



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 06:54 PM GMT

PDB ID : 1D2I  
Title : CRYSTAL STRUCTURE OF RESTRICTION ENDONUCLEASE BGLII  
COMPLEXED WITH DNA 16-MER  
Authors : Lukacs, C.M.; Kucera, R.; Schildkraut, I.; Aggarwal, A.K.  
Deposited on : 1999-09-23  
Resolution : 1.70 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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

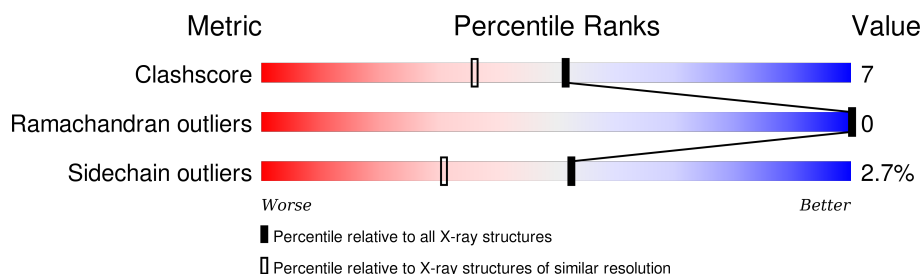
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	3585 (1.70-1.70)
Ramachandran outliers	100387	3527 (1.70-1.70)
Sidechain outliers	100360	3527 (1.70-1.70)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	C	16	 56% 38% 6%
1	D	16	 56% 44%
2	A	223	 85% 13% •
2	B	223	 86% 10% ••

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4785 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 DNA (5'-D(\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*GP\*AP\*TP\*CP\*TP\*AP\*TP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	16	Total	C	N	O	P	0	0	0
			325	159	57	94	15			
1	D	16	Total	C	N	O	P	0	0	0
			325	159	57	94	15			

- Molecule 2 is a protein called PROTEIN (RESTRICTION ENDONUCLEASE BGLII).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	223	Total	C	N	O	Se	0	0	0
			1817	1155	308	349	5			
2	B	218	Total	C	N	O	Se	0	0	0
			1773	1130	296	342	5			

- 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	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	188	Total	O	0	0
			188	188		
4	B	262	Total	O	0	0
			262	262		
4	C	50	Total	O	0	0
			50	50		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	43	Total	O	0	0
			43	43		

### 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.

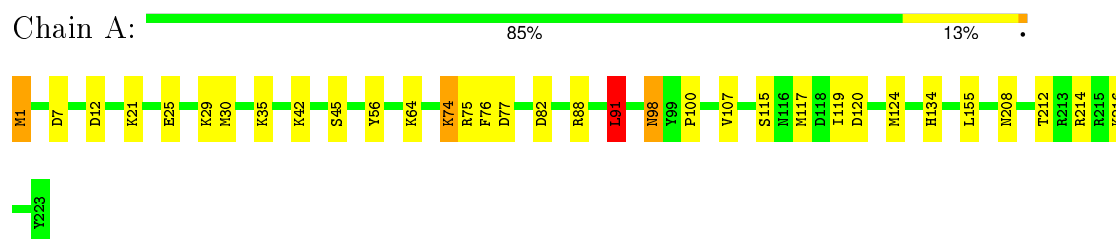
- Molecule 1: DNA (5'-D(\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*GP\*AP\*TP\*CP\*TP\*AP\*TP\*AP\*A)-3')



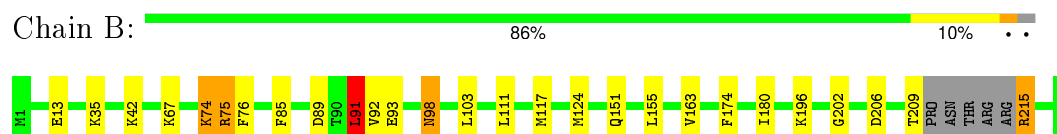
- Molecule 1: DNA (5'-D(\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*GP\*AP\*TP\*CP\*TP\*AP\*TP\*AP\*A)-3')



- Molecule 2: PROTEIN (RESTRICTION ENDONUCLEASE BGLII)



- Molecule 2: PROTEIN (RESTRICTION ENDONUCLEASE BGLII)



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.80 Å   101.70 Å   116.90 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	20.00 – 1.70	Depositor
% Data completeness (in resolution range)	99.9 (20.00-1.70)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.182 , 0.200	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4785	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	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: 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	C	1.08	0/364	1.24	4/560 (0.7%)
1	D	0.99	0/364	1.09	1/560 (0.2%)
2	A	0.78	1/1850 (0.1%)	0.91	2/2498 (0.1%)
2	B	0.84	0/1804	0.95	4/2434 (0.2%)
All	All	0.85	1/4382 (0.0%)	0.98	11/6052 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	2
1	D	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1	MSE	SE-CE	-5.11	1.65	1.95

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	10	DT	C1'-O4'-C4'	-7.27	102.83	110.10
2	A	91	LEU	CA-CB-CG	7.23	131.94	115.30
1	C	4	DT	C5'-C4'-C3'	-7.17	101.18	114.10
1	D	26	DT	C1'-O4'-C4'	-6.62	103.48	110.10
2	B	89	ASP	CB-CG-OD1	6.40	124.06	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	11	DC	Sidechain
1	C	13	DA	Sidechain
1	D	27	DC	Sidechain

## 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	C	325	0	185	3	0
1	D	325	0	185	6	0
2	A	1817	0	1815	27	0
2	B	1773	0	1768	23	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	188	0	0	1	0
4	B	262	0	0	2	0
4	C	50	0	0	2	0
4	D	43	0	0	1	0
All	All	4785	0	3953	56	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 7.

The worst 5 of 56 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)
2:A:42:LYS:NZ	4:A:730:HOH:O	1.79	1.10
2:A:119:ILE:HD11	2:A:124:MSE:HG3	1.28	1.07
4:C:132:HOH:O	2:A:35:LYS:HD2	1.71	0.89
2:B:215:ARG:HH11	2:B:215:ARG:HG2	1.34	0.89
2:A:74:LYS:HE2	2:A:74:LYS:H	1.46	0.80

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	
2	A	221/223 (99%)	215 (97%)	6 (3%)	0	100	100
2	B	214/223 (96%)	208 (97%)	6 (3%)	0	100	100
All	All	435/446 (98%)	423 (97%)	12 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	
2	A	206/201 (102%)	202 (98%)	4 (2%)	65	46
2	B	201/201 (100%)	194 (96%)	7 (4%)	43	20
All	All	407/402 (101%)	396 (97%)	11 (3%)	52	31

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	13	GLU
2	B	35	LYS
2	B	98	ASN
2	A	98	ASN
2	B	91	LEU

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

Mol	Chain	Res	Type
2	A	153	ASN
2	B	69	ASN
2	B	149	GLN
2	A	150	ASN
2	B	98	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](#)

Of 2 ligands modelled in this entry, 2 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 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.