wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2023 – 01:13 am GMT

PDB ID	:	1W2Y
Title	:	The crystal structure of a complex of Campylobacter jejuni dUTPase with substrate analogue dUpNHp
Authors	:	Moroz, O.V.; Harkiolaki, M.; Galperin, M.Y.; Vagin, A.A.; Gonzalez-Pacanowska, D.; Wilson, K.S.
Deposited on	:	2004-07-09
Resolution	:	1.65 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:

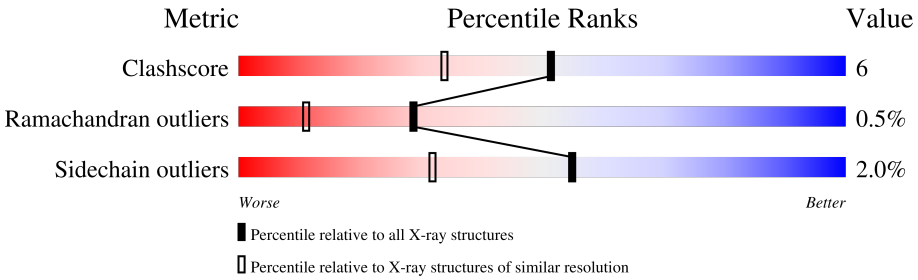
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.65 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))
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)

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%

Mol	Chain	Length	Quality of chain
1	A	229	82% 12% • • •
1	B	229	80% 10% • 9%

2 Entry composition (i)

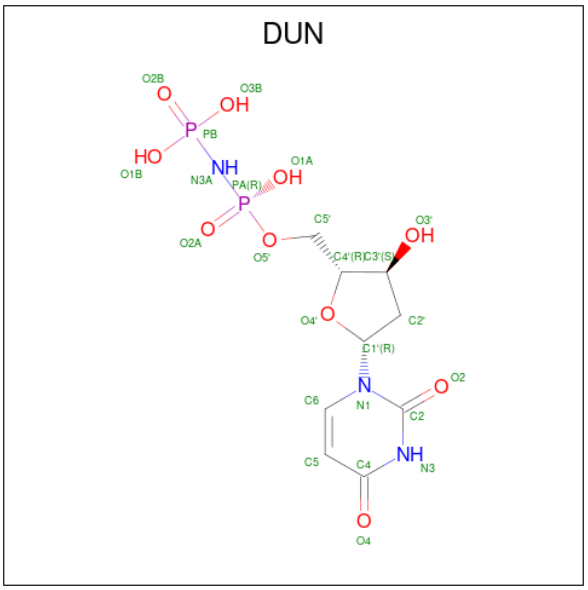
There are 4 unique types of molecules in this entry. The entry contains 4067 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDE HYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	83	4	0
			1895	1220	306	360	9			
1	B	209	Total	C	N	O	S	59	2	0
			1735	1121	285	320	9			

- Molecule 2 is 2'-DEOXYURIDINE 5'-ALPHA,BETA-IMIDO-DIPHOSPHATE (three-letter code: DUN) (formula: $C_9H_{15}N_3O_{10}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			24	9	3	10	2		
2	B	1	Total	C	N	O	P	0	0
			24	9	3	10	2		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	3	Total Mg 3 3	0	0
3	B	3	Total Mg 3 3	0	0

- Molecule 4 is water.

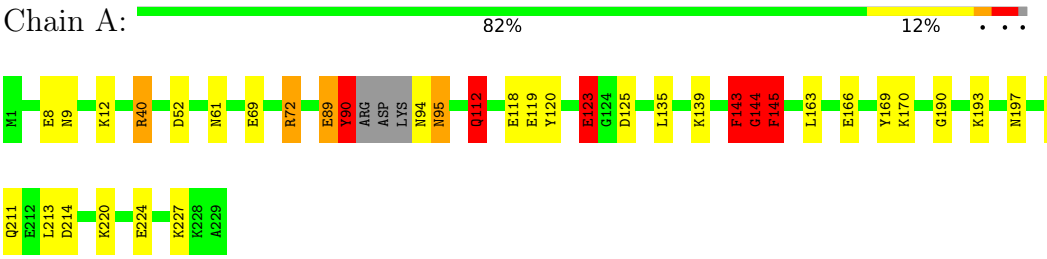
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	195	Total O 195 195	0	0
4	B	188	Total O 188 188	0	0

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. 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. 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.

