



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 29, 2024 – 02:41 pm BST

PDB ID : 2WZG
Title : Legionella glucosyltransferase (Lgt1) crystal structure
Authors : Hurtado-Guerrero, R.; Zusman, T.; Pathak, S.; Ibrahim, A.F.M.; Shepherd, S.; Prescott, A.; Segal, G.; Van Aalten, D.M.F.
Deposited on : 2009-11-29
Resolution : 1.90 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 ⓘ 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	:	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

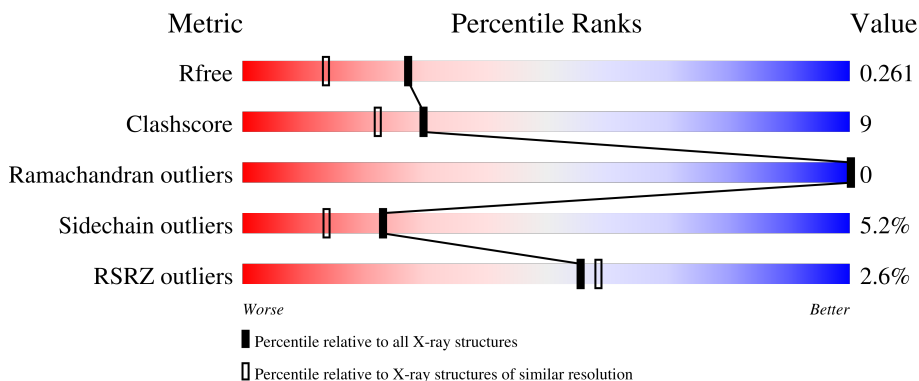
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 $\leq 5\%$. 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	525	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4470 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.

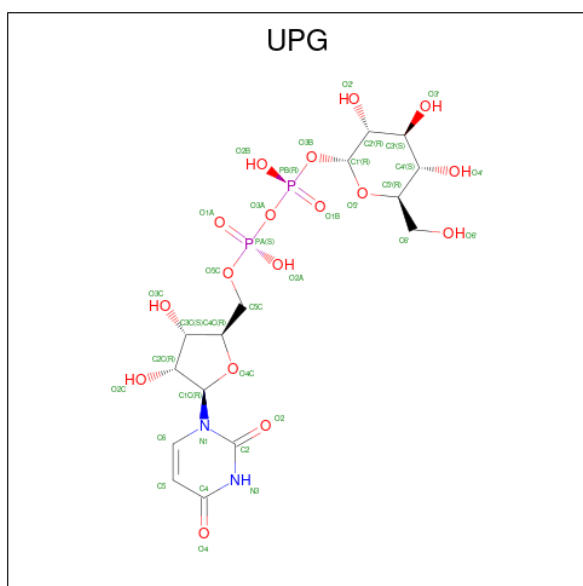
- Molecule 1 is a protein called GLUCOSYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	502	Total	C	N	O	S	0	21	0
			4115	2634	692	771	18			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		

- Molecule 3 is URIDINE-5'-DIPHOSPHATE-GLUCOSE (three-letter code: UPG) (formula: C₁₅H₂₄N₂O₁₇P₂).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	318	Total 318	O 318	0	0

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.

Chain A:

