



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 11:44 AM GMT

PDB ID : 1TX3  
Title : HINCII BOUND TO COGNATE DNA  
Authors : Horton, N.C.; Dorner, L.F.; Perona, J.J.  
Deposited on : 2004-07-01  
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>

---

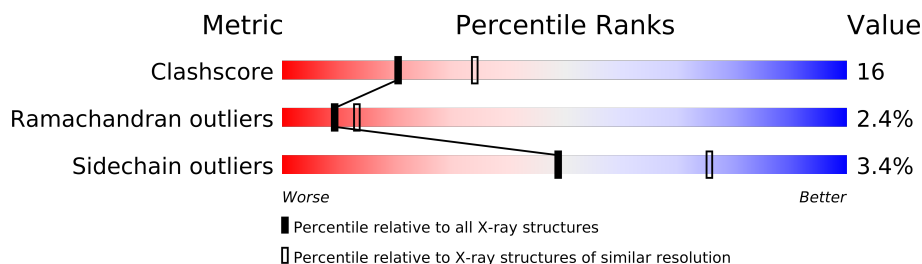
The following versions of software and data (see [references](#)) were used in the production of this report:

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	E	13	
1	F	13	
1	G	13	
1	H	13	
2	A	257	
2	B	257	
2	C	257	
2	D	257	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9904 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 5'-D(\*GP\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	13	Total	C	N	O	P	0	0	0
			265	125	52	76	12			
1	F	13	Total	C	N	O	P	0	0	0
			265	125	52	76	12			
1	G	13	Total	C	N	O	P	0	0	0
			265	125	52	76	12			
1	H	13	Total	C	N	O	P	0	0	0
			265	125	52	76	12			

- Molecule 2 is a protein called Type II restriction enzyme HindII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	249	Total	C	N	O	S	0	0	0
			2020	1315	325	374	6			
2	B	254	Total	C	N	O	S	0	0	0
			2073	1347	336	384	6			
2	C	255	Total	C	N	O	S	0	0	0
			2004	1304	324	370	6			
2	D	244	Total	C	N	O	S	0	0	0
			1970	1282	316	366	6			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	67	ASN	LYS	CONFLICT	UNP P44413
B	67	ASN	LYS	CONFLICT	UNP P44413
C	67	ASN	LYS	CONFLICT	UNP P44413
D	67	ASN	LYS	CONFLICT	UNP P44413

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Na 1	0	0
3	A	1	Total 1	Na 1	0	0
3	D	1	Total 1	Na 1	0	0
3	C	1	Total 1	Na 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	230	Total 230	O 230	0	0
4	B	234	Total 234	O 234	0	0
4	C	65	Total 65	O 65	0	0
4	D	102	Total 102	O 102	0	0
4	E	45	Total 45	O 45	0	0
4	F	40	Total 40	O 40	0	0
4	G	26	Total 26	O 26	0	0
4	H	31	Total 31	O 31	0	0

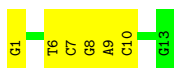
### 3 Residue-property plots

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: 5'-D(\*GP\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain E: 



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

Chain F: 



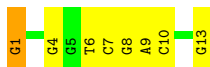
- Molecule 1: 5'-D(\*GP\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain G: 



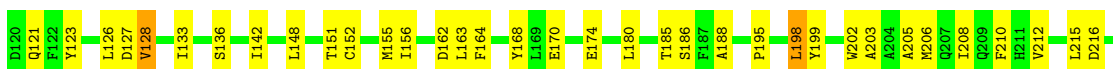
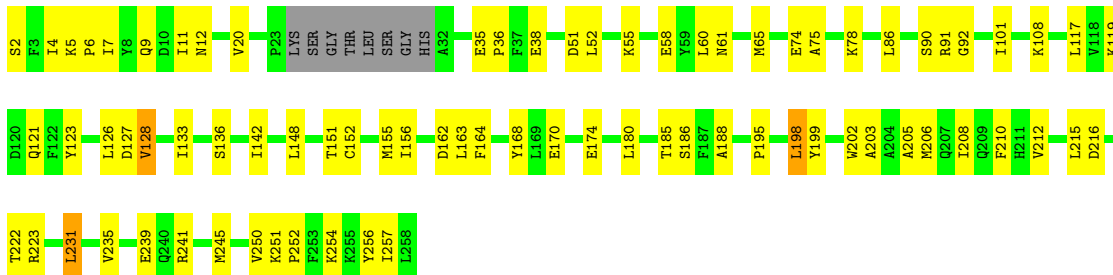
- Molecule 1: 5'-D(\*GP\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain H: 



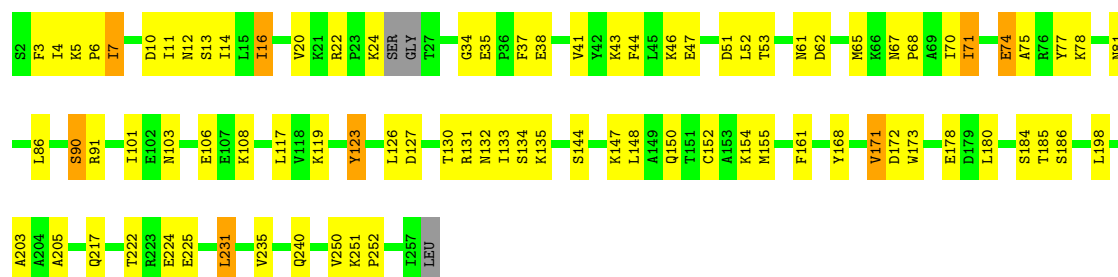
- Molecule 2: Type II restriction enzyme HindII

Chain A: 



