



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 28, 2016 – 10:42 AM EDT

PDB ID : 4YDO  
Title : CRYSTAL STRUCTURE OF CANDIDA ALBICANS PROTEIN FARNE-  
SYLTRANSFERASE IN APO FORM  
Authors : Kumar, S.; Mabanglo, M.F.; Hast, M.A.; Shi, Y.; Beese, L.S.  
Deposited on : 2015-02-22  
Resolution : 3.00 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  
Mogul : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027790  
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) : rb-20027790

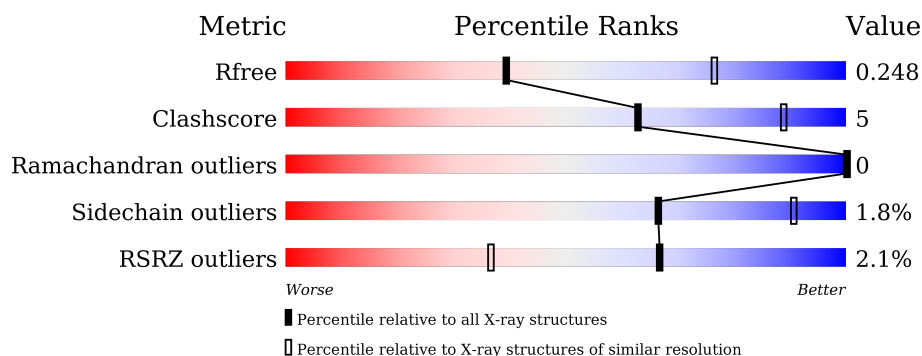
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	329	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 81%, grey 10%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>81%</span> <span>10%</span> <span>8%</span> </div> </div>
2	B	587	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 2%, yellow 12%, green 57%, grey 30%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>57%</span> <span>12%</span> <span>30%</span> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5929 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 Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	0	0
			2566	1646	417	496	7			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	initiating methionine	UNP Q9Y765
A	-22	GLY	-	expression tag	UNP Q9Y765
A	-21	SER	-	expression tag	UNP Q9Y765
A	-20	SER	-	expression tag	UNP Q9Y765
A	-19	HIS	-	expression tag	UNP Q9Y765
A	-18	HIS	-	expression tag	UNP Q9Y765
A	-17	HIS	-	expression tag	UNP Q9Y765
A	-16	HIS	-	expression tag	UNP Q9Y765
A	-15	HIS	-	expression tag	UNP Q9Y765
A	-14	HIS	-	expression tag	UNP Q9Y765
A	-13	SER	-	expression tag	UNP Q9Y765
A	-12	GLN	-	expression tag	UNP Q9Y765
A	-11	ASP	-	expression tag	UNP Q9Y765
A	-10	PRO	-	expression tag	UNP Q9Y765
A	-9	ASN	-	expression tag	UNP Q9Y765
A	-8	SER	-	expression tag	UNP Q9Y765
A	-7	SER	-	expression tag	UNP Q9Y765
A	-6	SER	-	expression tag	UNP Q9Y765
A	-5	ALA	-	expression tag	UNP Q9Y765
A	-4	ARG	-	expression tag	UNP Q9Y765
A	-3	LEU	-	expression tag	UNP Q9Y765
A	-2	GLN	-	expression tag	UNP Q9Y765
A	-1	VAL	-	expression tag	UNP Q9Y765
A	0	ASP	-	expression tag	UNP Q9Y765
A	38	ILE	LEU	conflict	UNP Q9Y765
A	41	SER	ALA	conflict	UNP Q9Y765

- Molecule 2 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	412	Total	C	N	O	S	0	0	0
			3316	2138	539	624	15			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca	0	0
			1	1		
3	A	1	Total	Ca	0	0
			1	1		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	18	Total	O	0	0
			18	18		
5	B	26	Total	O	0	0
			26	26		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.82Å 96.82Å 189.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.30 – 3.00 86.18 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (43.30-3.00) 98.8 (86.18-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1839)	Depositor
R, $R_{free}$	0.226 , 0.249 0.224 , 0.248	Depositor DCC
$R_{free}$ test set	1848 reflections (9.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	81.2	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 38.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5929	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	0.27	0/2630	0.44	0/3569
2	B	0.25	0/3395	0.46	0/4602
All	All	0.26	0/6025	0.45	0/8171

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

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	2566	0	2472	23	0
2	B	3316	0	3223	43	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	B	1	0	0	0	0
5	A	18	0	0	0	0
5	B	26	0	0	0	0
All	All	5929	0	5695	61	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 5.

The worst 5 of 61 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)
2:B:259:TYR:HA	2:B:319:ILE:HD11	1.69	0.74
2:B:424:ARG:HB2	2:B:427:LYS:HG3	1.70	0.72
1:A:71:ILE:HD13	2:B:252:LEU:HD23	1.74	0.69
2:B:359:PHE:HB2	2:B:372:THR:HA	1.76	0.67
1:A:151:LYS:HE3	2:B:364:ASN:HB2	1.83	0.61

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	301/329 (92%)	293 (97%)	8 (3%)	0	100	100
2	B	400/587 (68%)	388 (97%)	12 (3%)	0	100	100
All	All	701/916 (76%)	681 (97%)	20 (3%)	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	289/313 (92%)	285 (99%)	4 (1%)	74	93

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	367/530 (69%)	359 (98%)	8 (2%)	60	88
All	All	656/843 (78%)	644 (98%)	12 (2%)	66	91

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	215	ILE
2	B	237	LEU
2	B	413	ARG
2	B	143	HIS
2	B	326	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 ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 3 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	303/329 (92%)	-0.05	2 (0%) 89 70	46, 66, 102, 121	0
2	B	412/587 (70%)	0.21	13 (3%) 51 23	55, 82, 119, 146	0
All	All	715/916 (78%)	0.10	15 (2%) 67 36	46, 77, 113, 146	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	457	GLN	7.6
2	B	515	GLN	3.6
2	B	461	VAL	3.5
2	B	576	ILE	3.5
2	B	516	ASP	3.1

### 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
4	ZN	B	601	1/1	0.93	0.18	-0.26	58,58,58,58	0
3	CA	A	401	1/1	0.98	0.10	-	57,57,57,57	1
3	CA	B	602	1/1	0.97	0.22	-	69,69,69,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.