



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2017 – 05:54 PM EDT

PDB ID : 3DMP  
Title : 2.6 Å crystal structure of uracil phosphoribosyltransferase from Burkholderia pseudomallei  
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : unknown  
Resolution : 2.60 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

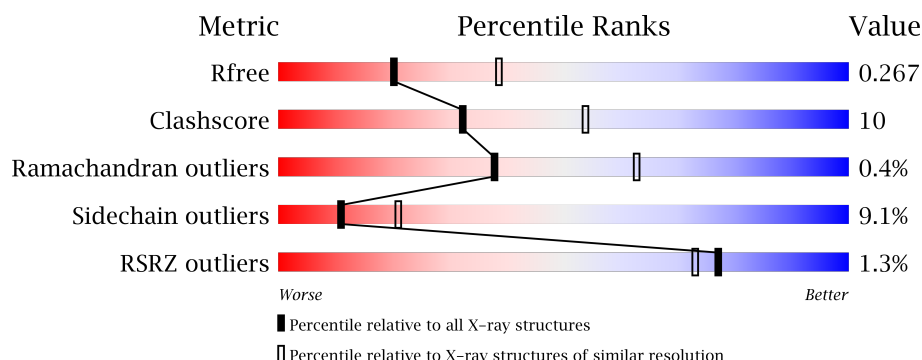
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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.

Mol	Chain	Length	Quality of chain
1	A	217	<div> <div>0.1%</div> <div> <div></div> <div>73%</div> <div>21%</div> <div>5%</div> </div> </div>
1	B	217	<div> <div>2%</div> <div> <div></div> <div>74%</div> <div>20%</div> <div></div> </div> </div>
1	C	217	<div> <div></div> <div> <div></div> <div>77%</div> <div>18%</div> <div></div> </div> </div>
1	D	217	<div> <div>0.1%</div> <div> <div></div> <div>77%</div> <div>18%</div> <div></div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6777 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 Uracil phosphoribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	214	Total	C	N	O	S	0	0	0
			1676	1070	297	302	7			
1	B	214	Total	C	N	O	S	0	0	0
			1676	1070	297	302	7			
1	C	213	Total	C	N	O	S	0	0	0
			1666	1064	294	301	7			
1	D	214	Total	C	N	O	S	0	0	0
			1676	1070	297	302	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP Q63VS8
B	0	SER	-	EXPRESSION TAG	UNP Q63VS8
C	0	SER	-	EXPRESSION TAG	UNP Q63VS8
D	0	SER	-	EXPRESSION TAG	UNP Q63VS8

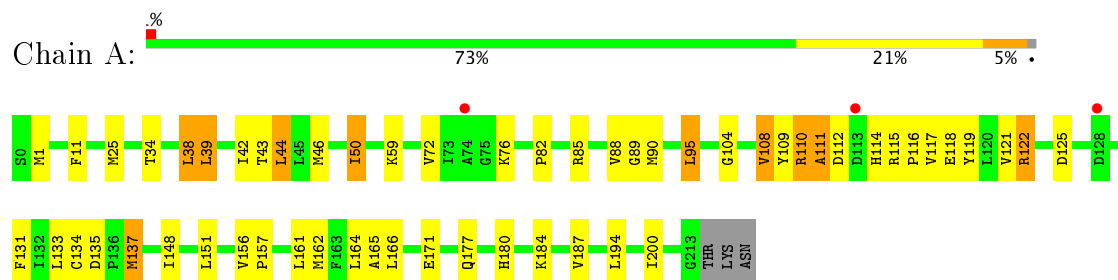
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	15	Total	O	0	0
			15	15		
2	B	19	Total	O	0	0
			19	19		
2	C	23	Total	O	0	0
			23	23		
2	D	26	Total	O	0	0
			26	26		

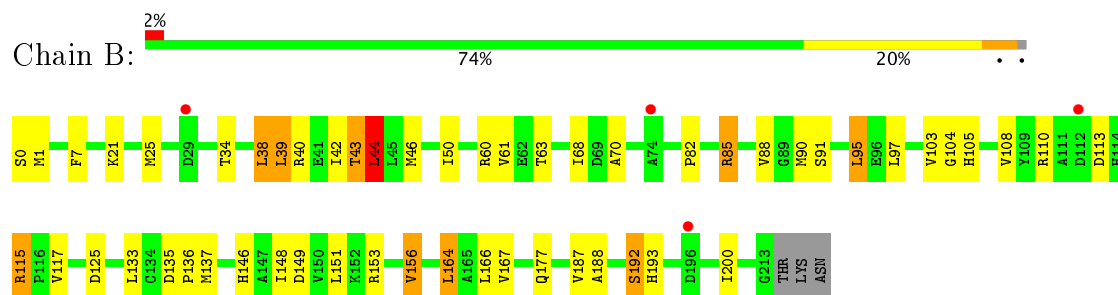
### 3 Residue-property plots [i](#)

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.