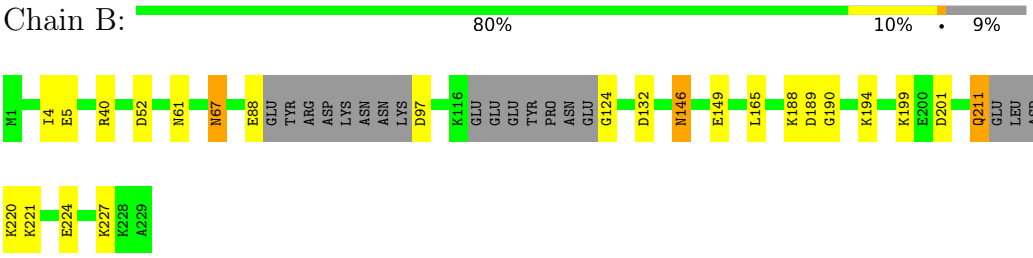
• Molecule 1: DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDE HYDROLASE

Chain A:



• Molecule 1: DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDE HYDROLASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.96Å 70.63Å 92.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.90 – 1.65 48.59 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.0 (55.90-1.65) 99.0 (48.59-1.65)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, R_{free}	0.151 , 0.194 0.237 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	25.0	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 54.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4067	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.77% 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: MG, DUN

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	1.41	18/1953 (0.9%)	1.56	24/2633 (0.9%)
1	B	1.85	8/1777 (0.5%)	1.14	13/2391 (0.5%)
All	All	1.64	26/3730 (0.7%)	1.38	37/5024 (0.7%)

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	A	2	5
1	B	1	0
All	All	3	5

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	211	GLN	CB-CG	-41.82	0.39	1.52
1	B	224	GLU	CB-CG	32.54	2.13	1.52
1	B	146	ASN	CG-OD1	29.28	1.88	1.24
1	B	146	ASN	CG-ND2	-28.83	0.60	1.32
1	A	90	TYR	CB-CG	-20.79	1.20	1.51

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	143	PHE	O-C-N	-50.35	37.60	123.20
1	A	143	PHE	CA-C-N	23.65	163.50	116.20
1	B	146	ASN	CB-CG-OD1	-19.89	81.83	121.60
1	B	146	ASN	CB-CG-ND2	16.51	156.32	116.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	221	LYS	N-CA-CB	-16.38	81.12	110.60

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	94	ASN	CA
1	A	227	LYS	CA
1	B	227	LYS	CA

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	123	GLU	Sidechain
1	A	143	PHE	Mainchain,Peptide
1	A	144	GLY	Mainchain,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	1895	0	1865	26	0
1	B	1735	0	1736	17	0
2	A	24	0	12	0	0
2	B	24	0	12	0	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
4	A	195	0	0	10	0
4	B	188	0	0	11	0
All	All	4067	0	3625	39	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 39 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:217:THR:N	4:B:2178:HOH:O	1.96	0.97
1:B:189:ASP:HB2	4:B:2170:HOH:O	1.67	0.92
1:A:9:ASN:HB2	4:A:2011:HOH:O	1.85	0.77
1:A:144:GLY:O	4:A:2147:HOH:O	2.01	0.77
1:A:61:ASN:ND2	1:B:190:GLY:HA2	2.00	0.76

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	226/229 (99%)	223 (99%)	1 (0%)	2 (1%)	17 4
1	B	203/229 (89%)	202 (100%)	1 (0%)	0	100 100
All	All	429/458 (94%)	425 (99%)	2 (0%)	2 (0%)	29 11

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	143	PHE
1	A	145	PHE

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	210/209 (100%)	207 (99%)	3 (1%)	67 46
1	B	191/209 (91%)	186 (97%)	5 (3%)	46 21
All	All	401/418 (96%)	393 (98%)	8 (2%)	55 32

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	211	GLN
1	B	188	LYS
1	B	146	ASN
1	B	67	ASN
1	B	165	LEU

