



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 02:40 pm GMT

PDB ID : 1KC6  
Title : HincII Bound to Cognate DNA  
Authors : Horton, N.C.; Dorner, L.F.; Perona, J.J.  
Deposited on : 2001-11-07  
Resolution : 2.60 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  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

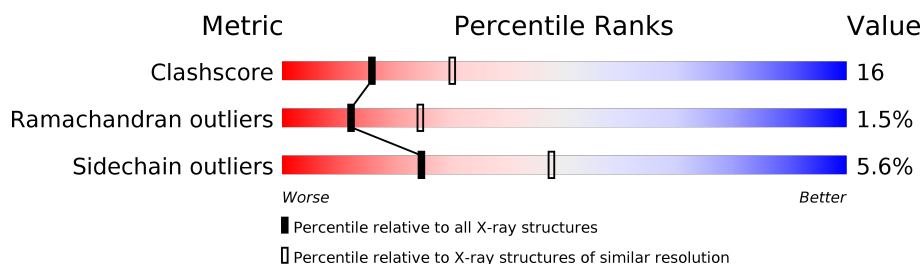
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)

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	E	12	
1	F	12	
1	G	12	
1	H	12	
2	A	257	
2	B	257	
2	C	257	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	257	<div><div></div><div>60%</div><div>29%</div><div>• 8%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9131 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(P\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	12	Total	C	N	O	P	0	0	0
			246	115	47	72	12			
1	F	12	Total	C	N	O	P	0	0	0
			246	115	47	72	12			
1	G	12	Total	C	N	O	P	0	0	0
			246	115	47	72	12			
1	H	12	Total	C	N	O	P	0	0	0
			246	115	47	72	12			

- Molecule 2 is a protein called TYPE II RESTRICTION ENZYME HINCII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	249	Total	C	N	O	S	0	0	0
			1993	1299	322	366	6			
2	B	247	Total	C	N	O	S	0	0	0
			1984	1295	320	363	6			
2	C	233	Total	C	N	O	S	0	0	0
			1794	1174	289	326	5			
2	D	236	Total	C	N	O	S	0	0	0
			1855	1216	297	336	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	130	THR	ARG	SEE REMARK 999	UNP P17743
A	173	TRP	SER	SEE REMARK 999	UNP P17743
B	130	THR	ARG	SEE REMARK 999	UNP P17743
B	173	TRP	SER	SEE REMARK 999	UNP P17743
C	130	THR	ARG	SEE REMARK 999	UNP P17743
C	173	TRP	SER	SEE REMARK 999	UNP P17743
D	130	THR	ARG	SEE REMARK 999	UNP P17743
D	173	TRP	SER	SEE REMARK 999	UNP P17743

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Na 1	0	0
3	E	1	Total 1	Na 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	145	Total 145	O 145	0	0
4	B	152	Total 152	O 152	0	0
4	C	46	Total 46	O 46	0	0
4	D	77	Total 77	O 77	0	0
4	E	26	Total 26	O 26	0	0
4	F	42	Total 42	O 42	0	0
4	G	16	Total 16	O 16	0	0
4	H	15	Total 15	O 15	0	0

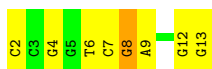
### 3 Residue-property plots [i](#)

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 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.

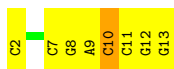
- Molecule 1: 5'-D(P\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain E: 



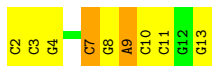
- Molecule 1: 5'-D(P\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain F: 

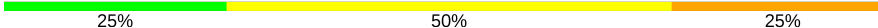


- Molecule 1: 5'-D(P\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain G: 



