



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 10:29 pm GMT

PDB ID : 4H1G  
Title : Structure of *Candida albicans* Kar3 motor domain fused to maltose-binding protein  
Authors : Allingham, J.A.; Duan, D.; Delorme, C.; Joshi, M.  
Deposited on : 2012-09-10  
Resolution : 2.15 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](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

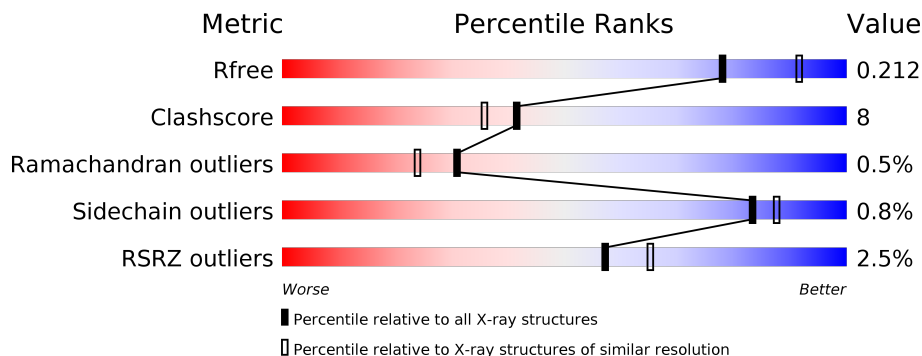
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1287 (2.16-2.16)
Clashscore	122126	1390 (2.16-2.16)
Ramachandran outliers	120053	1368 (2.16-2.16)
Sidechain outliers	120020	1367 (2.16-2.16)
RSRZ outliers	108989	1262 (2.16-2.16)

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	715	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5578 atoms, of which 6 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 Maltose binding protein-CaKar3 motor domain fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	670	5105	3264	842	986	13	0	3	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-370	MET	-	INITIATING METHIONINE	UNP D8A942
A	-8	ALA	-	LINKER	UNP D8A942
A	-7	ALA	-	LINKER	UNP D8A942
A	-6	ALA	-	LINKER	UNP D8A942
A	-5	GLN	-	LINKER	UNP D8A942
A	-4	THR	-	LINKER	UNP D8A942
A	-3	ASN	-	LINKER	UNP D8A942
A	-2	ALA	-	LINKER	UNP D8A942
A	-1	ALA	-	LINKER	UNP D8A942
A	0	ALA	-	LINKER	UNP D8A942
A	405	LEU	SER	SEE REMARK 999	UNP C4YKK8
A	683	THR	ALA	SEE REMARK 999	UNP C4YKK8

- Molecule 2 is MALTOTETRAOSE (three-letter code: MTT) (formula: C<sub>24</sub>H<sub>42</sub>O<sub>21</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	27	10	5	10	2	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
5	A	1	10	2	6	2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	390	Total 390 O 390	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.46Å 74.93Å 98.67Å 90.00° 96.44° 90.00°	Depositor
Resolution (Å)	19.90 – 2.15 19.90 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.0 (19.90-2.15) 99.0 (19.90-2.15)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.85 (at 2.15Å)	Xtrriage
Refinement program	PHENIX 1.8_1069	Depositor
R, $R_{free}$	0.169 , 0.211 0.170 , 0.212	Depositor DCC
$R_{free}$ test set	1931 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.2	Xtrriage
Anisotropy	0.282	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 53.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5578	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.96% 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: MG, MTT, ADP, EDO

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	0.31	0/5214	0.47	0/7086

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	5105	0	4929	77	0
2	A	45	0	42	1	0
3	A	1	0	0	0	0
4	A	27	0	12	0	0
5	A	4	6	6	0	0
6	A	390	0	0	13	0
All	All	5572	6	4989	77	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 8.

