wwPDB X-ray Structure Validation Summary Report (i)

Apr 27, 2024 – 04:36 pm BST

PDB ID	:	2VL8
Title	:	CRYSTAL STRUCTURE OF THE CATALYTIC DOMAIN OF LETHAL TOXIN FROM CLOSTRIDIUM SORDELLII IN COMPLEX WITH UDP, CASTANOSPERMINE AND CALCIUM ION
Authors	:	Jank, T.; Ziegler, M.O.P.; Schulz, G.E.; Aktories, K.
Deposited on	:	2008-01-09
Resolution	:	2.31 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

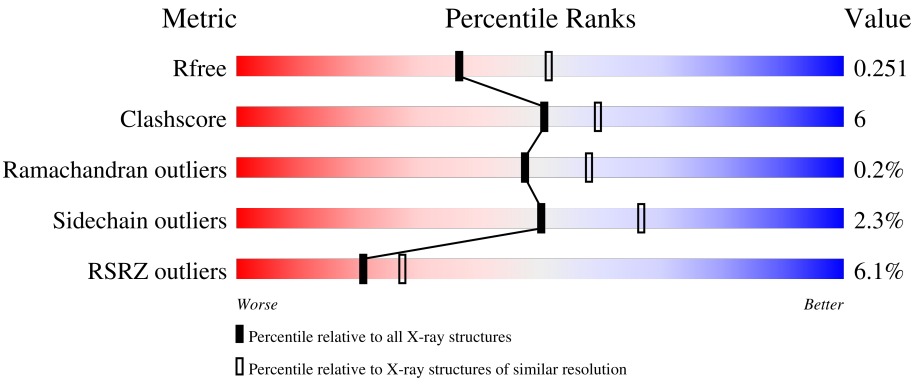
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.31 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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	546	4% 87% 12%
1	B	546	7% 83% 13%
1	C	546	7% 86% 11%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13720 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.

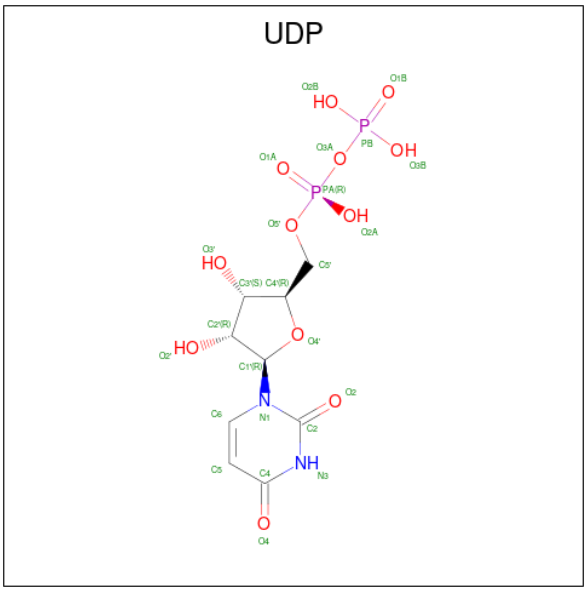
- Molecule 1 is a protein called CYTOTOXIN L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	539	Total	C	N	O	S	0	0	0
			4419	2820	716	867	16			
1	B	533	Total	C	N	O	S	0	0	0
			4374	2792	708	859	15			
1	C	537	Total	C	N	O	S	0	0	0
			4408	2814	714	864	16			

There are 6 discrepancies between the modelled and reference sequences:

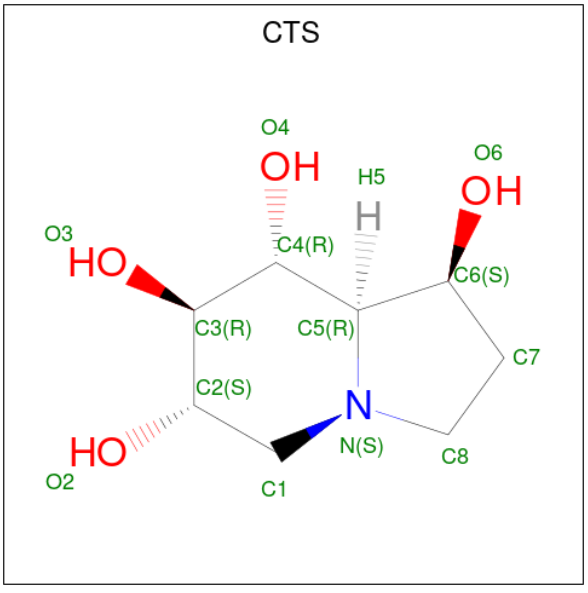
Chain	Residue	Modelled	Actual	Comment	Reference
A	13	ALA	VAL	engineered mutation	UNP Q46342
A	289	MET	ILE	engineered mutation	UNP Q46342
B	13	ALA	VAL	engineered mutation	UNP Q46342
B	289	MET	ILE	engineered mutation	UNP Q46342
C	13	ALA	VAL	engineered mutation	UNP Q46342
C	289	MET	ILE	engineered mutation	UNP Q46342

