wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2023 – 01:25 PM EST

PDB ID	:	2NNW
Title	:	Alternative conformations of Nop56/58-fibrillarin complex and implication for induced-fit assenly of box C/D RNPs
Authors	:	Oruganti, S.; Zhang, Y.; Terns, R.; Terns, M.P.; Li, H.
Deposited on	:	2006-10-24
Resolution	:	2.70 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

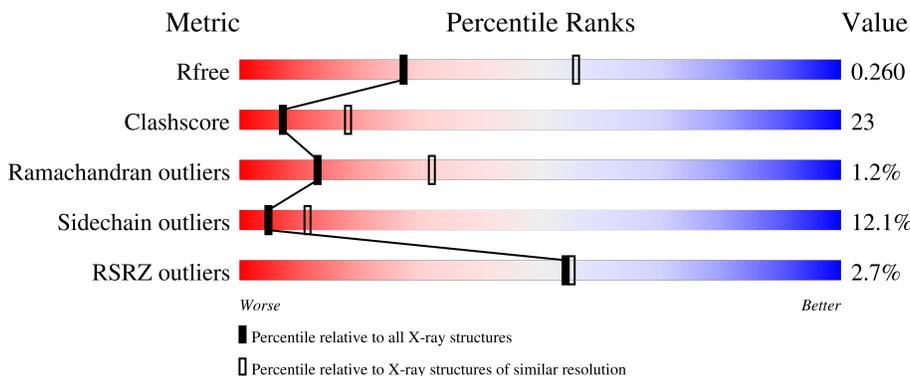|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.70 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	Percentile Ranks	Value
Rfree		0.260
Clashscore		23
Ramachandran outliers		1.2%
Sidechain outliers		12.1%
RSRZ outliers		2.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 >=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 <=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	376	2% / 61% / 28% / · / 7%
1	C	376	2% / 57% / 31% / · / 7%
2	B	234	4% / 54% / 33% / 9% / ·
2	D	234	2% / 53% / 36% / 9% / ·

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 9336 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 NOP5/NOP56 related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	350	Total	C	N	O	S	0	0	0
			2846	1815	491	533	7			
1	C	350	Total	C	N	O	S	0	0	0
			2846	1815	491	533	7			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	MET	-	cloning artifact	UNP Q8U4M1
A	-4	HIS	-	expression tag	UNP Q8U4M1
A	-3	HIS	-	expression tag	UNP Q8U4M1
A	-2	HIS	-	expression tag	UNP Q8U4M1
A	-1	HIS	-	expression tag	UNP Q8U4M1
A	0	HIS	-	expression tag	UNP Q8U4M1
A	1	HIS	-	expression tag	UNP Q8U4M1
A	2	VAL	-	cloning artifact	UNP Q8U4M1
A	3	MET	-	cloning artifact	UNP Q8U4M1
A	4	ILE	-	cloning artifact	UNP Q8U4M1
C	-5	MET	-	cloning artifact	UNP Q8U4M1
C	-4	HIS	-	expression tag	UNP Q8U4M1
C	-3	HIS	-	expression tag	UNP Q8U4M1
C	-2	HIS	-	expression tag	UNP Q8U4M1
C	-1	HIS	-	expression tag	UNP Q8U4M1
C	0	HIS	-	expression tag	UNP Q8U4M1
C	1	HIS	-	expression tag	UNP Q8U4M1
C	2	VAL	-	cloning artifact	UNP Q8U4M1
C	3	MET	-	cloning artifact	UNP Q8U4M1
C	4	ILE	-	cloning artifact	UNP Q8U4M1

- Molecule 2 is a protein called Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase.

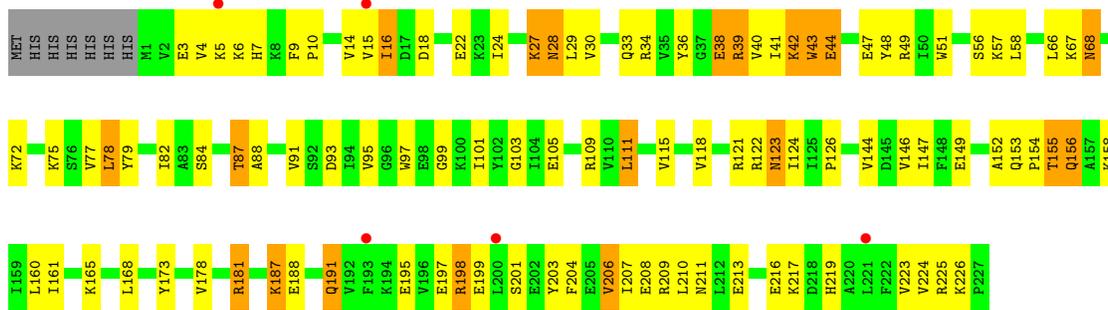| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	227	Total	C	N	O	S	0	0	0
			1822	1174	312	334	2			
2	D	227	Total	C	N	O	S	0	0	0
			1822	1174	312	334	2			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	MET	-	cloning artifact	UNP Q8U4M2
B	-5	HIS	-	expression tag	UNP Q8U4M2
B	-4	HIS	-	expression tag	UNP Q8U4M2
B	-3	HIS	-	expression tag	UNP Q8U4M2
B	-2	HIS	-	expression tag	UNP Q8U4M2
B	-1	HIS	-	expression tag	UNP Q8U4M2
B	0	HIS	-	expression tag	UNP Q8U4M2
D	-6	MET	-	cloning artifact	UNP Q8U4M2
D	-5	HIS	-	expression tag	UNP Q8U4M2
D	-4	HIS	-	expression tag	UNP Q8U4M2
D	-3	HIS	-	expression tag	UNP Q8U4M2
D	-2	HIS	-	expression tag	UNP Q8U4M2
D	-1	HIS	-	expression tag	UNP Q8U4M2
D	0	HIS	-	expression tag	UNP Q8U4M2