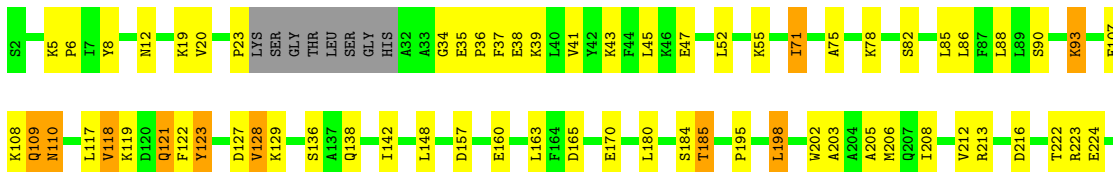
- Molecule 1: 5'-D(P\*CP\*CP\*GP\*GP\*TP\*CP\*GP\*AP\*CP\*CP\*GP\*G)-3'

Chain H: 



- Molecule 2: TYPE II RESTRICTION ENZYME HINCII

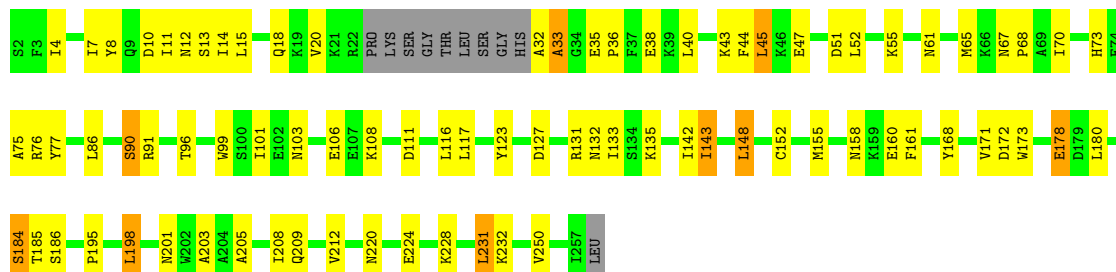
Chain A: 





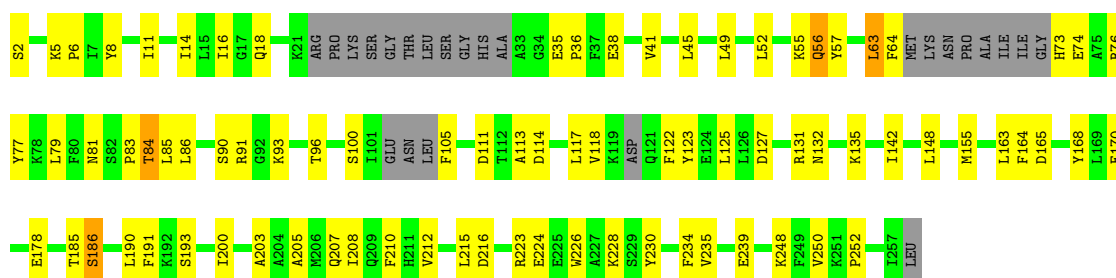
• Molecule 2: TYPE II RESTRICTION ENZYME HINCII

Chain B: 64% 28%



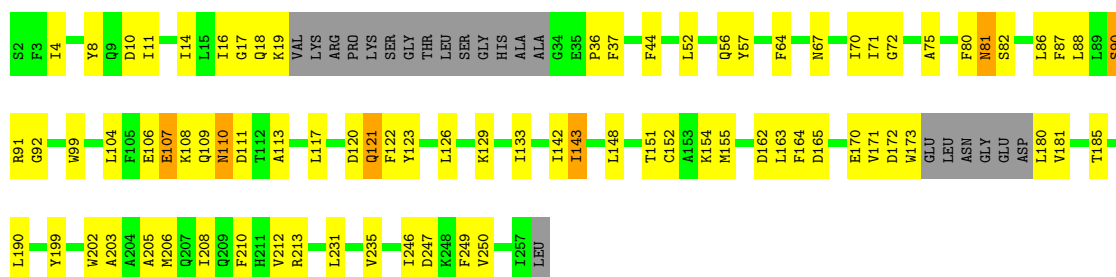
• Molecule 2: TYPE II RESTRICTION ENZYME HINCII

