



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 12, 2017 – 07:30 AM EDT

PDB ID : 4YSM  
Title : Calcium-Dependent Protein Kinase from Eimeria tenella  
Authors : Merritt, E.A.  
Deposited on : unknown  
Resolution : 3.19 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20029824

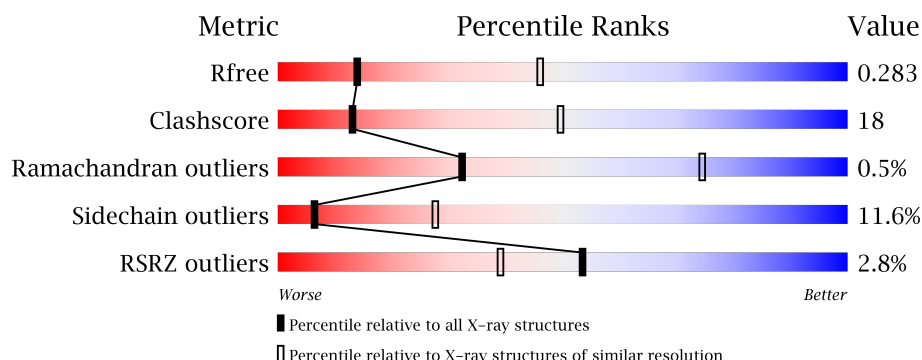
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.19 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))
$R_{free}$	100719	1015 (3.22-3.18)
Clashscore	112137	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)
RSRZ outliers	101464	1020 (3.22-3.18)

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. 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	491	<div> <div>4%</div> <div> <div></div> <div>61%</div> <div>33%</div> <div>• •</div> </div> </div>
1	B	491	<div> <div>2%</div> <div> <div></div> <div>62%</div> <div>31%</div> <div>• •</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7528 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 Calmodulin-like domain protein kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	475	Total	C	N	O	S	0	0	0
			3765	2401	623	725	16			
1	B	475	Total	C	N	O	S	0	0	0
			3753	2393	620	724	16			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q3HNM4
A	-2	PRO	-	expression tag	UNP Q3HNM4
A	-1	GLY	-	expression tag	UNP Q3HNM4
A	0	SER	-	expression tag	UNP Q3HNM4
A	392	ASP	ALA	See Remark 999	UNP Q3HNM4
B	-3	GLY	-	expression tag	UNP Q3HNM4
B	-2	PRO	-	expression tag	UNP Q3HNM4
B	-1	GLY	-	expression tag	UNP Q3HNM4
B	0	SER	-	expression tag	UNP Q3HNM4
B	392	ASP	ALA	See Remark 999	UNP Q3HNM4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	4	Total	Ca	0	0
			4	4		
2	A	4	Total	Ca	0	0
			4	4		

- Molecule 3 is water.