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



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

- Molecule 3 is CASTANOSPERMINE (three-letter code: CTS) (formula: $C_8H_{15}NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			13	8	1	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C N O 13 8 1 4	0	0
3	C	1	Total C N O 13 8 1 4	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	3	Total Ca 3 3	0	0
4	B	1	Total Ca 1 1	0	0
4	C	1	Total Ca 1 1	0	0

- Molecule 5 is water.

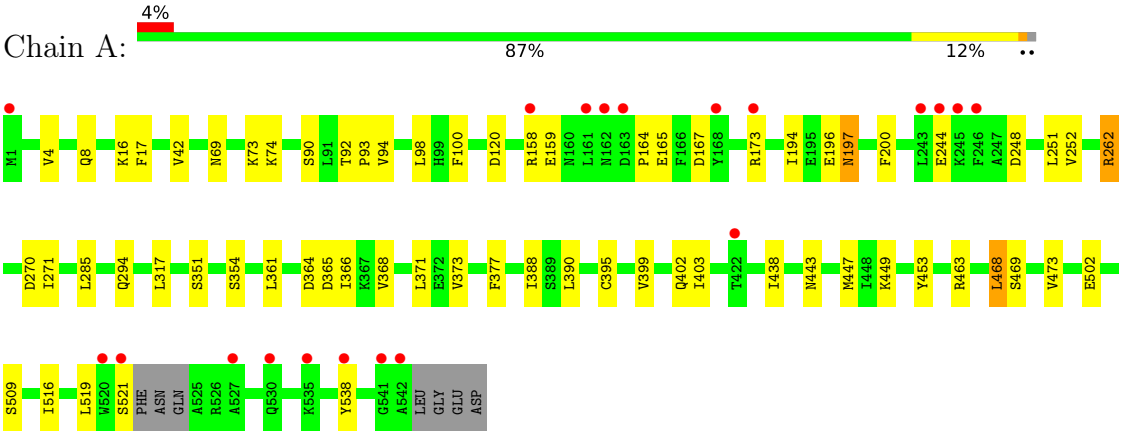
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	139	Total O 139 139	0	0
5	B	129	Total O 129 129	0	0
5	C	132	Total O 132 132	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

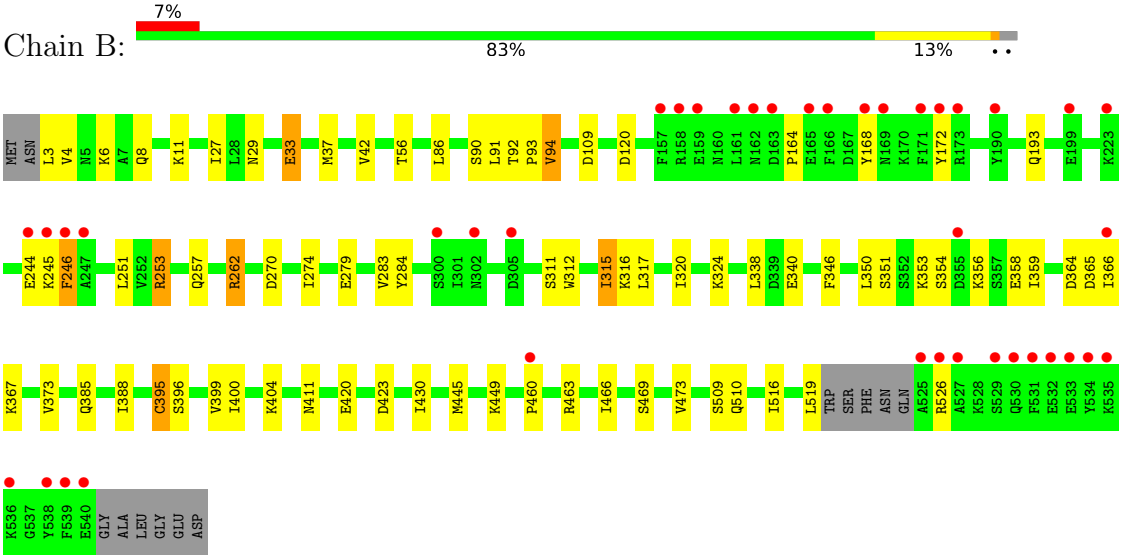
• Molecule 1: CYTOTOXIN L

Chain A:



