



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 17, 2021 – 07:37 PM JST

PDB ID : 5Z7D  
Title : p204HINab-dsDNA complex structure  
Authors : Jin, T.; Jiang, J.; Xiao, T.S.  
Deposited on : 2018-01-28  
Resolution : 4.50 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

<https://www.wwpdb.org/validation/2017/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)	:	1.13
EDS	:	2.18
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.18

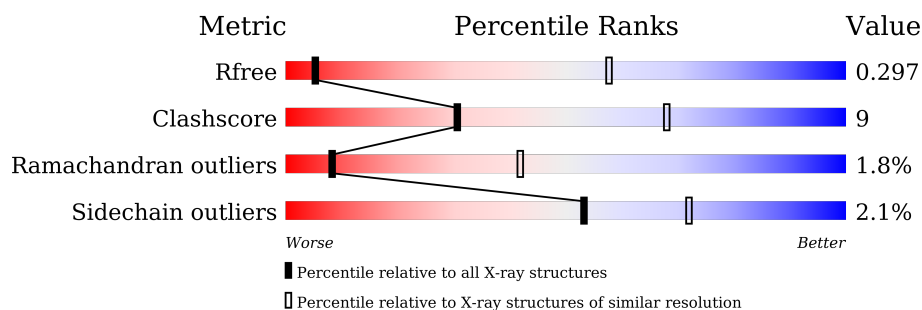
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.50 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))
$R_{free}$	130704	1055 (5.20-3.80)
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	412	
1	B	412	
1	C	412	
2	I	16	
2	K	16	
3	J	16	
3	L	16	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10624 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 protein called Interferon-activable protein 204.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	395	Total	C	N	O	S	0	0	0
			3087	1972	529	571	15			
1	B	397	Total	C	N	O	S	0	0	0
			3127	1993	534	584	16			
1	C	396	Total	C	N	O	S	0	0	0
			3098	1978	534	571	15			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	212	GLY	-	expression tag	UNP P0DOV2
A	213	SER	-	expression tag	UNP P0DOV2
A	214	VAL	-	expression tag	UNP P0DOV2
A	215	ASP	-	expression tag	UNP P0DOV2
A	620	ALA	-	expression tag	UNP P0DOV2
A	621	ALA	-	expression tag	UNP P0DOV2
A	622	ALA	-	expression tag	UNP P0DOV2
A	623	SER	-	expression tag	UNP P0DOV2
B	212	GLY	-	expression tag	UNP P0DOV2
B	213	SER	-	expression tag	UNP P0DOV2
B	214	VAL	-	expression tag	UNP P0DOV2
B	215	ASP	-	expression tag	UNP P0DOV2
B	620	ALA	-	expression tag	UNP P0DOV2
B	621	ALA	-	expression tag	UNP P0DOV2
B	622	ALA	-	expression tag	UNP P0DOV2
B	623	SER	-	expression tag	UNP P0DOV2
C	212	GLY	-	expression tag	UNP P0DOV2
C	213	SER	-	expression tag	UNP P0DOV2
C	214	VAL	-	expression tag	UNP P0DOV2
C	215	ASP	-	expression tag	UNP P0DOV2
C	620	ALA	-	expression tag	UNP P0DOV2
C	621	ALA	-	expression tag	UNP P0DOV2
C	622	ALA	-	expression tag	UNP P0DOV2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	623	SER	-	expression tag	UNP P0DOV2

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	16	Total	C	N	O	P	0	0	0
			331	156	69	90	16			
2	K	16	Total	C	N	O	P	0	0	0
			331	156	69	90	16			

