



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 08:10 PM GMT

PDB ID : 4R47  
Title : Racemic crystal structure of a bimolecular DNA G-quadruplex (P21/n)  
Authors : Mandal, P.K.; Collie, G.W.; Kauffmann, B.; Huc, I.  
Deposited on : 2014-08-19  
Resolution : 1.85 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  
<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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : **FAILED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

# 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 1.85 Å.


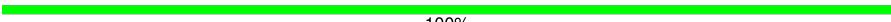
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1898 (1.86-1.86)

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.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	12	 92%                      8%
1	B	12	 100%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 669 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 DNA chain called 5'-D(\*GP\*GP\*GP\*GP\*TP\*TP\*TP\*TP\*GP\*GP\*GP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	12	Total	C	N	O	P	0	0	0
			253	120	48	74	11			
1	B	12	Total	C	N	O	P	0	0	0
			253	120	48	74	11			

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

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

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	2	Total	Mg	0	0
			2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	74	Total	O	0	3
			77	77		
4	B	77	Total	O	0	1
			78	78		

### 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 failed to run properly.

- Molecule 1: 5'-D(\*GP\*GP\*GP\*GP\*TP\*TP\*TP\*TP\*GP\*GP\*GP\*G)-3'

Chain A:  92% 8%



- Molecule 1: 5'-D(\*GP\*GP\*GP\*GP\*TP\*TP\*TP\*TP\*GP\*GP\*GP\*G)-3'

Chain B:  100%

There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

EDS failed to run properly - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21/n 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	28.50Å 47.30Å 46.80Å 90.00° 97.02° 90.00°	Depositor
Resolution (Å)	24.28 – 1.85	Depositor
% Data completeness (in resolution range)	95.8 (24.28-1.85)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.09 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.283 , 0.304	Depositor
Wilson B-factor (Å <sup>2</sup> )	49.2	Xtriage
Anisotropy	-0.970	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle =$ (Not available), $\langle L^2 \rangle =$ (Not available)	Xtriage
Outliers	103 of 9892 reflections (1.041%)	Xtriage
Total number of atoms	669	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *None*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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: K, MG

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	0.67	0/284	0.97	1/440 (0.2%)
1	B	0.70	0/284	0.93	0/440
All	All	0.68	0/568	0.95	1/880 (0.1%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1009	DG	O5'-P-OP1	-6.71	99.66	105.70

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	253	0	138	0	0
1	B	253	0	138	0	0
2	A	4	0	0	0	0
2	B	1	0	0	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	77	0	0	0	1
4	B	78	0	0	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	669	0	276	0	1

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 0.

There are no clashes within the asymmetric unit.

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:1261:HOH:O	4:B:2244:HOH:O[4_545]	2.18	0.02

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

### 5.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

### 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](#)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

## 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 failed to run properly - this section will therefore be empty.

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

EDS failed to run properly - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS failed to run properly - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS failed to run properly - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS failed to run properly - this section will therefore be empty.