



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 22, 2024 – 11:35 PM EDT

PDB ID : 6N1Z  
Title : Importin-9 bound to H2A-H2B  
Authors : Tomchick, D.R.; Chook, Y.M.; Padavannil, A.  
Deposited on : 2018-11-12  
Resolution : 2.70 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

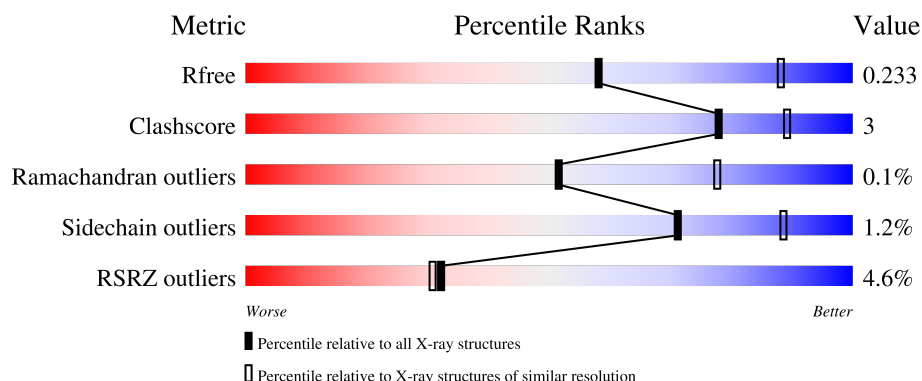
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 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  $\leq 5\%$ . 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	A	1041	 3% 86% 7% 8%
1	D	1041	 3% 84% 7% 9%
2	B	130	 9% 62% 5% 34%
2	E	130	 3% 62% • 35%
3	C	123	 6% 67% 7% • 25%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	123	 A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: a red segment at the beginning labeled '10%', a large green segment labeled '74%', a small yellow segment labeled '5%', and a grey segment at the end labeled '21%'.

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 36201 atoms, of which 18040 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 Importin-9.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	961	Total	C	H	N	O	S	0	1	0
			15105	4818	7575	1251	1415	46			
1	D	949	Total	C	H	N	O	S	0	1	0
			14945	4766	7503	1242	1388	46			

- Molecule 2 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	86	Total	C	H	N	O	0	0	0
			1381	421	708	134	118			
2	E	85	Total	C	H	N	O	0	0	0
			1367	417	701	133	116			

- Molecule 3 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	92	Total	C	H	N	O	S	0	0	0
			1471	456	746	132	135	2			
3	F	97	Total	C	H	N	O	S	0	0	0
			1577	484	807	143	141	2			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MET	-	initiating methionine	UNP P02281
C	30	THR	SER	engineered mutation	UNP P02281
F	1	MET	-	initiating methionine	UNP P02281
F	30	THR	SER	engineered mutation	UNP P02281

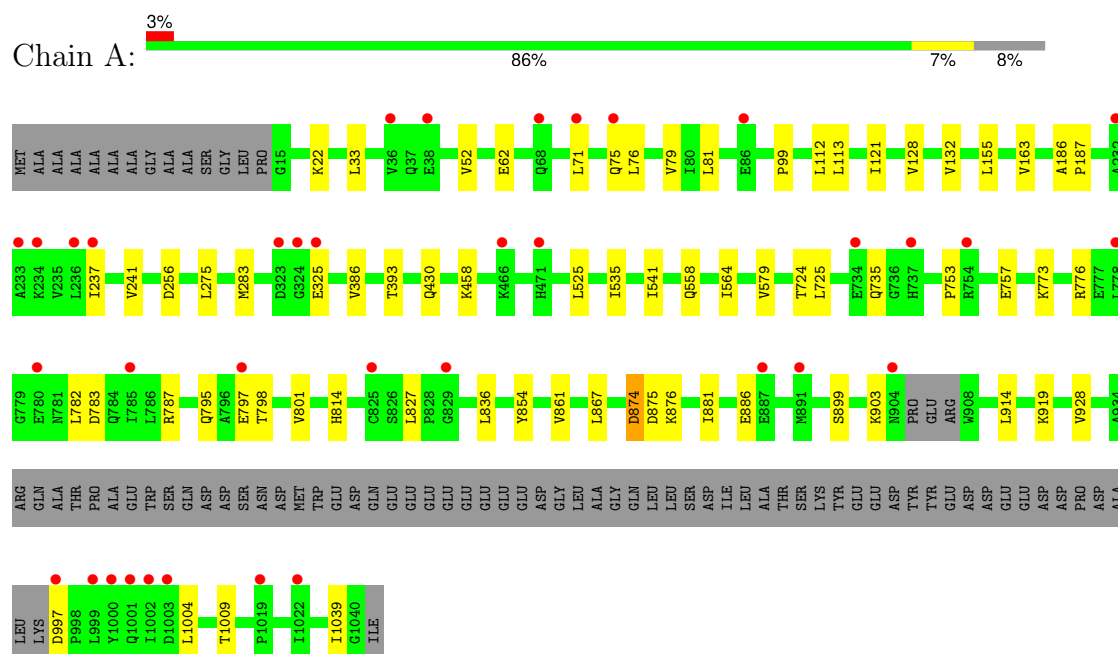
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	163	Total 163	O 163	0	0
4	B	2	Total 2	O 2	0	0
4	C	5	Total 5	O 5	0	0
4	D	179	Total 179	O 179	0	0
4	E	3	Total 3	O 3	0	0
4	F	3	Total 3	O 3	0	0