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: NOP5/NOP56 related protein

- Molecule 1: NOP5/NOP56 related protein

• Molecule 2: Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase

Chain B:

• Molecule 2: Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 1 2	Depositor
Cell constants a, b, c, α, β, γ	120.71Å 120.71Å 297.74Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.33 – 2.70 49.32 – 2.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (49.33-2.70) 92.9 (49.32-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.58 (at 2.61Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.234 , 0.265 0.233 , 0.260	Depositor DCC
R_{free} test set	3605 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	52.6	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 36.7	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.467 for -h,-k,l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9336	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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	0.67	1/2896 (0.0%)	0.72	1/3896 (0.0%)
1	C	0.80	2/2896 (0.1%)	0.81	0/3896
2	B	0.62	0/1861	0.74	1/2515 (0.0%)
2	D	0.76	0/1861	0.81	1/2515 (0.0%)
All	All	0.72	3/9514 (0.0%)	0.77	3/12822 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	134	MET	SD-CE	-6.28	1.42	1.77
1	A	134	MET	SD-CE	-5.45	1.47	1.77
1	C	253	MET	SD-CE	-5.36	1.47	1.77

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	134	MET	CG-SD-CE	5.26	108.62	100.20
2	B	109	ARG	NE-CZ-NH2	-5.20	117.70	120.30
2	D	109	ARG	NE-CZ-NH2	-5.12	117.74	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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	2846	0	2877	107	2
1	C	2846	0	2877	116	2
2	B	1822	0	1869	101	0
2	D	1822	0	1869	107	0
All	All	9336	0	9492	426	4

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 23.

The worst 5 of 426 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:A:135:VAL:HG11	1:A:253:MET:HE1	1.17	1.13
2:D:181:ARG:H	2:D:181:ARG:HD3	1.14	1.10
2:B:211:ASN:HD21	2:B:213:GLU:HG2	1.24	0.99
1:C:330:LYS:HD3	1:C:368:ILE:HD12	1.38	0.98
1:A:157:ARG:HH11	1:A:174:HIS:HD2	1.17	0.92

All (4) 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:234:GLU:OE2	1:A:234:GLU:OE2[4_775]	1.73	0.47
1:A:147:LYS:NZ	1:A:147:LYS:NZ[4_775]	1.87	0.33
1:C:147:LYS:NZ	1:C:147:LYS:NZ[6_576]	1.95	0.25
1:C:234:GLU:OE2	1:C:234:GLU:OE2[6_576]	2.00	0.20

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	346/376 (92%)	326 (94%)	17 (5%)	3 (1%)	17	40

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	346/376 (92%)	321 (93%)	22 (6%)	3 (1%)	17	40
2	B	225/234 (96%)	204 (91%)	16 (7%)	5 (2%)	6	17
2	D	225/234 (96%)	200 (89%)	22 (10%)	3 (1%)	12	30
All	All	1142/1220 (94%)	1051 (92%)	77 (7%)	14 (1%)	13	32

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	156	GLN
1	A	312	GLY
2	B	122	ARG
2	B	156	GLN
2	B	206	VAL

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	299/321 (93%)	269 (90%)	30 (10%)	7	18
1	C	299/321 (93%)	267 (89%)	32 (11%)	6	15
2	B	197/204 (97%)	167 (85%)	30 (15%)	3	7
2	D	197/204 (97%)	169 (86%)	28 (14%)	3	8
All	All	992/1050 (94%)	872 (88%)	120 (12%)	5	11

5 of 120 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	198	ARG
2	D	123	ASN
1	C	139	ILE
2	D	111	LEU
2	D	201	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	123	ASN
2	D	156	GLN
2	B	156	GLN
2	B	123	ASN
2	D	211	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 ⓘ

There are no monosaccharides 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	350/376 (93%)	0.45	8 (2%)	60	62	30, 58, 102, 115	0
1	C	350/376 (93%)	0.46	9 (2%)	56	57	32, 59, 103, 118	0
2	B	227/234 (97%)	0.47	9 (3%)	38	37	34, 64, 92, 127	0
2	D	227/234 (97%)	0.48	5 (2%)	62	63	34, 64, 94, 128	0
All	All	1154/1220 (94%)	0.47	31 (2%)	54	55	30, 61, 101, 128	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	82	PHE	4.2
1	A	54	LEU	4.0
2	D	221	LEU	3.3
2	D	200	LEU	3.3
1	C	42	LEU	3.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.