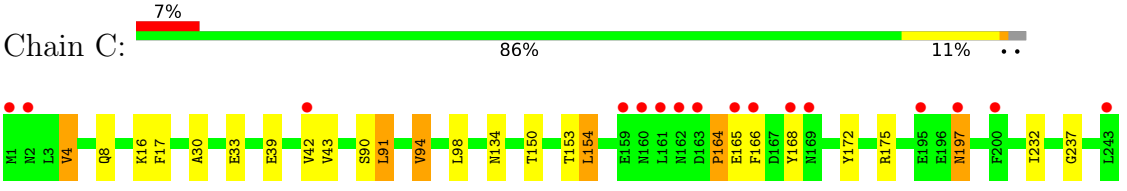
• Molecule 1: CYTOTOXIN L

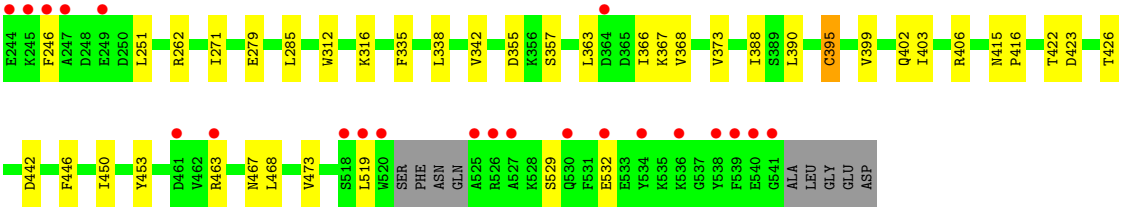
Chain B:



• Molecule 1: CYTOTOXIN L

Chain C:




E244
K245
F246
A247
D248
E249
D250
L251
R262
I271
E279
L285
W312
K316
F335
L338
V342
D355
K356
S357
L363
D364
D365
I366
K367
V368
V373
I388
S389
L390
C395
V399
Q402
I403
R406
N415
P416
T422
D423
T426
D442
F446
I450
Y453
D461
V462
R463
N467
L468
V473
S518
L519
W520
SER
PHE
ASN
GLN
A525
R526
A527
K528
S529
Q530
F531
E532
E533
Y534
K535
K536
G537
Y538
F539
E540
G541
ALA
LEU
GLY
GLU
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	57.69Å 191.33Å 205.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	95.78 – 2.31 95.67 – 2.31	Depositor EDS
% Data completeness (in resolution range)	99.6 (95.78-2.31) 99.6 (95.67-2.31)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.64 (at 2.32Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.217 , 0.255 0.216 , 0.251	Depositor DCC
R_{free} test set	3013 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å^2)	29.0	Xtriage
Anisotropy	0.533	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 42.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	13720	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 31.01 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1886e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: UDP, CA, CTS

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	0.66	0/4500	0.67	1/6072 (0.0%)
1	B	0.66	0/4453	0.67	2/6008 (0.0%)
1	C	0.67	0/4489	0.67	0/6057
All	All	0.66	0/13442	0.67	3/18137 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	468	LEU	CA-CB-CG	5.44	127.81	115.30
1	B	253	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	B	253	ARG	NE-CZ-NH2	-5.01	117.80	120.30

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	4419	0	4382	46	0
1	B	4374	0	4340	60	0
1	C	4408	0	4372	45	0
2	A	25	0	11	0	0
2	B	25	0	11	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	25	0	11	0	0
3	A	13	0	15	0	0
3	B	13	0	15	0	0
3	C	13	0	15	0	0
4	A	3	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	139	0	0	4	0
5	B	129	0	0	3	0
5	C	132	0	0	0	0
All	All	13720	0	13172	151	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 6.

The worst 5 of 151 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:262:ARG:HH11	1:B:262:ARG:HG3	0.89	1.05
1:B:262:ARG:HG3	1:B:262:ARG:NH1	1.66	0.96
1:B:262:ARG:HH11	1:B:262:ARG:CG	1.80	0.95
1:A:4:VAL:HG22	1:A:8:GLN:HB2	1.62	0.81
1:A:196:GLU:HB3	1:A:197:ASN:ND2	1.95	0.81

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	535/546 (98%)	517 (97%)	17 (3%)	1 (0%)	47 58
1	B	529/546 (97%)	505 (96%)	23 (4%)	1 (0%)	47 58

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	533/546 (98%)	517 (97%)	15 (3%)	1 (0%)	47 58
All	All	1597/1638 (98%)	1539 (96%)	55 (3%)	3 (0%)	47 58

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	164	PRO
1	C	164	PRO
1	A	164	PRO

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	499/505 (99%)	487 (98%)	12 (2%)	49 65
1	B	495/505 (98%)	483 (98%)	12 (2%)	49 65
1	C	498/505 (99%)	487 (98%)	11 (2%)	52 68
All	All	1492/1515 (98%)	1457 (98%)	35 (2%)	50 66