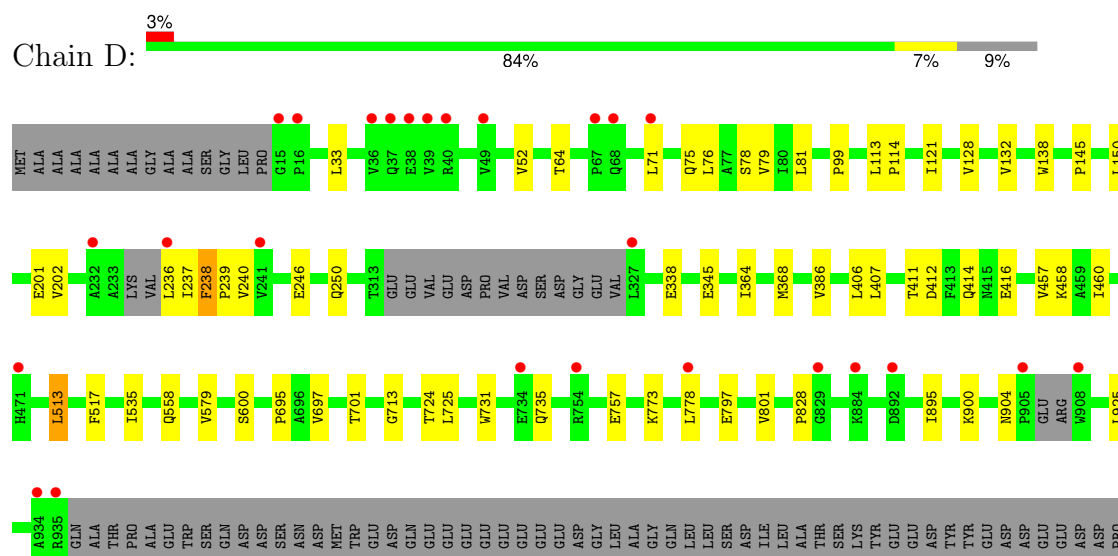
### 3 Residue-property plots [i](#)

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 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.

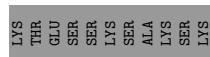
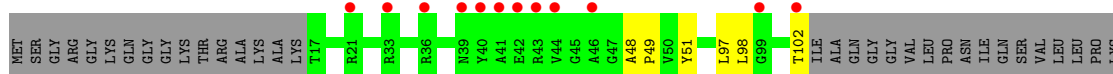
#### • Molecule 1: Importin-9



