



# Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 11:10 pm BST

PDB ID : 1PKJ  
Title : Structural basis for recognition and catalysis by the bifunctional dCTP deaminase and dUTPase from *Methanococcus jannaschii*  
Authors : Huffman, J.L.; Li, H.; White, R.H.; Tainer, J.A.  
Deposited on : 2003-06-05  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

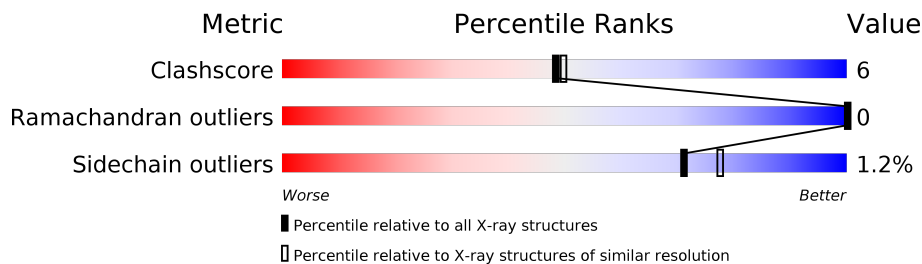
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (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  $\geq 3$ , 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  $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	204	 75% 12% 12%
1	B	204	 74% 12% 14%

## 2 Entry composition [i](#)

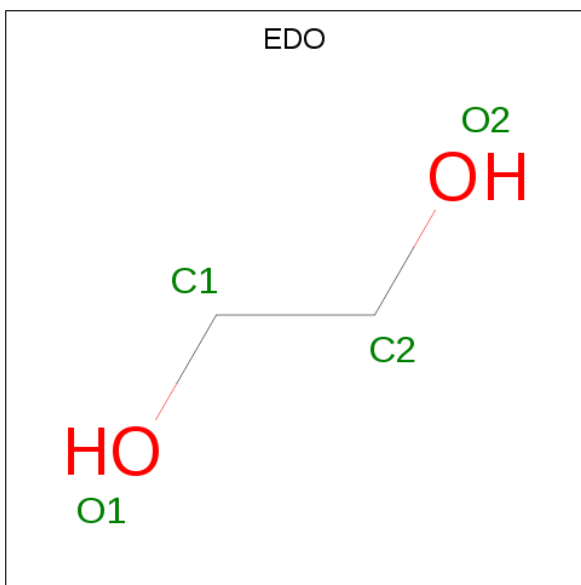
There are 4 unique types of molecules in this entry. The entry contains 3161 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bifunctional deaminase/diphosphatase.

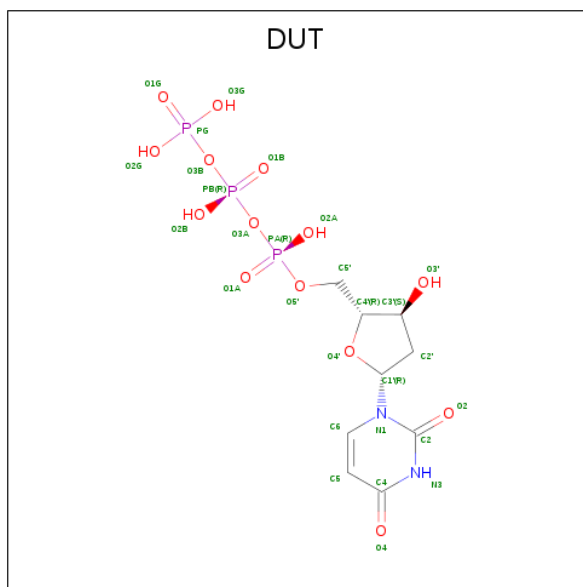
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	180	Total 1456	C 943	N 232	O 278	S 3	0	0	0
1	B	175	Total 1414	C 918	N 224	O 269	S 3	0	0	0

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 3 is DEOXYURIDINE-5'-TRIPHOSPHATE (three-letter code: DUT) (formula: C<sub>9</sub>H<sub>15</sub>N<sub>2</sub>O<sub>14</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	B	1	28	9	2	14	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	133	Total	O	0	0
			133	133		
4	B	122	Total	O	0	0
			122	122		



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.74Å 110.74Å 110.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.14 – 2.10	Depositor
% Data completeness (in resolution range)	97.0 (20.14-2.10)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.193 , 0.233	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	3161	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	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: DUT, EDO

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	0.36	0/1485	0.60	0/2008
1	B	0.34	0/1442	0.60	0/1951
All	All	0.35	0/2927	0.60	0/3959

There are no bond length outliers.

