



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:43 PM GMT

PDB ID : 1GUQ  
Title : STRUCTURE OF NUCLEOTIDYLTRANSFERASE COMPLEXED WITH  
UDP-GLUCOSE  
Authors : Thoden, J.B.; Rayment, I.; Holden, H.  
Deposited on : 1996-10-23  
Resolution : 1.80 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  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

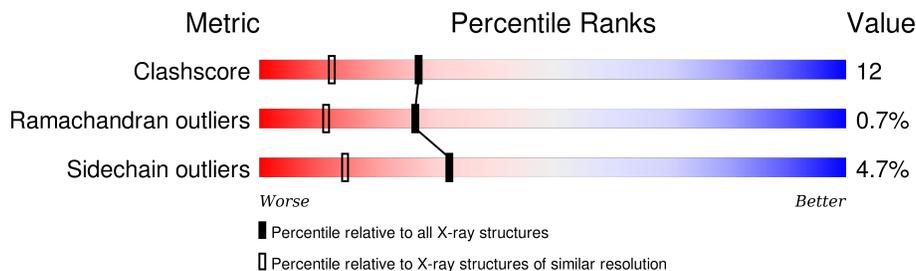
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	348	
1	B	348	
1	C	348	
1	D	348	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12597 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 GALACTOSE-1-PHOSPHATE URIDYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	347	2784	1763	492	516	13	0	0	0
1	B	344	2766	1753	489	511	13	0	0	0
1	C	344	2766	1753	489	511	13	0	0	0
1	D	344	2766	1753	489	511	13	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	166	GLY	HIS	ENGINEERED	UNP P09148
B	166	GLY	HIS	ENGINEERED	UNP P09148
C	166	GLY	HIS	ENGINEERED	UNP P09148
D	166	GLY	HIS	ENGINEERED	UNP P09148

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		

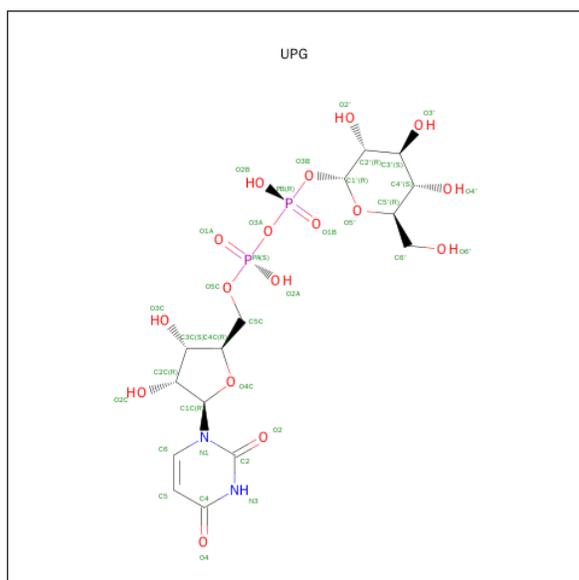
- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Fe 1 1	0	0
3	A	1	Total Fe 1 1	0	0
3	D	1	Total Fe 1 1	0	0
3	C	1	Total Fe 1 1	0	0

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total K 1 1	0	0
4	A	1	Total K 1 1	0	0
4	D	1	Total K 1 1	0	0
4	C	1	Total K 1 1	0	0

- Molecule 5 is URIDINE-5'-DIPHOSPHATE-GLUCOSE (three-letter code: UPG) (formula: C<sub>15</sub>H<sub>24</sub>N<sub>2</sub>O<sub>17</sub>P<sub>2</sub>).



*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
5	B	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
5	C	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
5	D	1	Total	C	N	O	P	0	0
			36	15	2	17	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	368	Total	O	0	0
			368	368		
6	B	363	Total	O	0	0
			363	363		
6	C	351	Total	O	0	0
			351	351		
6	D	277	Total	O	0	0
			277	277		





## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.40Å 57.50Å 188.90Å 90.00° 100.13° 90.00°	Depositor
Resolution (Å)	30.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-1.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.174 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	12597	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

## 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, ZN, K, FE

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	1.04	21/2869 (0.7%)	1.48	40/3912 (1.0%)
1	B	1.03	19/2851 (0.7%)	1.50	52/3889 (1.3%)
1	C	1.04	18/2851 (0.6%)	1.51	48/3889 (1.2%)
1	D	1.01	18/2851 (0.6%)	1.51	60/3889 (1.5%)
All	All	1.03	76/11422 (0.7%)	1.50	200/15579 (1.3%)