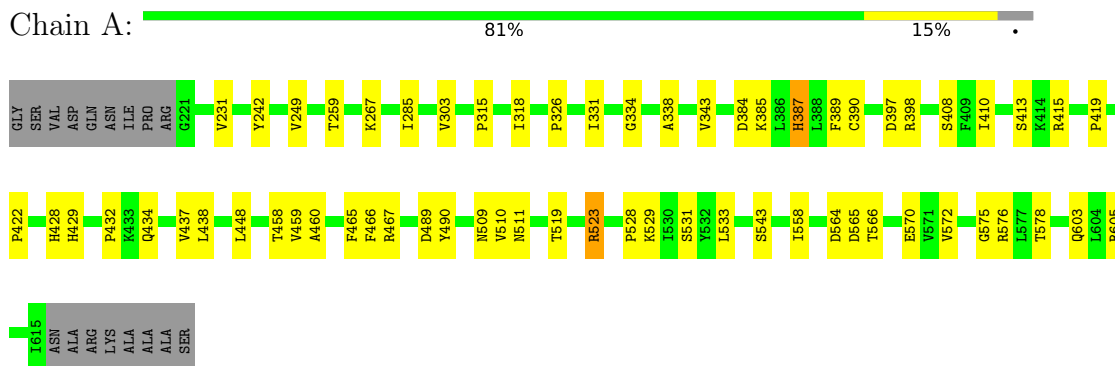
- Molecule 3 is a DNA chain called DNA (5'-D(P\*CP\*CP\*AP\*TP\*CP\*AP\*GP\*AP\*AP\*AP\*GP\*AP\*GP\*AP\*GP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	16	Total	C	N	O	P	0	0	0
			325	156	51	102	16			
3	L	16	Total	C	N	O	P	0	0	0
			325	156	51	102	16			

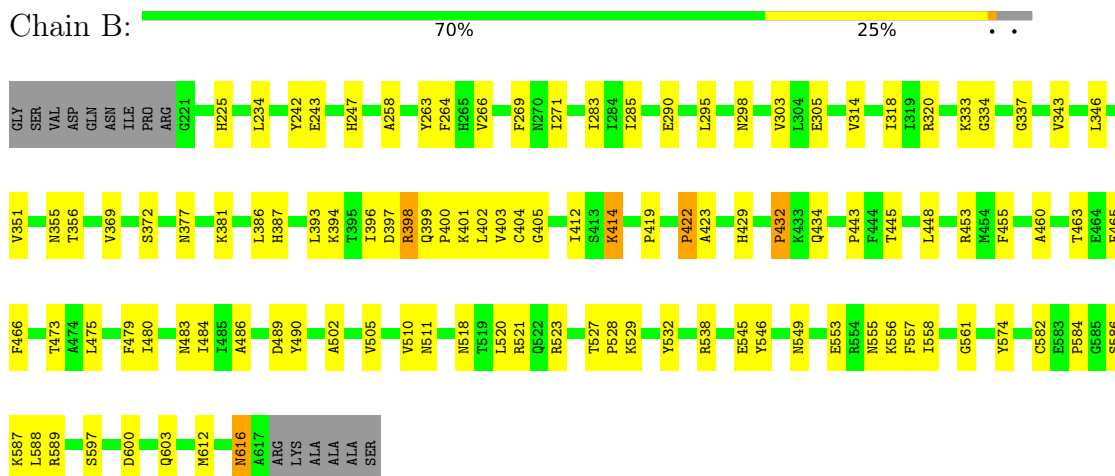
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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.

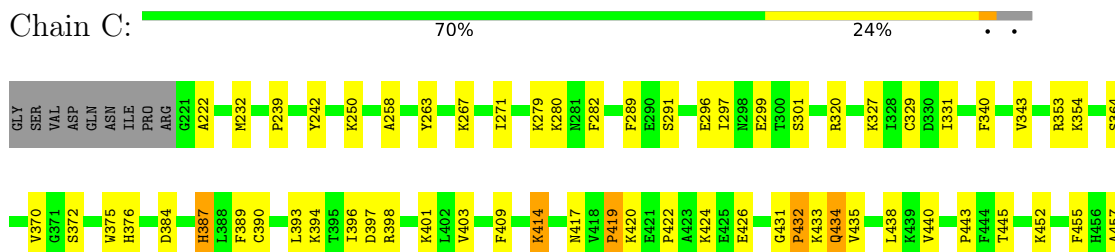
- Molecule 1: Interferon-activable protein 204

