



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 7, 2022 – 12:04 PM EDT

PDB ID : 7MVW
Title : Crystal structure of Chaetomium thermophilum Nup188 NTD (residues 1-1134)
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 : 2.76 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.28.1
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) : 2.28.1

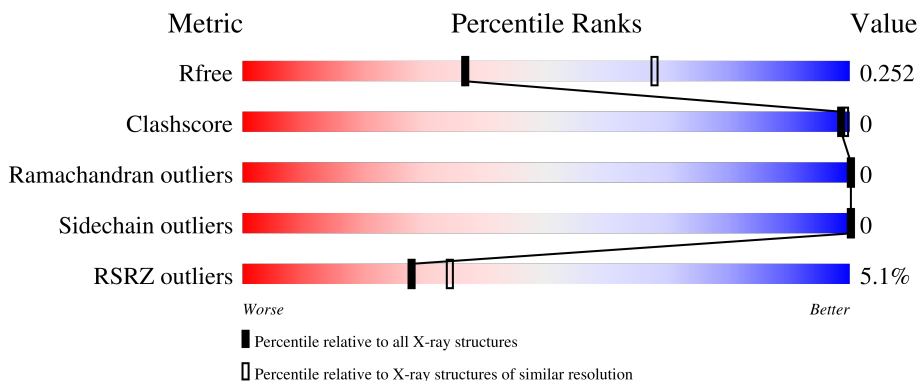
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.76 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 $\leq 5\%$. 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	1138	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 16292 atoms, of which 8151 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 NUP188.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	1038	16221	5167	8132	1383	1508	31	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP G0SFH5
A	-2	PRO	-	expression tag	UNP G0SFH5
A	-1	HIS	-	expression tag	UNP G0SFH5
A	0	ASN	-	expression tag	UNP G0SFH5

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	12	3	6	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			12	3	6	3		
2	A	1	Total	C	H	O	0	0
			13	3	7	3		

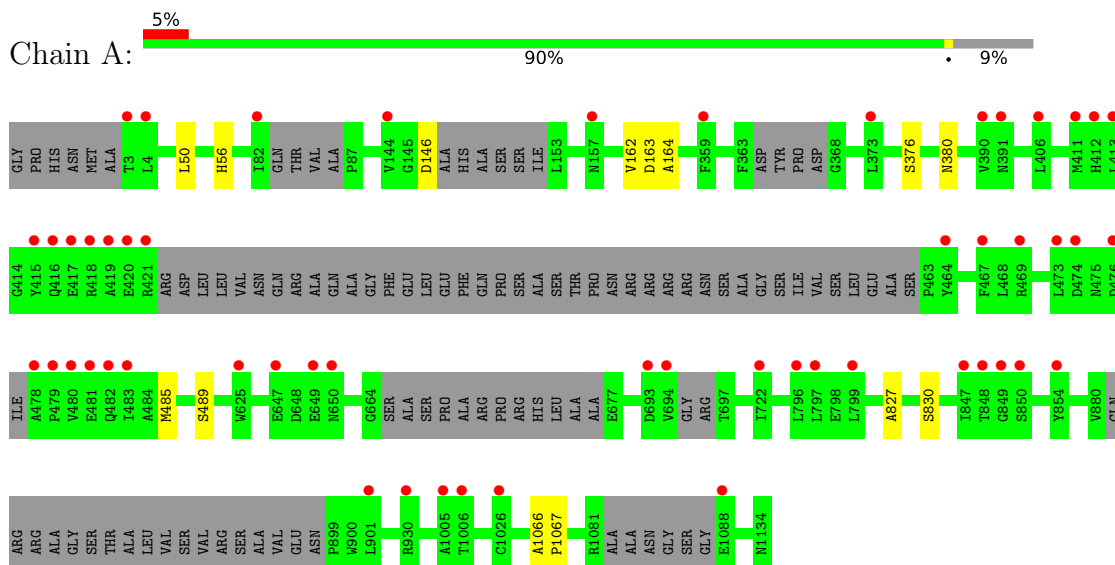
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	34	Total	O	0	0
			34	34		

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: Nucleoporin NUP188



4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α , β , γ	173.54Å 173.54Å 111.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.39 – 2.76 44.68 – 2.76	Depositor EDS
% Data completeness (in resolution range)	99.8 (43.39-2.76) 99.9 (44.68-2.76)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.66 (at 2.77Å)	Xtrriage
Refinement program	PHENIX 1.19.1_4122+SVN	Depositor
R, R_{free}	0.223 , 0.254 0.220 , 0.252	Depositor DCC
R_{free} test set	2449 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	69.3	Xtrriage
Anisotropy	0.487	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 38.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.028 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	16292	wwPDB-VP
Average B, all atoms (Å ²)	89.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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	0/8251	0.39	0/11224

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	8089	8132	8131	8	0
2	A	18	19	24	0	0
3	A	34	0	0	0	0
All	All	8141	8151	8155	8	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 0.

