



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2018 – 01:59 am GMT

PDB ID : 4C0T  
Title : Candida albicans PKh Kinase Domain  
Authors : Schulze, J.O.; Pastor-Flores, D.; Biondi, R.M.  
Deposited on : 2013-08-07  
Resolution : 3.16 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

<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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

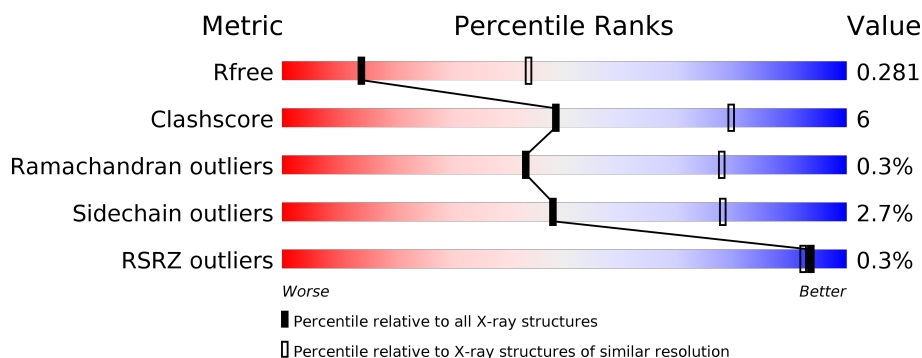
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.16 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))
$R_{free}$	111664	1389 (3.20-3.12)
Clashscore	122126	1522 (3.20-3.12)
Ramachandran outliers	120053	1493 (3.20-3.12)
Sidechain outliers	120020	1492 (3.20-3.12)
RSRZ outliers	108989	1344 (3.20-3.12)

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	974	<div> <div style="width: 24%; background-color: green;"></div> <div style="width: 6%; background-color: yellow;"></div> <div style="width: 70%; background-color: grey;"></div> <div>24% 6% 70%</div> </div>

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 2375 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 LIKELY PROTEIN KINASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	N	O	P	S	0	0	0
			2375	1534	398	436	1	6			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	expression tag	UNP Q5A3P6
A	-26	SER	-	expression tag	UNP Q5A3P6
A	-25	TYR	-	expression tag	UNP Q5A3P6
A	-24	TYR	-	expression tag	UNP Q5A3P6
A	-23	HIS	-	expression tag	UNP Q5A3P6
A	-22	HIS	-	expression tag	UNP Q5A3P6
A	-21	HIS	-	expression tag	UNP Q5A3P6
A	-20	HIS	-	expression tag	UNP Q5A3P6
A	-19	HIS	-	expression tag	UNP Q5A3P6
A	-18	HIS	-	expression tag	UNP Q5A3P6
A	-17	ASP	-	expression tag	UNP Q5A3P6
A	-16	TYR	-	expression tag	UNP Q5A3P6
A	-15	ASP	-	expression tag	UNP Q5A3P6
A	-14	ILE	-	expression tag	UNP Q5A3P6
A	-13	PRO	-	expression tag	UNP Q5A3P6
A	-12	THR	-	expression tag	UNP Q5A3P6
A	-11	THR	-	expression tag	UNP Q5A3P6
A	-10	GLU	-	expression tag	UNP Q5A3P6
A	-9	ASN	-	expression tag	UNP Q5A3P6
A	-8	LEU	-	expression tag	UNP Q5A3P6
A	-7	TYR	-	expression tag	UNP Q5A3P6
A	-6	PHE	-	expression tag	UNP Q5A3P6
A	-5	GLN	-	expression tag	UNP Q5A3P6
A	-4	GLY	-	expression tag	UNP Q5A3P6
A	-3	ALA	-	expression tag	UNP Q5A3P6
A	-2	MET	-	expression tag	UNP Q5A3P6
A	-1	GLY	-	expression tag	UNP Q5A3P6

