



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 07:05 am BST

PDB ID : 5HB4  
Title : Crystal structure of Chaetomium thermophilum Nup192  
Authors : Stuwe, T.; Lin, D.H.; Hoelz, A.  
Deposited on : 2015-12-31  
Resolution : 3.20 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](mailto: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  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

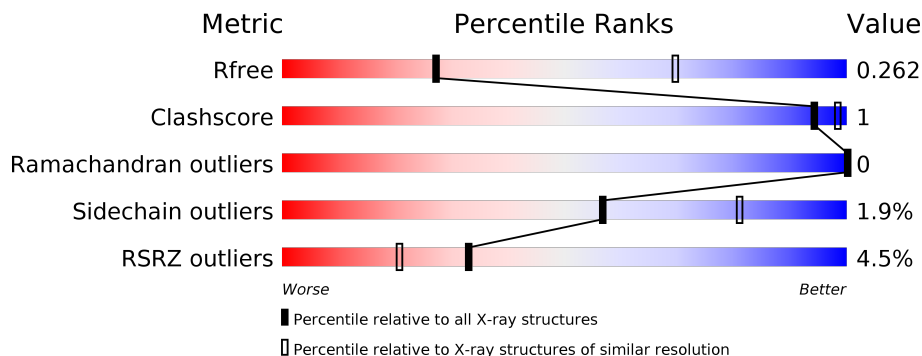
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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  $\geq 3$ , 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	B	1596	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 22100 atoms, of which 11065 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 Nup192,Nucleoporin NUP192.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	B	1388	22095	7062	11065	1868	2042	58	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	181	GLY	-	linker	PDB ?
B	182	SER	-	linker	PDB ?
B	183	GLY	-	linker	PDB ?
B	184	SER	-	linker	PDB ?

- Molecule 2 is OSMIUM ION (three-letter code: OS) (formula: Os).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	5	Total	Os	0	0
			5	5		



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	190.74Å 53.50Å 171.79Å 90.00° 108.33° 90.00°	Depositor
Resolution (Å)	47.74 – 3.20 47.74 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.74-3.20) 99.1 (47.74-3.20)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.41 (at 3.19Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.233 , 0.265 0.235 , 0.262	Depositor DCC
$R_{free}$ test set	1377 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	126.4	Xtrriage
Anisotropy	0.329	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 91.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	22100	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	148.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0/11248	0.38	0/15246

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	B	11030	11065	11059	28	0
2	B	5	0	0	0	0
All	All	11035	11065	11059	28	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 (28) 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:1043:LEU:O	1:B:1065:SER:OG	2.24	0.55
1:B:803:ASP:OD2	1:B:1259:ARG:NH2	2.40	0.55

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:200:LEU:HD22	1:B:232:ARG:HG2	1.94	0.49
1:B:522:ASN:OD1	1:B:523:GLU:N	2.46	0.48
1:B:1081:GLU:O	1:B:1084:GLN:NE2	2.47	0.47
1:B:1421:ARG:NH2	1:B:1482:ASP:OD2	2.47	0.47
1:B:1243:GLY:N	1:B:1244:PRO:HD3	2.30	0.46
1:B:1569:ALA:N	1:B:1573:GLN:OE1	2.49	0.46
1:B:791:ASP:OD1	1:B:875:ARG:NH1	2.49	0.46
1:B:1170:PHE:HB2	1:B:1316:ALA:HB1	1.98	0.46
1:B:1506:ARG:NH1	1:B:1585:GLN:O	2.47	0.44
1:B:254:HIS:O	1:B:254:HIS:ND1	2.50	0.44
1:B:200:LEU:HD12	1:B:276:PHE:HZ	1.84	0.43
1:B:1243:GLY:N	1:B:1244:PRO:CD	2.82	0.43
1:B:299:ASP:O	1:B:303:ALA:N	2.45	0.43
1:B:588:PRO:O	1:B:590:GLU:N	2.52	0.43
1:B:445:ARG:HG3	1:B:446:THR:N	2.34	0.42
1:B:1256:TYR:CG	1:B:1297:LEU:HD12	2.55	0.42
1:B:200:LEU:HD23	1:B:200:LEU:HA	1.93	0.42
1:B:437:LEU:N	1:B:438:PRO:HD3	2.35	0.42
1:B:299:ASP:HB2	1:B:302:GLN:HB2	2.02	0.41
1:B:954:GLU:HG2	1:B:1024:PHE:HA	2.01	0.41
1:B:1231:SER:OG	1:B:1256:TYR:OH	2.38	0.41
1:B:1386:ASP:N	1:B:1386:ASP:OD1	2.53	0.41
1:B:1059:PRO:O	1:B:1065:SER:N	2.53	0.41
1:B:1226:TRP:HB2	1:B:1307:LEU:HD11	2.02	0.41
1:B:1194:ASP:O	1:B:1198:GLY:N	2.52	0.40
1:B:1597:LEU:HD12	1:B:1653:ARG:HG3	2.04	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	1360/1596 (85%)	1294 (95%)	66 (5%)	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	1203/1366 (88%)	1180 (98%)	23 (2%)	57	81

