



wwPDB X-ray Structure Validation Summary Report

Sep 14, 2022 – 11:35 am BST

PDB ID : 7PPR
Title : The structure of UDP-glucose pyrophosphorylase from *Aspergillus fumigatus*
Authors : Morton, S.; Raimi, O.G.; Yan, K.; van Aalten, D.M.F.
Deposited on : 2021-09-14
Resolution : 2.57 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 : 2.30
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.30

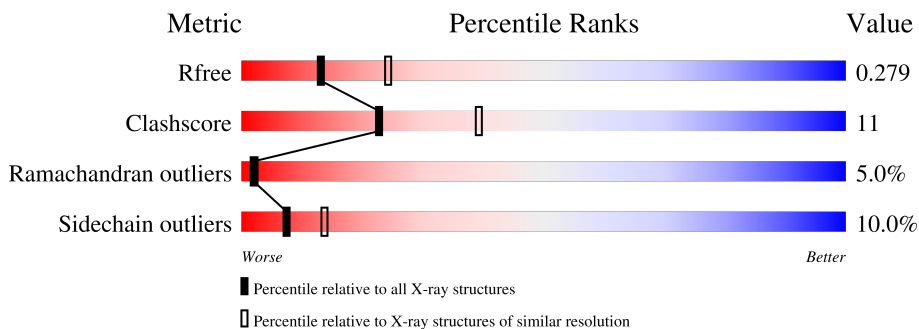
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.57 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	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	492	
1	B	492	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7187 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 UTP--glucose-1-phosphate uridylyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	450	3452	2193	606	641	12	0	6	0
1	B	464	3527	2242	606	666	13	0	4	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	GLY	-	expression tag	UNP E9R9Z5
A	21	PRO	-	expression tag	UNP E9R9Z5
A	22	LEU	-	expression tag	UNP E9R9Z5
A	23	GLY	-	expression tag	UNP E9R9Z5
A	24	SER	-	expression tag	UNP E9R9Z5
B	20	GLY	-	expression tag	UNP E9R9Z5
B	21	PRO	-	expression tag	UNP E9R9Z5
B	22	LEU	-	expression tag	UNP E9R9Z5
B	23	GLY	-	expression tag	UNP E9R9Z5
B	24	SER	-	expression tag	UNP E9R9Z5

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Cl	0	0
			2	2		

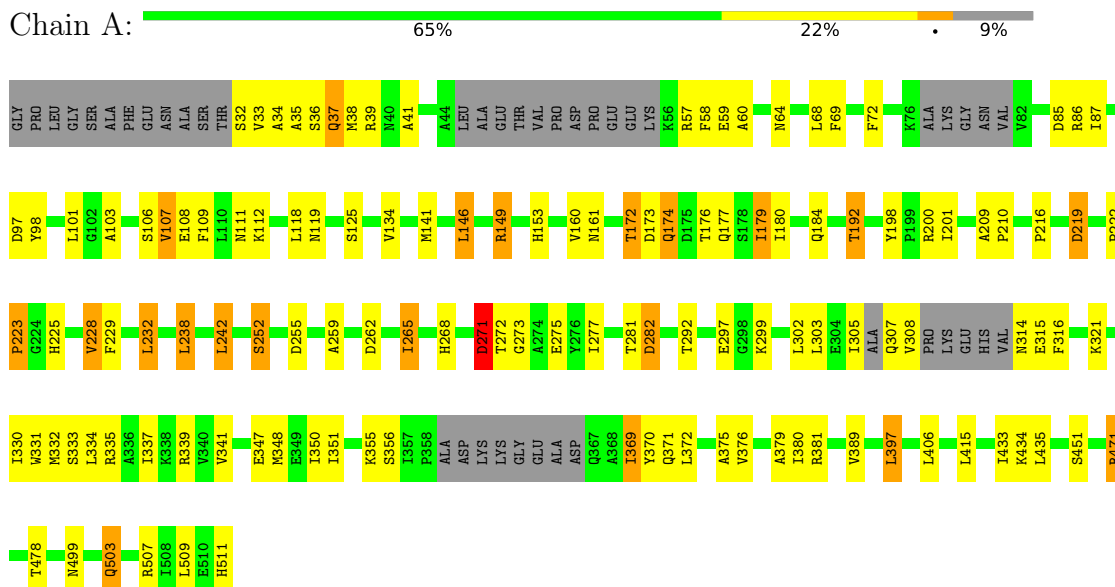
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	61	Total	O	0	0
			61	61		
4	B	70	Total	O	0	0
			70	70		