*Continued on next page...*

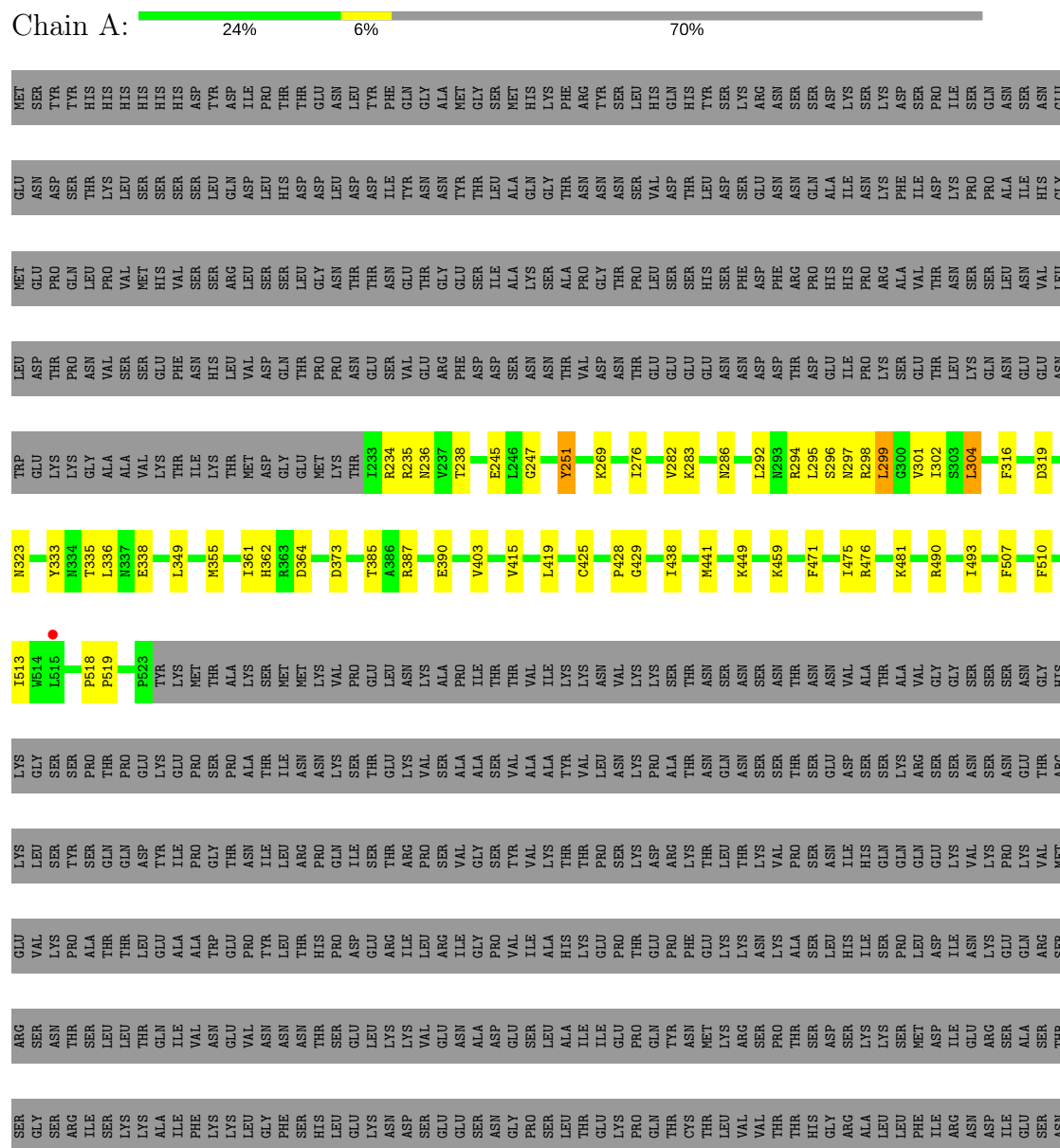
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q5A3P6
A	945	HIS	-	expression tag	UNP Q5A3P6
A	946	MET	-	expression tag	UNP Q5A3P6

### 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: LIKELY PROTEIN KINASE



TYR	LEU	LEU	ALA	GLU	ILE	LYS	LEU	LYS	TYR	PRO	PHE	ILE	HIS	PHE	GLN	GLU	LEU	ILE	VAL	SER	GLN	THR	LYS	PHE	SER	LYS	LEU	VAL	PRO	SER	PRO	VAL	VAL	GLY	VAL	PHE	VAL	ILE	SER	SER	SER	ASP	ASN	LEU	LEU	ILE	PHE	GLU	VAL	VAL	LYS	PHE	GLU	VAL	VAL	THR	TRP	GLU	ALA
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

LEU  
ALA  
LYS  
SER  
LYS  
TYR  
ASN  
GLU  
ILE  
HIS  
MET

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.21Å 100.21Å 60.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.86 – 3.16 70.86 – 3.16	Depositor EDS
% Data completeness (in resolution range)	99.9 (70.86-3.16) 99.9 (70.86-3.16)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.07 (at 3.19Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.228 , 0.280 0.229 , 0.281	Depositor DCC
$R_{free}$ test set	280 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.5	Xtriage
Anisotropy	0.511	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 37.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	2375	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.93% 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: SEP

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	0.21	0/2417	0.38	0/3259

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	2375	0	2396	31	0
All	All	2375	0	2396	31	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 31 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:295:LEU:HD21	1:A:355:MET:HG2	1.74	0.68
1:A:428:PRO:HB2	1:A:490:ARG:HD2	1.85	0.58
1:A:234:ARG:HG2	1:A:235:ARG:H	1.71	0.56

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:449:LYS:O	1:A:459:LYS:NZ	2.34	0.55
1:A:236:ASN:OD1	1:A:238:THR:OG1	2.21	0.50

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	288/974 (30%)	264 (92%)	23 (8%)	1 (0%)	43 78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	296	SER

### 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	258/889 (29%)	251 (97%)	7 (3%)	48 78

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	299	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	510	PHE
1	A	304	LEU
1	A	283	LYS
1	A	385	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
1	SEP	A	407	1	9,9,10	1.56	2 (22%)	9,12,14	1.50	2 (22%)

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. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	SEP	A	407	1	-	0/5/8/10	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	407	SEP	CA-C	2.05	1.52	1.50
1	A	407	SEP	P-O1P	3.23	1.61	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	407	SEP	P-OG-CB	-2.69	110.87	118.30
1	A	407	SEP	OG-CB-CA	2.61	110.74	108.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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	290/974 (29%)	-0.14	1 (0%) 93 92	10, 28, 60, 71	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	515	LEU	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	SEP	A	407	10/11	0.89	0.16	26,39,76,81	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

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

There are no such residues in this entry.