There are no bond angle outliers.

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	1456	0	1467	21	0
1	B	1414	0	1431	18	0
2	A	4	0	6	0	0
2	B	4	0	6	0	0
3	B	28	0	11	1	0
4	A	133	0	0	3	0
4	B	122	0	0	2	0
All	All	3161	0	2921	35	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 6.

All (35) 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:105:GLU:HG2	1:A:139:LYS:HG2	1.54	0.88
1:A:18:ILE:HB	1:A:105:GLU:HB2	1.58	0.86
1:A:129:GLN:HB2	4:A:508:HOH:O	1.89	0.72
1:B:18:ILE:HB	1:B:105:GLU:HB2	1.72	0.71
1:A:31:SER:HB3	1:A:165:ILE:HD13	1.78	0.66
1:A:46:GLU:HG2	1:A:47:VAL:HG13	1.78	0.64
1:A:180:ARG:HH12	1:B:60:LYS:HB2	1.66	0.60
1:B:111:SER:HB2	1:B:169:LEU:HD21	1.88	0.55
1:A:57:LYS:HD3	1:A:57:LYS:N	2.23	0.54
1:B:79:ILE:HG22	1:B:83:LYS:HE3	1.91	0.52
1:A:74:LEU:HD13	1:A:91:VAL:HG12	1.92	0.51
1:B:43:TYR:HD2	1:B:48:TYR:HH	1.60	0.48
1:A:179:GLU:HG2	1:B:59:ILE:CD1	2.43	0.48
1:A:125:LEU:C	1:A:125:LEU:HD13	2.34	0.47
1:B:151:LYS:HE3	4:B:509:HOH:O	2.14	0.47
1:B:79:ILE:CG2	1:B:83:LYS:HE3	2.45	0.46
1:A:129:GLN:HE22	1:A:145:GLU:HB3	1.81	0.46
1:B:134:ILE:HG23	3:B:250:DUT:H2'2	1.98	0.45
1:A:129:GLN:NE2	4:A:508:HOH:O	2.46	0.45
1:B:35:THR:O	1:B:100:THR:HA	2.17	0.45
1:A:62:LYS:HG3	4:A:534:HOH:O	2.17	0.44
1:B:96:VAL:HG13	1:B:146:ILE:HB	1.99	0.44
1:A:62:LYS:NZ	1:A:62:LYS:HB3	2.32	0.44
1:A:96:VAL:HG13	1:A:146:ILE:HB	2.00	0.43
1:B:105:GLU:HG2	1:B:139:LYS:HG2	2.00	0.43
1:A:180:ARG:HH12	1:B:60:LYS:CB	2.32	0.43
1:A:67:VAL:O	1:A:152:PRO:HA	2.18	0.42
1:A:50:LEU:HD22	1:B:85:LYS:HE3	2.01	0.42
1:A:9:ILE:HD12	1:A:24:LYS:HE2	2.01	0.41
1:B:136:ALA:HB2	1:B:169:LEU:HD11	2.03	0.41
1:B:116:GLY:HA3	4:B:351:HOH:O	2.20	0.41
1:B:92:VAL:HG22	1:B:93:GLU:N	2.36	0.41
1:A:126:THR:O	1:A:146:ILE:HA	2.21	0.41
1:A:69:PRO:O	1:A:72:TYR:HB2	2.21	0.41
1:B:20:LYS:HA	1:B:21:PRO:HA	1.94	0.40

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	178/204 (87%)	173 (97%)	5 (3%)	0	100	100
1	B	173/204 (85%)	168 (97%)	5 (3%)	0	100	100
All	All	351/408 (86%)	341 (97%)	10 (3%)	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	163/186 (88%)	160 (98%)	3 (2%)	59	65
1	B	159/186 (86%)	158 (99%)	1 (1%)	86	90
All	All	322/372 (87%)	318 (99%)	4 (1%)	71	77

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

Mol	Chain	Res	Type
1	A	57	LYS
1	A	128	HIS
1	A	180	ARG
1	B	5	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	ASN
1	A	101	ASN
1	A	113	GLN
1	A	115	GLN
1	A	129	GLN
1	A	158	ASN
1	B	55	ASN
1	B	73	ASN
1	B	101	ASN
1	B	113	GLN
1	B	115	GLN
1	B	158	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](#)

3 ligands are modelled in this entry.

