Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 08:38 AM EDT

PDB ID	:	7MVT
Title	:	Crystal structure of the Chaetomium thermophilum Nup192-Nic96 complex (Nup192 residues 185-1756; Nic96 residues 187-301)
Authors	:	Petrovic, S.; Samanta, D.; Perriches, T.; Bley, C.J.; Thierbach, K.; Brown, B.; Nie, S.; Mobbs, G.W.; Stevens, T.A.; Liu, X.; Tomaleri, G.P.; Schaus, L.; Hoelz, A.
Deposited on	:	2021-05-15
Resolution	:	3.60 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 ⓘ 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 ⓘ) were used in the production of this report:

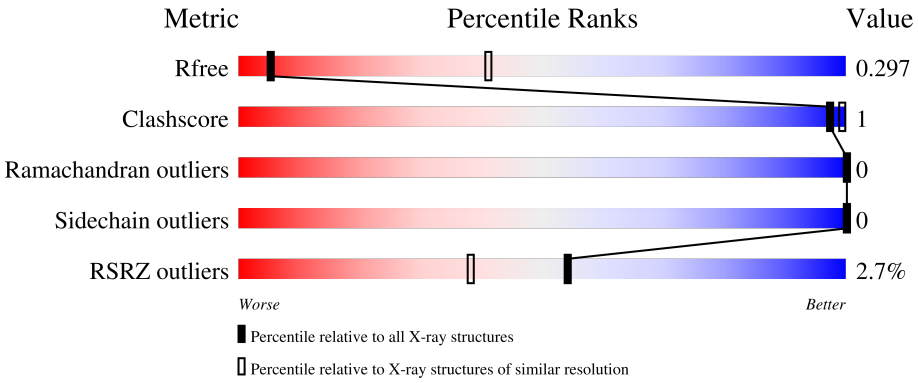
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

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

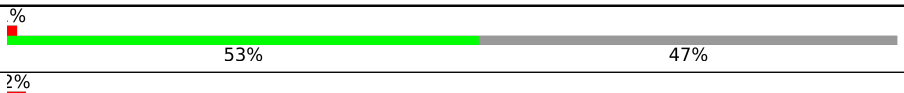
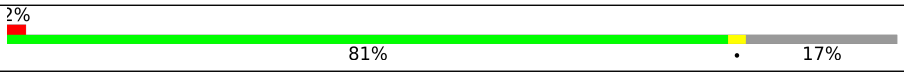
The reported resolution of this entry is 3.60 Å.

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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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	B	116	
2	A	1596	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 22203 atoms, of which 11130 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 NIC96.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	B	62	981	312	483	85	100	1	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	186	SER	-	expression tag	UNP G0S024

- Molecule 2 is a protein called Nucleoporin NUP192.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	A	1325	21222	6793	10647	1787	1939	56	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	161	GLY	-	expression tag	UNP G0S4T0
A	162	PRO	-	expression tag	UNP G0S4T0
A	163	HIS	-	expression tag	UNP G0S4T0
A	164	MET	-	expression tag	UNP G0S4T0
A	165	GLY	-	expression tag	UNP G0S4T0
A	166	LEU	-	expression tag	UNP G0S4T0
A	167	GLN	-	expression tag	UNP G0S4T0
A	168	GLU	-	expression tag	UNP G0S4T0
A	169	SER	-	expression tag	UNP G0S4T0
A	170	PHE	-	expression tag	UNP G0S4T0
A	171	GLY	-	expression tag	UNP G0S4T0
A	172	VAL	-	expression tag	UNP G0S4T0
A	173	ALA	-	expression tag	UNP G0S4T0
A	174	ALA	-	expression tag	UNP G0S4T0
A	175	GLU	-	expression tag	UNP G0S4T0

Continued on next page...

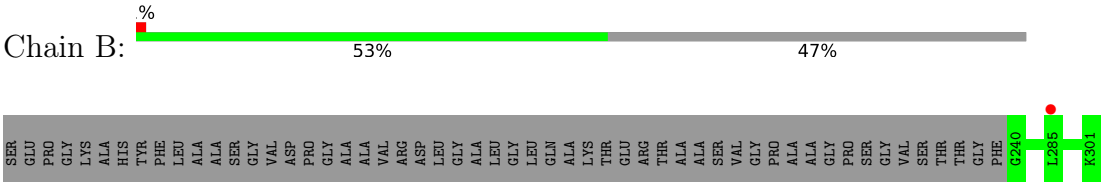
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	176	ASP	-	expression tag	UNP G0S4T0
A	177	LYS	-	expression tag	UNP G0S4T0
A	178	ILE	-	expression tag	UNP G0S4T0
A	179	PHE	-	expression tag	UNP G0S4T0
A	180	GLY	-	expression tag	UNP G0S4T0
A	181	GLY	-	expression tag	UNP G0S4T0
A	182	SER	-	expression tag	UNP G0S4T0
A	183	GLY	-	expression tag	UNP G0S4T0
A	184	SER	-	expression tag	UNP G0S4T0