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	254	HIS
1	B	260	PHE
1	B	331	TRP
1	B	337	ASN
1	B	339	PHE
1	B	416	HIS
1	B	623	ASP
1	B	679	PHE
1	B	754	ARG
1	B	823	ASN
1	B	1011	SER
1	B	1133	GLN
1	B	1182	GLN
1	B	1201	PHE
1	B	1272	TYR
1	B	1279	THR
1	B	1353	PHE
1	B	1406	LEU
1	B	1506	ARG
1	B	1547	LEU
1	B	1628	ARG
1	B	1658	ILE
1	B	1708	GLU



Some 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	B	1388/1596 (86%)	0.30	62 (4%) <span style="border: 1px solid red; padding: 2px;">33</span> <span style="border: 1px solid red; padding: 2px;">21</span>	86, 132, 180, 207	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	205	ASP	5.3
1	B	1202	VAL	5.2
1	B	1149	GLU	5.1
1	B	227	THR	4.8
1	B	1708	GLU	4.4
1	B	1666	THR	4.2
1	B	587	ASP	3.9
1	B	342	ASP	3.8
1	B	1681	GLY	3.6
1	B	1613	ALA	3.5
1	B	301	GLN	3.5
1	B	201	GLN	3.2
1	B	1625	GLY	3.2
1	B	1186	PRO	3.1
1	B	1137	GLN	3.1
1	B	452	ARG	3.0
1	B	1716	GLU	2.9
1	B	297	MET	2.9
1	B	453	GLN	2.9
1	B	1376	ASP	2.9
1	B	272	LYS	2.8
1	B	589	ALA	2.7
1	B	698	THR	2.7
1	B	1454	THR	2.7
1	B	983	LEU	2.7
1	B	331	TRP	2.6
1	B	1144	LEU	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	1724	GLU	2.6
1	B	353	LEU	2.6
1	B	1045	PHE	2.6
1	B	352	ASN	2.6
1	B	725	LEU	2.5
1	B	559	LEU	2.5
1	B	986	ASN	2.5
1	B	485	PHE	2.5
1	B	813	ILE	2.5
1	B	1665	LEU	2.4
1	B	291	PHE	2.4
1	B	1636	TYR	2.4
1	B	457	ASN	2.4
1	B	590	GLU	2.4
1	B	1228	GLU	2.4
1	B	1627	PRO	2.4
1	B	1052	LEU	2.3
1	B	302	GLN	2.3
1	B	376	ILE	2.3
1	B	985	ARG	2.3
1	B	330	ALA	2.3
1	B	1340	LEU	2.3
1	B	314	GLY	2.3
1	B	1358	SER	2.2
1	B	940	LEU	2.2
1	B	1712	GLN	2.2
1	B	1629	VAL	2.2
1	B	1671	LEU	2.1
1	B	1719	ILE	2.1
1	B	816	ASP	2.1
1	B	310	ILE	2.1
1	B	1468	ILE	2.1
1	B	1290	GLU	2.0
1	B	1203	ASP	2.0
1	B	1676	LEU	2.0

## 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 [i](#)

There are no carbohydrates 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	OS	B	1805	1/1	0.67	0.12	191,191,191,191	1
2	OS	B	1804	1/1	0.92	0.30	176,176,176,176	1
2	OS	B	1803	1/1	0.93	0.37	167,167,167,167	1
2	OS	B	1801	1/1	0.97	0.33	138,138,138,138	0
2	OS	B	1802	1/1	0.98	0.39	161,161,161,161	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.