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	C	1	0

The worst 5 of 76 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	205	GLU	CD-OE2	9.70	1.36	1.25
1	C	329	GLU	CD-OE1	9.11	1.35	1.25
1	B	152	GLU	CD-OE1	8.44	1.34	1.25
1	C	232	GLU	CD-OE1	8.09	1.34	1.25
1	A	121	GLU	CD-OE1	8.05	1.34	1.25

The worst 5 of 200 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	334	ARG	NE-CZ-NH1	11.25	125.92	120.30
1	B	336	ARG	NE-CZ-NH2	-11.07	114.77	120.30
1	C	107	ARG	NE-CZ-NH2	-10.98	114.81	120.30
1	B	324	ARG	NE-CZ-NH1	9.80	125.20	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	254	ASP	CB-CG-OD2	-9.71	109.56	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	44	VAL	CA

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 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	2784	0	2661	53	0
1	B	2766	0	2644	54	0
1	C	2766	0	2644	77	0
1	D	2766	0	2644	96	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	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
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	36	0	22	0	0
5	B	36	0	22	0	0
5	C	36	0	22	0	0
5	D	36	0	22	1	0
6	A	368	0	0	9	1
6	B	363	0	0	11	1
6	C	351	0	0	6	0
6	D	277	0	0	6	0
All	All	12597	0	10681	263	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 12.

The worst 5 of 263 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:289:GLU:HG2	1:A:291:GLN:NE2	1.47	1.26
1:A:289:GLU:CG	1:A:291:GLN:HE22	1.68	1.07
1:C:21:TRP:CH2	1:D:100:GLN:HG3	1.94	1.01
1:D:342:HIS:CE1	1:D:344:ARG:HB2	1.96	0.99
1:A:38:GLU:HB2	1:A:308:THR:HA	1.44	0.99

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)
6:A:471:HOH:O	6:B:675:HOH:O[1_655]	2.18	0.02

## 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	345/348 (99%)	327 (95%)	16 (5%)	2 (1%)	30 14
1	B	342/348 (98%)	324 (95%)	16 (5%)	2 (1%)	30 14
1	C	342/348 (98%)	319 (93%)	19 (6%)	4 (1%)	16 4
1	D	342/348 (98%)	326 (95%)	14 (4%)	2 (1%)	30 14
All	All	1371/1392 (98%)	1296 (94%)	65 (5%)	10 (1%)	26 11

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	44	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	315	GLY
1	C	43	GLN
1	D	314	VAL
1	A	314	VAL

### 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	297/298 (100%)	287 (97%)	10 (3%)	44	26
1	B	295/298 (99%)	285 (97%)	10 (3%)	44	26
1	C	295/298 (99%)	280 (95%)	15 (5%)	29	12
1	D	295/298 (99%)	275 (93%)	20 (7%)	20	6
All	All	1182/1192 (99%)	1127 (95%)	55 (5%)	32	14

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	42	LYS
1	C	236	LEU
1	D	253	SER
1	C	43	GLN
1	C	194	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	100	GLN
1	C	27	HIS
1	D	100	GLN
1	B	291	GLN
1	C	5	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 16 ligands modelled in this entry, 12 are monoatomic - leaving 4 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
5	UPG	A	352	-	29,38,38	1.26	4 (13%)	43,58,58	2.44	4 (9%)
5	UPG	B	352	-	29,38,38	1.45	3 (10%)	43,58,58	2.58	8 (18%)
5	UPG	C	352	-	29,38,38	1.50	3 (10%)	43,58,58	2.16	6 (13%)
5	UPG	D	352	-	29,38,38	1.54	3 (10%)	43,58,58	2.37	4 (9%)

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
5	UPG	A	352	-	-	0/19/59/59	0/3/3/3
5	UPG	B	352	-	-	0/19/59/59	0/3/3/3
5	UPG	C	352	-	-	0/19/59/59	0/3/3/3
5	UPG	D	352	-	-	0/19/59/59	0/3/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	352	UPG	C6-C5	-3.42	1.30	1.38
5	B	352	UPG	C6-C5	-3.39	1.30	1.38
5	C	352	UPG	C6-C5	-3.24	1.31	1.38
5	A	352	UPG	C6-C5	-2.37	1.32	1.38
5	A	352	UPG	C4'-C5'	2.03	1.57	1.53

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	352	UPG	C4C-O4C-C1C	-5.98	103.14	109.72
5	B	352	UPG	C4C-O4C-C1C	-5.62	103.54	109.72
5	D	352	UPG	C4C-O4C-C1C	-4.14	105.17	109.72
5	A	352	UPG	C5-C4-N3	-3.65	113.75	123.12
5	B	352	UPG	C5-C4-N3	-3.41	114.36	123.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	352	UPG	1	0

## 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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers

EDS was not executed - this section will therefore be empty.