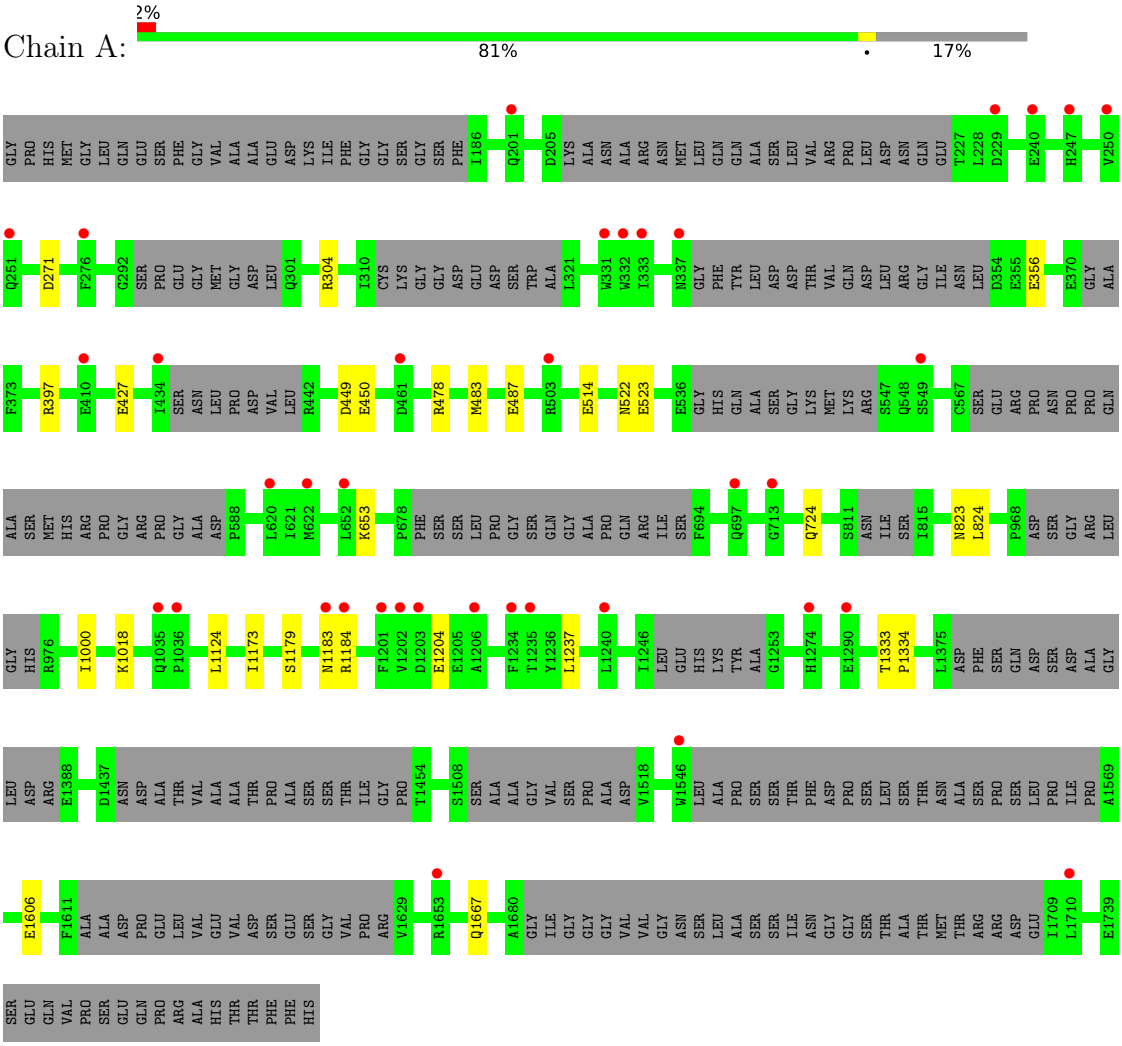
3 Residue-property plots (i)

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 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 NIC96



