



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 18, 2017 – 05:33 AM EDT

PDB ID : 5HB0
Title : Crystal structure of Chaetomium thermophilum Nup170 CTD-Nup145N complex
Authors : Lin, D.H.; Mobbs, G.; Hoelz, A.
Deposited on : unknown
Resolution : 3.50 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

<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

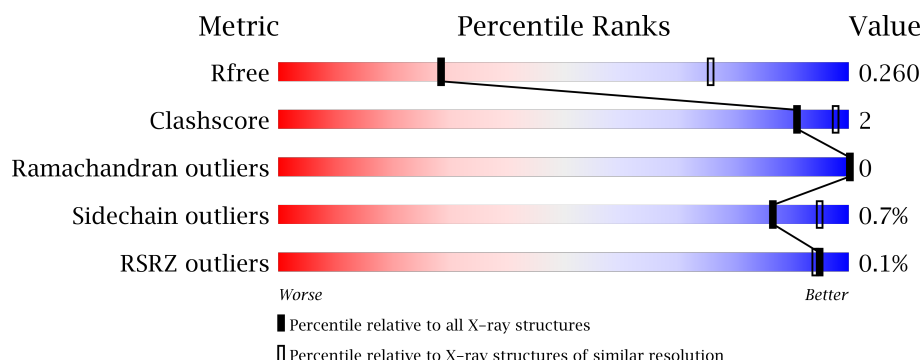
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1195 (3.60-3.40)
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)
RSRZ outliers	101464	1226 (3.60-3.40)

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. 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	549	
1	B	549	
1	C	549	
1	D	549	
2	E	22	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	22	 77% 23%
2	G	22	 5% 86% 14%
2	H	22	 64% 14% 23%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 33811 atoms, of which 16813 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 NUP170.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	520	Total	C	H	N	O	S	0	0	0
			8252	2623	4097	731	791	10			
1	A	508	Total	C	H	N	O	S	0	0	0
			8126	2585	4037	719	775	10			
1	C	519	Total	C	H	N	O	S	0	0	0
			8260	2629	4107	725	789	10			
1	D	502	Total	C	H	N	O	S	0	0	0
			8006	2546	3976	710	764	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	THR	deletion	UNP G0S7B6
B	?	-	LEU	deletion	UNP G0S7B6
B	?	-	ARG	deletion	UNP G0S7B6
A	?	-	THR	deletion	UNP G0S7B6
A	?	-	LEU	deletion	UNP G0S7B6
A	?	-	ARG	deletion	UNP G0S7B6
C	?	-	THR	deletion	UNP G0S7B6
C	?	-	LEU	deletion	UNP G0S7B6
C	?	-	ARG	deletion	UNP G0S7B6
D	?	-	THR	deletion	UNP G0S7B6
D	?	-	LEU	deletion	UNP G0S7B6
D	?	-	ARG	deletion	UNP G0S7B6

- Molecule 2 is a protein called Nucleoporin NUP145.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	E	17	Total	C	H	N	O	S	0	0	0
			282	88	144	24	25	1			
2	F	17	Total	C	H	N	O	S	0	0	0
			282	88	144	24	25	1			

Continued on next page...

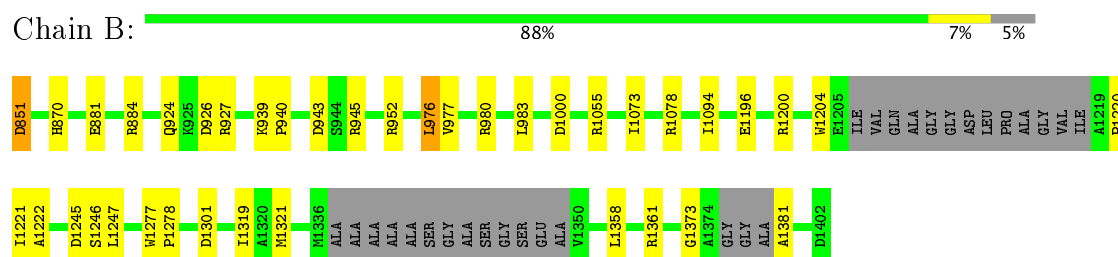
Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	G	19	Total	C	H	N	O	S	0	0	0
			321	100	164	29	27	1			
2	H	17	Total	C	H	N	O	S	0	0	0
			282	88	144	24	25	1			

