



# wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 07:59 PM GMT

PDB ID : 4KTQ  
Title : BINARY COMPLEX OF THE LARGE FRAGMENT OF DNA POLYMERASE I FROM T. AQUATICUS BOUND TO A PRIMER/TEMPLATE DNA  
Authors : Li, Y.; Waksman, G.  
Deposited on : 1998-09-09  
Resolution : 2.50 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

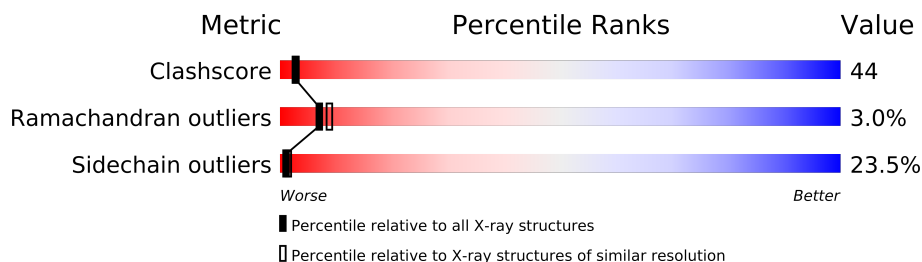
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.50 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	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	12	
2	C	13	
3	A	539	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4769 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA chain called DNA (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	12	Total	C	N	O	P	0	0	0
			240	114	48	67	11			

- Molecule 2 is a DNA chain called DNA (5'-D(\*GP\*GP\*GP\*CP\*GP\*CP\*CP\*GP\*TP\*GP\*GP\*TP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	13	Total	C	N	O	P	0	0	0
			267	126	51	78	12			

- Molecule 3 is a protein called PROTEIN (LARGE FRAGMENT OF DNA POLYMERASE I).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	539	Total	C	N	O	S	0	0	0
			4159	2652	740	754	13			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	84	Total	O	0	0
			84	84		
4	B	8	Total	O	0	0
			8	8		
4	C	11	Total	O	0	0
			11	11		

### 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 errors 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.

Note EDS was not executed.

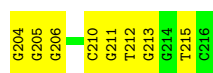
- Molecule 1: DNA (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3')

Chain B: 



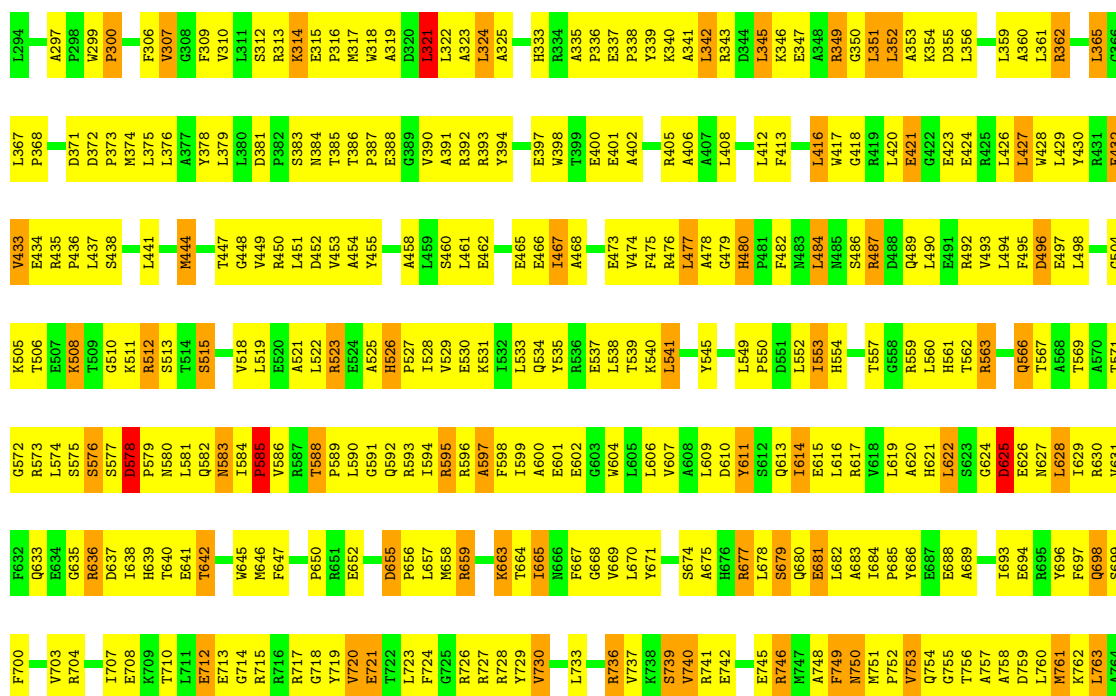
- Molecule 2: DNA (5'-D(\*GP\*GP\*GP\*CP\*GP\*CP\*CP\*GP\*TP\*GP\*GP\*TP\*C)-3')

Chain C: 



- Molecule 3: PROTEIN (LARGE FRAGMENT OF DNA POLYMERASE I)

Chain A: 





## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.84Å 110.84Å 90.84Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.50	Depositor
% Data completeness (in resolution range)	91.8 (30.00-2.50)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.227 , 0.298	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4769	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DOC

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	B	0.47	0/249	0.73	0/382
2	C	0.53	0/299	0.83	0/461
3	A	0.36	0/4249	0.58	0/5776
All	All	0.38	0/4797	0.61	0/6619

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	240	0	134	12	0
2	C	267	0	147	17	0
3	A	4159	0	4117	385	0
4	A	84	0	0	8	0
4	B	8	0	0	1	0
4	C	11	0	0	1	0
All	All	4769	0	4398	397	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including

hydrogens) of the entry. The overall clashscore for this entry is 44.

The worst 5 of 397 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:351:LEU:HD23	3:A:374:MET:HG2	1.34	1.06
3:A:780:LEU:HD11	3:A:790:GLU:HG3	1.40	1.02
3:A:614:ILE:HD11	3:A:760:LEU:HD12	1.49	0.95
3:A:779:MET:HA	3:A:789:LEU:HD12	1.50	0.93
3:A:677:ARG:HH12	3:A:746:ARG:HH12	1.17	0.91

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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
3	A	537/539 (100%)	419 (78%)	102 (19%)	16 (3%)	7 9

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	525	ALA
3	A	625	ASP
3	A	681	GLU
3	A	586	VAL
3	A	597	ALA

### 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
3	A	412/441 (93%)	315 (76%)	97 (24%)	1 1



5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	595	ARG
3	A	657	LEU
3	A	789	LEU
3	A	614	ILE
3	A	636	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	561	HIS
3	A	566	GLN
3	A	583	ASN
3	A	534	GLN
3	A	582	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
1	DOC	B	112	1,2	17,19,20	1.02	1 (5%)	20,26,29	0.91	1 (5%)

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
1	DOC	B	112	1,2	-	0/5/18/19	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	112	DOC	P-OP1	2.52	1.49	1.46

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	112	DOC	C2-N3-C4	3.14	120.12	115.57

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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.