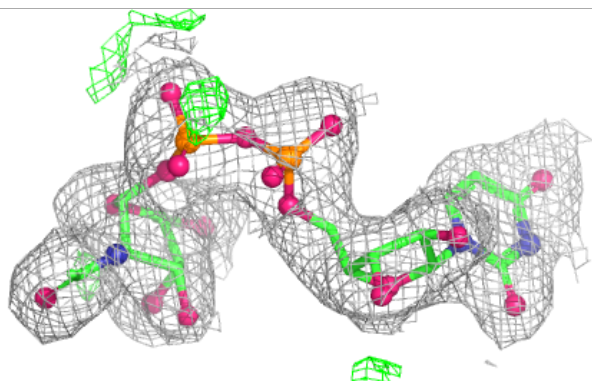
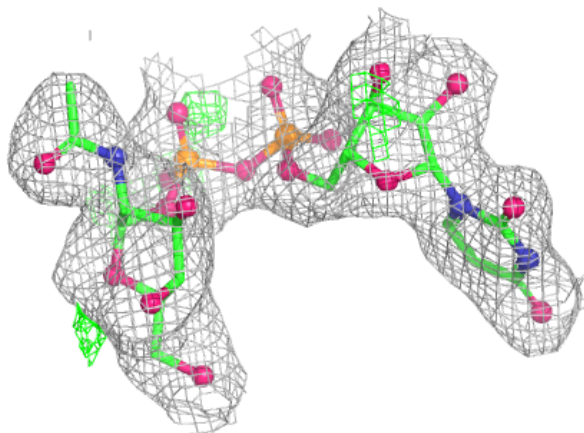
**Electron density around UD2 B 1573:**

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



**Electron density around UD2 E 1571:**

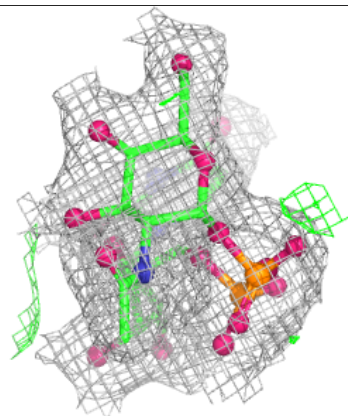
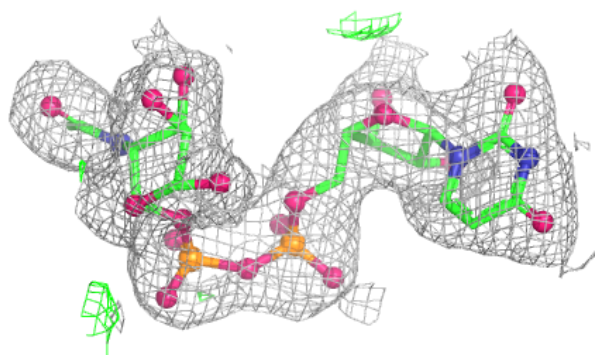
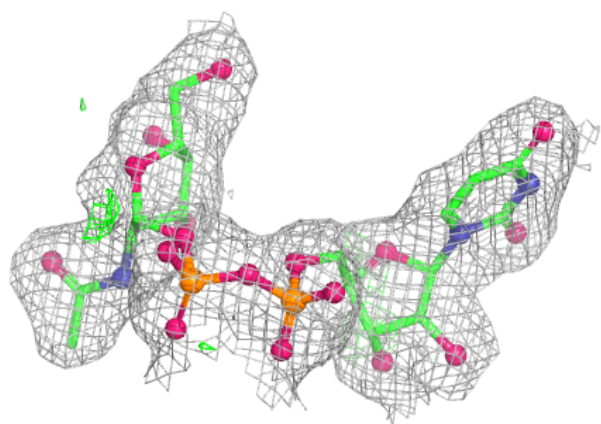
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



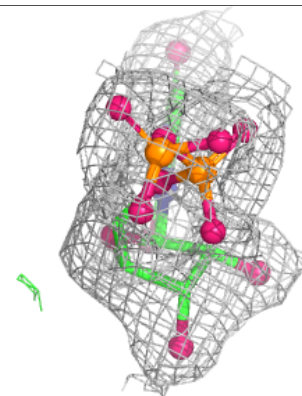
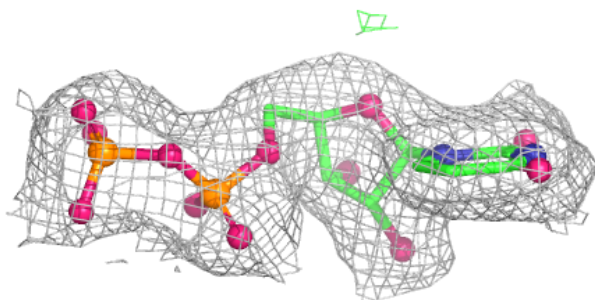
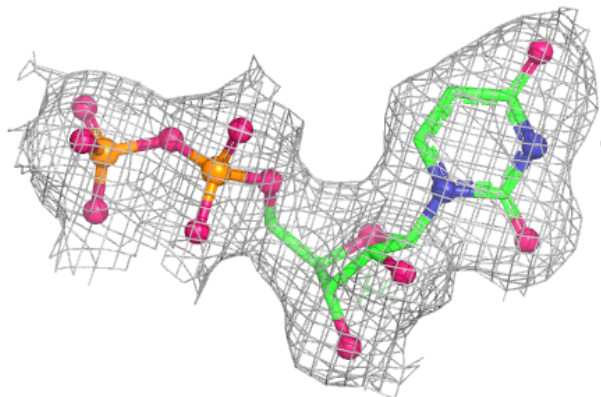


**Electron density around UD2 A 1574:**

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

**Electron density around UDP C 1571:**

$2mF_o-DF_c$  (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.