All (8) 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:163:ASP:OD1	1:A:164:ALA:N	2.41	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:827:ALA:O	1:A:830:SER:OG	2.34	0.44
1:A:162:VAL:HG22	1:A:163:ASP:H	1.83	0.44
1:A:376:SER:O	1:A:380:ASN:ND2	2.47	0.44
1:A:146:ASP:OD1	1:A:146:ASP:N	2.52	0.43
1:A:1066:ALA:HB3	1:A:1067:PRO:HD3	2.01	0.42
1:A:50:LEU:O	1:A:56:HIS:NE2	2.52	0.41
1:A:485:MET:O	1:A:489:SER:N	2.50	0.41

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	1018/1138 (90%)	1003 (98%)	15 (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	A	870/945 (92%)	870 (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. There are

no such sidechains identified.

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](#)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	GOL	A	1201	-	5,5,5	0.91	0	5,5,5	0.96	0
2	GOL	A	1203	-	5,5,5	0.95	0	5,5,5	0.97	0
2	GOL	A	1202	-	5,5,5	1.01	0	5,5,5	0.80	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	A	1201	-	-	1/4/4/4	-
2	GOL	A	1203	-	-	2/4/4/4	-
2	GOL	A	1202	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1202	GOL	C1-C2-C3-O3
2	A	1202	GOL	O2-C2-C3-O3
2	A	1203	GOL	O1-C1-C2-C3
2	A	1203	GOL	O1-C1-C2-O2
2	A	1201	GOL	C1-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	1038/1138 (91%)	0.39	53 (5%) 28 34	44, 75, 131, 172	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	797	LEU	6.3
1	A	3	THR	6.2
1	A	476	ASP	5.8
1	A	473	LEU	4.9
1	A	480	VAL	4.8
1	A	469	ARG	4.3
1	A	479	PRO	4.1
1	A	418	ARG	4.0
1	A	411	MET	3.8
1	A	412	HIS	3.7
1	A	848	THR	3.6
1	A	483	ILE	3.6
1	A	849	GLY	3.6
1	A	650	ASN	3.5
1	A	854	TYR	3.5
1	A	415	TYR	3.4
1	A	694	VAL	3.4
1	A	82	ILE	3.3
1	A	421	ARG	3.3
1	A	420	GLU	3.1
1	A	359	PHE	3.1
1	A	1088	GLU	3.1
1	A	4	LEU	3.1
1	A	478	ALA	3.1
1	A	419	ALA	2.9
1	A	649	GLU	2.9
1	A	482	GLN	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1006	THR	2.7
1	A	850	SER	2.7
1	A	481	GLU	2.6
1	A	144	VAL	2.6
1	A	413	LEU	2.6
1	A	901	LEU	2.6
1	A	474	ASP	2.5
1	A	390	VAL	2.5
1	A	391	ASN	2.5
1	A	467	PHE	2.4
1	A	417	GLU	2.4
1	A	693	ASP	2.4
1	A	647	GLU	2.4
1	A	157	ASN	2.4
1	A	847	ILE	2.3
1	A	796	LEU	2.3
1	A	930	ARG	2.3
1	A	1026	CYS	2.3
1	A	416	GLN	2.2
1	A	373	LEU	2.1
1	A	625	TRP	2.1
1	A	406	LEU	2.1
1	A	1005	ALA	2.1
1	A	464	TYR	2.1
1	A	722	ILE	2.1
1	A	799	LEU	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 [\(i\)](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GOL	A	1203	6/6	0.71	0.37	84,102,126,126	0
2	GOL	A	1202	6/6	0.73	0.39	77,95,103,114	0
2	GOL	A	1201	6/6	0.83	0.25	58,88,108,118	0

6.5 Other polymers [i](#)

There are no such residues in this entry.