

# wwPDB X-ray Structure Validation Summary Report (i)

#### May 25, 2020 – 01:20 pm BST

PDB ID	•	
$\operatorname{Title}$	:	DUTPase from phage phi11 of S.aureus: visualization of the species-specific
		insert
Authors	:	Leveles, I.; Harmat, V.; Nemeth, V.; Bendes, A.; Szabo, J.; Kadar, V.; Zagyva,
		I.; Rona, G.; Toth, J.; Vertessy, B.G.
Deposited on	:	2012-08-30
$\operatorname{Resolution}$	:	2.10 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.

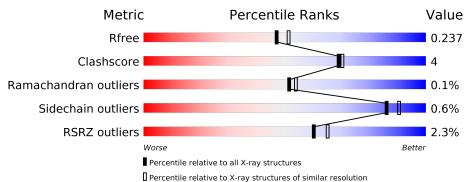
The following versions of software and data (see references (1)) were used in the production of this report:

# 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 2.10 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries},{ m resolution\ range}({ m \AA}))$
$R_{free}$	130704	5197(2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647(2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 on 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	А	169	85%	5%	9%
1	В	169	85%	6%	9%
1	С	169	% • 83%	8%	• 9%
1	D	169	77%	13%	10%
1	Е	169	84%	7%	9%
1	F	169	80%	10%	9%



# 2 Entry composition (i)

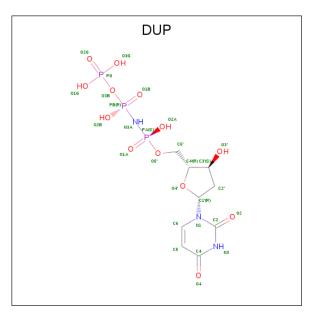
There are 4 unique types of molecules in this entry. The entry contains 7352 atoms, of which 0 are hydrogens and 0 are deuteriums.

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.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
1	А	153	Total	С	Ν	Ο	$\mathbf{S}$	0	1	0
	Л	100	1139	721	194	223	1	0	T	0
1	В	154	Total	С	Ν	Ο	S	0	0	0
	D	104	1159	731	193	234	1	0	0	0
1	С	154	Total	С	Ν	Ο	S	0	0	0
	U	104	1137	721	192	223	1	0		
1	р	152	Total	С	Ν	Ο	S	0	0	0
	D	152	1117	704	188	224	1	0	0	U
1	Е	154	Total	С	Ν	Ο	S	0	0	0
		E 104	1126	713	191	221	1	0	0	0
1	F	153	Total	С	Ν	Ο	S	0	1	0
	L L	100	1122	713	191	217	1			U

• Molecule 1 is a protein called DUTPase.

• Molecule 2 is 2'-DEOXYURIDINE 5'-ALPHA, BETA-IMIDO-TRIPHOSPHATE (three-letter code: DUP) (formula:  $C_9H_{16}N_3O_{13}P_3$ ).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	А	1	Total	С	Ν	Ο	Р	0	0
	Л	T	28	9	3	13	3	0	0
2	В	1	Total	С	Ν	Ο	Р	0	0
2	D	I	28	9	3	13	3	0	0
2	С	1	Total	С	Ν	Ο	Р	0	0
	0		28	9	3	13	3	0	0
2	П	1	Total	С	Ν	Ο	Р	0	0
	D	, I	28	9	3	13	3	0	0
2	Е	1	Total	С	Ν	Ο	Р	0	0
		*	28	9	3	13	3	0	0
2	F	1	Total	С	Ν	Ο	Р	0	0
	T		28	9	3	13	3		0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	2	Total Mg 2 2	0	0
3	Е	1	Total Mg 1 1	0	0
3	В	1	Total Mg 1 1	0	0
3	С	1	Total Mg 1 1	0	0
3	А	2	Total Mg 2 2	0	0
3	F	1	Total Mg 1 1	0	0