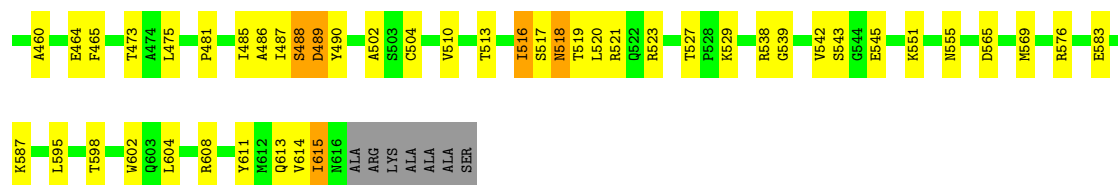


- Molecule 1: Interferon-activable protein 204



- Molecule 1: Interferon-activable protein 204





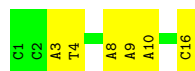
- Molecule 2: DNA (5'-D(P\*CP\*CP\*AP\*TP\*CP\*AP\*GP\*AP\*AP\*AP\*GP\*AP\*GP\*AP\*GP\*C)-3')

Chain I: 94% 6%



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

Chain K: 62% 38%



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

Chain J: 94% 6%



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

Chain L: 88% 12%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.39Å 101.39Å 783.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.26 – 4.50 47.26 – 4.50	Depositor EDS
% Data completeness (in resolution range)	91.9 (47.26-4.50) 91.9 (47.26-4.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 4.45Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.260 , 0.297 0.260 , 0.297	Depositor DCC
$R_{free}$ test set	709 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	112.1	Xtriage
Anisotropy	0.226	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.22 , 69.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.81	EDS
Total number of atoms	10624	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	162.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.29	0/3154	0.54	0/4266
1	B	0.27	0/3193	0.47	0/4317
1	C	0.28	0/3164	0.47	0/4277
2	I	0.54	0/373	0.79	0/573
2	K	0.50	0/373	0.78	0/573
3	J	0.48	0/361	1.00	0/555
3	L	0.49	0/361	1.00	0/555
All	All	0.32	0/10979	0.57	0/15116

There are no bond length outliers.

There are no bond angle outliers.

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	3087	0	3037	38	0
1	B	3127	0	3100	73	0
1	C	3098	0	3063	67	0
2	I	331	0	178	3	0
2	K	331	0	178	5	0
3	J	325	0	184	4	0
3	L	325	0	184	6	0
All	All	10624	0	9924	182	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 9.

The worst 5 of 182 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:C:434:GLN:HA	1:C:485:ILE:O	1.68	0.93
1:C:434:GLN:HA	1:C:486:ALA:HA	1.55	0.89
3:J:16:DG:H4'	3:L:1:DG:H5'	1.58	0.86
1:B:258:ALA:HA	1:B:263:TYR:HA	1.66	0.77
1:B:434:GLN:OE1	1:B:510:VAL:HG12	1.85	0.77

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	
1	A	393/412 (95%)	369 (94%)	21 (5%)	3 (1%)	19	60
1	B	395/412 (96%)	346 (88%)	44 (11%)	5 (1%)	12	48
1	C	394/412 (96%)	331 (84%)	50 (13%)	13 (3%)	4	30
All	All	1182/1236 (96%)	1046 (88%)	115 (10%)	21 (2%)	8	42

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	510	VAL
1	C	434	GLN
1	C	510	VAL
1	C	518	ASN
1	C	555	ASN

### 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	
1	A	334/366 (91%)	331 (99%)	3 (1%)	78	87
1	B	345/366 (94%)	337 (98%)	8 (2%)	50	70
1	C	336/366 (92%)	326 (97%)	10 (3%)	41	63
All	All	1015/1098 (92%)	994 (98%)	21 (2%)	53	72

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	414	LYS
1	C	489	ASP
1	C	615	ILE
1	C	523	ARG
1	C	432	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.