Chain C: 59% 30% 9%



• Molecule 2: TYPE II RESTRICTION ENZYME HINCII

Chain D: 60% 29% 8%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore 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 – 2.60	Depositor
% Data completeness (in resolution range)	87.2 (20.00-2.60)	Depositor
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.219 , 0.285	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9131	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	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	0.65	0/275	0.91	0/422
1	F	0.79	0/275	1.08	0/422
1	G	0.58	0/275	0.91	0/422
1	H	0.67	0/275	0.91	0/422
2	A	0.55	0/2038	0.67	2/2757 (0.1%)
2	B	0.53	0/2029	0.67	1/2744 (0.0%)
2	C	0.38	0/1834	0.51	0/2493
2	D	0.44	0/1898	0.59	1/2577 (0.0%)
All	All	0.51	0/8899	0.67	4/12259 (0.0%)

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	2
1	F	0	5
1	G	0	2
1	H	0	4
All	All	0	13

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	23	PRO	N-CA-CB	5.70	110.14	103.30
2	D	36	PRO	N-CA-CB	5.59	110.01	103.30
2	B	55	LYS	N-CA-C	-5.18	97.00	111.00
2	A	55	LYS	N-CA-C	-5.15	97.08	111.00

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	8	DG	Sidechain
1	E	9	DA	Sidechain
1	F	2	DC	Sidechain
1	F	7	DC	Sidechain
1	F	8	DG	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	E	246	0	134	15	0
1	F	246	0	134	5	0
1	G	246	0	134	11	0
1	H	246	0	134	17	0
2	A	1993	0	1930	49	0
2	B	1984	0	1936	64	0
2	C	1794	0	1658	61	0
2	D	1855	0	1750	65	0
3	A	1	0	0	0	0
3	E	1	0	0	0	0
4	A	145	0	0	4	0
4	B	152	0	0	6	0
4	C	46	0	0	0	0
4	D	77	0	0	7	0
4	E	26	0	0	4	0
4	F	42	0	0	0	0
4	G	16	0	0	5	0
4	H	15	0	0	2	0
All	All	9131	0	7810	267	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 16.

The worst 5 of 267 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:G:4:DG:H2'	4:G:166:HOH:O	1.36	1.25
1:H:8:DG:H2''	1:H:9:DA:H5''	1.37	1.06
1:H:3:DC:H2''	1:H:4:DG:H5''	1.51	0.92
1:G:13:DG:H2'	4:G:222:HOH:O	1.74	0.87
2:A:8:TYR:OH	2:A:185:THR:HG22	1.77	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	
2	A	245/257 (95%)	229 (94%)	13 (5%)	3 (1%)	15	32
2	B	243/257 (95%)	224 (92%)	17 (7%)	2 (1%)	22	44
2	C	223/257 (87%)	192 (86%)	26 (12%)	5 (2%)	8	14
2	D	230/257 (90%)	205 (89%)	21 (9%)	4 (2%)	11	21
All	All	941/1028 (92%)	850 (90%)	77 (8%)	14 (2%)	12	24

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	56	GLN
2	D	16	ILE
2	D	90	SER
2	B	33	ALA
2	C	100	SER

### 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	
2	A	207/229 (90%)	190 (92%)	17 (8%)	13	26
2	B	207/229 (90%)	196 (95%)	11 (5%)	26	50
2	C	174/229 (76%)	168 (97%)	6 (3%)	42	69
2	D	185/229 (81%)	176 (95%)	9 (5%)	29	54
All	All	773/916 (84%)	730 (94%)	43 (6%)	25	48

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	45	LEU
2	B	178	GLU
2	D	143	ILE
2	B	116	LEU
2	B	143	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	98	ASN
2	B	211	HIS
2	D	56	GLN
2	A	237	GLN
2	D	110	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

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

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.