Full wwPDB X-ray Structure Validation Report (i)

Aug 7, 2023 – 03:55 PM EDT

PDB ID	:	6X02
Title	:	Nup84-Nup133 (aa521-1157) from S. cerevisiae bound by VHH-SAN8
Authors	:	Nordeen, S.A.; Schwartz, T.U.
Deposited on	:	2020-05-15
Resolution	:	6.38 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (1)) were used in the production of this report:

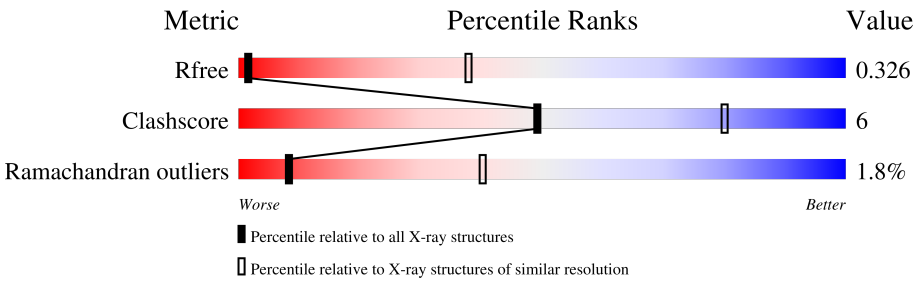
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 6.38 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	1009 (8.70-3.88)
Clashscore	141614	1058 (8.70-3.90)
Ramachandran outliers	138981	1006 (8.70-3.88)

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%

Mol	Chain	Length	Quality of chain
1	A	726	84% 5% 10%
2	B	643	86% 7% 6%
3	C	131	55% 12% • 31%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6667 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 Nucleoporin NUP84.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	650	3234	1934	650	650	0	0	0

- Molecule 2 is a protein called Nucleoporin NUP133.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	602	2993	1789	602	602	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	515	MET	-	expression tag	UNP P36161
B	516	ALA	-	expression tag	UNP P36161
B	517	ASP	-	expression tag	UNP P36161
B	518	PRO	-	expression tag	UNP P36161
B	519	GLY	-	expression tag	UNP P36161
B	520	PHE	-	expression tag	UNP P36161
B	1021	ILE	LEU	conflict	UNP P36161
B	1027	ASP	ASN	conflict	UNP P36161

- Molecule 3 is a protein called VHH-SAN8.

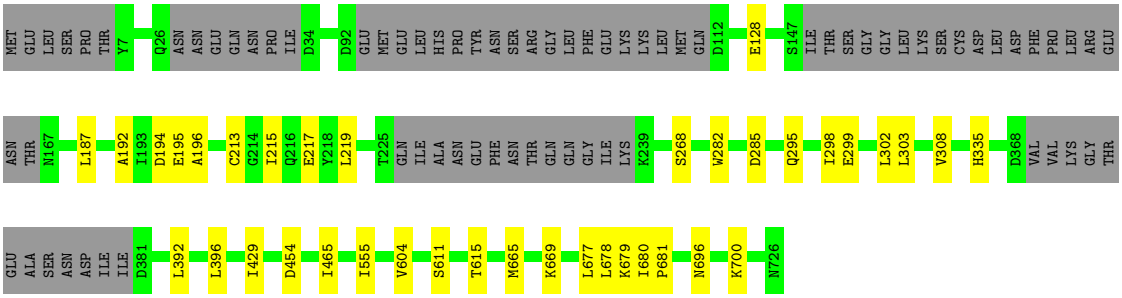
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	90	440	260	90	90	0	0	0

3 Residue-property plots (i)

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. 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. 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: Nucleoporin NUP84

Chain A:

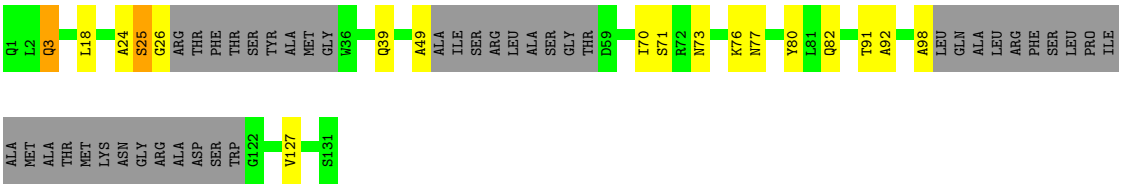


- Molecule 2: Nucleoporin NUP133

Chain B:

- Molecule 3: VHH-SAN8

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.26Å 287.65Å 297.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	103.38 – 6.38 103.38 – 6.38	Depositor EDS
% Data completeness (in resolution range)	98.5 (103.38-6.38) 98.5 (103.38-6.38)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.28 (at 6.19Å)	Xtriage
Refinement program	PHENIX 1.18_3845	Depositor
R, R_{free}	0.314 , 0.324 0.313 , 0.326	Depositor DCC
R_{free} test set	1350 reflections (10.03%)	wwPDB-VP
Wilson B-factor (Å^2)	614.5	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 999.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.018 for -h,l,k	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	6667	wwPDB-VP
Average B, all atoms (Å^2)	574.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% 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 (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	0/3228	0.42	0/4500
2	B	0.25	0/2989	0.49	1/4168 (0.0%)
3	C	0.26	0/436	0.57	0/598
All	All	0.25	0/6653	0.46	1/9266 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	658	PHE	C-N-CA	7.68	154.24	122.00

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	3234	0	1392	18	0
2	B	2993	0	1244	27	0
3	C	440	0	211	14	0
All	All	6667	0	2847	56	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 6.