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	EDO	B	299	-	3,3,3	0.64	0	2,2,2	0.19	0
2	EDO	A	298	-	3,3,3	0.64	0	2,2,2	0.19	0
3	DUT	B	250	-	22,29,29	2.36	9 (40%)	27,45,45	1.55	4 (14%)

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
2	EDO	B	299	-	-	1/1/1/1	-
2	EDO	A	298	-	-	1/1/1/1	-
3	DUT	B	250	-	-	3/19/34/34	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	250	DUT	PG-O1G	4.46	1.64	1.50
3	B	250	DUT	C4-N3	4.43	1.40	1.33
3	B	250	DUT	C6-N1	4.26	1.41	1.35
3	B	250	DUT	PB-O1B	4.03	1.65	1.50
3	B	250	DUT	PA-O1A	3.66	1.63	1.50
3	B	250	DUT	PG-O2G	2.74	1.65	1.54
3	B	250	DUT	PG-O3G	2.54	1.64	1.54
3	B	250	DUT	C2-N3	2.13	1.42	1.38
3	B	250	DUT	PA-O2A	2.13	1.65	1.55

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	250	DUT	PB-O3B-PG	-4.17	118.53	132.83
3	B	250	DUT	C5-C4-N3	-3.75	115.05	123.31
3	B	250	DUT	PB-O3A-PA	3.17	143.71	132.83
3	B	250	DUT	O3G-PG-O3B	2.11	111.70	104.64

There are no chirality outliers.

All (5) torsion outliers are listed below:

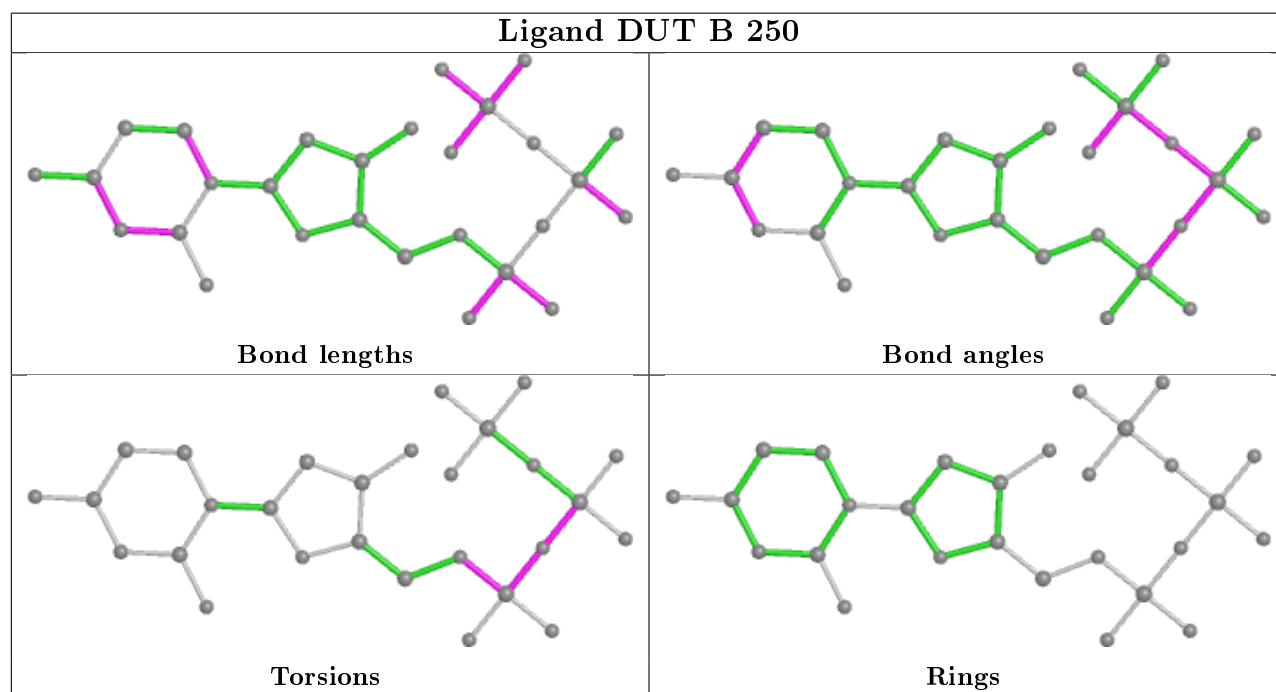
Mol	Chain	Res	Type	Atoms
2	B	299	EDO	O1-C1-C2-O2
2	A	298	EDO	O1-C1-C2-O2
3	B	250	DUT	PB-O3A-PA-O1A
3	B	250	DUT	PA-O3A-PB-O2B
3	B	250	DUT	C5'-O5'-PA-O1A

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	250	DUT	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 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.