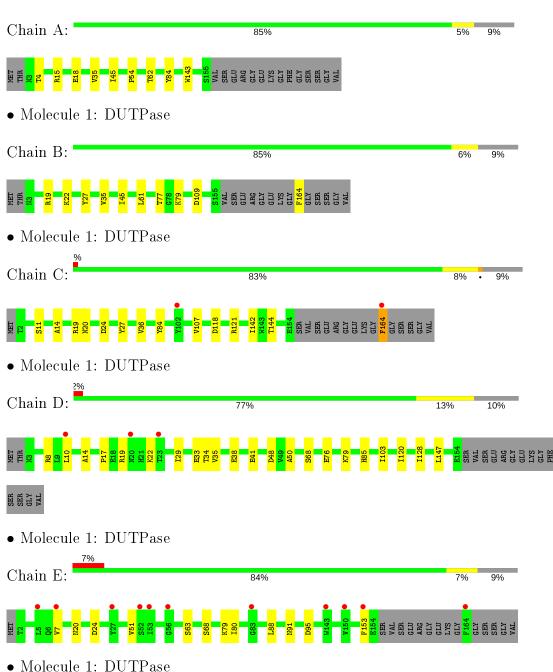
• Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	85	Total O 85 85	0	0
4	В	87	Total O	0	0
4	D	01	87 87	0	0
4	С	82	Total O 82 82	0	0
4	D	37	Total O 37 37	0	0
4	Е	43	Total         O           43         43	0	0
4	F	42	$\begin{array}{cc} \text{Total} & \text{O} \\ 42 & 42 \end{array}$	0	0



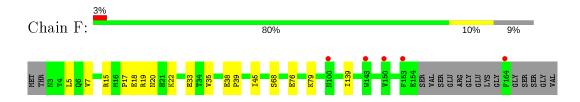
# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: DUTPase







## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants	109.66Å 109.66Å 169.27Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	47.10 - 2.10	Depositor
Resolution (A)	47.10 - 2.10	EDS
% Data completeness	$100.0 \ (47.10-2.10)$	Depositor
(in resolution range)	$100.0 \ (47.10-2.10)$	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.18 (at 2.10 \text{\AA})$	Xtriage
Refinement program	REFMAC $5.5.0110$	Depositor
$R, R_{free}$	0.183 , $0.232$	Depositor
It, It <i>free</i>	0.189 , $0.237$	DCC
$R_{free}$ test set	3140 reflections $(5.16%)$	wwPDB-VP
Wilson B-factor ( $Å^2$ )	41.4	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.32 , $60.6$	EDS
L-test for twinning <sup>2</sup>	$ < 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	7352	wwPDB-VP
Average B, all atoms $(Å^2)$	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: DUP, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with |Z| > 5 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bo	nd lengths	Bond angles		
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.94	1/1160~(0.1%)	0.90	1/1580~(0.1%)	
1	В	0.89	0/1177	0.91	1/1602~(0.1%)	
1	С	0.97	0/1154	0.95	4/1573~(0.3%)	
1	D	0.73	0/1134	0.80	0/1547	
1	Ε	0.75	0/1144	0.81	1/1560~(0.1%)	
1	F	0.78	0/1142	0.86	0/1557	
All	All	0.85	1/6911~(0.0%)	0.87	7/9419~(0.1%)	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
1	А	143	TRP	CD2-CE2	6.33	1.49	1.41

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	С	121	ARG	NE-CZ-NH1	5.95	123.27	120.30
1	С	121	ARG	NE-CZ-NH2	-5.80	117.40	120.30
1	А	15	ARG	NE-CZ-NH1	5.59	123.10	120.30
1	Е	95	ASP	CB-CG-OD2	5.33	123.09	118.30
1	С	118	ASP	CB-CG-OD1	5.10	122.89	118.30

There are no chirality outliers.

There are no planarity outliers.