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	94	VAL
1	C	134	ASN
1	C	395	CYS
1	B	6	LYS
1	A	538	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	238	ASN
1	C	242	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	415	ASN
1	B	53	ASN
1	A	510	GLN

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 11 ligands modelled in this entry, 5 are monoatomic - leaving 6 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	CTS	C	1544	-	14,14,14	1.28	2 (14%)	16,21,21	0.85	1 (6%)
2	UDP	A	1543	4	24,26,26	0.92	1 (4%)	37,40,40	1.67	7 (18%)
2	UDP	B	1543	4	24,26,26	0.95	1 (4%)	37,40,40	1.60	5 (13%)
3	CTS	B	1544	-	14,14,14	1.27	1 (7%)	16,21,21	1.04	1 (6%)
3	CTS	A	1544	-	14,14,14	1.12	0	16,21,21	0.76	0
2	UDP	C	1543	4	24,26,26	0.86	1 (4%)	37,40,40	1.81	7 (18%)

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	CTS	C	1544	-	-	-	1/2/2/2
2	UDP	A	1543	4	-	0/16/32/32	0/2/2/2
2	UDP	B	1543	4	-	0/16/32/32	0/2/2/2
3	CTS	B	1544	-	-	-	1/2/2/2
3	CTS	A	1544	-	-	-	1/2/2/2
2	UDP	C	1543	4	-	0/16/32/32	0/2/2/2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1544	CTS	C6-C5	2.57	1.56	1.54
2	B	1543	UDP	C6-C5	2.25	1.40	1.35
3	C	1544	CTS	C1-C2	2.23	1.55	1.52
2	A	1543	UDP	C6-C5	2.20	1.40	1.35
3	B	1544	CTS	C1-C2	2.06	1.55	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1543	UDP	C4-N3-C2	-5.64	119.14	126.58
2	B	1543	UDP	C4-N3-C2	-5.31	119.58	126.58
2	A	1543	UDP	C4-N3-C2	-4.95	120.05	126.58
2	C	1543	UDP	N3-C2-N1	4.21	120.48	114.89
2	B	1543	UDP	C5-C4-N3	4.13	121.02	114.84

There are no chirality outliers.

There are no torsion outliers.

All (3) ring outliers are listed below:

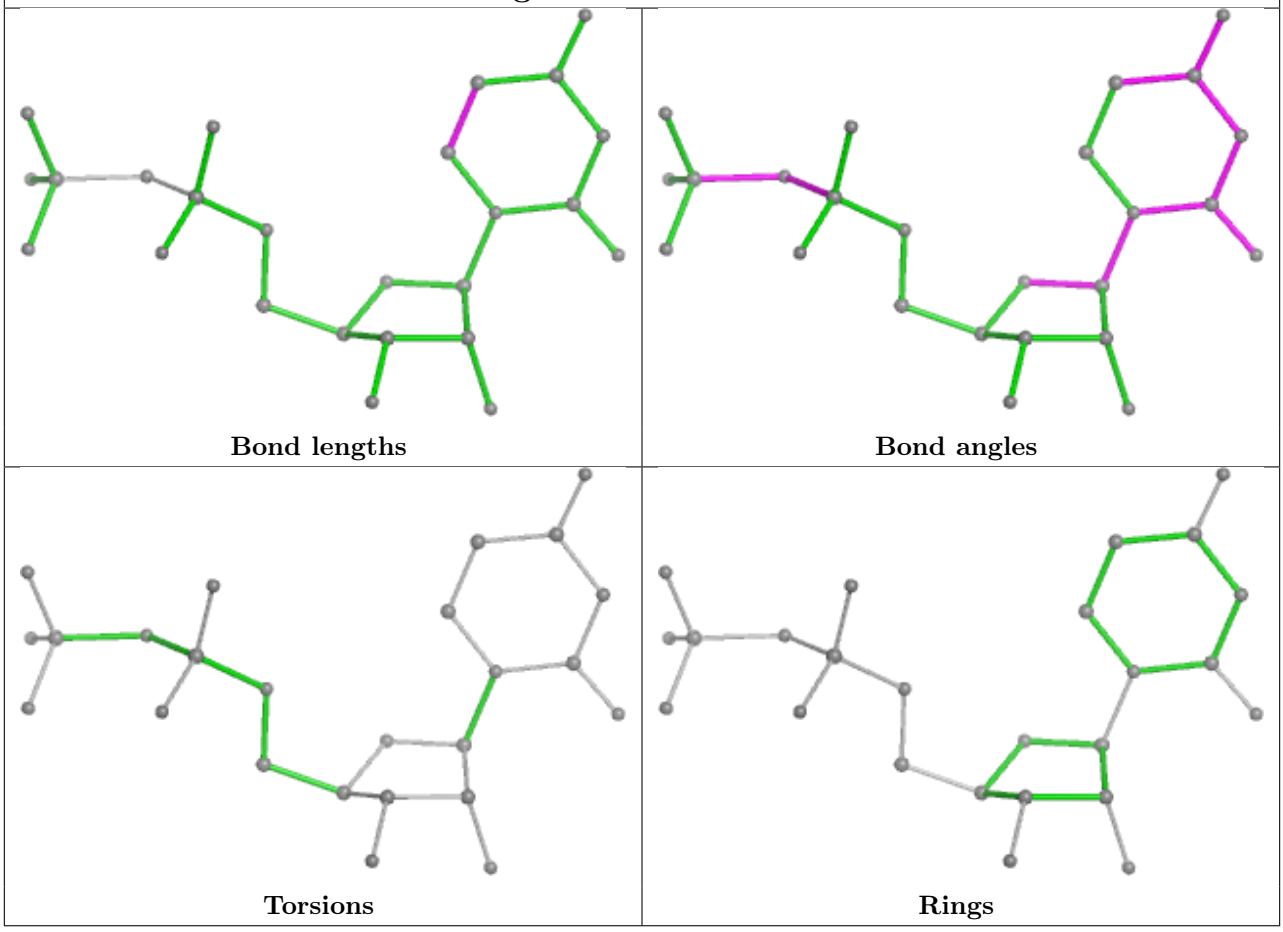
Mol	Chain	Res	Type	Atoms
3	C	1544	CTS	C1-C2-C3-C4-C5-N
3	A	1544	CTS	C1-C2-C3-C4-C5-N
3	B	1544	CTS	C1-C2-C3-C4-C5-N