Amino Acid	Frequency (%)
MET	2%
LYS	79%
ALA	79%
ARG	79%
ARG	79%
S6	79%
S7	79%
E8	79%
K11	79%
R15	79%
S18	79%
T30	79%
K36	79%
L39	79%
H46	79%
L47	79%
Q48	79%
E49	79%
G50	79%
S51	79%
V52	79%
T53	79%
Y54	79%
Q58	79%
L61	79%
H83	79%
P84	79%
G85	79%
L90	79%
T91	79%
P92	79%
T95	79%
N96	79%
E105	79%
L112	79%
T116	79%
F120	79%
F130	79%
I138	79%
K143	79%
K144	79%
E146	15%
E154	15%
R158	15%
R159	15%
Q162	15%
Y163	15%
K167	15%
I171	15%
Y172	15%
N178	15%
M182	15%
M185	15%
L196	15%
K213	15%
D246	15%
I247	15%
D248	15%
L249	15%
T250	15%
V251	15%
D252	15%
S253	15%
S254	15%
K255	15%
T256	15%
T262	15%
M292	15%
D295	15%
R303	15%
E304	15%
I319	15%
Q334	15%
N347	15%
I348	15%
R352	15%
Q356	15%
R360	15%
T367	15%
K375	15%
K383	15%
Q386	15%
I387	15%
K388	15%
Y389	15%
D392	15%
H393	15%
N410	15%
PRO	15%
SER	15%
GLU	15%
ILE	15%
ASN	15%
GLN	15%
HIS	15%
THR	15%
L419	15%
D420	15%
F421	15%
G422	15%
R423	15%
A432	15%
Y440	15%
L443	15%
I447	15%
Y454	15%
A460	15%
S461	15%
N462	15%
T466	15%
H467	15%
R468	15%
R469	15%
S470	15%
T471	15%
G493	15%
P494	15%
F495	15%
R508	15%
E509	15%
LEU	15%
GLY	15%
VAL	15%
LEU	15%
ASN	15%
ARG	15%
GLU	15%
GLY	15%
L518	15%
S523	15%
VAL	15%
GLY	15%

4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	122.75Å 122.75Å 103.42Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 1.90 19.72 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.5 (20.00-1.90) 93.9 (19.72-1.90)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.87 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.4.0069	Depositor
R, R_{free}	0.200 , 0.262 0.199 , 0.261	Depositor DCC
R_{free} test set	652 reflections (1.51%)	wwPDB-VP
Wilson B-factor (Å ²)	30.3	Xtriage
Anisotropy	0.477	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 50.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.021 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4470	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: UPG, 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	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.57	0/4271	0.65	4/5772 (0.1%)

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	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	248	ASP	CB-CG-OD1	5.82	123.54	118.30
1	A	246	ASP	CB-CG-OD1	5.45	123.20	118.30
1	A	295	ASP	CB-CG-OD1	5.24	123.02	118.30
1	A	196	LEU	CA-CB-CG	5.11	127.06	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	163	TYR	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	4115	0	4131	78	0
2	A	1	0	0	0	0
3	A	36	0	22	0	0
4	A	318	0	0	0	1
All	All	4470	0	4153	78	1

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 9.

The worst 5 of 78 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15[B]:ARG:NH1	1:A:48:GLN:HE21	1.51	1.08
1:A:469:ARG:HH11	1:A:469:ARG:HG2	1.16	1.04
1:A:469:ARG:HH11	1:A:469:ARG:CG	1.74	1.00
1:A:508:ARG:NH1	1:A:509:GLU:HB3	1.79	0.97
1:A:15[B]:ARG:HH12	1:A:48:GLN:NE2	1.67	0.91

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:2050:HOH:O	4:A:2112:HOH:O[9_444]	2.00	0.20

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	517/525 (98%)	508 (98%)	9 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	462/460 (100%)	435 (94%)	27 (6%)	20 10

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	182[B]	ASN
1	A	303	ARG
1	A	469	ARG
1	A	249	LEU
1	A	347[A]	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	499	ASN
1	A	490	HIS
1	A	356	GLN
1	A	467	HIS
1	A	334	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 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	UPG	A	1525	2	35,38,38	0.93	1 (2%)	53,58,58	1.39	6 (11%)

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	UPG	A	1525	2	-	2/23/59/59	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1525	UPG	C6-C5	2.44	1.40	1.35

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1525	UPG	C4-N3-C2	-4.88	120.15	126.58
3	A	1525	UPG	N3-C2-N1	3.85	120.00	114.89
3	A	1525	UPG	C5-C4-N3	3.68	120.34	114.84
3	A	1525	UPG	O2-C2-N1	-2.71	119.19	122.79
3	A	1525	UPG	O4-C4-C5	-2.56	120.65	125.16

There are no chirality outliers.