#### 5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	А	1139	0	1106	6	0
1	В	1159	0	1114	7	0
1	С	1137	0	1091	7	0
1	D	1117	0	1052	18	0
1	Е	1126	0	1054	8	0
1	F	1122	0	1067	14	0
2	А	28	0	12	1	0
2	В	28	0	12	0	0
2	С	28	0	12	1	0
2	D	28	0	12	0	0
2	Е	28	0	12	0	0
2	F	28	0	12	0	0
3	А	2	0	0	0	0
3	В	1	0	0	0	0
3	С	1	0	0	0	0
3	D	2	0	0	0	0
3	Е	1	0	0	0	0
3	F	1	0	0	0	0
4	А	85	0	0	0	0
4	В	87	0	0	0	0
4	С	82	0	0	1	0
4	D	37	0	0	2	0
4	Е	43	0	0	1	0
4	F	42	0	0	0	0
All	All	7352	0	6556	50	0

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.

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 4.

The worst 5 of 50 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:103:ILE:HD12	1:D:120:ILE:HG22	1.76	0.68
1:D:103:ILE:HD12	1:D:120:ILE:CG2	2.27	0.63
1:E:20:ASN:HB3	1:E:24:ASP:OD2	1.99	0.63
1:E:80:ILE:HD12	1:E:88:LEU:HD22	1.82	0.62
1:C:36:VAL:HG12	4:C:343:HOH:O	2.00	0.61

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	А	152/169~(90%)	151~(99%)	1 (1%)	0	100 100
1	В	151/169~(89%)	150~(99%)	1 (1%)	0	100 100
1	С	151/169~(89%)	148 (98%)	3(2%)	0	100 100
1	D	150/169~(89%)	148 (99%)	2(1%)	0	100 100
1	Ε	151/169~(89%)	145~(96%)	6 (4%)	0	100 100
1	F	151/169~(89%)	147 (97%)	3~(2%)	1 (1%)	22 18
All	All	906/1014~(89%)	889~(98%)	16~(2%)	1 (0%)	51 54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	20	ASN

#### 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	А	117/142~(82%)	115~(98%)	2(2%)	60 67
1	В	122/142~(86%)	121~(99%)	1 (1%)	81 86
1	С	115/142~(81%)	114 (99%)	1 (1%)	78 84
1	D	110/142~(78%)	110~(100%)	0	100 100
1	Е	109/142~(77%)	109~(100%)	0	100 100
1	F	110/142~(78%)	110~(100%)	0	100 100

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	683/852~(80%)	679~(99%)	4 (1%)	86 90	

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

Mol	Chain	Res	Type
1	А	4	THR
1	А	54	PRO
1	В	164	PHE
1	С	164	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	В	136	GLN
1	Е	91	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 carbohydrates in this entry.

#### 5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 8 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	Turne	Chain	Res Link		B	ond leng	gths	Bond angles		
	Type	Cham	nes	Link	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	DUP	В	201	3	28,29,29	3.12	10 (35%)	37,45,45	2.09	11 (29%)
2	DUP	Е	201	3	28,29,29	2.99	10 (35%)	37,45,45	2.33	14 (37%)
2	DUP	D	201	3	28,29,29	3.13	11 (39%)	37,45,45	1.90	9 (24%)
2	DUP	F	201	3	28,29,29	<b>3.03</b>	10 (35%)	37,45,45	1.97	11 (29%)
2	DUP	А	201	3	28,29,29	2.82	11 (39%)	37,45,45	2.15	11 (29%)
2	DUP	С	201	3	28, 29, 29	3.05	11 (39%)	37,45,45	2.36	14 (37%)

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	$\mathbf{Res}$	$\mathbf{Link}$	Chirals	Torsions	Rings
2	DUP	В	201	3	-	2/19/34/34	0/2/2/2
2	DUP	Е	201	3	-	6/19/34/34	0/2/2/2
2	DUP	D	201	3	-	2/19/34/34	0/2/2/2
2	DUP	F	201	3	-	0/19/34/34	0/2/2/2
2	DUP	А	201	3	-	1/19/34/34	0/2/2/2
2	DUP	С	201	3	-	2/19/34/34	0/2/2/2

