



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 12:58 pm BST

PDB ID : 1QZQ  
Title : human Tyrosyl DNA phosphodiesterase  
Authors : Raymond, A.C.; Rideout, M.C.; Staker, B.; Hjerrild, K.; Burgin Jr., A.B.  
Deposited on : 2003-09-17  
Resolution : 2.40 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

<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  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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

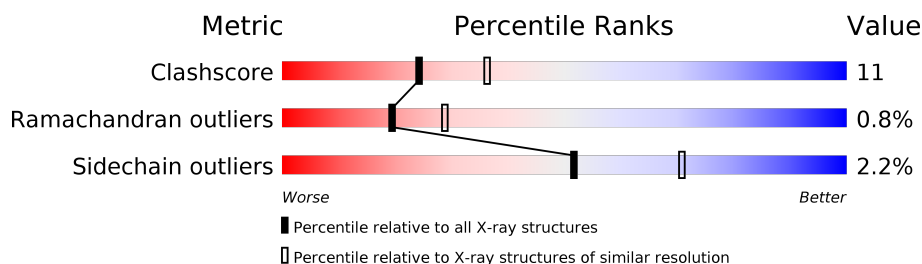
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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	483	
1	B	483	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7250 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 tyrosyl-DNA phosphodiesterase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	439	Total	C	N	O	S	0	0	0
			3517	2282	593	630	12			
1	B	437	Total	C	N	O	S	0	0	0
			3495	2270	589	623	13			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	126	MET	-	CLONING ARTIFACT	UNP Q9NUW8
A	127	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
A	128	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
A	129	TYR	-	CLONING ARTIFACT	UNP Q9NUW8
A	130	MET	-	CLONING ARTIFACT	UNP Q9NUW8
A	131	PRO	-	CLONING ARTIFACT	UNP Q9NUW8
A	132	THR	-	CLONING ARTIFACT	UNP Q9NUW8
A	133	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
A	134	HIS	-	EXPRESSION TAG	UNP Q9NUW8
A	135	HIS	-	EXPRESSION TAG	UNP Q9NUW8
A	136	HIS	-	EXPRESSION TAG	UNP Q9NUW8
A	137	HIS	-	EXPRESSION TAG	UNP Q9NUW8
A	138	HIS	-	EXPRESSION TAG	UNP Q9NUW8
A	139	HIS	-	EXPRESSION TAG	UNP Q9NUW8
A	140	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
A	141	ASN	-	CLONING ARTIFACT	UNP Q9NUW8
A	142	LEU	-	CLONING ARTIFACT	UNP Q9NUW8
A	143	TYR	-	CLONING ARTIFACT	UNP Q9NUW8
A	144	PHE	-	CLONING ARTIFACT	UNP Q9NUW8
A	145	GLN	-	CLONING ARTIFACT	UNP Q9NUW8
A	146	GLY	-	CLONING ARTIFACT	UNP Q9NUW8
A	147	THR	-	CLONING ARTIFACT	UNP Q9NUW8
A	148	SER	-	CLONING ARTIFACT	UNP Q9NUW8
B	126	MET	-	CLONING ARTIFACT	UNP Q9NUW8
B	127	GLU	-	CLONING ARTIFACT	UNP Q9NUW8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	128	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
B	129	TYR	-	CLONING ARTIFACT	UNP Q9NUW8
B	130	MET	-	CLONING ARTIFACT	UNP Q9NUW8
B	131	PRO	-	CLONING ARTIFACT	UNP Q9NUW8
B	132	THR	-	CLONING ARTIFACT	UNP Q9NUW8
B	133	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
B	134	HIS	-	EXPRESSION TAG	UNP Q9NUW8
B	135	HIS	-	EXPRESSION TAG	UNP Q9NUW8
B	136	HIS	-	EXPRESSION TAG	UNP Q9NUW8
B	137	HIS	-	EXPRESSION TAG	UNP Q9NUW8
B	138	HIS	-	EXPRESSION TAG	UNP Q9NUW8
B	139	HIS	-	EXPRESSION TAG	UNP Q9NUW8
B	140	GLU	-	CLONING ARTIFACT	UNP Q9NUW8
B	141	ASN	-	CLONING ARTIFACT	UNP Q9NUW8
B	142	LEU	-	CLONING ARTIFACT	UNP Q9NUW8
B	143	TYR	-	CLONING ARTIFACT	UNP Q9NUW8
B	144	PHE	-	CLONING ARTIFACT	UNP Q9NUW8
B	145	GLN	-	CLONING ARTIFACT	UNP Q9NUW8
B	146	GLY	-	CLONING ARTIFACT	UNP Q9NUW8
B	147	THR	-	CLONING ARTIFACT	UNP Q9NUW8
B	148	SER	-	CLONING ARTIFACT	UNP Q9NUW8

- Molecule 2 is water.