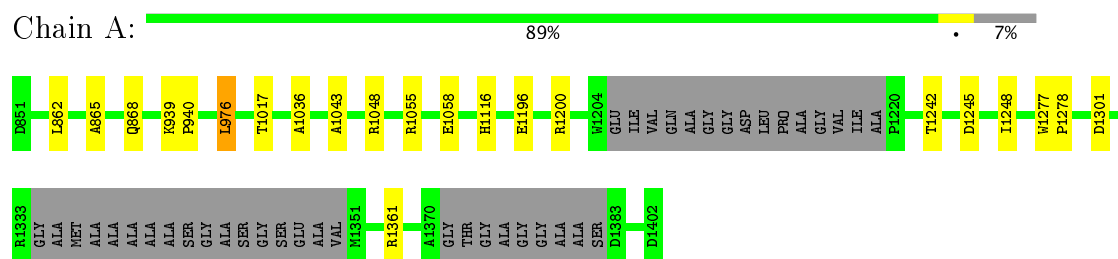
3 Residue-property plots

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.

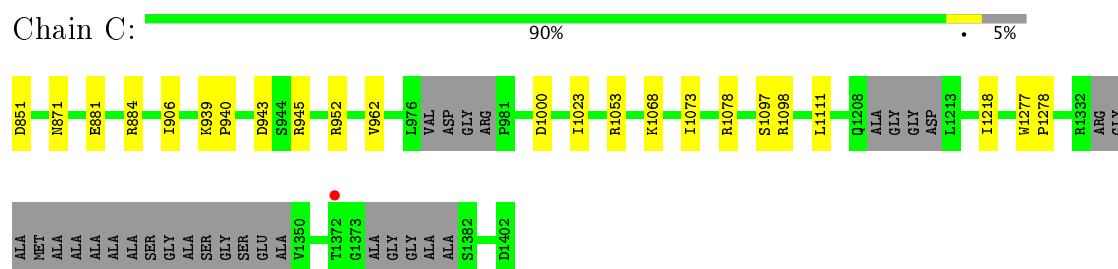
• Molecule 1: Nucleoporin NUP170



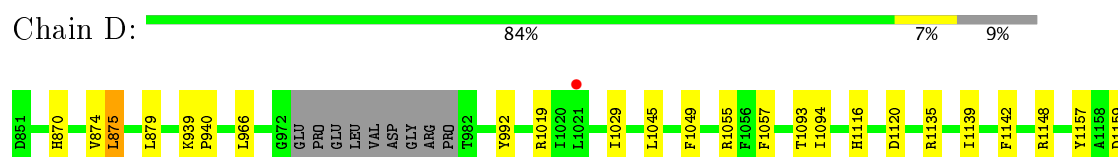
• Molecule 1: Nucleoporin NUP170

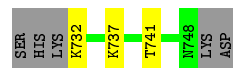


• Molecule 1: Nucleoporin NUP170



• Molecule 1: Nucleoporin NUP170





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	75.59Å 111.47Å 111.75Å 91.82° 92.47° 91.36°	Depositor
Resolution (Å)	45.36 – 3.50 49.19 – 3.50	Depositor EDS
% Data completeness (in resolution range)	98.7 (45.36-3.50) 97.9 (49.19-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.37 (at 3.48Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.210 , 0.260 0.209 , 0.260	Depositor DCC
R_{free} test set	2009 reflections (4.42%)	DCC
Wilson B-factor (Å ²)	81.1	Xtriage
Anisotropy	0.599	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 31.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.054 for h,-k,-l 0.037 for -h,l,k 0.038 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	33811	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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

