



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2023 – 01:13 AM EDT

PDB ID : 7MNI  
Title : Crystal structure of the N-terminal domain of NUP88 in complex with NUP98 C-terminal Autoproteolytic Domain  
Authors : Bley, C.J.; Nie, S.; Mobbs, G.W.; Petrovic, S.; Gres, A.T.; Liu, X.; Mukherjee, S.; Harvey, S.; Huber, F.M.; Lin, D.H.; Brown, B.; Tang, A.W.; Rundlet, E.J.; Correia, A.R.; Chen, S.; Regmi, S.G.; Stevens, T.A.; Jette, C.A.; Dasso, M.; Patke, A.; Palazzo, A.F.; Kossiakoff, A.A.; Hoelz, A.  
Deposited on : 2021-05-01  
Resolution : 2.00 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  
Xtrriage (Phenix) : 1.13  
EDS : 2.35.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Rfmac : 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) : 2.35.1

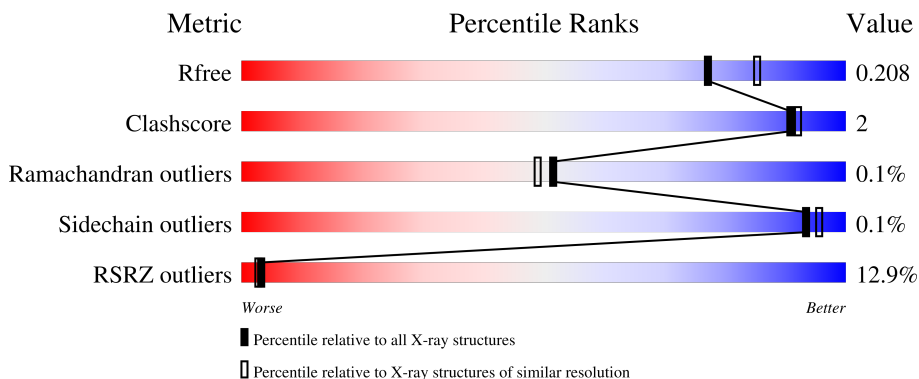
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	494	 10% 80% 17%
1	C	494	 10% 80% 17%
2	B	153	 20% 95% 7%
2	D	153	 10% 92% 7%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 18544 atoms, of which 9102 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 Nuclear pore complex protein Nup88.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	409	6599	2134	3319	540	583	23	0	14	0
1	C	412	6716	2174	3374	550	597	21	0	18	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q99567
C	0	SER	-	expression tag	UNP Q99567

- Molecule 2 is a protein called Nuclear pore complex protein Nup98.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	B	150	2395	760	1198	205	228	4	0	1	0
2	D	150	2416	765	1211	208	228	4	0	2	0

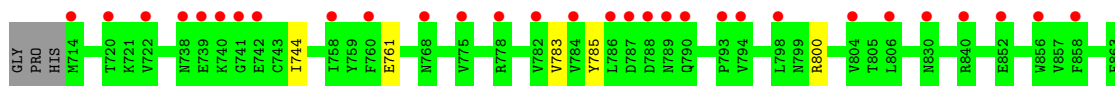
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	711	GLY	-	expression tag	UNP P52948-2
B	712	PRO	-	expression tag	UNP P52948-2
B	713	HIS	-	expression tag	UNP P52948-2
B	714	MET	-	expression tag	UNP P52948-2
D	711	GLY	-	expression tag	UNP P52948-2
D	712	PRO	-	expression tag	UNP P52948-2
D	713	HIS	-	expression tag	UNP P52948-2
D	714	MET	-	expression tag	UNP P52948-2

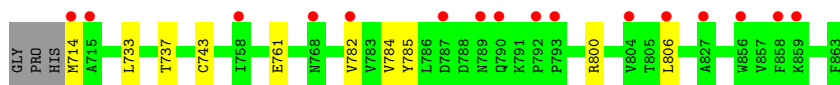
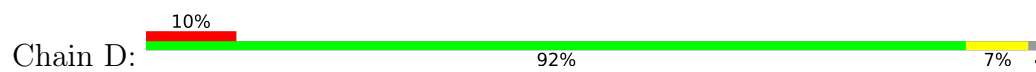
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	173	Total 174	O 174	0	1
3	B	27	Total 27	O 27	0	0
3	C	161	Total 161	O 161	0	0
3	D	56	Total 56	O 56	0	0





- Molecule 2: Nuclear pore complex protein Nup98



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.75Å 72.08Å 73.04Å 92.53° 103.69° 100.39°	Depositor
Resolution (Å)	20.00 – 2.00 20.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.7 (20.00-2.00) 97.7 (20.00-2.00)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.59 (at 2.01Å)	Xtrriage
Refinement program	PHENIX 1.17.1	Depositor
R, $R_{free}$	0.175 , 0.208 0.175 , 0.208	Depositor DCC
$R_{free}$ test set	3640 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.7	Xtrriage
Anisotropy	0.182	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 46.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.010 for -h,-l,-k	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	18544	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.68% 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	0.25	0/3407	0.45	0/4641
1	C	0.25	0/3469	0.45	0/4731
2	B	0.24	0/1223	0.43	0/1653
2	D	0.25	0/1234	0.43	0/1667
All	All	0.25	0/9333	0.45	0/12692

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	3280	3319	3315	10	0
1	C	3342	3374	3361	12	0
2	B	1197	1198	1198	3	0
2	D	1205	1211	1211	8	0
3	A	174	0	0	0	0
3	B	27	0	0	0	0
3	C	161	0	0	1	0
3	D	56	0	0	1	0
All	All	9442	9102	9085	31	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 2.



The worst 5 of 31 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:D:782[A]:VAL:HG13	2:D:806:LEU:HD21	1.68	0.73
1:C:61[B]:VAL:HG22	1:C:70:LEU:HD23	1.77	0.66
2:D:782[A]:VAL:CG1	2:D:806:LEU:HD21	2.28	0.63
1:A:393:VAL:HG13	1:A:407[A]:CYS:SG	2.40	0.61
1:A:61[A]:VAL:HG12	1:A:70:LEU:HD23	1.83	0.60

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	409/494 (83%)	402 (98%)	7 (2%)	0	100	100
1	C	417/494 (84%)	412 (99%)	5 (1%)	0	100	100
2	B	149/153 (97%)	144 (97%)	4 (3%)	1 (1%)	22	16
2	D	150/153 (98%)	148 (99%)	2 (1%)	0	100	100
All	All	1125/1294 (87%)	1106 (98%)	18 (2%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	785	TYR

### 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	373/432 (86%)	372 (100%)	1 (0%)	92	95
1	C	380/432 (88%)	380 (100%)	0	100	100
2	B	132/133 (99%)	132 (100%)	0	100	100
2	D	133/133 (100%)	133 (100%)	0	100	100
All	All	1018/1130 (90%)	1017 (100%)	1 (0%)	93	95

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	309	TYR

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

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	409/494 (82%)	0.55	50 (12%) 4 3	26, 45, 100, 140	0
1	C	412/494 (83%)	0.51	49 (11%) 4 4	26, 45, 95, 152	0
2	B	150/153 (98%)	0.96	30 (20%) 1 0	44, 71, 114, 149	0
2	D	150/153 (98%)	0.37	16 (10%) 6 5	31, 51, 90, 125	0
All	All	1121/1294 (86%)	0.57	145 (12%) 3 3	26, 49, 101, 152	0

The worst 5 of 145 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	458	LEU	8.6
1	C	359	LEU	7.8
1	C	257	ASN	7.4
1	A	258	GLY	7.0
1	C	167	SER	6.7

### 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.