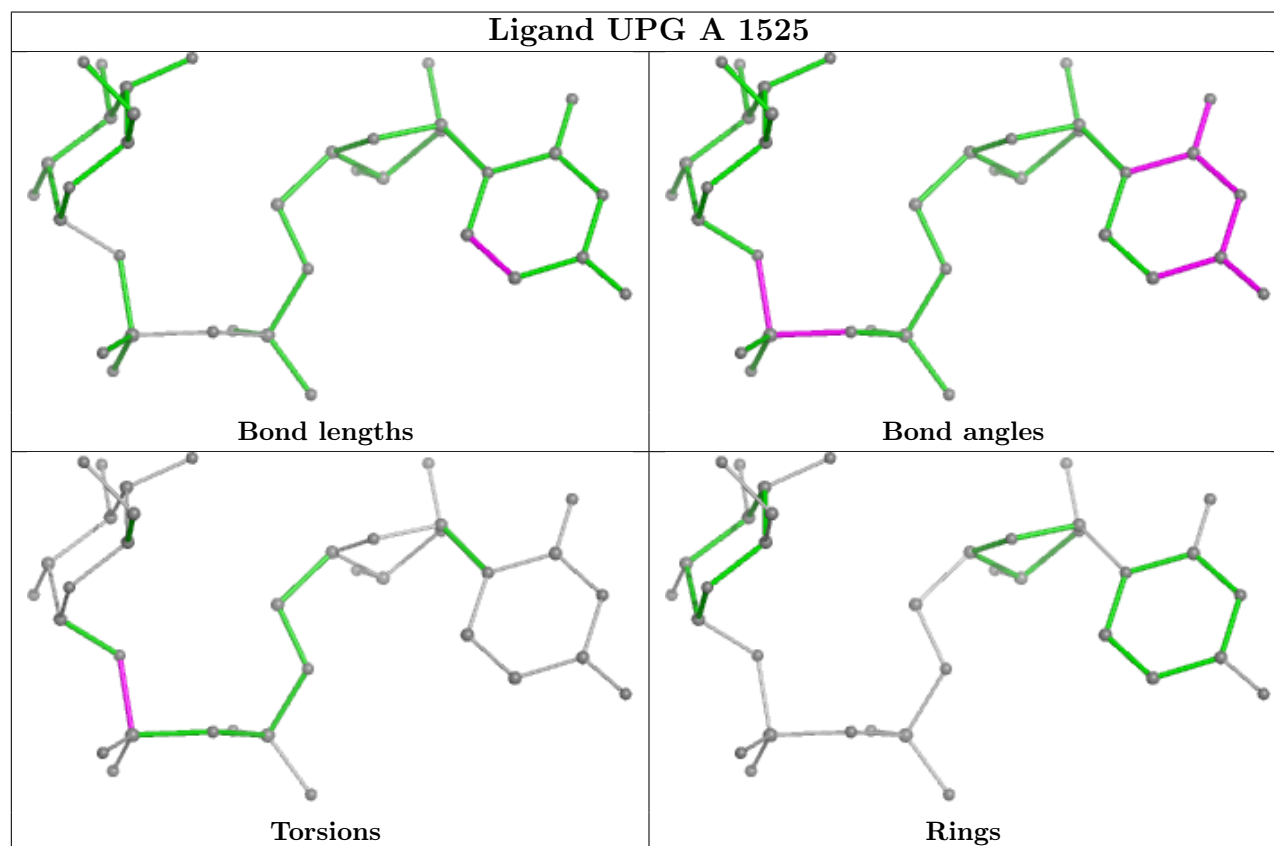
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1525	UPG	C1'-O3B-PB-O3A
3	A	1525	UPG	C1'-O3B-PB-O2B

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 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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	502/525 (95%)	0.04	13 (2%) 56 58	21, 33, 50, 66	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	470	SER	3.7
1	A	460	ALA	3.6
1	A	419	LEU	3.4
1	A	469	ARG	3.2
1	A	471	THR	3.2

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](#)