The worst 5 of 63 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
2	D	201	DUP	O2-C2	8.28	1.38	1.23
2	F	201	DUP	O2-C2	7.76	1.37	1.23
2	С	201	DUP	O2-C2	7.51	1.36	1.23
2	Е	201	DUP	O2-C2	7.14	1.36	1.23
2	А	201	DUP	O4-C4	6.99	1.37	1.23

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
2	С	201	DUP	C4-N3-C2	-6.45	120.44	125.79
2	А	201	DUP	C4-N3-C2	-6.25	120.61	125.79
2	А	201	DUP	O2A-PA-O1A	6.09	122.69	109.92
2	Е	201	DUP	C4-N3-C2	-6.08	120.75	125.79
2	F	201	DUP	C2'-C1'-N1	-5.70	108.64	115.61

There are no chirality outliers.

5 of 13 torsion outliers are listed below:



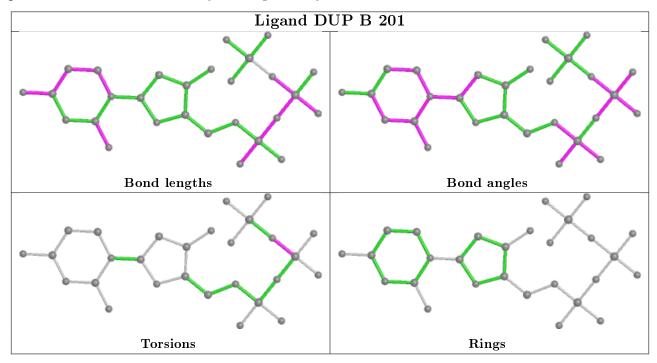
Mol	Chain	Res	Type	Atoms
2	В	201	DUP	PG-O3B-PB-O1B
2	В	201	DUP	PG-O3B-PB-O2B
2	Е	201	DUP	PG-O3B-PB-O1B
2	Е	201	DUP	PG-O3B-PB-O2B
2	D	201	DUP	PB-N3A-PA-O1A

There are no ring outliers.

2 monomers are involved in 2 short contacts:

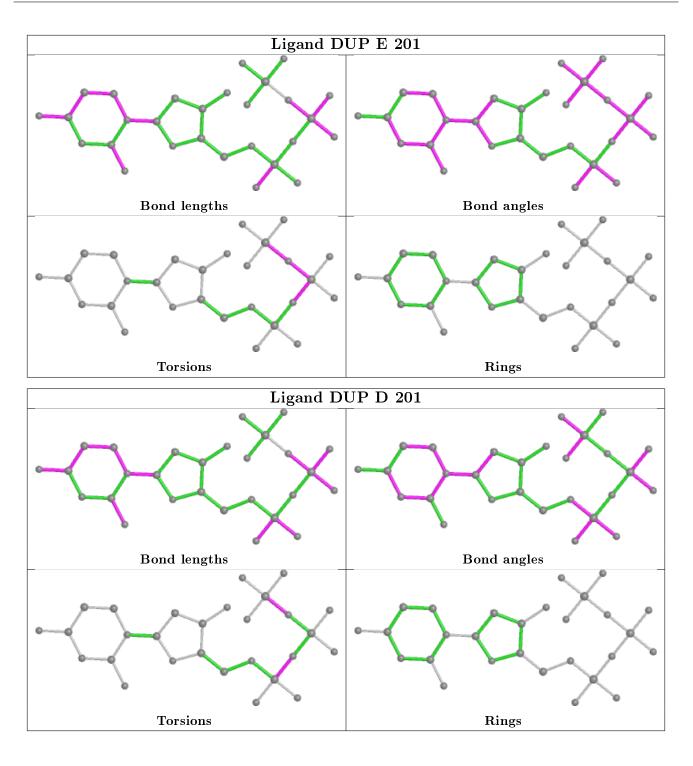
Mol	Chain	$\mathbf{Res}$	Type	Clashes	Symm-Clashes
2	А	201	DUP	1	0
2	С	201	DUP	1	0

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.