All (77) 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:489:GLU:CB	1:A:498:LEU:HD11	2.01	0.91
1:A:679[A]:LYS:HE2	6:A:1159:HOH:O	1.71	0.89
1:A:-281:LEU:HD23	1:A:-263:PRO:HG2	1.62	0.79
1:A:-327:LEU:CD1	1:A:-310:ILE:HD11	2.11	0.79
1:A:489:GLU:HB3	1:A:498:LEU:HD11	1.64	0.79
1:A:489:GLU:HB2	1:A:498:LEU:HD11	1.67	0.76
1:A:504:ASP:O	1:A:507:THR:HG22	1.88	0.74
1:A:-327:LEU:HD13	1:A:-310:ILE:HD11	1.73	0.71
1:A:666:ASP:OD2	6:A:1151:HOH:O	2.10	0.70
1:A:454:HIS:CE1	1:A:456:THR:HG22	2.29	0.67
1:A:544:ASN:O	1:A:546:LYS:N	2.27	0.67
1:A:-198:GLU:O	1:A:-197:ASN:HB2	1.94	0.67
1:A:-302:GLY:HA3	1:A:-38:ASN:O	1.95	0.66
1:A:679[B]:LYS:HE2	6:A:1159:HOH:O	1.96	0.65
1:A:-298:GLN:HG2	6:A:1115:HOH:O	1.97	0.63
1:A:-325:GLU:O	1:A:-321:GLN:HG3	1.97	0.63
1:A:-327:LEU:HD13	1:A:-310:ILE:CD1	2.28	0.63
1:A:433:GLY:HA3	6:A:1149:HOH:O	2.00	0.61
1:A:416:ASN:HB3	1:A:457:ASN:O	2.04	0.58
1:A:354:ILE:HG12	1:A:660:ILE:HD11	1.86	0.58
1:A:454:HIS:ND1	1:A:455:PRO:HD2	2.19	0.57
1:A:347[A]:ASN:ND2	1:A:348:ILE:HD12	2.20	0.56
1:A:-303:PHE:HB3	1:A:-266:ILE:HD13	1.88	0.55
1:A:390:VAL:HG22	1:A:403:ARG:HG2	1.90	0.53
1:A:395:ILE:O	1:A:398:ASN:N	2.33	0.53
1:A:398:ASN:OD1	1:A:399:PHE:N	2.43	0.51
1:A:-325:GLU:OE2	2:A:701:MTT:H111	2.11	0.51
1:A:-62:GLU:OE2	6:A:1176:HOH:O	2.19	0.51
1:A:431:LEU:HD23	1:A:569:LEU:HD13	1.92	0.51
1:A:-151:LYS:HE3	6:A:908:HOH:O	2.11	0.50
1:A:454:HIS:CE1	1:A:456:THR:CG2	2.96	0.49
1:A:523:THR:O	1:A:524:ASN:HB2	2.13	0.48
1:A:-323:PHE:N	1:A:-322:PRO:HD2	2.28	0.48
1:A:-327:LEU:C	1:A:-327:LEU:HD12	2.34	0.47
1:A:-114:LYS:HD2	6:A:899:HOH:O	2.14	0.47
1:A:-65:LYS:O	1:A:-61:GLU:HG3	2.14	0.47
1:A:634:PRO:HB2	1:A:637:ASN:OD1	2.15	0.47
1:A:454:HIS:ND1	1:A:456:THR:HG22	2.30	0.47
1:A:664:THR:O	1:A:667:LEU:HB2	2.16	0.46
1:A:658:VAL:HG21	1:A:677:ALA:HB2	1.97	0.46
1:A:-322:PRO:HA	1:A:-295:LEU:CD1	2.46	0.46
1:A:439:PHE:CE1	1:A:587:ILE:HD13	2.50	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:352:CYS:HB2	1:A:406:PHE:CE2	2.50	0.46
1:A:559:SER:O	1:A:560:ARG:HD2	2.16	0.45
1:A:-281:LEU:HD22	1:A:-276:TRP:CZ2	2.51	0.45
1:A:-323:PHE:HB3	1:A:-322:PRO:CD	2.47	0.45
1:A:346:GLY:HA2	6:A:1041:HOH:O	2.16	0.45
1:A:-296:GLY:HA2	6:A:1117:HOH:O	2.16	0.45
1:A:670:THR:O	1:A:674:LEU:HG	2.17	0.44
1:A:411:GLU:HG3	1:A:412:ARG:H	1.83	0.44
1:A:-270:ASN:ND2	6:A:1116:HOH:O	2.44	0.44
1:A:431:LEU:CD2	1:A:569:LEU:HD13	2.47	0.44
1:A:348:ILE:HD13	1:A:626:ASN:OD1	2.19	0.43
1:A:-198:GLU:O	1:A:-197:ASN:CB	2.64	0.43
1:A:-278:PHE:CZ	1:A:-260:VAL:HG21	2.54	0.43
1:A:500:ASN:HA	1:A:501:PRO:HD2	1.87	0.43
1:A:-111:VAL:HB	1:A:-41:ILE:HA	2.01	0.43
1:A:396:ASN:HA	1:A:397:ASN:HA	1.65	0.43
1:A:557:HIS:C	1:A:557:HIS:CD2	2.91	0.43
1:A:-114:LYS:HB3	1:A:-114:LYS:HE2	1.90	0.43
1:A:-141:PRO:HA	1:A:-138:TRP:CE2	2.53	0.42
1:A:528:ILE:O	1:A:528:ILE:HG13	2.18	0.42
1:A:559:SER:C	1:A:560:ARG:HD2	2.40	0.42
1:A:-212:TRP:N	1:A:-211:PRO:CD	2.83	0.42
1:A:493:GLU:HG3	6:A:1100:HOH:O	2.20	0.42
1:A:347[A]:ASN:CG	1:A:348:ILE:HD12	2.39	0.42
1:A:-317:THR:O	1:A:-317:THR:HG22	2.18	0.41
1:A:639:LYS:HD3	1:A:639:LYS:HA	1.82	0.41
1:A:-323:PHE:HB3	1:A:-322:PRO:HD3	2.02	0.41
1:A:488:ILE:C	1:A:498:LEU:HD13	2.41	0.41
1:A:491:TYR:OH	1:A:607:GLU:OE2	2.32	0.41
1:A:-211:PRO:HG3	1:A:-113:PRO:HA	2.03	0.41
1:A:344:LEU:HD12	6:A:1004:HOH:O	2.20	0.41
1:A:535:GLN:O	1:A:539:ILE:HD12	2.20	0.41
1:A:392:THR:HA	1:A:400:SER:O	2.20	0.40
1:A:441:TYR:C	1:A:441:TYR:CD1	2.94	0.40
1:A:492:ASN:O	1:A:493:GLU:CB	2.70	0.40