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

*Continued on next page...*

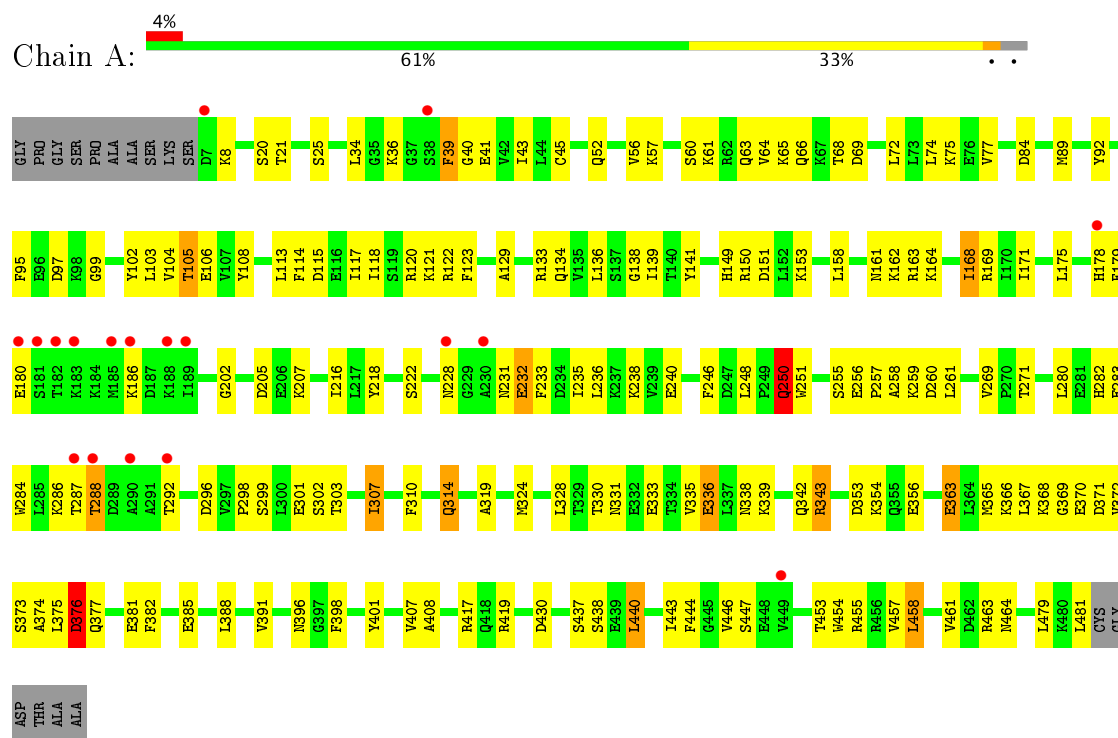
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	O	0	0
			1	1		

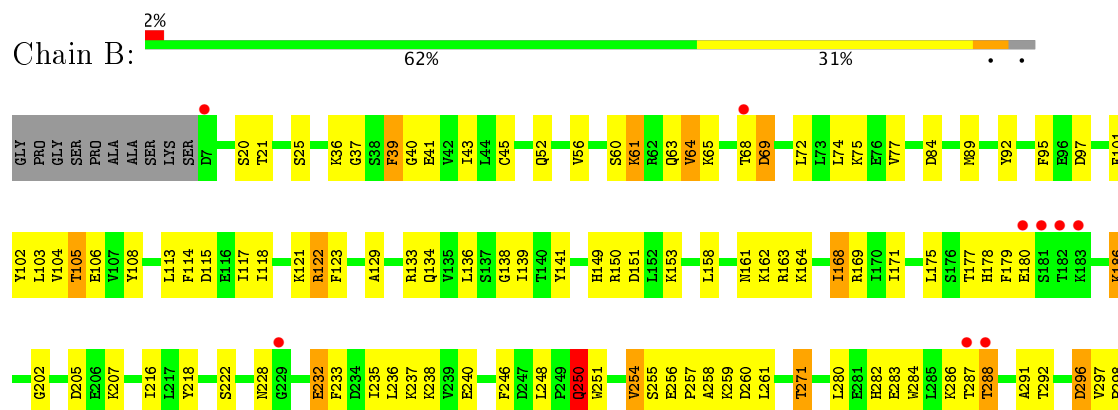
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Calmodulin-like domain protein kinase



- Molecule 1: Calmodulin-like domain protein kinase



S299	
S302	
T303	
E400	
Y401	
V407	
R417	
Q418	
R419	
S437	
S438	
E439	
L440	
I443	
F444	
G445	
V446	
S447	
E448	
V449	
T453	
W454	
W457	
L458	
V461	
D462	
R463	
N464	
N465	
E472	
R475	
L479	
K480	
L481	
CYS	
GLY	
ASP	
THR	
ALA	
ALA	
E381	
F382	
E385	
L388	
V391	
S299	
S302	
T303	
E400	
Y401	
V407	
R417	
Q418	
R419	
S437	
S438	
E439	
L440	
I443	
F444	
G445	
V446	
S447	
E448	
V449	
T453	
W454	
W457	
L458	
V461	
D462	
R463	
N464	
N465	
E472	
R475	
L479	
K480	
L481	
CYS	
GLY	
ASP	
THR	
ALA	
ALA	
E381	
F382	
E385	
L388	
V391	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.67Å 109.56Å 77.73Å 90.00° 92.17° 90.00°	Depositor
Resolution (Å)	38.84 – 3.19 38.84 – 3.19	Depositor EDS
% Data completeness (in resolution range)	97.7 (38.84-3.19) 97.8 (38.84-3.19)	Depositor EDS
$R_{merge}$	0.50	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.44 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.228 , 0.280 0.230 , 0.283	Depositor DCC
$R_{free}$ test set	910 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	74.2	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.047 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7528	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.40% 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: CA