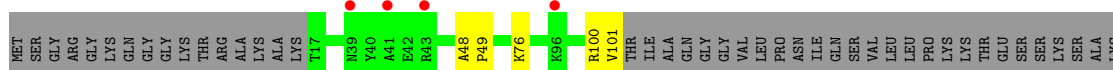
#### • Molecule 1: Importin-9



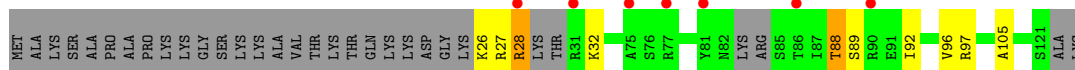
- Molecule 2: Histone H2A



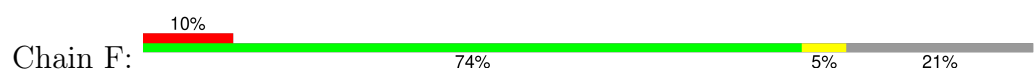
- Molecule 2: Histone H2A



- Molecule 3: Histone H2B 1.1



- Molecule 3: Histone H2B 1.1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.42Å 223.29Å 131.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.44 – 2.70 45.44 – 2.70	Depositor EDS
% Data completeness (in resolution range)	90.5 (45.44-2.70) 90.6 (45.44-2.70)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 2.69Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.209 , 0.238 0.209 , 0.233	Depositor DCC
$R_{free}$ test set	1978 reflections (2.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.6	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 39.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.004 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	36201	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.34 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.2394e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.25	0/7685	0.41	0/10454
1	D	0.25	0/7595	0.41	0/10326
2	B	0.26	0/681	0.45	0/918
2	E	0.26	0/674	0.41	0/908
3	C	0.24	0/734	0.40	0/984
3	F	0.25	0/781	0.42	0/1047
All	All	0.25	0/18150	0.41	0/24637

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	7530	7575	7574	41	0
1	D	7442	7503	7503	48	0
2	B	673	708	708	3	0
2	E	666	701	701	9	0
3	C	725	746	746	6	0
3	F	770	807	807	8	0
4	A	163	0	0	0	0
4	B	2	0	0	0	0
4	C	5	0	0	0	0
4	D	179	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	3	0	0	0	0
4	F	3	0	0	0	0
All	All	18161	18040	18039	104	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 3.

The worst 5 of 104 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:E:101:VAL:HG11	3:F:66:ASP:HB3	1.40	1.01
1:D:999:LEU:HD23	1:D:1002:ILE:HD12	1.61	0.82
1:D:701:THR:HG21	1:D:713:GLY:HA3	1.72	0.70
1:D:71:LEU:CD2	1:D:121:ILE:HD13	2.23	0.68
1:D:237:ILE:O	1:D:240:VAL:HG22	1.94	0.67

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	956/1041 (92%)	932 (98%)	23 (2%)	1 (0%)	51	78
1	D	940/1041 (90%)	912 (97%)	26 (3%)	2 (0%)	47	73
2	B	84/130 (65%)	81 (96%)	3 (4%)	0	100	100
2	E	83/130 (64%)	81 (98%)	2 (2%)	0	100	100
3	C	86/123 (70%)	84 (98%)	2 (2%)	0	100	100
3	F	95/123 (77%)	90 (95%)	5 (5%)	0	100	100
All	All	2244/2588 (87%)	2180 (97%)	61 (3%)	3 (0%)	51	78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	875	ASP
1	D	238	PHE
1	D	828	PRO

### 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	
1	A	832/894 (93%)	824 (99%)	8 (1%)	76	91
1	D	821/894 (92%)	812 (99%)	9 (1%)	73	90
2	B	68/102 (67%)	66 (97%)	2 (3%)	42	71
2	E	67/102 (66%)	67 (100%)	0	100	100
3	C	79/103 (77%)	76 (96%)	3 (4%)	33	62
3	F	84/103 (82%)	83 (99%)	1 (1%)	71	88
All	All	1951/2198 (89%)	1928 (99%)	23 (1%)	71	88

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	412	ASP
1	D	460	ILE
1	D	458	LYS
1	D	513	LEU
1	A	874	ASP

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

### 5.3.3 RNA ⓘ

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 [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	A	961/1041 (92%)	0.15	36 (3%) 41 41	12, 47, 98, 153	0
1	D	949/1041 (91%)	0.11	33 (3%) 44 44	12, 42, 99, 151	0
2	B	86/130 (66%)	0.71	12 (13%) 2 1	39, 62, 114, 162	0
2	E	85/130 (65%)	0.24	4 (4%) 31 30	29, 49, 100, 149	0
3	C	92/123 (74%)	0.53	7 (7%) 13 12	36, 70, 126, 154	0
3	F	97/123 (78%)	0.55	12 (12%) 4 3	24, 61, 137, 178	0
All	All	2270/2588 (87%)	0.19	104 (4%) 32 31	12, 47, 105, 178	0

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1002	ILE	9.1
1	D	1000	TYR	8.3
2	B	42	GLU	6.8
1	D	327	LEU	6.5
1	D	38	GLU	5.6

### 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 monosaccharides 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.