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

Mol	Chain	Res	Type
1	A	9	ASN
1	A	61	ASN
1	B	61	ASN
1	B	67	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 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
2	DUN	B	1230	3	22,25,25	1.80	4 (18%)	23,38,38	2.39	5 (21%)
2	DUN	A	1230	3	22,25,25	1.25	2 (9%)	23,38,38	1.75	5 (21%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	DUN	B	1230	3	-	0/10/28/28	0/2/2/2
2	DUN	A	1230	3	-	2/10/28/28	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)
2	B	1230	DUN	PA-O2A	5.89	1.55	1.46
2	A	1230	DUN	PA-O2A	3.10	1.51	1.46
2	B	1230	DUN	PB-O1B	-2.92	1.48	1.56
2	B	1230	DUN	PB-O2B	2.79	1.50	1.46
2	B	1230	DUN	PA-O1A	-2.22	1.50	1.56

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1230	DUN	O2B-PB-N3A	-7.35	100.95	111.77
2	B	1230	DUN	O2A-PA-N3A	-5.79	103.24	111.77
2	B	1230	DUN	O1A-PA-O2A	4.48	119.32	109.92
2	A	1230	DUN	O1A-PA-O2A	4.07	118.46	109.92
2	A	1230	DUN	O2A-PA-N3A	-3.59	106.48	111.77

There are no chirality outliers.

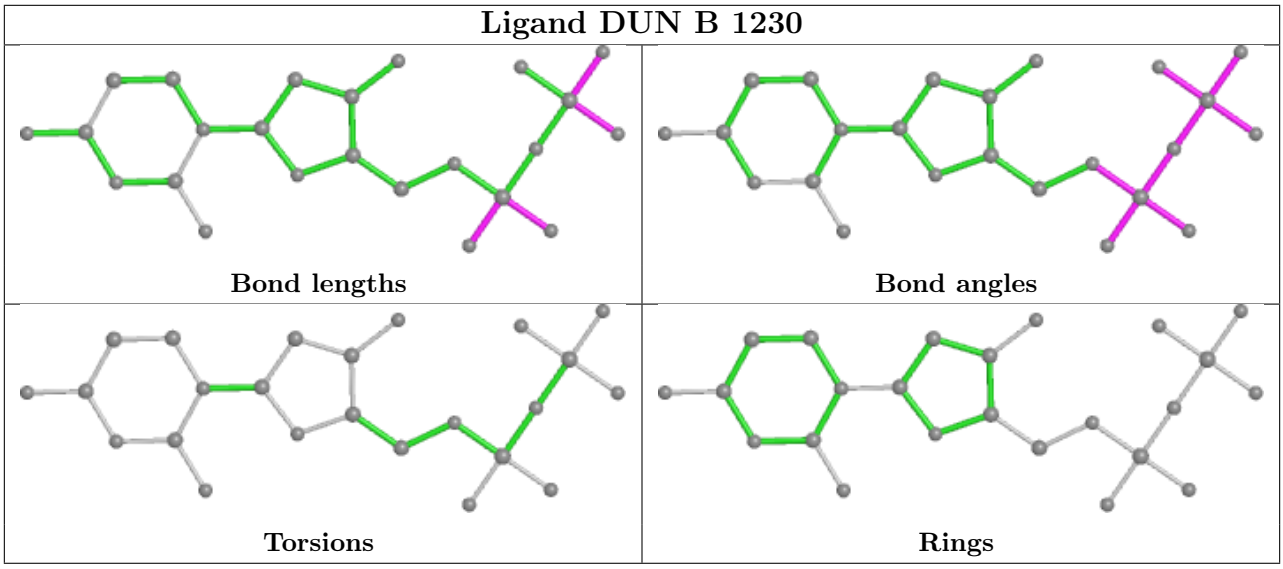
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1230	DUN	C5'-O5'-PA-O2A
2	A	1230	DUN	C5'-O5'-PA-O1A

