



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:00 pm BST

PDB ID : 2OT8  
Title : Karyopherin Beta2/Transportin-hnRNPM NLS Complex  
Authors : Cansizoglu, A.E.; Chook, Y.M.  
Deposited on : 2007-02-07  
Resolution : 3.10 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

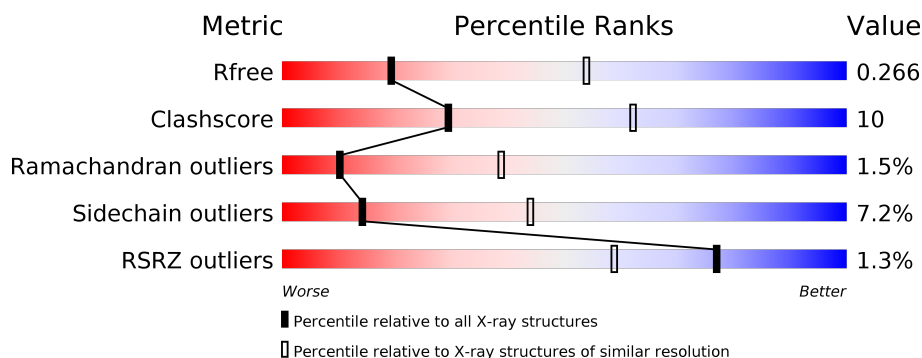
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	A	852	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 20%, green 74%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>74%</span> <span>20%</span> <span>• •</span> </div> </div>
1	B	852	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 1%, yellow 22%, green 71%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>71%</span> <span>22%</span> <span>• 5%</span> </div> </div>
2	C	30	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, orange 1%, yellow 13%, green 43%, grey 43%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>43%</span> <span>13%</span> <span>43%</span> </div> </div>
2	D	30	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, orange 0%, yellow 10%, green 53%, grey 33%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span></span> <span>53%</span> <span>10%</span> <span>• 33%</span> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12802 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 Transportin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	824	Total	C	N	O	S	0	0	0
			6310	4040	1052	1167	51			
1	B	813	Total	C	N	O	S	0	0	0
			6196	3963	1038	1145	50			

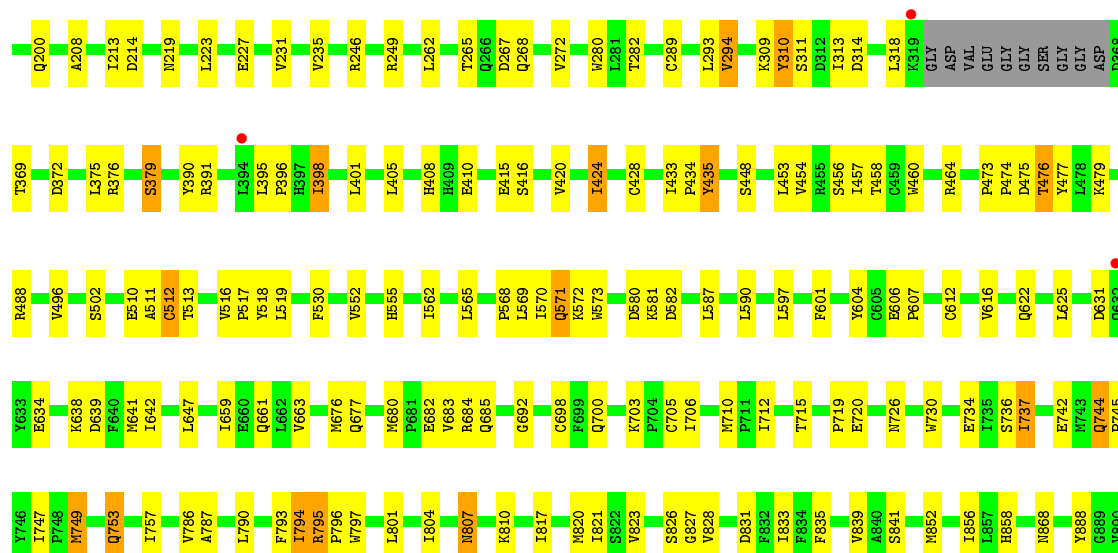
There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	324	GLY	-	linker	UNP Q92973
A	325	GLY	-	linker	UNP Q92973
A	326	SER	-	linker	UNP Q92973
A	327	GLY	-	linker	UNP Q92973
A	328	GLY	-	linker	UNP Q92973
B	324	GLY	-	linker	UNP Q92973
B	325	GLY	-	linker	UNP Q92973
B	326	SER	-	linker	UNP Q92973
B	327	GLY	-	linker	UNP Q92973
B	328	GLY	-	linker	UNP Q92973