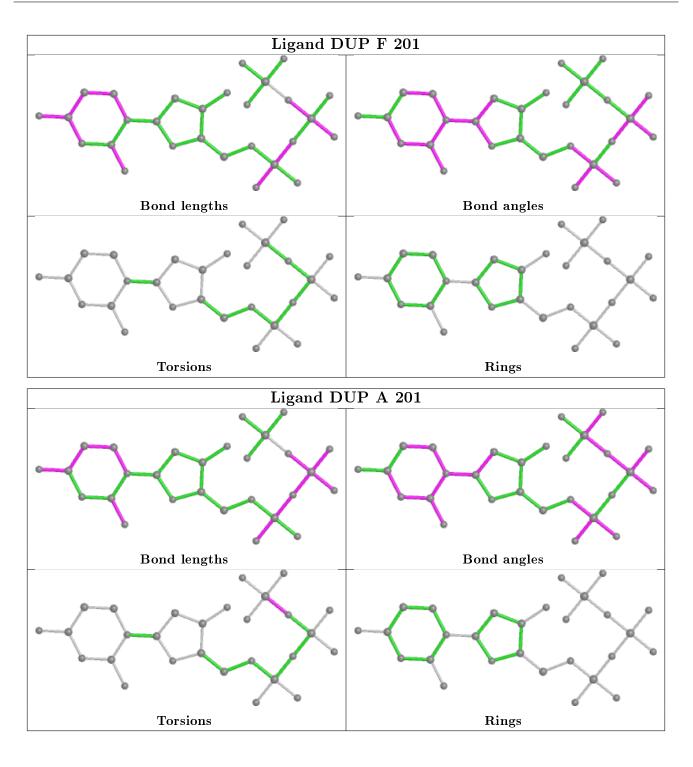




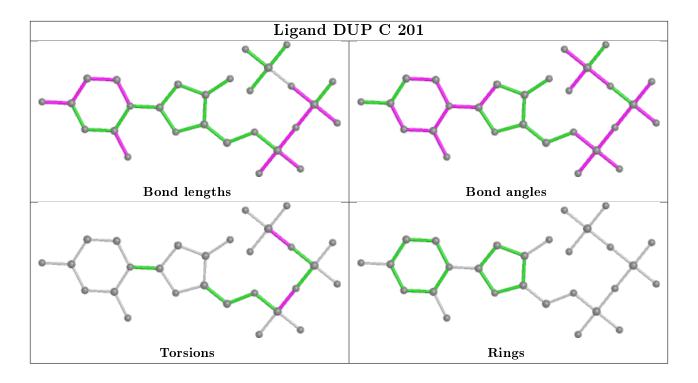












### 5.7 Other polymers (i)

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,  $95^{th}$  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 $>$	$\# RSRZ {>}2$	$\mathbf{OWAB}(\mathbf{\AA}^2)$	Q<0.9
1	А	153/169~(90%)	-0.51	0 100 100	37,  48,  71,  100	0
1	В	154/169~(91%)	-0.39	0 100 100	36,  50,  78,  101	0
1	С	154/169~(91%)	-0.30	2 (1%) 77 80	35,46,80,95	0
1	D	152/169~(89%)	-0.00	3 (1%) 65 69	53, 79, 111, 136	0
1	Ε	154/169~(91%)	0.13	11 (7%) 16 20	47, 68, 112, 130	0
1	F	153/169~(90%)	-0.05	5 (3%) 46 53	46, 64, 100, 112	0
All	All	920/1014~(90%)	-0.19	21 (2%) 60 65	35, 58, 101, 136	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	RSRZ
1	F	143	TRP	4.1
1	Е	150	VAL	3.9
1	D	23	THR	3.7
1	Е	52	SER	3.6
1	Е	153	PHE	3.4