No monomer is involved in short contacts.

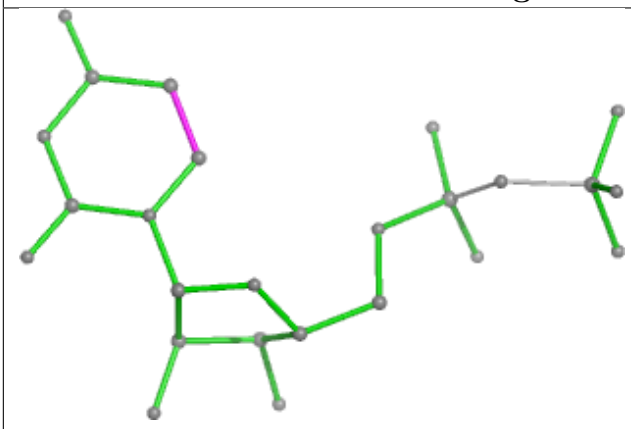
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

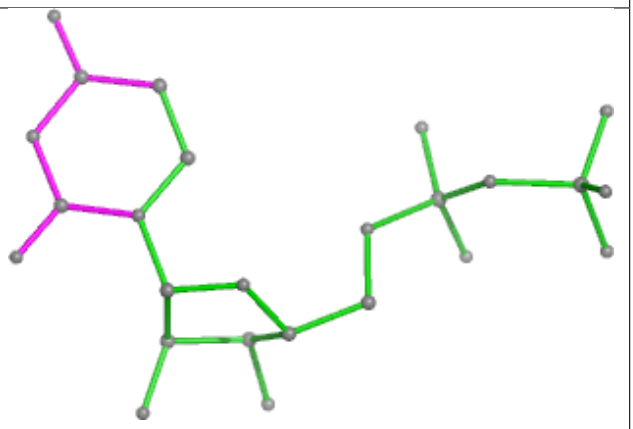
Ligand UDP A 1543



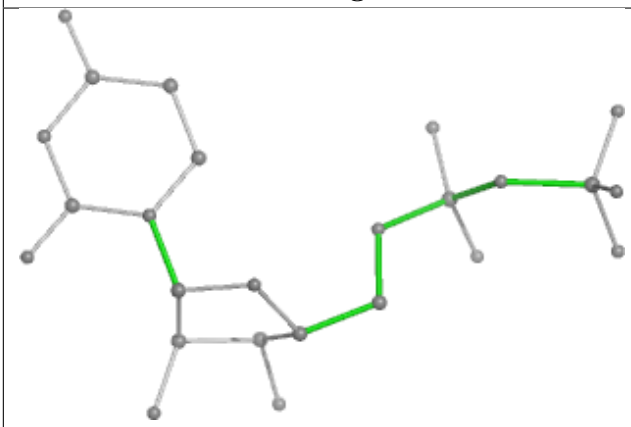
Ligand UDP B 1543



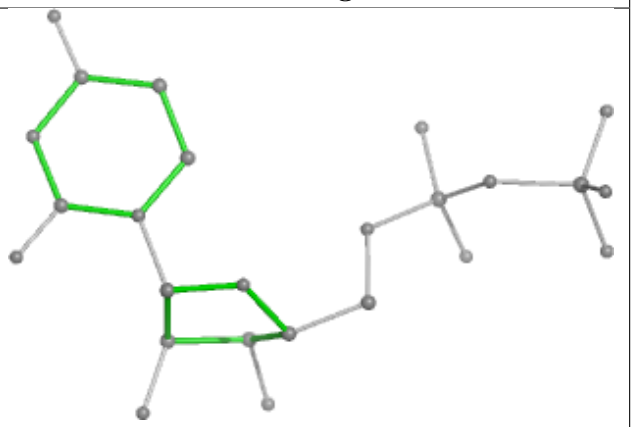
Bond lengths



Bond angles

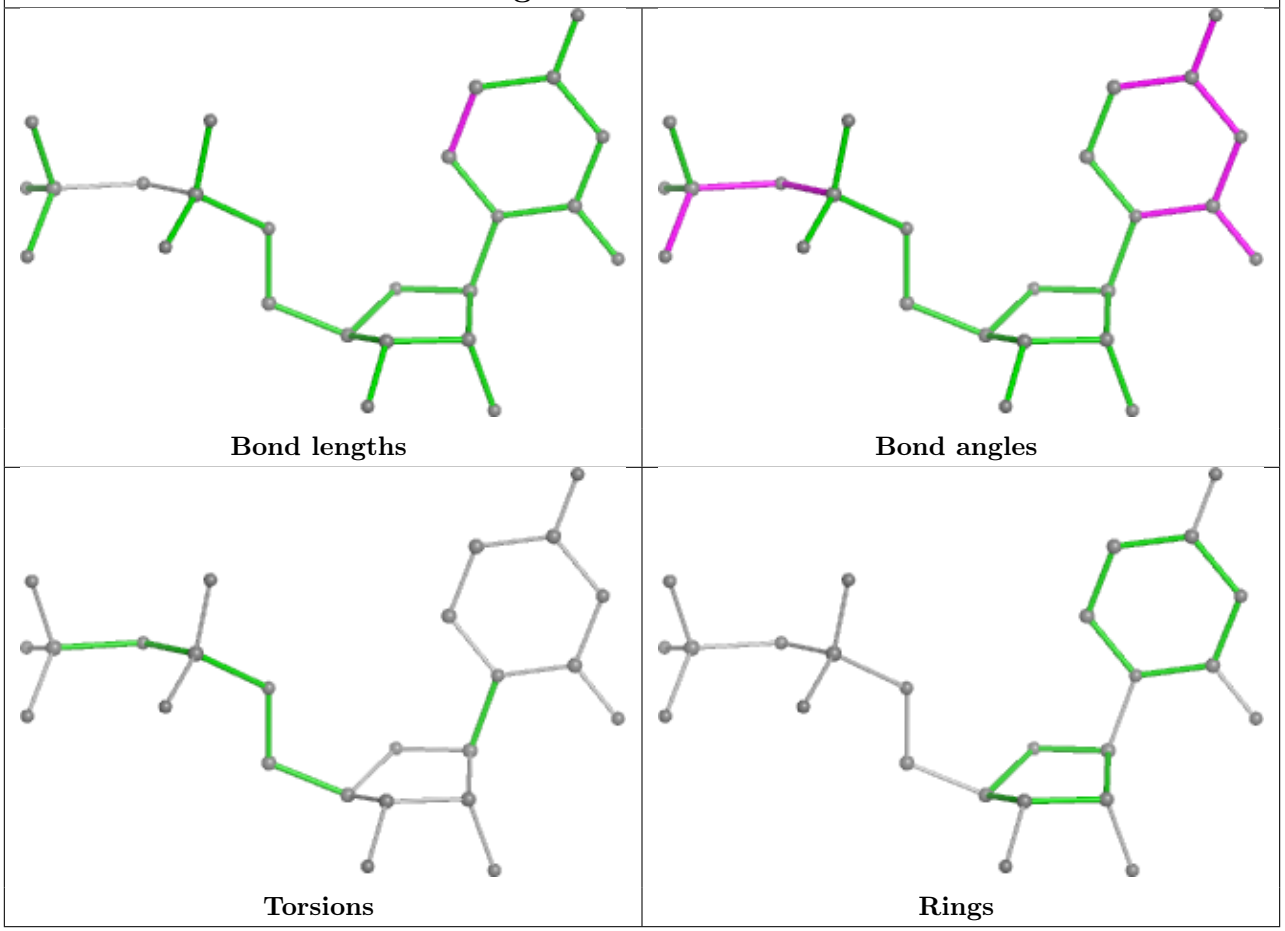


Torsions



Rings

Ligand UDP C 1543



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

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, 95^{th} 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	539/546 (98%)	0.36	20 (3%)	41	48	26, 39, 65, 90	0
1	B	533/546 (97%)	0.47	40 (7%)	14	19	24, 38, 74, 122	0
1	C	537/546 (98%)	0.50	38 (7%)	16	21	26, 37, 70, 88	0
All	All	1609/1638 (98%)	0.44	98 (6%)	21	27	24, 38, 71, 122	0