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	0.31	0/4169	0.44	0/5664
1	B	0.30	0/4235	0.42	0/5755
1	C	0.31	0/4233	0.41	0/5753
1	D	0.29	0/4106	0.43	0/5575
2	E	0.33	0/140	0.51	0/188
2	F	0.39	0/140	0.50	0/188
2	G	0.30	0/160	0.53	0/214
2	H	0.35	0/140	0.49	0/188
All	All	0.30	0/17323	0.43	0/23525

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	976	LEU	Peptide

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	4089	4037	4035	11	0
1	B	4155	4097	4092	23	0
1	C	4153	4107	4104	12	0
1	D	4030	3976	3971	20	0
2	E	138	144	144	1	0
2	F	138	144	144	0	0
2	G	157	164	164	0	0
2	H	138	144	144	2	0
All	All	16998	16813	16798	67	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 67 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:B:1301:ASP:OD1	1:B:1361:ARG:NH1	2.13	0.82
1:C:1023:ILE:O	1:C:1053:ARG:NH1	2.28	0.67
1:C:1073:ILE:O	1:C:1078:ARG:NH2	2.29	0.65
1:A:1301:ASP:OD2	1:A:1361:ARG:NH1	2.29	0.65
1:B:1373:GLY:O	1:B:1381:ALA:N	2.30	0.65

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	500/549 (91%)	483 (97%)	17 (3%)	0	100	100
1	B	512/549 (93%)	489 (96%)	23 (4%)	0	100	100
1	C	509/549 (93%)	484 (95%)	25 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	492/549 (90%)	472 (96%)	20 (4%)	0	100	100
2	E	15/22 (68%)	11 (73%)	4 (27%)	0	100	100
2	F	15/22 (68%)	11 (73%)	4 (27%)	0	100	100
2	G	17/22 (77%)	12 (71%)	5 (29%)	0	100	100
2	H	15/22 (68%)	11 (73%)	4 (27%)	0	100	100
All	All	2075/2284 (91%)	1973 (95%)	102 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	443/460 (96%)	440 (99%)	3 (1%)	87	95
1	B	447/460 (97%)	445 (100%)	2 (0%)	93	97
1	C	450/460 (98%)	448 (100%)	2 (0%)	93	97
1	D	435/460 (95%)	430 (99%)	5 (1%)	78	91
2	E	17/22 (77%)	17 (100%)	0	100	100
2	F	17/22 (77%)	17 (100%)	0	100	100
2	G	19/22 (86%)	19 (100%)	0	100	100
2	H	17/22 (77%)	16 (94%)	1 (6%)	23	61
All	All	1845/1928 (96%)	1832 (99%)	13 (1%)	87	95

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	851	ASP
1	C	871	ASN
1	D	1116	HIS
1	A	1116	HIS
1	D	966	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	923	GLN

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, 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	508/549 (92%)	-0.32	0 100 100	49, 80, 120, 158	0
1	B	520/549 (94%)	-0.32	0 100 100	45, 90, 127, 166	0
1	C	519/549 (94%)	-0.32	1 (0%) 94 93	46, 78, 133, 167	0
1	D	502/549 (91%)	-0.26	1 (0%) 94 93	61, 93, 136, 156	0
2	E	17/22 (77%)	-0.24	0 100 100	74, 97, 122, 128	0
2	F	17/22 (77%)	-0.26	0 100 100	72, 90, 114, 124	0
2	G	19/22 (86%)	-0.16	1 (5%) 27 22	62, 77, 131, 131	0
2	H	17/22 (77%)	-0.28	0 100 100	69, 90, 120, 128	0
All	All	2119/2284 (92%)	-0.30	3 (0%) 95 94	45, 86, 131, 167	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	730	HIS	4.6
1	C	1372	THR	4.2
1	D	1021	LEU	2.4

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.