



wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 04:05 am BST

PDB ID : 3I4R
Title : Nup107(aa658-925)/Nup133(aa517-1156) complex, H.sapiens
Authors : Whittle, J.R.R.; Schwartz, T.U.
Deposited on : 2009-07-02
Resolution : 3.53 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

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

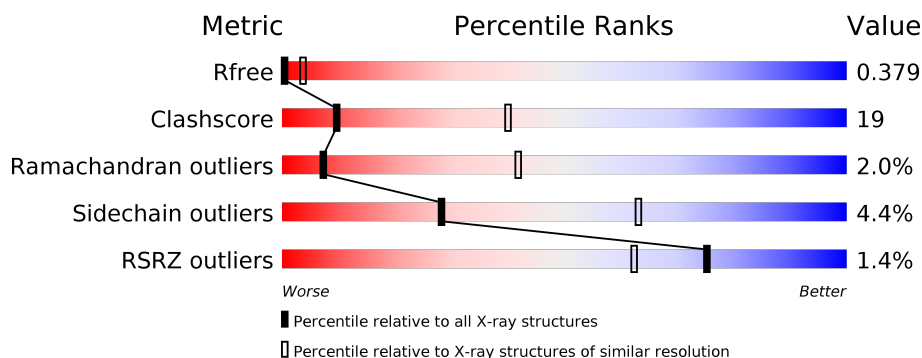
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.53 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	1028 (3.60-3.48)
Clashscore	141614	1109 (3.60-3.48)
Ramachandran outliers	138981	1073 (3.60-3.48)
Sidechain outliers	138945	1074 (3.60-3.48)
RSRZ outliers	127900	1079 (3.62-3.46)

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 ≥ 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	277	
2	B	644	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5161 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 Nuclear pore complex protein Nup107.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	247	Total	C	N	O	S	0	0	0
			1993	1282	343	355	13			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	656	GLY	-	EXPRESSION TAG	UNP P57740
A	657	SER	-	EXPRESSION TAG	UNP P57740
A	926	SER	-	EXPRESSION TAG	UNP P57740
A	927	GLY	-	EXPRESSION TAG	UNP P57740
A	928	ARG	-	EXPRESSION TAG	UNP P57740
A	929	ILE	-	EXPRESSION TAG	UNP P57740
A	930	VAL	-	EXPRESSION TAG	UNP P57740
A	931	THR	-	EXPRESSION TAG	UNP P57740
A	932	ASP	-	EXPRESSION TAG	UNP P57740

- Molecule 2 is a protein called Nuclear pore complex protein Nup133.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	528	Total	C	N	O	S	0	0	0
			3168	1934	580	648	6			

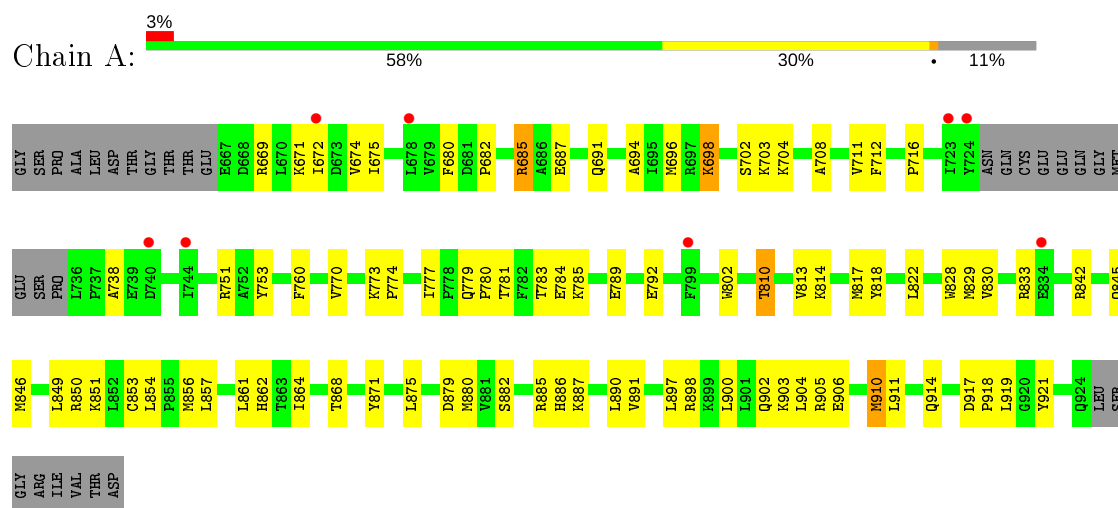
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	513	GLY	-	EXPRESSION TAG	UNP Q8WUM0
B	514	SER	-	EXPRESSION TAG	UNP Q8WUM0
B	515	HIS	-	EXPRESSION TAG	UNP Q8WUM0
B	516	MET	-	EXPRESSION TAG	UNP Q8WUM0

