Full wwPDB X-ray Structure Validation Report (i)

Nov 25, 2024 – 04:39 PM EST

PDB ID : 2IW0

Title : Structure of the chitin deacetylase from the fungal pathogen Colletotrichum lindemuthianum

Authors : Blair, D.E.; Hekmat, O.; Schuttelkopf, A.W.; Shrestha, B.; Tokuyasu, K.; Withers, S.G.; van Aalten, D.M.F.

Deposited on : 2006-06-23

Resolution : 1.81 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 (i) 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 (i)) were used in the production of this report:

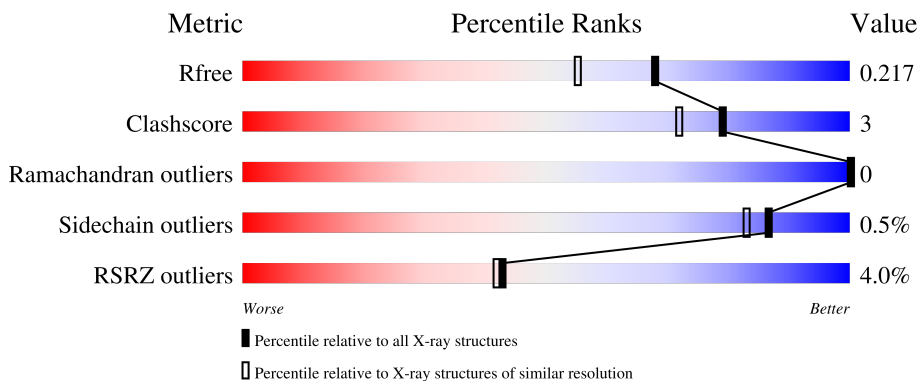|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1 Overall quality at a glance (i)

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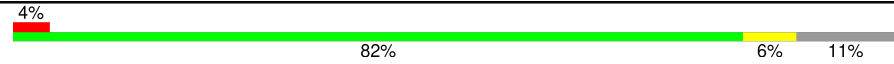*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.81 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   Percentile Ranks   Value

Rfree — 0.217
Clashscore — 3
Ramachandran outliers — 0
Sidechain outliers — 0.5%
RSRZ outliers — 4.0%

Worse   *Better*

▐ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	9242 (1.84-1.80)
Clashscore	180529	1080 (1.82-1.82)
Ramachandran outliers	177936	1073 (1.82-1.82)
Sidechain outliers	177891	1073 (1.82-1.82)
RSRZ outliers	164620	9241 (1.84-1.80)

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%

Mol	Chain	Length	Quality of chain
1	A	254	4% / 82% / 6% / 11%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 1960 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 CHITIN DEACETYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total 1766	C 1098	N 316	O 345	S 7	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	125	GLN	HIS	SEE REMARK 999	UNP Q6DWK3
A	126	LEU	VAL	SEE REMARK 999	UNP Q6DWK3
A	249	HIS	-	expression tag	UNP Q6DWK3
A	250	HIS	-	expression tag	UNP Q6DWK3
A	251	HIS	-	expression tag	UNP Q6DWK3
A	252	HIS	-	expression tag	UNP Q6DWK3
A	253	HIS	-	expression tag	UNP Q6DWK3
A	254	HIS	-	expression tag	UNP Q6DWK3

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 4	C 2	O 2	0	0
3	A	1	Total 4	C 2	O 2	0	0
3	A	1	Total 4	C 2	O 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Cl 1	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

PO4

O3

O⁻

O2 O⁻ — P = O O1

P

O⁻

O4

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 5	O 4	P 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	175	Total 175	O 175	0	0