## 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 carbohydrates in this entry.



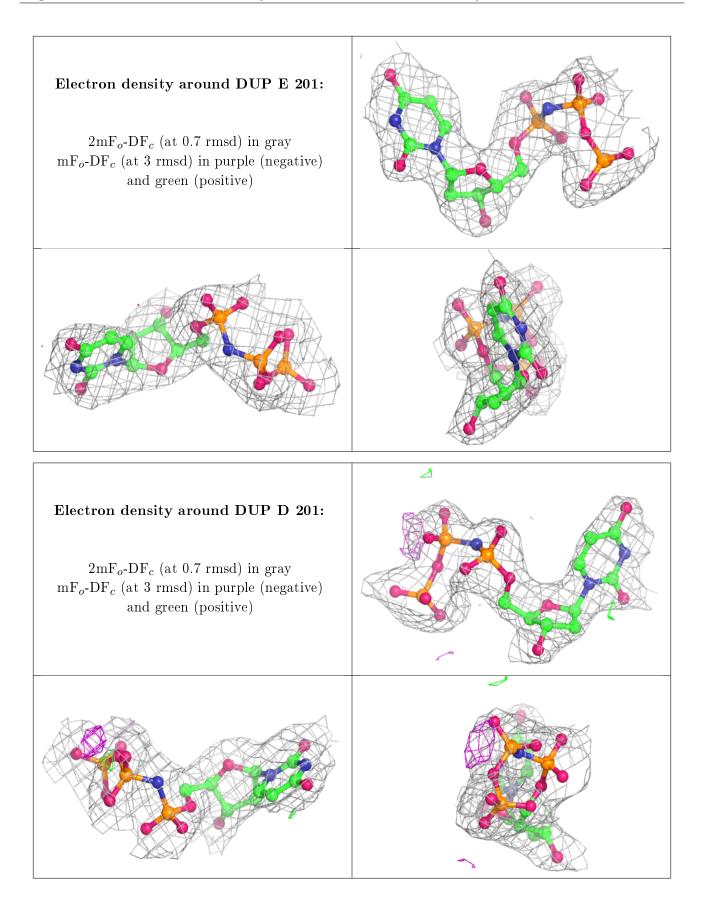
### 6.4 Ligands (i)

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.

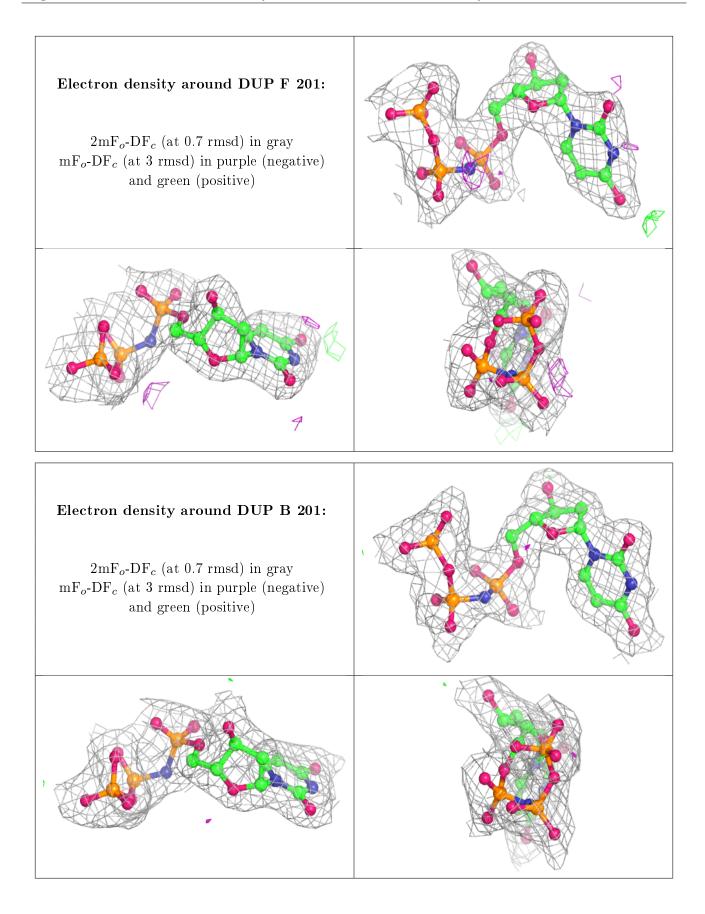
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	$Q{<}0.9$
3	MG	Е	202	1/1	0.91	0.07	$73,\!73,\!73,\!73$	0
3	MG	F	202	1/1	0.96	0.03	$65,\!65,\!65,\!65$	0
3	MG	D	202	1/1	0.96	0.12	80,80,80,80	0
2	DUP	Е	201	28/28	0.97	0.09	$64,\!69,\!81,\!82$	0
2	DUP	D	201	28/28	0.97	0.10	$69,\!74,\!78,\!83$	0
2	DUP	F	201	28/28	0.97	0.07	52,62,67,70	0
2	DUP	В	201	28/28	0.98	0.07	40,44,54,61	0
3	MG	В	202	1/1	0.98	0.07	47,47,47,47	0
3	MG	С	202	1/1	0.99	0.05	$43,\!43,\!43,\!43$	0
3	MG	D	203	1/1	0.99	0.21	32,32,32,32	0
3	MG	А	202	1/1	0.99	0.07	42,42,42,42	0
2	DUP	А	201	28/28	0.99	0.08	$34,\!40,\!49,\!56$	0
2	DUP	С	201	28/28	0.99	0.07	$34,\!39,\!50,\!54$	0
3	MG	А	203	1/1	1.00	0.26	$25,\!25,\!25,\!25$	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.