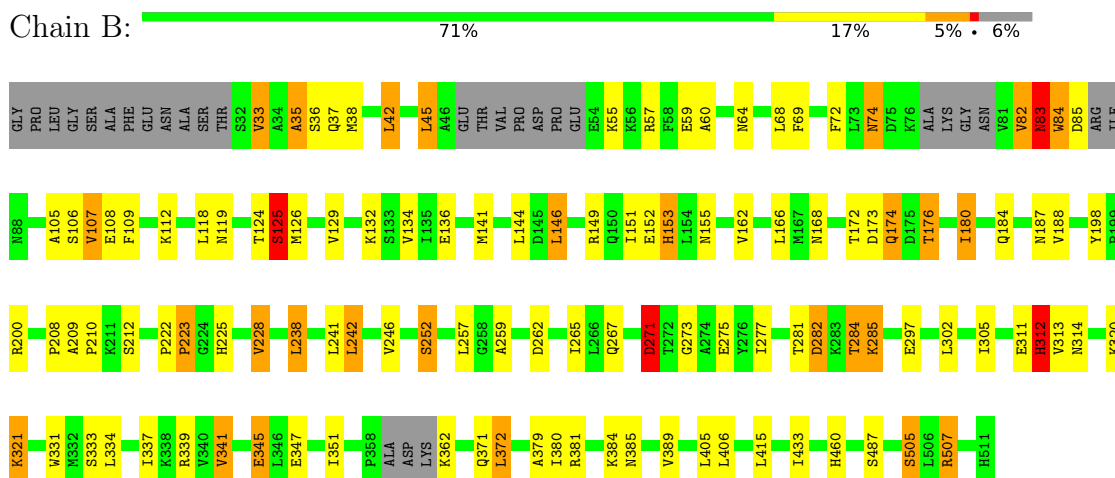
3 Residue-property plots

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: UTP--glucose-1-phosphate uridylyltransferase



- Molecule 1: UTP--glucose-1-phosphate uridylyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	134.16Å 152.61Å 160.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	80.00 – 2.57 80.00 – 2.57	Depositor EDS
% Data completeness (in resolution range)	99.8 (80.00-2.57) 99.8 (80.00-2.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 2.58Å)	Xtrriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.227 , 0.277 0.231 , 0.279	Depositor DCC
R_{free} test set	2699 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å ²)	60.7	Xtrriage
Anisotropy	0.181	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.015 for -h,-l,-k	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7187	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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: CL, SO4

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	0.74	0/3518	0.91	2/4765 (0.0%)
1	B	0.74	0/3596	0.88	1/4882 (0.0%)
All	All	0.74	0/7114	0.90	3/9647 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	255	ASP	CB-CA-C	-9.59	91.23	110.40
1	B	271	ASP	CB-CA-C	6.31	123.02	110.40
1	A	271	ASP	CB-CA-C	5.25	120.89	110.40

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 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	3452	0	3361	77	0
1	B	3527	0	3428	83	0
2	A	45	0	0	2	0
2	B	30	0	0	1	0
3	B	2	0	0	0	0
4	A	61	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	70	0	0	0	0
All	All	7187	0	6789	159	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 11.

The worst 5 of 159 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:112:LYS:O	1:B:246:VAL:O	1.60	1.17
1:B:57[B]:ARG:HG2	1:B:57[B]:ARG:HH11	1.12	1.07
1:A:119:ASN:O	1:A:172:THR:HG21	1.60	1.01
1:A:111:ASN:OD1	1:A:160:VAL:HG13	1.62	0.97
1:B:57[B]:ARG:HH11	1:B:57[B]:ARG:CG	1.83	0.92

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	443/492 (90%)	387 (87%)	38 (9%)	18 (4%)	3	3
1	B	457/492 (93%)	390 (85%)	38 (8%)	29 (6%)	1	1
All	All	900/984 (92%)	777 (86%)	76 (8%)	47 (5%)	2	2

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ALA
1	A	37	GLN
1	A	38	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	58	PHE
1	A	107	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	354/425 (83%)	316 (89%)	38 (11%)	6	11
1	B	366/425 (86%)	332 (91%)	34 (9%)	9	16
All	All	720/850 (85%)	648 (90%)	72 (10%)	7	14

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	238	LEU
1	B	507	ARG
1	B	252	SER
1	B	312	HIS
1	A	314	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	371	GLN
1	B	468	ASN
1	B	153	HIS
1	B	155	ASN
1	B	168	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 17 ligands modelled in this entry, 2 are monoatomic - leaving 15 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	SO4	A	607	-	4,4,4	0.35	0	6,6,6	0.04	0
2	SO4	A	606	-	4,4,4	0.34	0	6,6,6	0.15	0
2	SO4	A	609	-	4,4,4	0.35	0	6,6,6	0.11	0
2	SO4	A	602	-	4,4,4	0.38	0	6,6,6	0.12	0
2	SO4	A	603	-	4,4,4	0.35	0	6,6,6	0.12	0
2	SO4	B	601	-	4,4,4	0.31	0	6,6,6	0.07	0
2	SO4	B	602	-	4,4,4	0.25	0	6,6,6	0.17	0
2	SO4	B	603	-	4,4,4	0.49	0	6,6,6	0.17	0
2	SO4	B	606	-	4,4,4	0.35	0	6,6,6	0.09	0
2	SO4	A	605	-	4,4,4	0.35	0	6,6,6	0.16	0
2	SO4	A	604	-	4,4,4	0.29	0	6,6,6	0.17	0
2	SO4	A	601	-	4,4,4	0.33	0	6,6,6	0.05	0
2	SO4	B	605	-	4,4,4	0.32	0	6,6,6	0.12	0
2	SO4	A	608	-	4,4,4	0.35	0	6,6,6	0.08	0
2	SO4	B	604	-	4,4,4	0.33	0	6,6,6	0.09	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	602	SO4	1	0
2	A	603	SO4	1	0
2	B	602	SO4	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

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.

6.5 Other polymers

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