



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:38 PM GMT

PDB ID : 4YKE
Title : Crystal structure of eukaryotic Mre11 catalytic domain from *Chaetomium thermophilum*
Authors : Seifert, F.U.; Lammens, K.; Hopfner, K.-P.
Deposited on : 2015-03-04
Resolution : 2.78 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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

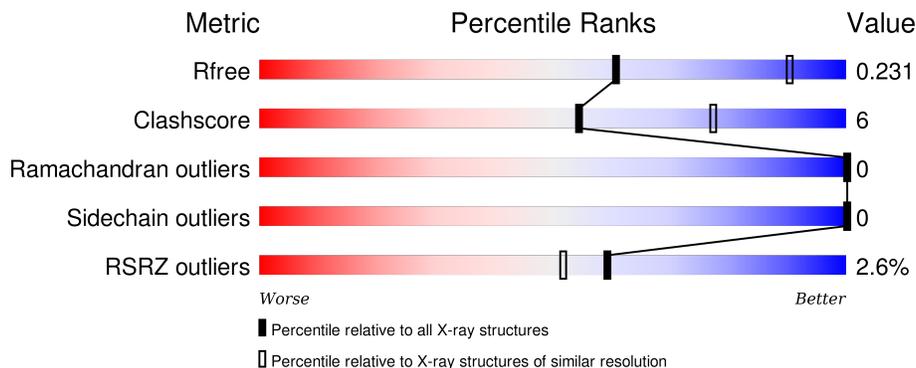
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.78 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}	91344	3004 (2.80-2.76)
Clashscore	102246	3480 (2.80-2.76)
Ramachandran outliers	100387	3423 (2.80-2.76)
Sidechain outliers	100360	3425 (2.80-2.76)
RSRZ outliers	91569	3016 (2.80-2.76)

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	548	 2% 62% 12% 26%
1	B	548	 2% 62% 12% 26%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6612 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 Mre11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	404	3271	2068	579	609	15	127	0	0
1	B	405	3277	2071	582	609	15	158	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	538	ALA	-	expression tag	UNP G0RYR3
A	539	ALA	-	expression tag	UNP G0RYR3
A	540	ALA	-	expression tag	UNP G0RYR3
A	541	LEU	-	expression tag	UNP G0RYR3
A	542	GLU	-	expression tag	UNP G0RYR3
A	543	HIS	-	expression tag	UNP G0RYR3
A	544	HIS	-	expression tag	UNP G0RYR3
A	545	HIS	-	expression tag	UNP G0RYR3
A	546	HIS	-	expression tag	UNP G0RYR3
A	547	HIS	-	expression tag	UNP G0RYR3
A	548	HIS	-	expression tag	UNP G0RYR3
B	538	ALA	-	expression tag	UNP G0RYR3
B	539	ALA	-	expression tag	UNP G0RYR3
B	540	ALA	-	expression tag	UNP G0RYR3
B	541	LEU	-	expression tag	UNP G0RYR3
B	542	GLU	-	expression tag	UNP G0RYR3
B	543	HIS	-	expression tag	UNP G0RYR3
B	544	HIS	-	expression tag	UNP G0RYR3
B	545	HIS	-	expression tag	UNP G0RYR3
B	546	HIS	-	expression tag	UNP G0RYR3
B	547	HIS	-	expression tag	UNP G0RYR3
B	548	HIS	-	expression tag	UNP G0RYR3

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Mn 2 2	0	0
2	A	2	Total Mn 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	34	Total O 34 34	0	0
3	B	26	Total O 26 26	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	56.72Å 56.55Å 304.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.52 – 2.78 49.52 – 2.78	Depositor EDS
% Data completeness (in resolution range)	98.2 (49.52-2.78) 97.9 (49.52-2.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.99 (at 2.77Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.198 , 0.231 0.198 , 0.231	Depositor DCC
R_{free} test set	1245 reflections (4.95%)	DCC
Wilson B-factor (Å ²)	66.1	Xtrriage
Anisotropy	0.675	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 72.6	EDS
Estimated twinning fraction	0.438 for k,h,-l	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtrriage
Outliers	0 of 25153 reflections	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6612	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MN

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	0/3348	0.42	0/4528
1	B	0.23	0/3355	0.42	0/4539
All	All	0.24	0/6703	0.42	0/9067

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	3271	0	3218	40	0
1	B	3277	0	3228	42	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	34	0	0	0	0
3	B	26	0	0	1	0
All	All	6612	0	6446	79	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.

The worst 5 of 79 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:179:ASN:HD22	1:B:213:HIS:HD2	1.23	0.85
1:B:92:SER:HB3	1:B:164:LEU:H	1.53	0.73
1:A:284:SER:HB2	1:A:291:GLU:HB2	1.74	0.68
1:A:17:ASP:HB2	1:A:264:SER:HA	1.79	0.64
1:B:321:LYS:HA	1:B:380:THR:HB	1.79	0.63

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	398/548 (73%)	381 (96%)	17 (4%)	0	100	100
1	B	401/548 (73%)	384 (96%)	17 (4%)	0	100	100
All	All	799/1096 (73%)	765 (96%)	34 (4%)	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	364/490 (74%)	364 (100%)	0	100	100
1	B	365/490 (74%)	365 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	729/980 (74%)	729 (100%)	0	100	100

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

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 4 ligands modelled in this entry, 4 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 [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	403/548 (73%)	0.14	9 (2%) 65 58	34, 76, 135, 213	30 (7%)
1	B	405/548 (73%)	0.19	12 (2%) 54 46	40, 80, 131, 183	42 (10%)
All	All	808/1096 (73%)	0.17	21 (2%) 59 52	34, 78, 132, 213	72 (8%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	403	ASP	4.9
1	B	384	VAL	4.0
1	B	394	PHE	3.9
1	A	182	ASP	3.4
1	A	378	GLU	3.3

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 [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	MN	A	601	1/1	0.99	0.15	-0.40	57,57,57,57	0
2	MN	B	601	1/1	0.99	0.14	-0.67	49,49,49,49	0
2	MN	B	602	1/1	0.99	0.14	-0.77	56,56,56,56	0
2	MN	A	602	1/1	0.99	0.12	-1.37	76,76,76,76	0

6.5 Other polymers [i](#)

There are no such residues in this entry.