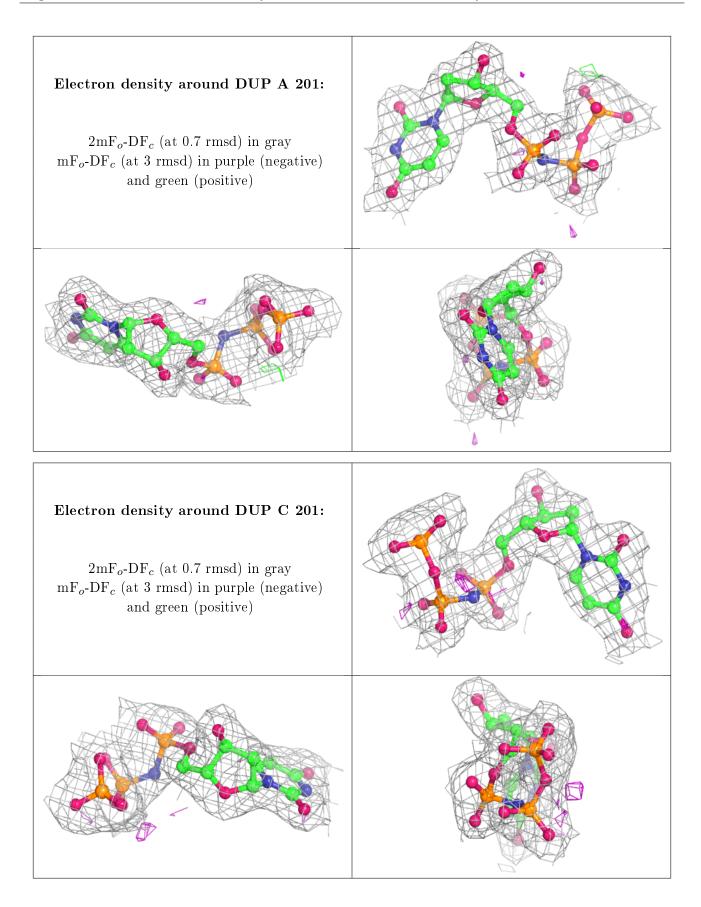














## 6.5 Other polymers (i)

There are no such residues in this entry.