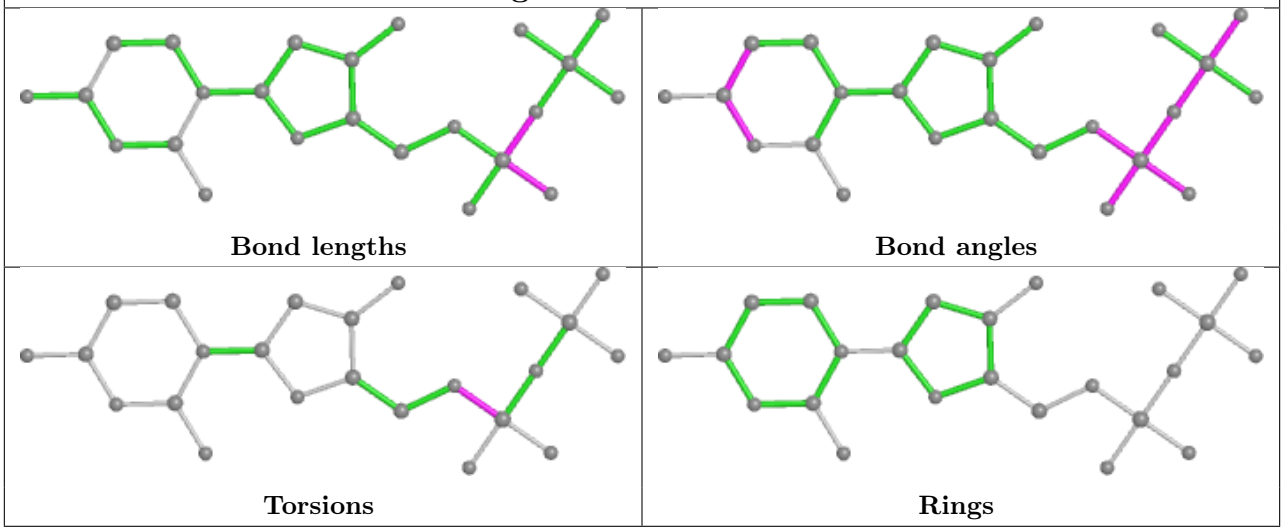
There are no ring outliers.

No monomer is involved in short contacts.

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



Ligand DUN A 1230



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	143:PHE	C	144:GLY	N	1.01
1	A	144:GLY	C	145:PHE	N	0.98

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

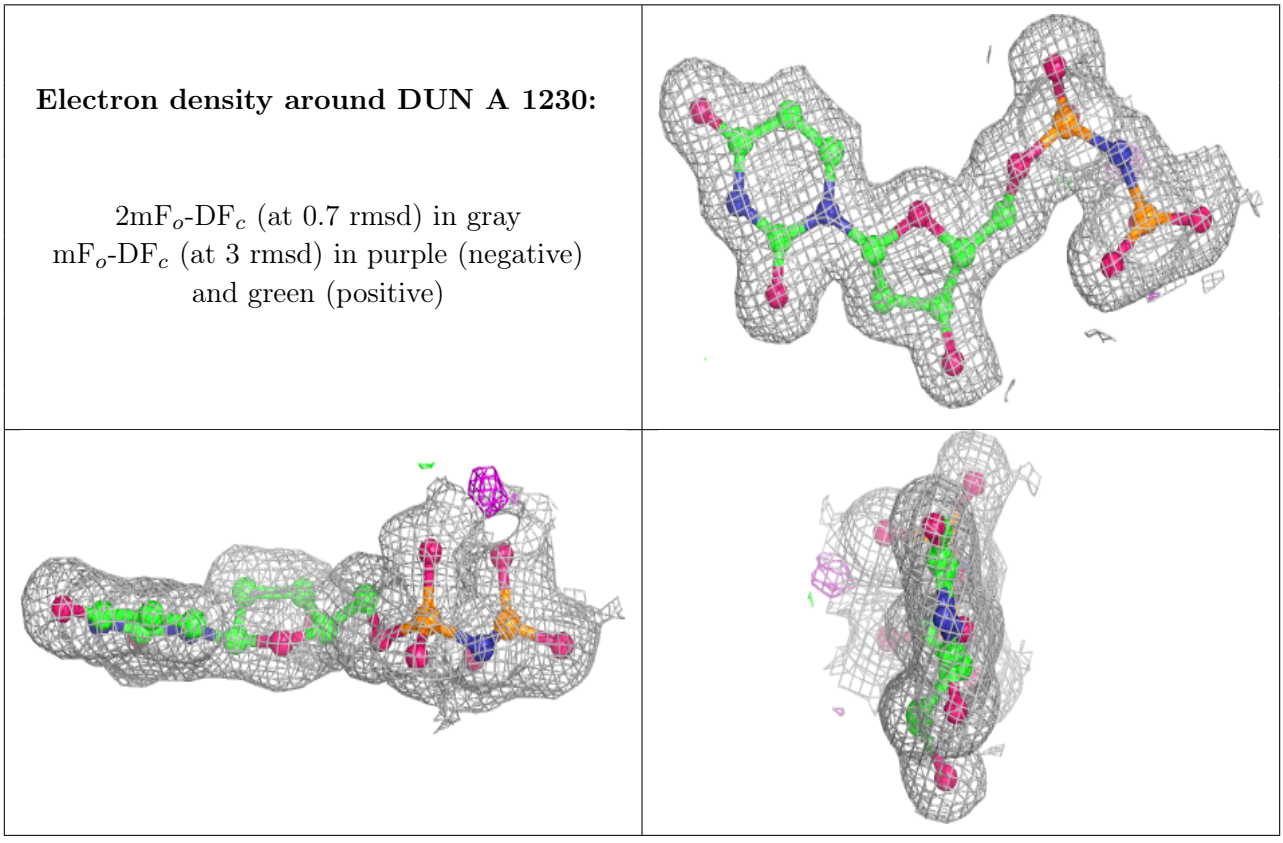
6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