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	247	ALA	9.8
1	A	542	ALA	9.0
1	C	525	ALA	8.6
1	B	530	GLN	8.5
1	B	161	LEU	8.3

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

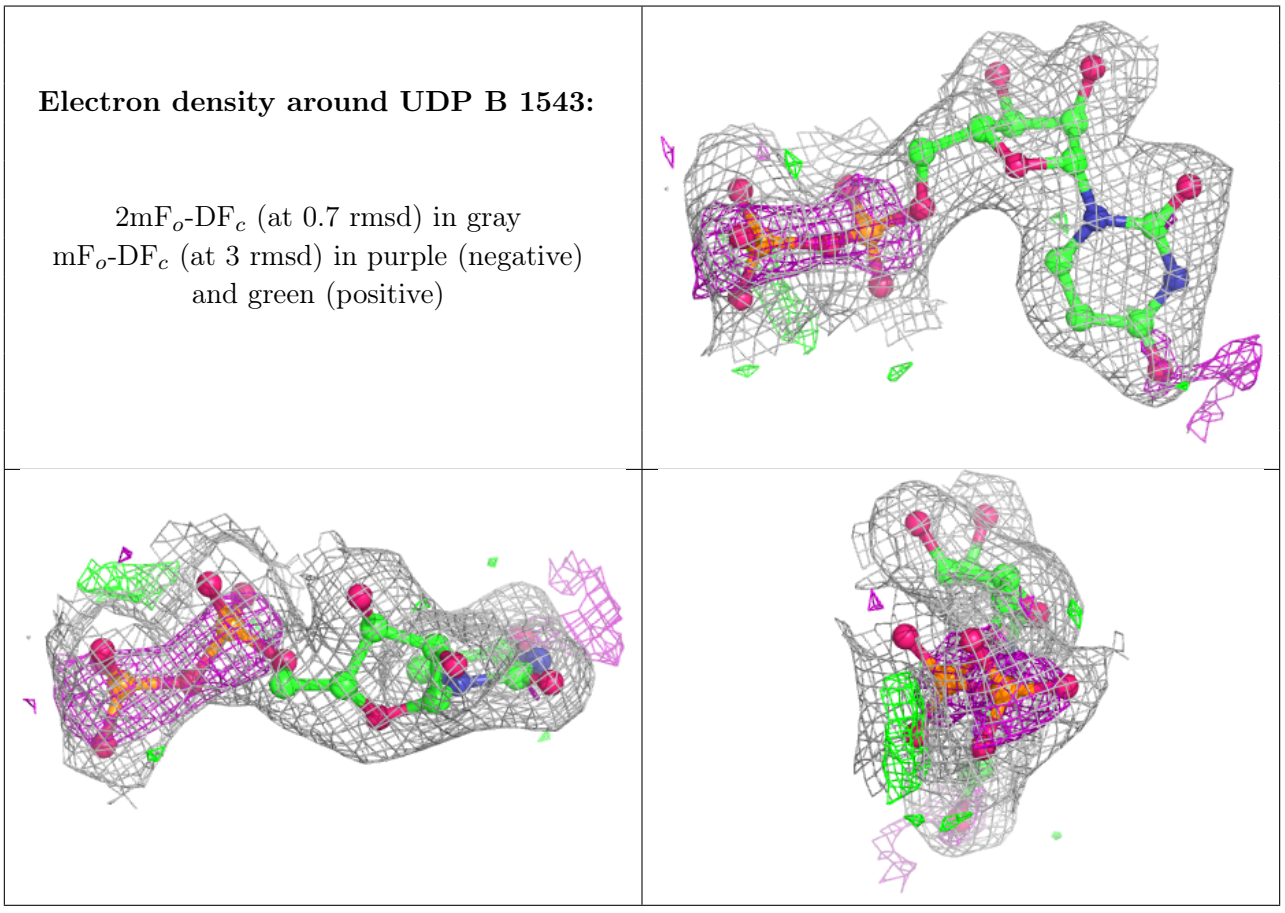
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	CTS	B	1544	13/13	0.87	0.13	44,48,49,49	0
3	CTS	C	1544	13/13	0.88	0.19	39,45,46,47	0
3	CTS	A	1544	13/13	0.91	0.16	37,40,41,42	0
2	UDP	B	1543	25/25	0.95	0.11	33,38,48,50	0
4	CA	A	1546	1/1	0.96	0.09	42,42,42,42	0
2	UDP	A	1543	25/25	0.97	0.09	32,36,48,50	0
2	UDP	C	1543	25/25	0.97	0.12	34,39,52,52	0
4	CA	A	1547	1/1	0.97	0.05	48,48,48,48	0
4	CA	A	1545	1/1	0.99	0.13	33,33,33,33	0
4	CA	B	1545	1/1	0.99	0.07	40,40,40,40	0
4	CA	C	1545	1/1	1.00	0.11	34,34,34,34	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.

