



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 12:37 am BST

PDB ID : 2I9T  
Title : Structure of NF-kB p65-p50 heterodimer bound to PRDII element of B-interferon promoter  
Authors : Escalante, C.R.; Shen, L.; Thanos, D.; Aggarwal, A.K.  
Deposited on : 2006-09-06  
Resolution : 2.80 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	:	2.11
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)	:	2.11

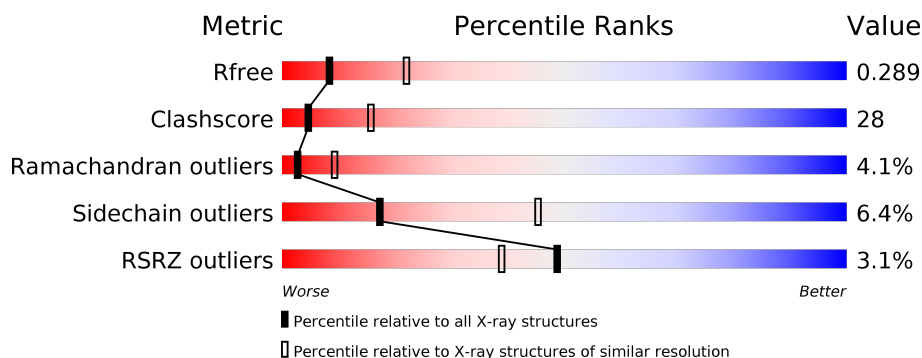
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 on 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  $\leq 5\%$ . 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.

Mol	Chain	Length	Quality of chain
1	C	17	<div> <div>53%</div> <div>41%</div> <div>6%</div> </div>
2	D	17	<div> <div>6%</div> <div>18%</div> <div>82%</div> </div>
3	A	279	<div> <div>%</div> <div>62%</div> <div>32%</div> <div>• •</div> </div>
4	B	313	<div> <div>5%</div> <div>55%</div> <div>40%</div> <div>5%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5226 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(\*AP\*GP\*TP\*GP\*GP\*GP\*AP\*AP\*AP\*TP\*TP\*CP\*CP\*TP\*CP\*TP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	17	Total	C	N	O	P	0	0	0
			348	167	64	101	16			

- Molecule 2 is a DNA chain called 5'-D(\*CP\*AP\*GP\*AP\*GP\*GP\*AP\*AP\*TP\*TP\*TP\*CP\*CP\*CP\*AP\*CP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	17	Total	C	N	O	P	0	0	0
			343	165	63	99	16			

- Molecule 3 is a protein called Transcription factor p65.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	275	Total	C	N	O	S	0	0	0
			2122	1324	393	394	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	GLY	-	CLONING ARTIFACT	UNP Q04207
A	14	SER	-	CLONING ARTIFACT	UNP Q04207
A	15	HIS	-	CLONING ARTIFACT	UNP Q04207
A	16	MET	-	CLONING ARTIFACT	UNP Q04207
A	17	HIS	-	CLONING ARTIFACT	UNP Q04207
A	18	ALA	-	CLONING ARTIFACT	UNP Q04207

- Molecule 4 is a protein called Nuclear factor NF-kappa-B p105 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	312	Total	C	N	O	S	0	0	0
			2392	1517	423	440	12			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	338	MET	-	INITIATING METHIONINE	UNP P25799

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	21	Total	O	0	0
			21	21		

### 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.

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

Chain C: 



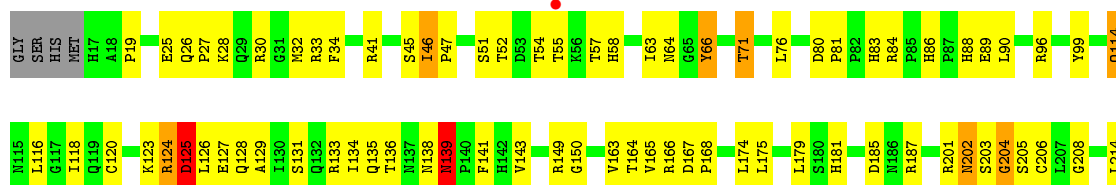
- Molecule 2: 5'-D(\*CP\*AP\*GP\*AP\*GP\*GP\*AP\*AP\*TP\*TP\*TP\*CP\*CP\*CP\*AP\*CP\*T)-3',

Chain D: 