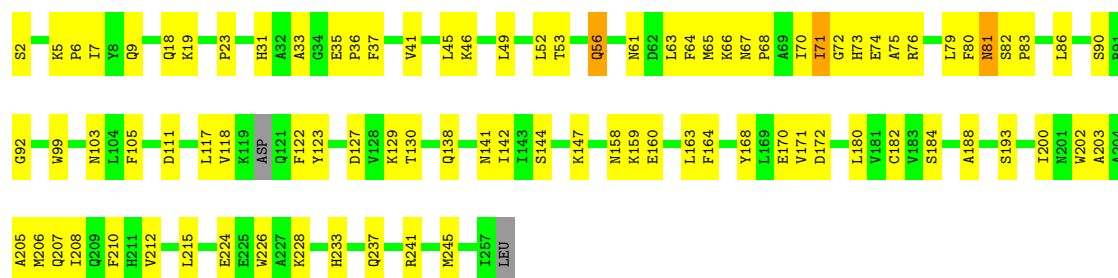
- Molecule 2: Type II restriction enzyme HindII

## Chain B:



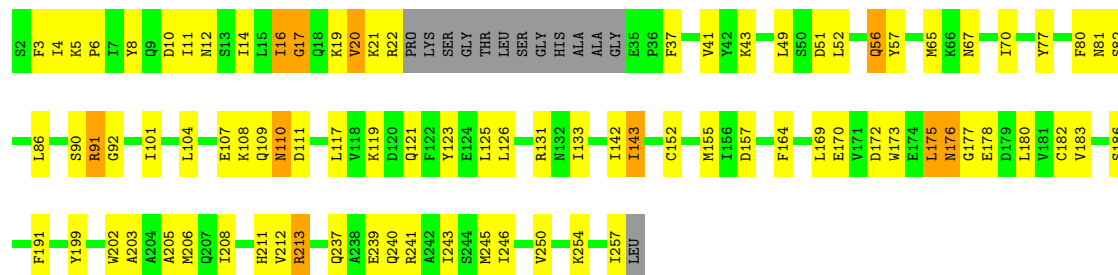
- Molecule 2: Type II restriction enzyme HindII

## Chain C:



- Molecule 2: Type II restriction enzyme HindII

## Chain D:



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.12Å 177.67Å 256.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50	Depositor
% Data completeness (in resolution range)	86.6 (20.00-2.50)	Depositor
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.204 , 0.271	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9904	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	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:  
NA

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	E	0.39	0/297	0.73	0/457
1	F	0.46	0/297	0.77	0/457
1	G	0.35	0/297	0.68	0/457
1	H	0.35	0/297	0.71	0/457
2	A	0.35	0/2066	0.53	0/2794
2	B	0.35	0/2119	0.53	0/2859
2	C	0.30	0/2051	0.45	0/2785
2	D	0.31	0/2014	0.48	0/2728
All	All	0.34	0/9438	0.54	0/12994

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	E	0	3
1	F	0	6
1	G	0	3
1	H	0	6
All	All	0	18

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 18 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	7	DC	Sidechain

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	E	8	DG	Sidechain
1	E	9	DA	Sidechain
1	F	1	DG	Sidechain
1	F	2	DC	Sidechain

## 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	E	265	0	146	4	0
1	F	265	0	146	6	0
1	G	265	0	146	13	0
1	H	265	0	146	4	0
2	A	2020	0	1979	61	0
2	B	2073	0	2061	71	0
2	C	2004	0	1911	67	1
2	D	1970	0	1894	65	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	230	0	0	13	1
4	B	234	0	0	9	0
4	C	65	0	0	1	0
4	D	102	0	0	8	0
4	E	45	0	0	2	0
4	F	40	0	0	0	0
4	G	26	0	0	1	0
4	H	31	0	0	3	0
All	All	9904	0	8429	276	2

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 16.

The worst 5 of 276 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:G:8:DG:H2"	1:G:9:DA:H5"	1.41	1.00
2:D:110:ASN:HB2	4:D:848:HOH:O	1.79	0.82
1:E:1:DG:H5"	4:H:592:HOH:O	1.82	0.78
2:C:61:ASN:HD21	2:C:105:PHE:HB2	1.48	0.78
2:D:86:LEU:HD12	2:D:90:SER:HB3	1.66	0.77

All (2) 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	Distance(Å)	Clash(Å)
2:C:81:ASN:OD1	2:C:81:ASN:OD1[7_657]	1.93	0.27
4:A:825:HOH:O	4:A:825:HOH:O[6_555]	2.16	0.04

## 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	245/257 (95%)	232 (95%)	13 (5%)	0	100	100
2	B	250/257 (97%)	228 (91%)	17 (7%)	5 (2%)	11	17
2	C	251/257 (98%)	215 (86%)	27 (11%)	9 (4%)	5	6
2	D	240/257 (93%)	210 (88%)	20 (8%)	10 (4%)	4	4
All	All	986/1028 (96%)	885 (90%)	77 (8%)	24 (2%)	9	13

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	34	GLY
2	C	56	GLN
2	C	72	GLY
2	C	73	HIS
2	C	74	GLU

### 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	215/229 (94%)	207 (96%)	8 (4%)	45	72
2	B	224/229 (98%)	214 (96%)	10 (4%)	38	63
2	C	205/229 (90%)	202 (98%)	3 (2%)	76	93
2	D	206/229 (90%)	198 (96%)	8 (4%)	43	70
All	All	850/916 (93%)	821 (97%)	29 (3%)	49	75

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	148	LEU
2	B	231	LEU
2	D	143	ILE
2	B	171	VAL
2	B	240	GLN

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

Mol	Chain	Res	Type
2	C	61	ASN
2	C	81	ASN
2	D	110	ASN
2	C	18	GLN
2	D	132	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 4 ligands modelled in this entry, 4 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.

## 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.