All (56) 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:662:ASN:O	2:B:666:LYS:CB	1.83	1.25
3:C:73:ASN:CB	3:C:77:ASN:O	1.84	1.25
1:A:298:ILE:O	1:A:302:LEU:CB	1.99	1.10
2:B:662:ASN:O	2:B:666:LYS:N	1.97	0.97
3:C:18:LEU:O	3:C:82:GLN:CB	2.18	0.91
3:C:3:GLN:O	3:C:25:SER:CB	2.24	0.86
2:B:663:GLN:O	2:B:667:LEU:N	2.08	0.86
2:B:537:GLU:O	2:B:541:SER:O	1.94	0.85
1:A:282:TRP:CB	1:A:335:HIS:CB	2.57	0.82
1:A:696:ASN:HA	1:A:700:LYS:HA	1.61	0.80
2:B:662:ASN:O	2:B:666:LYS:CA	2.30	0.79
1:A:195:GLU:CB	3:C:26:GLY:O	2.32	0.78
2:B:601:ASN:O	2:B:605:ARG:N	2.18	0.75
1:A:268:SER:O	1:A:295:GLN:CB	2.34	0.75
1:A:192:ALA:O	1:A:196:ALA:N	2.20	0.73
1:A:299:GLU:O	1:A:303:LEU:CB	2.42	0.67
2:B:962:ASN:O	2:B:966:LYS:CB	2.45	0.65
3:C:49:ALA:HB1	3:C:70:ILE:CB	2.30	0.62
2:B:657:LEU:O	2:B:660:SER:O	2.18	0.61
2:B:662:ASN:CB	2:B:731:GLU:CB	2.79	0.61
3:C:3:GLN:O	3:C:25:SER:N	2.33	0.61
2:B:604:ILE:O	2:B:608:ASP:N	2.27	0.60
2:B:847:LEU:O	2:B:850:TYR:N	2.31	0.60
1:A:298:ILE:O	1:A:302:LEU:N	2.36	0.59
1:A:213:CYS:O	1:A:217:GLU:N	2.34	0.58
3:C:3:GLN:CB	3:C:25:SER:CB	2.81	0.58
2:B:847:LEU:O	2:B:850:TYR:CB	2.52	0.58
2:B:919:GLU:O	2:B:923:HIS:N	2.32	0.57
2:B:900:MET:O	2:B:904:LYS:N	2.30	0.56
3:C:24:ALA:O	3:C:26:GLY:N	2.39	0.56
1:A:392:LEU:O	1:A:396:LEU:N	2.36	0.55
2:B:886:ALA:HB1	2:B:901:ASN:CB	2.37	0.54
2:B:897:ALA:O	2:B:901:ASN:N	2.39	0.54
1:A:194:ASP:HA	3:C:76:LYS:HA	1.90	0.54
2:B:1144:LYS:O	2:B:1146:TYR:N	2.43	0.51
1:A:215:ILE:O	1:A:219:LEU:N	2.44	0.51
2:B:657:LEU:C	2:B:660:SER:H	2.13	0.51
2:B:820:THR:O	2:B:824:ASN:N	2.44	0.50
2:B:916:SER:O	2:B:920:CYS:N	2.28	0.50
2:B:600:PHE:O	2:B:604:ILE:N	2.37	0.50
2:B:663:GLN:O	2:B:667:LEU:CB	2.61	0.49
1:A:665:MET:O	1:A:669:LYS:N	2.46	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:538:ILE:HA	2:B:541:SER:O	2.14	0.47
2:B:909:ASP:HA	2:B:913:LYS:CB	2.46	0.46
3:C:91:THR:HA	3:C:127:VAL:O	2.16	0.46
1:A:677:LEU:O	1:A:679:LYS:N	2.45	0.45
3:C:3:GLN:O	3:C:25:SER:CA	2.63	0.45
1:A:128:GLU:CB	3:C:98:ALA:O	2.65	0.45
1:A:611:SER:O	1:A:615:THR:N	2.44	0.44
2:B:847:LEU:C	2:B:850:TYR:H	2.21	0.44
3:C:71:SER:CB	3:C:80:TYR:HA	2.48	0.43
2:B:884:HIS:C	2:B:886:ALA:H	2.21	0.43
2:B:657:LEU:O	2:B:660:SER:N	2.43	0.43
1:A:187:LEU:O	1:A:192:ALA:HB3	2.20	0.42
3:C:39:GLN:O	3:C:92:ALA:HB1	2.21	0.41
1:A:298:ILE:O	1:A:302:LEU:CA	2.65	0.40

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	638/726 (88%)	576 (90%)	52 (8%)	10 (2%)	9 44
2	B	594/643 (92%)	520 (88%)	62 (10%)	12 (2%)	7 38
3	C	82/131 (63%)	73 (89%)	7 (8%)	2 (2%)	6 33
All	All	1314/1500 (88%)	1169 (89%)	121 (9%)	24 (2%)	8 40

All (24) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	638	SER
2	B	639	PRO
2	B	773	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	879	ALA
2	B	880	PRO
1	A	285	ASP
1	A	308	VAL
1	A	604	VAL
2	B	552	THR
2	B	980	ILE
2	B	1145	ASN
3	C	3	GLN
3	C	25	SER
2	B	775	ILE
2	B	895	SER
1	A	429	ILE
1	A	678	LEU
2	B	608	ASP
2	B	842	GLU
1	A	465	ILE
1	A	555	ILE
1	A	454	ASP
1	A	681	PRO
1	A	680	ILE

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.