



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 01:49 am BST

PDB ID : 5XBK  
Title : Crystal structure of human Importin4  
Authors : Song, J.J.; Yoon, J.  
Deposited on : 2017-03-20  
Resolution : 3.22 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

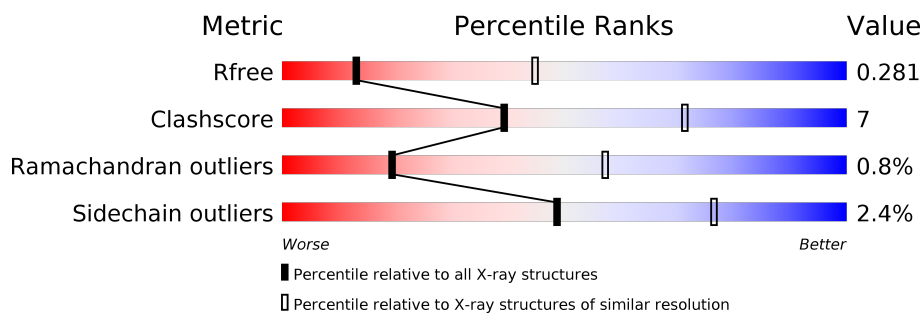
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.22 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	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	416	
1	B	416	
1	C	416	
1	D	416	
2	E	18	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11162 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 Importin-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	369	Total	C	N	O	S	0	0	0
			2785	1778	484	505	18			
1	A	366	Total	C	N	O	S	0	0	0
			2762	1769	477	499	17			
1	B	373	Total	C	N	O	S	0	0	0
			2817	1799	489	511	18			
1	C	368	Total	C	N	O	S	0	0	0
			2776	1770	483	505	18			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	666	GLY	-	expression tag	UNP Q8TEX9
D	667	SER	-	expression tag	UNP Q8TEX9
A	666	GLY	-	expression tag	UNP Q8TEX9
A	667	SER	-	expression tag	UNP Q8TEX9
B	666	GLY	-	expression tag	UNP Q8TEX9
B	667	SER	-	expression tag	UNP Q8TEX9
C	666	GLY	-	expression tag	UNP Q8TEX9
C	667	SER	-	expression tag	UNP Q8TEX9

- Molecule 2 is a protein called histone H3.

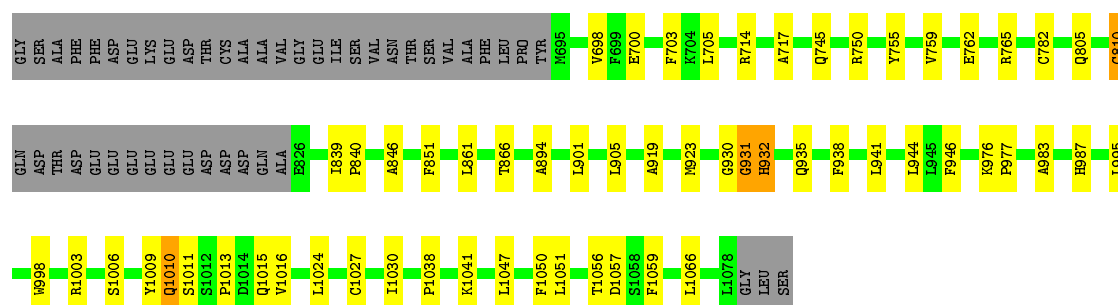
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	4	Total	C	N	O	0	0	0
			22	13	5	4			

### 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 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.

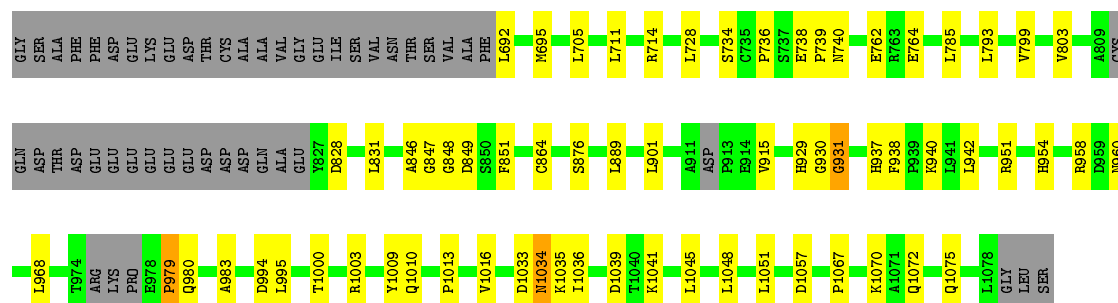
#### • Molecule 1: Importin-4

Chain D: 