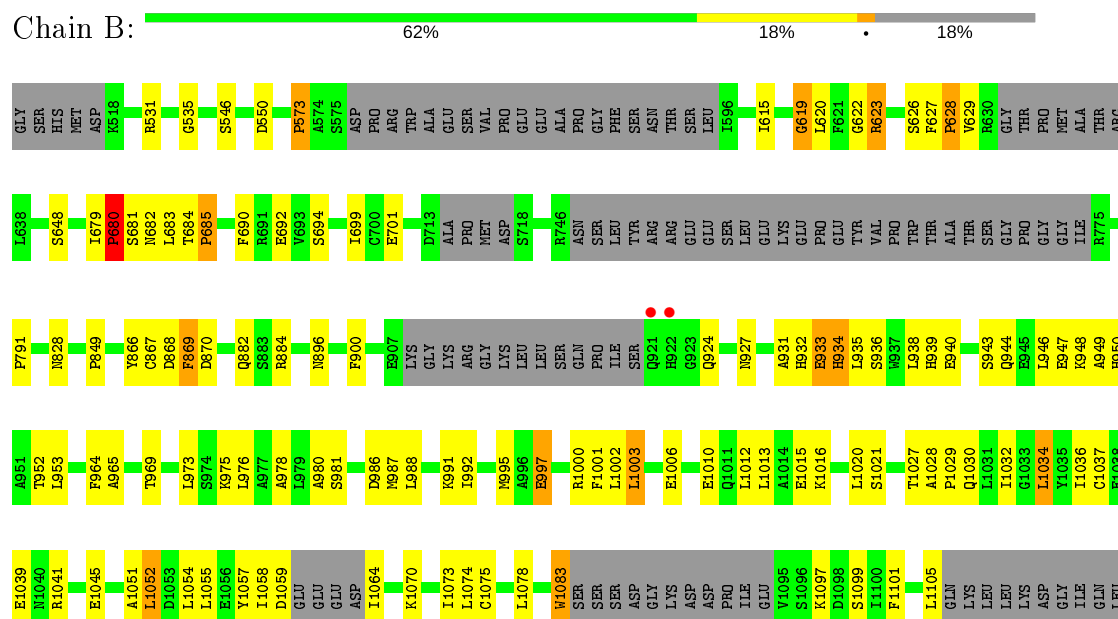
3 Residue-property plots

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 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: Nuclear pore complex protein Nup107



• Molecule 2: Nuclear pore complex protein Nup133



SER	GLU	TYR	L1119	P1120	E1121	V1122	K1123	D1124	L1126	Q1127	A1128	D1129	GLN	LEU	GLY	SER	LEU	LYS	SER	ASN	PRO	TYR	PHE	E1141	F1142	V1143	A1146	E1149	Y1150	Y1151	G1154	Q1155	I1156
-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	115.55Å 133.05Å 176.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.63 – 3.53 37.63 – 3.50	Depositor EDS
% Data completeness (in resolution range)	68.9 (37.63-3.53) 85.4 (37.63-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.4 _4	Depositor
R, R_{free}	0.315 , 0.370 0.340 , 0.379	Depositor DCC
R_{free} test set	2968 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å ²)	129.7	Xtriage
Anisotropy	0.632	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 192.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	5161	wwPDB-VP
Average B, all atoms (Å ²)	174.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.24	0/2037	0.41	0/2749
2	B	0.26	0/3179	0.53	6/4354 (0.1%)
All	All	0.25	0/5216	0.48	6/7103 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	B	685	PRO	N-CA-CB	8.52	113.52	103.30
2	B	680	PRO	N-CA-CB	7.01	111.71	103.30
2	B	573	PRO	N-CA-CB	6.78	111.44	103.30
2	B	849	PRO	N-CA-CB	6.23	110.78	103.30
2	B	791	PRO	N-CA-CB	6.18	110.71	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	1993	0	1981	71	0
2	B	3168	0	2224	117	0
All	All	5161	0	4205	182	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 19.

The worst 5 of 182 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:B:1032:ILE:HD11	2:B:1054:LEU:HB2	1.26	1.11
2:B:1052:LEU:HD21	2:B:1073:ILE:HD13	1.50	0.94
2:B:1016:LYS:HE2	2:B:1039:GLU:HG3	1.53	0.91
2:B:683:LEU:N	2:B:684:THR:HA	1.90	0.85
2:B:1120:PRO:HD3	2:B:1155:GLN:HE22	1.41	0.85

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	243/277 (88%)	235 (97%)	8 (3%)	0	100	100
2	B	508/644 (79%)	437 (86%)	56 (11%)	15 (3%)	4	33
All	All	751/921 (82%)	672 (90%)	64 (8%)	15 (2%)	7	42

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	619	GLY
2	B	623	ARG
2	B	627	PHE
2	B	680	PRO
2	B	685	PRO

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	211/246 (86%)	206 (98%)	5 (2%)	49	76
2	B	174/577 (30%)	162 (93%)	12 (7%)	15	48
All	All	385/823 (47%)	368 (96%)	17 (4%)	28	63

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	694	SER
2	B	934	HIS
2	B	1052	LEU
2	B	648	SER
2	B	1064	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	944	GLN
2	B	1155	GLN
2	B	958	ASN
1	A	902	GLN
2	B	950	HIS

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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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, 95th 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 ²)	Q<0.9
1	A	247/277 (89%)	-0.07	8 (3%)	47 35	82, 170, 288, 312	0
2	B	528/644 (81%)	-0.62	3 (0%)	89 81	75, 164, 260, 386	0
All	All	775/921 (84%)	-0.45	11 (1%)	75 62	75, 165, 272, 386	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	740	ASP	4.2
1	A	724	TYR	3.7
2	B	1154	GLY	3.5
1	A	723	ILE	3.1
1	A	744	ILE	2.9

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.