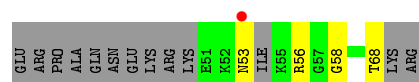
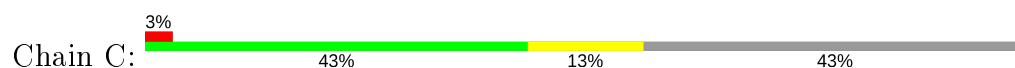
- Molecule 2 is a protein called Heterogeneous nuclear ribonucleoprotein M.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	17	Total	C	N	O	0	0	0
			135	82	27	26			
2	D	20	Total	C	N	O	0	0	0
			161	100	32	29			





• Molecule 2: Heterogeneous nuclear ribonucleoprotein M



• Molecule 2: Heterogeneous nuclear ribonucleoprotein M



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	153.20Å 155.00Å 141.53Å 90.00° 92.57° 90.00°	Depositor
Resolution (Å)	50.00 – 3.10 48.46 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.00-3.10) 96.1 (48.46-3.00)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 3.01Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.255 , 0.290 0.228 , 0.266	Depositor DCC
$R_{free}$ test set	3234 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	97.4	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 57.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l 0.000 for -k,-h,-l 0.008 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12802	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	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 20.65 % 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 8.2968e-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	0.42	0/6441	0.58	0/8775
1	B	0.42	2/6326 (0.0%)	0.56	1/8624 (0.0%)
2	C	0.50	0/137	0.57	0/182
2	D	0.50	0/164	0.62	0/218
All	All	0.42	2/13068 (0.0%)	0.57	1/17799 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	227	GLU	C-N	-5.57	1.21	1.34
1	B	570	ILE	C-N	5.27	1.46	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	25	PRO	N-CA-CB	5.74	110.19	103.30

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	6310	0	6131	115	0
1	B	6196	0	5960	124	0
2	C	135	0	119	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	161	0	157	3	0
All	All	12802	0	12367	240	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 10.

The worst 5 of 240 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:642:ILE:HD11	1:B:680:MET:HE3	1.38	1.04
1:A:746:TYR:HA	1:A:749:MET:HE3	1.61	0.83
1:B:821:ILE:HG12	1:B:828:VAL:HG21	1.63	0.80
1:A:112:ILE:O	1:A:116:THR:HG23	1.82	0.80
1:B:833:ILE:H	1:B:833:ILE:HD12	1.46	0.78

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	814/852 (96%)	743 (91%)	62 (8%)	9 (1%)	14	46
1	B	803/852 (94%)	721 (90%)	69 (9%)	13 (2%)	9	37
2	C	13/30 (43%)	10 (77%)	2 (15%)	1 (8%)	1	5
2	D	18/30 (60%)	11 (61%)	5 (28%)	2 (11%)	0	2
All	All	1648/1764 (93%)	1485 (90%)	138 (8%)	25 (2%)	10	39

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	84	PRO

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	162	ILE
1	B	24	SER
1	B	512	CYS
2	D	55	LYS

### 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	673/762 (88%)	624 (93%)	49 (7%)	14	43
1	B	652/762 (86%)	606 (93%)	46 (7%)	14	44
2	C	13/26 (50%)	10 (77%)	3 (23%)	1	3
2	D	16/26 (62%)	16 (100%)	0	100	100
All	All	1354/1576 (86%)	1256 (93%)	98 (7%)	14	44

5 of 98 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	808	GLU
1	B	131	PRO
1	B	795	ARG
1	A	844	ASN
1	B	74	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	868	ASN
1	B	97	ASN
1	B	858	HIS
1	B	74	ASN
1	B	173	ASN

### 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	A	824/852 (96%)	-0.07	6 (0%) 87 75	49, 86, 136, 177	0
1	B	813/852 (95%)	0.04	14 (1%) 70 49	47, 87, 168, 200	0
2	C	17/30 (56%)	0.22	1 (5%) 22 10	66, 113, 147, 153	0
2	D	20/30 (66%)	0.17	0 100 100	67, 124, 165, 169	0
All	All	1674/1764 (94%)	-0.01	21 (1%) 77 59	47, 87, 156, 200	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	41	ASN	5.0
1	B	24	SER	4.3
1	B	27	THR	4.2
1	B	83	PHE	3.8
1	B	82	ASN	3.6

### 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 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.