



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 25, 2017 – 11:46 PM EDT

PDB ID : 3JRO  
Title : NUP84-NUP145C-SEC13 edge element of the NPC lattice  
Authors : Brohawn, S.G.; Schwartz, T.U.  
Deposited on : unknown  
Resolution : 4.00 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

<http://wwpdb.org/validation/2016/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) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

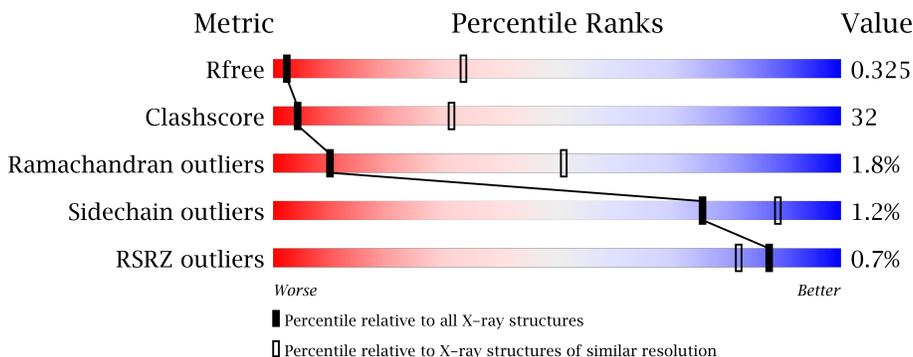
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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))
$R_{free}$	100719	1088 (4.40-3.60)
Clashscore	112137	1187 (4.40-3.60)
Ramachandran outliers	110173	1139 (4.40-3.60)
Sidechain outliers	110143	1126 (4.40-3.60)
RSRZ outliers	101464	1099 (4.40-3.60)

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. 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	753	
2	C	426	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8671 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 Fusion Protein of Protein Transport Protein SEC13 and Nucleoporin NUP145.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	701	5623	3598	943	1068	9	5	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1100	GLY	-	LINKER	UNP P49687
A	1101	GLY	-	LINKER	UNP P49687
A	1102	GLY	-	LINKER	UNP P49687
A	1103	GLY	-	LINKER	UNP P49687
A	1104	SER	-	LINKER	UNP P49687
A	1105	GLY	-	LINKER	UNP P49687
A	1106	GLY	-	LINKER	UNP P49687
A	1107	GLY	-	LINKER	UNP P49687
A	1108	GLY	-	LINKER	UNP P49687

- Molecule 2 is a protein called Nucleoporin NUP84.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
2	C	379	3048	1949	504	584	4	7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	EXPRESSION TAG	UNP P52891
C	0	SER	-	EXPRESSION TAG	UNP P52891



K421	R344	A270	W201	V132	V62
G424	V345	I271	L202	M133	I63
	L346	P272	M205	E134	S64
	M347	M273	I206	R135	S65
	I351	V276	S207	P136	K66
	L355	S280	I208	LYS	D67
	P356	D281	C209	ASN	W68
	I359	W282	M210	VAL	A72
	H360	E283	I211	T141	F73
	V363	S284	G214	S142	R74
	V370	D285	I215	K143	H75
LYS		L286	Q216	W144	H76
GLY		H287	E217	L485	
THR		I288	Y218	M146	E79
GLU		H289	L219	S147	L80
ALA		I290	M220	G151	L81
SER		M291	V221	G152	L82
ASN		Q292	V222	L153	V83
ASP		I293	I223	R85	F84
ILE		L294	I227	S155	N86
ILE		Q295		C156	L89
D381		I298	F231	D157	
R387		E299	H232	L158	N94
I388		M300	T233	D159	E95
V389		Y301	Q234	F160	R102
T390		E304	Q236	P161	G103
H391		M305	G236	L162	D170
L392		N306	I237	R163	V171
A393		Q307	H238	E164	L104
I394		V308	K239	V168	F105
G395		E312	E240	L169	E106
L396		L312	S241	D170	K107
D397		L313	S242	V171	K108
I398		L314	H243	K172	L109
I399		I314	R244	F173	M110
N400		L315	R245	K174	K114
P401		P316	T246	D177	Q115
V404		L317	S249		L116
E405		P318	L250	F180	Y117
V407		H320	S251	F181	Q118
D408		A321	Q252	K182	I119
		L322	Q253	L116	W120
		V331	A254	I184	I121
		A332	G255	Y185	V122
		S333	L256	F186	M123
		R334	E260	L187	L126
		H335	R261	I188	K127
		P336	R262		E128
		S337	A262		M129
		E338	I263		T130
		S339	S264		A200
		E340	S265		
		R341	Y266		
		P342	L267		
		L343	S268		
			G269		

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	170.47Å 170.47Å 270.73Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.89 – 4.00 49.89 – 4.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.89-4.00) 99.9 (49.89-4.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.77 (at 4.00Å)	Xtrriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.282 , 0.329 0.273 , 0.325	Depositor DCC
$R_{free}$ test set	1585 reflections (7.84%)	DCC
Wilson B-factor (Å <sup>2</sup> )	133.7	Xtrriage
Anisotropy	0.702	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 124.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	8671	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	188.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.09% 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	0.24	0/5744	0.41	0/7776
2	C	0.28	0/3101	0.49	0/4202
All	All	0.25	0/8845	0.44	0/11978

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	5623	0	5577	286	0
2	C	3048	0	3033	306	0
All	All	8671	0	8610	555	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 32.

The worst 5 of 555 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:C:104:LEU:HA	2:C:219:LEU:HD12	1.19	1.09
2:C:215:ILE:HG21	2:C:267:LEU:HD13	1.31	1.08

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1318:HIS:HB2	2:C:160:PHE:HB2	1.44	0.99
2:C:412:LEU:HB3	2:C:416:TYR:HE2	1.32	0.94
1:A:1251:THR:H	1:A:1257:LYS:HE2	1.35	0.91

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	695/753 (92%)	608 (88%)	78 (11%)	9 (1%)	14 57
2	C	373/426 (88%)	297 (80%)	66 (18%)	10 (3%)	6 43
All	All	1068/1179 (91%)	905 (85%)	144 (14%)	19 (2%)	10 51

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	218	TYR
2	C	281	ASP
2	C	55	ASN
2	C	171	VAL
2	C	222	VAL

### 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	621/655 (95%)	618 (100%)	3 (0%)	91	96
2	C	345/384 (90%)	336 (97%)	9 (3%)	51	78
All	All	966/1039 (93%)	954 (99%)	12 (1%)	75	89

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	163	ARG
2	C	170	ASP
2	C	301	TYR
2	C	99	TYR
2	C	280	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	234	GLN
2	C	360	HIS
2	C	273	ASN
2	C	118	GLN
2	C	240	HIS

### 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	696/753 (92%)	-0.23	7 (1%) 82 74	159, 188, 244, 325	0
2	C	372/426 (87%)	-0.37	1 (0%) 93 91	134, 171, 248, 370	0
All	All	1068/1179 (90%)	-0.28	8 (0%) 87 82	134, 183, 246, 370	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	70	LEU	3.0
1	A	71	ALA	2.7
1	A	194	TYR	2.6
1	A	82	TRP	2.5
1	A	69	ILE	2.3

### 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.