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	0.53	0/3829	0.74	2/5156 (0.0%)
1	B	0.55	0/3817	0.75	2/5142 (0.0%)
All	All	0.54	0/7646	0.75	4/10298 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	430	ASP	CB-CG-OD1	5.60	123.34	118.30
1	A	376	ASP	CB-CG-OD2	-5.60	113.26	118.30
1	B	343	ARG	NE-CZ-NH2	5.32	122.96	120.30
1	B	310	PHE	N-CA-C	-5.15	97.10	111.00

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	3765	0	3739	145	2
1	B	3753	0	3709	132	2
2	A	4	0	0	0	0
2	B	4	0	0	0	0
3	A	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
All	All	7528	0	7448	274	2

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 18.

The worst 5 of 274 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:248:LEU:HD12	1:A:251:TRP:HE1	0.99	1.08
1:A:371:ASP:O	1:A:375:LEU:HD11	1.54	1.06
1:A:64:VAL:HG23	1:A:69:ASP:HB3	1.38	1.05
1:A:336:GLU:N	1:A:336:GLU:OE2	1.88	1.05
1:B:246:PHE:HE2	1:B:259:LYS:HB3	1.19	1.03

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:376:ASP:OD2	1:B:25:SER:OG[2_355]	1.95	0.25
1:A:25:SER:OG	1:B:376:ASP:OD2[2_454]	1.99	0.21

## 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	473/491 (96%)	443 (94%)	28 (6%)	2 (0%)	38	77
1	B	473/491 (96%)	443 (94%)	27 (6%)	3 (1%)	28	72
All	All	946/982 (96%)	886 (94%)	55 (6%)	5 (0%)	32	74

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	250	GLN
1	A	376	ASP
1	B	64	VAL
1	B	121	LYS
1	B	250	GLN

### 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	406/426 (95%)	360 (89%)	46 (11%)	7	29
1	B	402/426 (94%)	354 (88%)	48 (12%)	6	26
All	All	808/852 (95%)	714 (88%)	94 (12%)	6	27

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	463	ARG
1	B	84	ASP
1	B	437	SER
1	A	464	ASN
1	B	39	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	331	ASN
1	B	331	ASN
1	B	52	GLN
1	A	306	ASN
1	B	19	HIS

### 5.3.3 RNA ⓘ

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 8 ligands modelled in this entry, 8 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, 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	A	475/491 (96%)	-0.03	18 (3%) 41 27	43, 72, 119, 139	0
1	B	475/491 (96%)	-0.16	9 (1%) 67 52	40, 68, 113, 149	0
All	All	950/982 (96%)	-0.10	27 (2%) 53 39	40, 70, 116, 149	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	7	ASP	7.1
1	A	288	THR	5.3
1	A	287	THR	4.5
1	A	181	SER	4.3
1	A	7	ASP	4.3

### 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	B	503	1/1	0.84	0.11	-0.90	91,91,91,91	0
2	CA	A	504	1/1	0.91	0.16	-0.91	104,104,104,104	0
2	CA	A	503	1/1	0.87	0.09	-1.20	83,83,83,83	0
2	CA	B	504	1/1	0.89	0.10	-1.29	84,84,84,84	0
2	CA	B	502	1/1	0.95	0.11	-1.49	58,58,58,58	0
2	CA	A	502	1/1	0.96	0.11	-1.60	58,58,58,58	0
2	CA	B	501	1/1	0.99	0.08	-3.41	53,53,53,53	0
2	CA	A	501	1/1	0.97	0.10	-4.36	56,56,56,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.