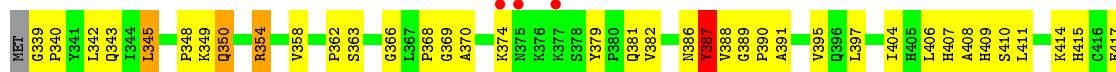
- Molecule 3: Transcription factor p65

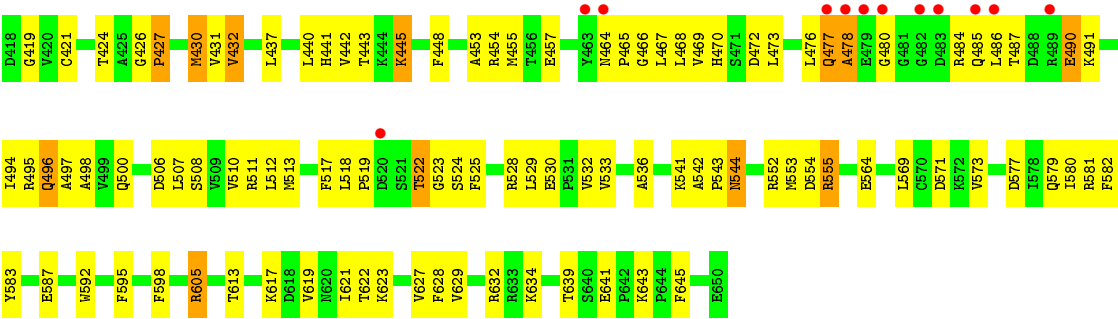
Chain A: 



- Molecule 4: Nuclear factor NF-kappa-B p105 subunit

Chain B: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.62Å 166.11Å 60.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.46 – 2.80 34.46 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.3 (34.46-2.80) 99.4 (34.46-2.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.27 (at 2.81Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.233 , 0.288 0.237 , 0.289	Depositor DCC
$R_{free}$ test set	2438 reflections (9.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.0	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 68.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	5226	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% 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

### 5.1 Standard geometry

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	0.56	0/390	1.40	4/601 (0.7%)
2	D	0.50	0/384	0.84	0/590
3	A	0.40	0/2176	0.65	0/2958
4	B	0.39	0/2444	0.63	0/3304
All	All	0.42	0/5394	0.75	4/7453 (0.1%)

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	5

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	710	DT	N1-C1'-C2'	5.87	123.75	112.60
1	C	712	DC	N1-C1'-C2'	5.40	122.87	112.60
1	C	711	DT	N1-C1'-C2'	5.26	122.59	112.60
1	C	709	DA	N9-C1'-C2'	5.08	122.26	112.60

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	702	DG	Sidechain
1	C	703	DT	Sidechain
1	C	709	DA	Sidechain
1	C	714	DT	Sidechain
1	C	715	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	348	0	194	27	0
2	D	343	0	193	31	0
3	A	2122	0	2011	101	0
4	B	2392	0	2344	130	0
5	B	21	0	0	0	3
All	All	5226	0	4742	280	3

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 28.

The worst 5 of 280 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)
4:B:470:HIS:HB3	4:B:473:LEU:HB2	1.30	1.08
2:D:721:DA:H2''	2:D:722:DG:H5''	1.39	1.04
2:D:732:DA:H2''	2:D:733:DC:H5''	1.43	0.98
4:B:342:LEU:HD12	4:B:532:VAL:HG11	1.47	0.96
2:D:732:DA:C2'	2:D:733:DC:H5''	1.96	0.95

All (3) 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)
5:B:808:HOH:O	5:B:822:HOH:O[2_555]	0.58	1.62
5:B:810:HOH:O	5:B:815:HOH:O[2_555]	1.53	0.67
5:B:816:HOH:O	5:B:826:HOH:O[2_555]	1.98	0.22

## 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	
3	A	273/279 (98%)	232 (85%)	29 (11%)	12 (4%)	2	8
4	B	310/313 (99%)	261 (84%)	37 (12%)	12 (4%)	3	10
All	All	583/592 (98%)	493 (85%)	66 (11%)	24 (4%)	3	9

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	125	ASP
4	B	374	LYS
4	B	387	TYR
4	B	478	ALA
3	A	134	ILE

### 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	
3	A	223/247 (90%)	211 (95%)	12 (5%)	22	53
4	B	249/269 (93%)	231 (93%)	18 (7%)	14	38
All	All	472/516 (92%)	442 (94%)	30 (6%)	17	45

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	354	ARG
4	B	417	GLU
4	B	605	ARG
4	B	387	TYR
4	B	430	MET

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

Mol	Chain	Res	Type
3	A	241	GLN
3	A	287	GLN
4	B	496	GLN
3	A	220	GLN
4	B	500	GLN

### 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](#)

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 [i](#)

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	C	17/17 (100%)	-0.34	0 <span>100</span> <span>100</span>	32, 55, 73, 78	0
2	D	17/17 (100%)	-0.05	1 (5%) <span>22</span> <span>14</span>	27, 45, 86, 102	0
3	A	275/279 (98%)	-0.24	3 (1%) <span>80</span> <span>75</span>	22, 45, 85, 103	0
4	B	312/313 (99%)	-0.12	15 (4%) <span>30</span> <span>21</span>	22, 48, 102, 104	0
All	All	621/626 (99%)	-0.18	19 (3%) <span>49</span> <span>39</span>	22, 47, 97, 104	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	478	ALA	6.7
4	B	483	ASP	5.8
4	B	482	GLY	4.4
3	A	245	HIS	4.3
2	D	734	DT	4.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.