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. 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. 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: CHITIN DEACETYLASE
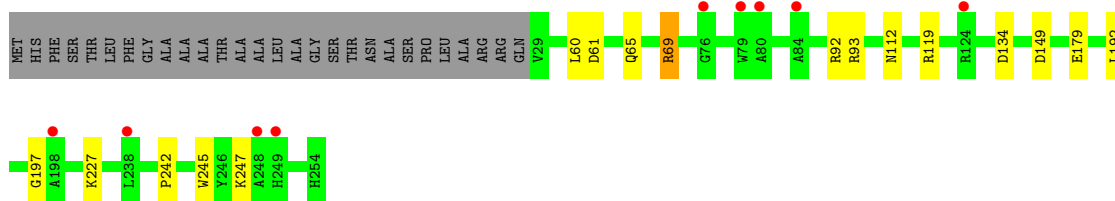
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	71.28Å　　55.02Å　　60.06Å 90.00°　　90.00°　　90.00°	Depositor
Resolution (Å)	20.00　–　1.81 20.00　–　1.81	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-1.81) 99.8 (20.00-1.81)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.72 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.172　,　0.215 0.173　,　0.217	Depositor DCC
R_{free} test set	719 reflections (3.25%)	wwPDB-VP
Wilson B-factor (Å^2)	23.0	Xtriage
Anisotropy	0.836	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 58.3	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	1960	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.28% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CL, PO4, ZN, ACT

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	0.97	2/1808 (0.1%)	0.87	3/2460 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	197	GLY	C-N	6.34	1.48	1.34
1	A	197	GLY	C-O	5.70	1.32	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	92	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	A	119	ARG	NE-CZ-NH2	-6.17	117.22	120.30
1	A	69	ARG	NE-CZ-NH2	-5.07	117.76	120.30

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	A	1766	0	1680	9	0
2	A	1	0	0	0	0
3	A	12	0	9	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	1	0	0	0	0
5	A	5	0	0	0	0
6	A	175	0	0	2	0
All	All	1960	0	1689	9	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 3.

All (9) 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:93:ARG:HG3	3:A:1258:ACT:O	1.36	1.21
1:A:112:ASN:HD21	1:A:149:ASP:H	1.35	0.75
1:A:247:LYS:HE2	6:A:2169:HOH:O	1.90	0.72
1:A:192:LEU:HD23	1:A:227:LYS:HD2	1.74	0.68
1:A:61:ASP:O	1:A:65:GLN:HG2	2.01	0.61
1:A:60:LEU:HD13	1:A:93:ARG:HD3	1.89	0.54
1:A:179:GLU:OE1	1:A:179:GLU:N	2.46	0.45
1:A:242:PRO:HA	1:A:245:TRP:CE2	2.53	0.43
1:A:134:ASP:OD1	6:A:2085:HOH:O	2.21	0.42

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	224/254 (88%)	212 (95%)	12 (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	192/210 (91%)	191 (100%)	1 (0%)	86 83

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	69	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	ASN
1	A	87	ASN
1	A	112	ASN
1	A	244	ASN
1	A	250	HIS
1	A	252	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

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	#$|Z| > 2$
3	ACT	A	1258	-	3,3,3	0.87	0	3,3,3	1.50	0
3	ACT	A	1256	2	3,3,3	0.67	0	3,3,3	1.65	1 (33%)
5	PO4	A	1260	-	4,4,4	0.78	0	6,6,6	0.44	0
3	ACT	A	1257	-	3,3,3	0.78	0	3,3,3	1.68	1 (33%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1257	ACT	OXT-C-CH3	2.23	124.42	115.05
3	A	1256	ACT	OXT-C-O	-2.13	114.13	122.03

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1258	ACT	1	0

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 ⓘ

Warning: The R factor obtained from EDS is 0.2463, which does not match the depositor's R factor of 0.172. Please interpret the results in this section carefully.

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	226/254 (88%)	0.74	9 (3%)	43	42	24, 29, 36, 45	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	ALA	2.9
1	A	248	ALA	2.8
1	A	198	ALA	2.7
1	A	249	HIS	2.6
1	A	80	ALA	2.6
1	A	79	TRP	2.4
1	A	238	LEU	2.3
1	A	124	ARG	2.3
1	A	76	GLY	2.2

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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(Å^2)	Q<0.9
3	ACT	A	1258	4/4	0.55	0.22	53,53,54,54	0
3	ACT	A	1257	4/4	0.72	0.17	46,47,49,50	0
5	PO4	A	1260	5/5	0.82	0.12	51,52,53,53	0
4	CL	A	1259	1/1	0.87	0.13	61,61,61,61	0
3	ACT	A	1256	4/4	0.93	0.09	27,27,27,29	0
2	ZN	A	1255	1/1	0.99	0.07	31,31,31,31	0

6.5 Other polymers ⓘ

There are no such residues in this entry.