• Molecule 2: Nucleoporin NUP192



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	76.78Å 76.78Å 712.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.57 – 3.60 38.57 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.57-3.60) 93.2 (38.57-3.60)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.13 (at 3.57Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.258 , 0.297 0.257 , 0.297	Depositor DCC
R_{free} test set	2330 reflections (8.87%)	wwPDB-VP
Wilson B-factor (Å^2)	131.4	Xtriage
Anisotropy	0.485	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 87.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	22203	wwPDB-VP
Average B, all atoms (Å^2)	187.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% 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	B	0.24	0/505	0.44	0/680
2	A	0.23	0/10780	0.37	0/14603
All	All	0.23	0/11285	0.37	0/15283

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

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	B	498	483	482	0	0
2	A	10575	10647	10647	17	0
All	All	11073	11130	11129	17	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 1.

All (17) 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:A:449:ASP:OD1	2:A:450:GLU:N	2.24	0.70
2:A:427:GLU:OE1	2:A:478:ARG:NH2	2.33	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:304:ARG:NH2	2:A:356:GLU:OE2	2.35	0.60
2:A:653:LYS:NZ	2:A:724:GLN:OE1	2.35	0.59
2:A:1183:ASN:OD1	2:A:1184:ARG:N	2.36	0.58
2:A:522:ASN:OD1	2:A:523:GLU:N	2.37	0.56
2:A:823:ASN:OD1	2:A:824:LEU:N	2.38	0.56
2:A:1179:SER:O	2:A:1183:ASN:ND2	2.44	0.51
2:A:483:MET:O	2:A:487:GLU:N	2.48	0.45
2:A:1333:THR:HB	2:A:1334:PRO:HD3	1.99	0.43
2:A:1237:LEU:HD12	2:A:1237:LEU:O	2.19	0.42
2:A:397:ARG:N	2:A:514:GLU:OE1	2.52	0.42
2:A:1124:LEU:HD13	2:A:1173:ILE:HG12	2.01	0.42
2:A:271:ASP:N	2:A:271:ASP:OD1	2.54	0.41
2:A:1204:GLU:OE1	2:A:1204:GLU:N	2.49	0.41
2:A:1606:GLU:OE2	2:A:1667:GLN:NE2	2.49	0.41
2:A:1000:ILE:HA	2:A:1018:LYS:HG2	2.03	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	B	60/116 (52%)	59 (98%)	1 (2%)	0	100	100
2	A	1287/1596 (81%)	1262 (98%)	25 (2%)	0	100	100
All	All	1347/1712 (79%)	1321 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	B	53/87 (61%)	53 (100%)	0	100 100
2	A	1154/1366 (84%)	1154 (100%)	0	100 100
All	All	1207/1453 (83%)	1207 (100%)	0	100 100

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

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

Mol	Chain	Res	Type
2	A	416	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 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

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, 95^{th} 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	B	62/116 (53%)	-0.26	1 (1%) 72 57	112, 180, 233, 242	0
2	A	1325/1596 (83%)	-0.06	37 (2%) 53 37	93, 162, 249, 286	0
All	All	1387/1712 (81%)	-0.06	38 (2%) 54 38	93, 163, 248, 286	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	250	VAL	4.7
2	A	337	ASN	4.5
2	A	251	GLN	4.0
2	A	333	ILE	4.0
2	A	1202	VAL	3.9
2	A	247	HIS	3.5
2	A	697	GLN	3.2
2	A	1206	ALA	2.8
2	A	276	PHE	2.7
2	A	1201	PHE	2.7
2	A	1203	ASP	2.6
1	B	285	LEU	2.5
2	A	1710	LEU	2.5
2	A	201	GLN	2.5
2	A	240	GLU	2.5
2	A	332	TRP	2.5
2	A	1235	THR	2.4
2	A	1035	GLN	2.4
2	A	1234	PHE	2.4
2	A	434	ILE	2.4
2	A	503	ARG	2.4
2	A	1546	TRP	2.3
2	A	1240	LEU	2.3
2	A	229	ASP	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	A	331	TRP	2.2
2	A	549	SER	2.2
2	A	461	ASP	2.1
2	A	652	LEU	2.1
2	A	1184	ARG	2.1
2	A	410	GLU	2.1
2	A	1036	PRO	2.1
2	A	1290	GLU	2.0
2	A	713	GLY	2.0
2	A	1653	ARG	2.0
2	A	620	LEU	2.0
2	A	622	MET	2.0
2	A	1274	HIS	2.0
2	A	1183	ASN	2.0

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 monosaccharides 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.