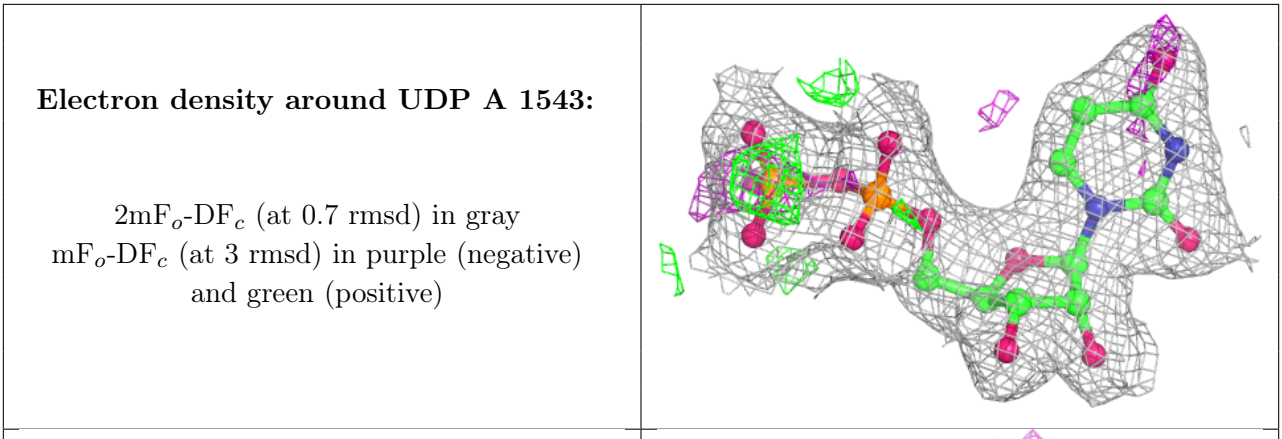
Electron density around UDP B 1543:

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



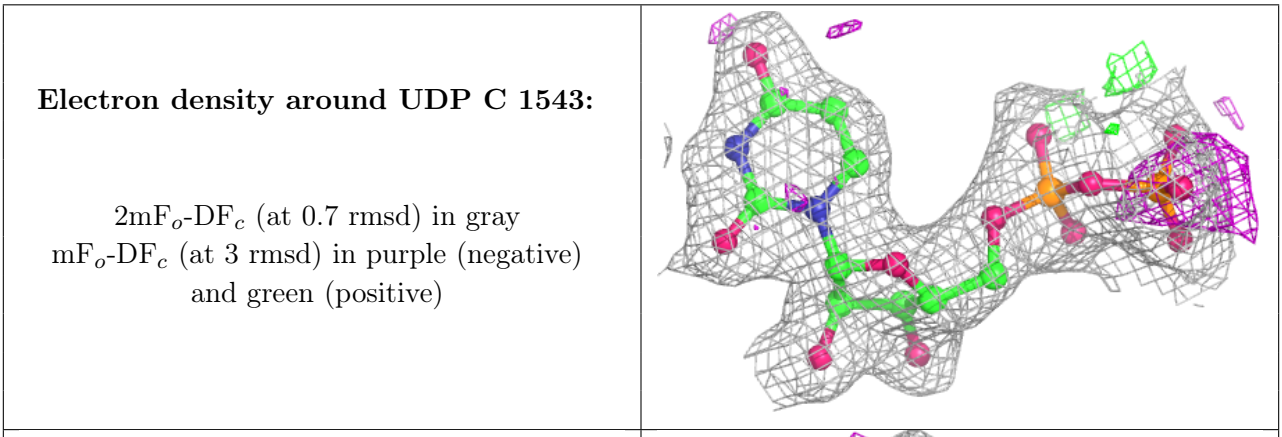
Electron density around UDP A 1543:

$2mF_o$-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 1543:

$2mF_o$-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.