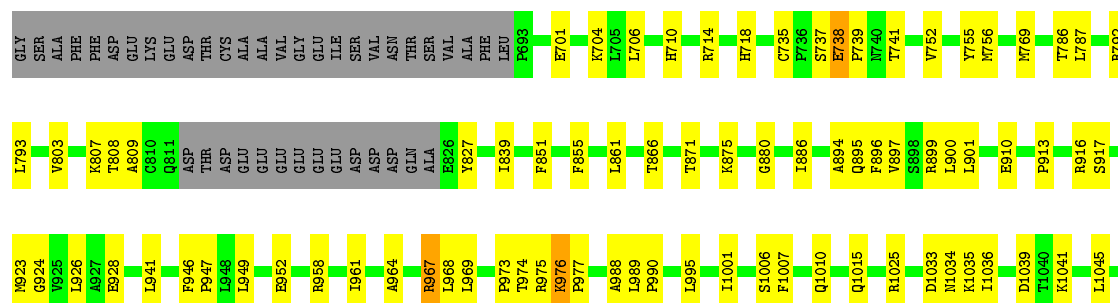
#### • Molecule 1: Importin-4

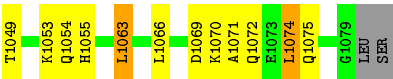
Chain A: 



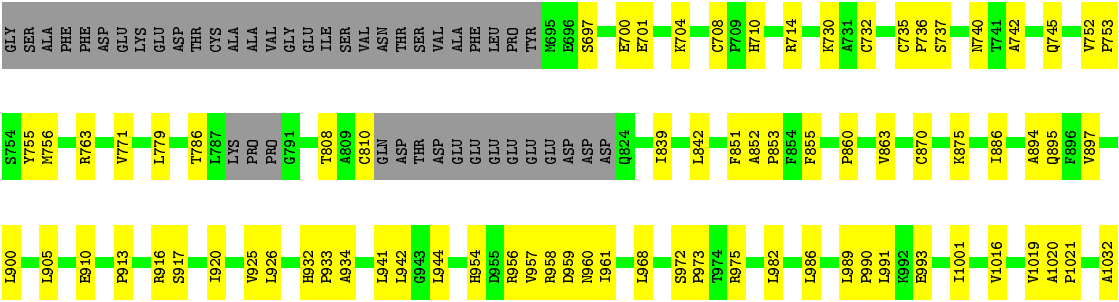
#### • Molecule 1: Importin-4

Chain B: 





• Molecule 1: Importin-4



• Molecule 2: histone H3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.02Å 108.54Å 124.86Å 90.00° 103.41° 90.00°	Depositor
Resolution (Å)	49.98 – 3.22 49.98 – 3.22	Depositor EDS
% Data completeness (in resolution range)	98.1 (49.98-3.22) 98.1 (49.98-3.22)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.57 (at 3.25Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, $R_{free}$	0.218 , 0.281 0.218 , 0.281	Depositor DCC
$R_{free}$ test set	2672 reflections (9.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	90.8	Xtriage
Anisotropy	0.302	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 33.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.074 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11162	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% 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	0.27	0/2813	0.43	0/3818
1	B	0.28	0/2871	0.43	0/3898
1	C	0.29	0/2825	0.45	0/3833
1	D	0.28	0/2837	0.44	0/3852
2	E	0.25	0/21	0.33	0/25
All	All	0.28	0/11367	0.44	0/15426

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	2762	0	2844	33	1
1	B	2817	0	2894	56	0
1	C	2776	0	2851	52	1
1	D	2785	0	2866	34	0
2	E	22	0	23	1	0
All	All	11162	0	11478	168	1

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 7.

The worst 5 of 168 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:B:1025:ARG:HA	1:B:1066:LEU:HD21	1.71	0.71
1:C:917:SER:OG	1:C:960:ASN:ND2	2.22	0.71
1:A:714:ARG:HH22	1:A:764:GLU:HG3	1.57	0.70
1:B:894:ALA:HB3	1:C:895:GLN:HG3	1.75	0.67
1:C:958:ARG:HA	1:C:961:ILE:HD12	1.75	0.67

All (1) 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)
1:A:1009:TYR:OH	1:C:1039:ASP:OD1[1_565]	2.19	0.01

## 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	358/416 (86%)	325 (91%)	30 (8%)	3 (1%)	19	57
1	B	369/416 (89%)	347 (94%)	19 (5%)	3 (1%)	19	57
1	C	362/416 (87%)	335 (92%)	24 (7%)	3 (1%)	19	57
1	D	365/416 (88%)	341 (93%)	21 (6%)	3 (1%)	19	57
2	E	2/18 (11%)	1 (50%)	1 (50%)	0	100	100
All	All	1456/1682 (87%)	1349 (93%)	95 (6%)	12 (1%)	19	57

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	932	HIS
1	D	931	GLY
1	A	930	GLY
1	B	809	ALA
1	A	979	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	292/334 (87%)	283 (97%)	9 (3%)	40	71
1	B	298/334 (89%)	290 (97%)	8 (3%)	44	74
1	C	293/334 (88%)	290 (99%)	3 (1%)	76	89
1	D	295/334 (88%)	287 (97%)	8 (3%)	44	74
2	E	1/13 (8%)	1 (100%)	0	100	100
All	All	1179/1349 (87%)	1151 (98%)	28 (2%)	49	76

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	938	PHE
1	A	1010	GLN
1	C	732	CYS
1	A	940	LYS
1	A	951	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	960	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 [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 ⓘ

### 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.