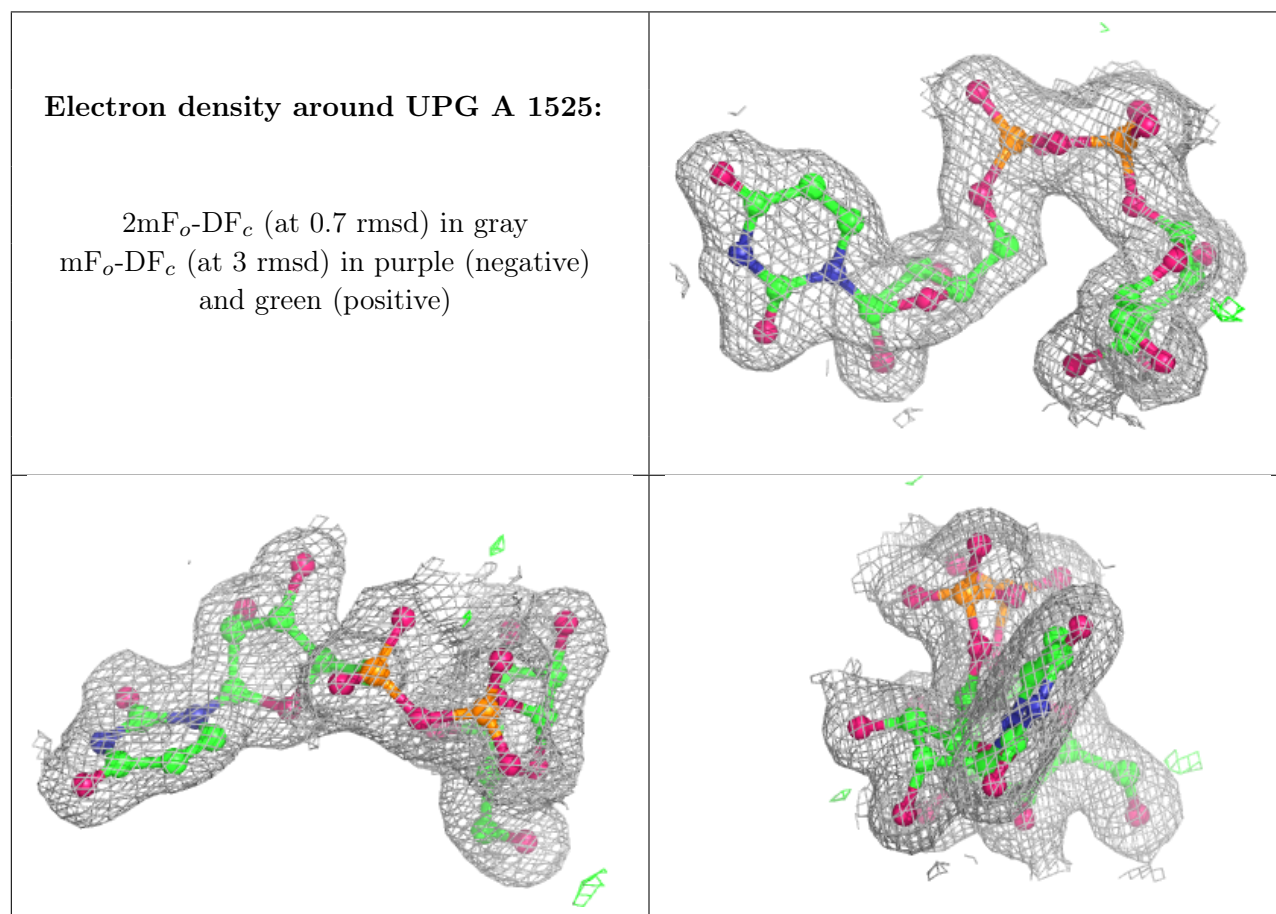
There are no monosaccharides in this entry.

6.4 Ligands [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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	MG	A	1524	1/1	0.94	0.06	37,37,37,37	0
3	UPG	A	1525	36/36	0.97	0.07	25,28,34,34	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.



6.5 Other polymers ⓘ

There are no such residues in this entry.