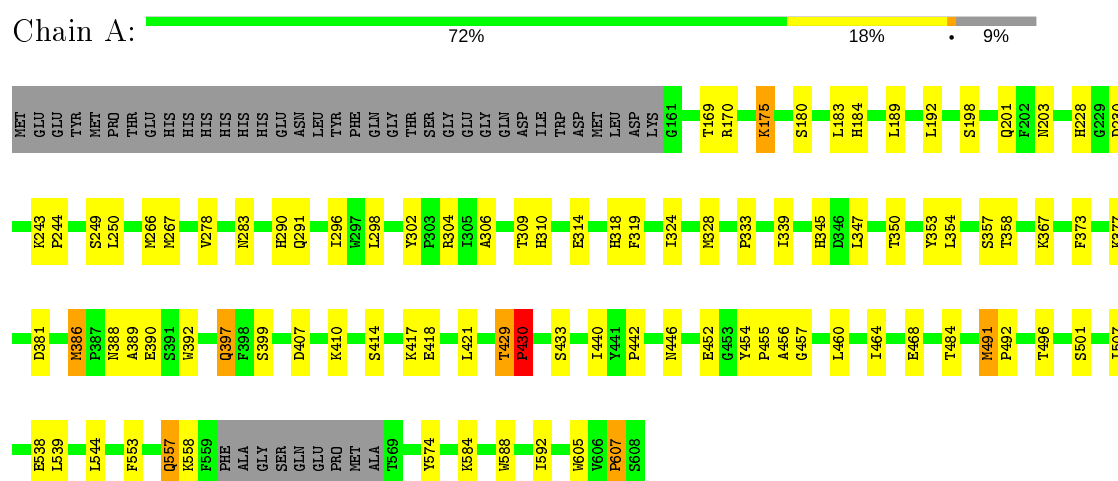
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	138	Total	O	0	0
			138	138		
2	B	100	Total	O	0	0
			100	100		

### 3 Residue-property plots

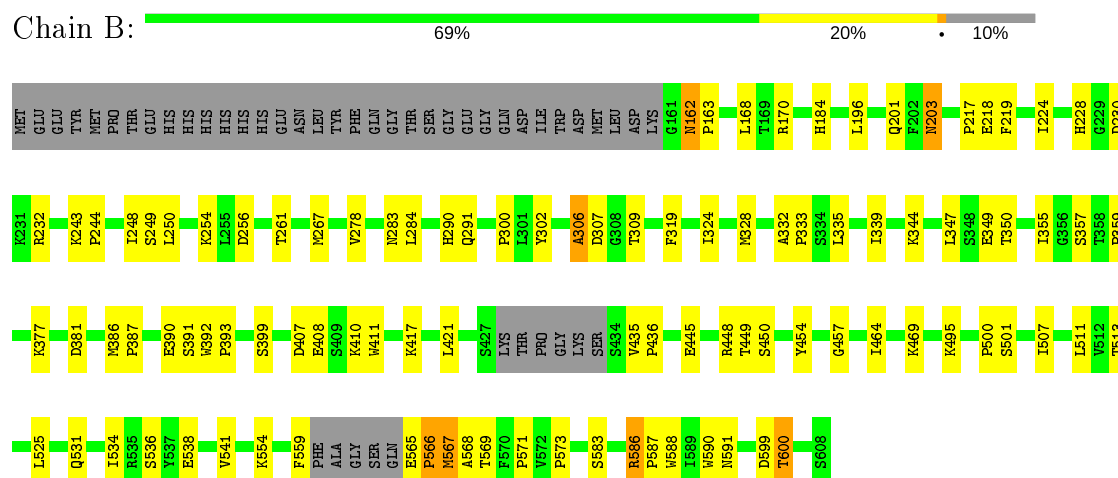
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

Note EDS was not executed.

- Molecule 1: tyrosyl-DNA phosphodiesterase 1



- Molecule 1: tyrosyl-DNA phosphodiesterase 1



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.02Å 105.13Å 194.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.54 – 2.40	Depositor
% Data completeness (in resolution range)	99.4 (48.54-2.40)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
Refinement program	CNX 2002	Depositor
R, $R_{free}$	0.185 , 0.234	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7250	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	0.47	0/3633	0.71	0/4934
1	B	0.43	0/3610	0.68	0/4904
All	All	0.45	0/7243	0.70	0/9838

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 ⓘ

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	3517	0	3463	77	0
1	B	3495	0	3430	76	0
2	A	138	0	0	2	0
2	B	100	0	0	0	0
All	All	7250	0	6893	153	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 153 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:201:GLN:HE22	1:A:267:MET:HE3	1.28	0.96

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:557:GLN:HE21	1:A:558:LYS:H	1.08	0.92
1:A:201:GLN:HE22	1:A:267:MET:CE	1.83	0.91
1:A:442:PRO:HB3	1:A:491:MET:HE1	1.55	0.88
1:B:261:THR:H	1:B:536:SER:HB3	1.41	0.85

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	435/483 (90%)	416 (96%)	16 (4%)	3 (1%)	22	32
1	B	431/483 (89%)	409 (95%)	18 (4%)	4 (1%)	17	25
All	All	866/966 (90%)	825 (95%)	34 (4%)	7 (1%)	19	29

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	607	PRO
1	B	306	ALA
1	B	567	MET
1	A	430	PRO
1	A	429	THR

### 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	383/421 (91%)	373 (97%)	10 (3%)	46	66
1	B	378/421 (90%)	371 (98%)	7 (2%)	57	75
All	All	761/842 (90%)	744 (98%)	17 (2%)	52	71

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	491	MET
1	A	557	GLN
1	B	408	GLU
1	A	430	PRO
1	B	586	ARG

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

Mol	Chain	Res	Type
1	A	557	GLN
1	B	162	ASN
1	B	363	GLN
1	A	388	ASN
1	B	388	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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 ⓘ

### 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.