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.

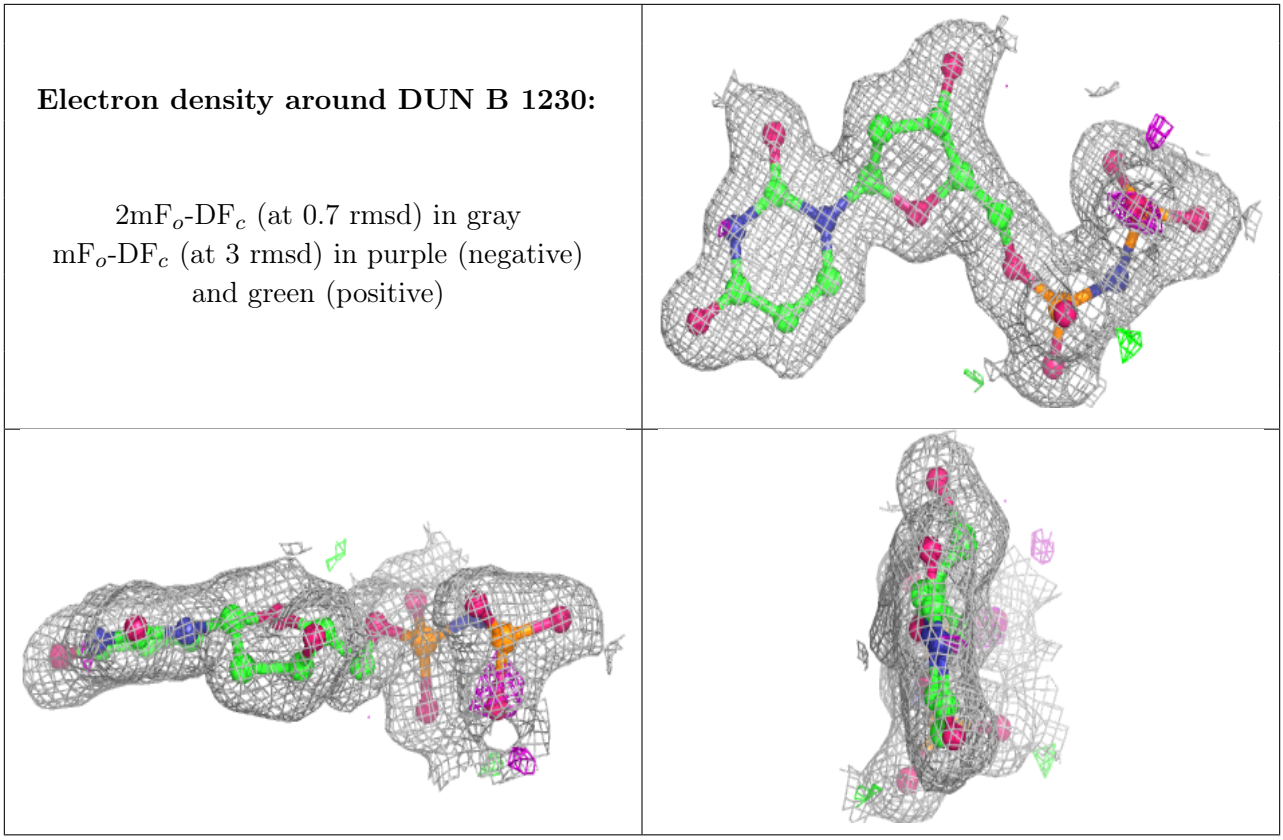
Electron density around DUN A 1230:

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



Electron density around DUN B 1230:

$2mF_o$-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)

Unable to reproduce the depositors R factor - this section is therefore empty.