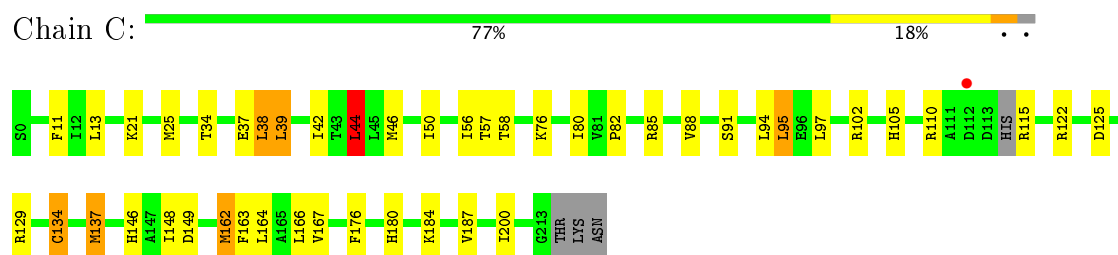
#### • Molecule 1: Uracil phosphoribosyltransferase



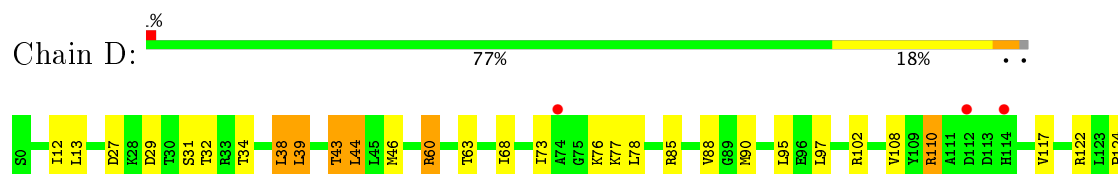
#### • Molecule 1: Uracil phosphoribosyltransferase



#### • Molecule 1: Uracil phosphoribosyltransferase



#### • Molecule 1: Uracil phosphoribosyltransferase



D125	
R129	
I130	
F131	
I132	
L133	
C134	
M137	
I148	
L151	
V156	
L161	
W162	
F163	
L164	
V167	
I200	
D206	
D209	
G213	
THR	
LYS	
ASN	

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.22Å 80.64Å 141.83Å 90.00° 99.11° 90.00°	Depositor
Resolution (Å)	49.33 – 2.60 49.32 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.33-2.60) 99.6 (49.32-2.60)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 2.61Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.219 , 0.271 0.217 , 0.267	Depositor DCC
$R_{free}$ test set	1758 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	44.9	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 42.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6777	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.17 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.9436e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<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.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.53	1/1711 (0.1%)	0.70	1/2323 (0.0%)
1	B	0.51	0/1711	0.71	1/2323 (0.0%)
1	C	0.54	1/1699 (0.1%)	0.69	1/2305 (0.0%)
1	D	0.51	1/1711 (0.1%)	0.70	1/2323 (0.0%)
All	All	0.52	3/6832 (0.0%)	0.70	4/9274 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	134	CYS	CB-SG	-8.51	1.67	1.82
1	A	134	CYS	CB-SG	-5.88	1.72	1.81
1	D	134	CYS	CB-SG	-5.49	1.72	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	44	LEU	CA-CB-CG	6.59	130.47	115.30
1	D	44	LEU	CA-CB-CG	6.22	129.60	115.30
1	C	44	LEU	CA-CB-CG	5.78	128.60	115.30
1	A	44	LEU	CA-CB-CG	5.72	128.46	115.30

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 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	1676	0	1717	51	0
1	B	1676	0	1717	36	0
1	C	1666	0	1709	32	0
1	D	1676	0	1717	24	0
2	A	15	0	0	0	0
2	B	19	0	0	2	0
2	C	23	0	0	1	0
2	D	26	0	0	1	0
All	All	6777	0	6860	135	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 10.

The worst 5 of 135 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:109:TYR:N	1:A:110:ARG:HB2	1.69	1.06
1:B:39:LEU:O	1:B:43:THR:HG22	1.59	1.02
1:C:21:LYS:HB2	1:C:42:ILE:HD11	1.45	0.96
1:C:21:LYS:CB	1:C:42:ILE:HD11	2.01	0.90
1:A:109:TYR:H	1:A:110:ARG:HB2	1.33	0.86

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	212/217 (98%)	198 (93%)	11 (5%)	3 (1%)	13	26
1	B	212/217 (98%)	202 (95%)	10 (5%)	0	100	100
1	C	209/217 (96%)	203 (97%)	6 (3%)	0	100	100
1	D	212/217 (98%)	203 (96%)	9 (4%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	845/868 (97%)	806 (95%)	36 (4%)	3 (0%)	38 63

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	110	ARG
1	A	111	ALA
1	A	108	VAL

### 5.3.2 Protein sidechains ⓘ

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	182/185 (98%)	168 (92%)	14 (8%)	15 29
1	B	182/185 (98%)	163 (90%)	19 (10%)	8 15
1	C	181/185 (98%)	168 (93%)	13 (7%)	17 33
1	D	182/185 (98%)	162 (89%)	20 (11%)	7 13
All	All	727/740 (98%)	661 (91%)	66 (9%)	11 21

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	156	VAL
1	C	44	LEU
1	D	122	ARG
1	B	164	LEU
1	C	37	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	146	HIS
1	D	3	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	177	GLN
1	B	3	GLN
1	C	19	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	214/217 (98%)	-0.02	3 (1%) 75 71	32, 45, 62, 70	0
1	B	214/217 (98%)	-0.12	4 (1%) 67 61	32, 43, 56, 65	0
1	C	213/217 (98%)	-0.19	1 (0%) 90 89	33, 43, 55, 70	0
1	D	214/217 (98%)	-0.02	3 (1%) 75 71	32, 45, 59, 65	0
All	All	855/868 (98%)	-0.09	11 (1%) 77 73	32, 44, 59, 70	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	74	ALA	4.4
1	D	112	ASP	2.7
1	A	74	ALA	2.5
1	B	112	ASP	2.3
1	D	114	HIS	2.3

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.