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	663/715 (93%)	643 (97%)	17 (3%)	3 (0%)	31 24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	545	LYS
1	A	-197	ASN
1	A	396	ASN

### 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	522/607 (86%)	518 (99%)	4 (1%)	83 87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-112	PHE
1	A	388	GLU
1	A	626	ASN
1	A	667	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	557	HIS

### 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, 1 is monoatomic - leaving 3 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$
2	MTT	A	701	-	48,48,48	1.57	7 (14%)	71,71,71	1.07	2 (2%)
4	ADP	A	703	3	25,29,29	1.08	1 (4%)	25,45,45	1.70	3 (12%)
5	EDO	A	704	-	3,3,3	0.49	0	2,2,2	0.35	0

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
2	MTT	A	701	-	-	0/20/100/100	0/4/4/4
4	ADP	A	703	3	-	0/12/32/32	0/3/3/3
5	EDO	A	704	-	-	0/1/1/1	0/0/0/0

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	MTT	C13-C14	-4.22	1.41	1.52
2	A	701	MTT	C22-C23	-4.07	1.41	1.52
2	A	701	MTT	C32-C33	-3.48	1.43	1.52
2	A	701	MTT	C42-C43	-3.43	1.43	1.52
2	A	701	MTT	C43-C44	-2.66	1.45	1.52
2	A	701	MTT	O35-C31	2.54	1.48	1.41
2	A	701	MTT	O25-C21	2.79	1.49	1.41
4	A	703	ADP	C5-C4	3.29	1.47	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	703	ADP	N3-C2-N1	-5.83	123.87	128.86
4	A	703	ADP	C4-C5-N7	-2.89	106.61	109.41
4	A	703	ADP	PA-O3A-PB	-2.49	124.27	132.63
2	A	701	MTT	O44-C31-C32	2.14	113.76	108.08
2	A	701	MTT	O34-C21-C22	2.53	114.80	108.08

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
2	A	701	MTT	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 [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	670/715 (93%)	-0.15	17 (2%) 57 65	17, 28, 50, 65	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	378	PRO	4.5
1	A	395	ILE	4.1
1	A	-317	THR	4.1
1	A	396	ASN	4.0
1	A	503	ILE	3.6
1	A	344	LEU	3.4
1	A	399	PHE	2.9
1	A	-199	TYR	2.9
1	A	605	LEU	2.6
1	A	516	ILE	2.6
1	A	660	ILE	2.6
1	A	398	ASN	2.5
1	A	664	THR	2.3
1	A	455	PRO	2.3
1	A	-196	GLY	2.3
1	A	379	GLN	2.2
1	A	-197	ASN	2.0

### 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. 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
5	EDO	A	704	4/4	0.79	0.20	31,44,54,64	0
2	MTT	A	701	45/45	0.89	0.15	16,24,51,55	0
3	MG	A	702	1/1	0.95	0.06	33,33,33,33	0
4	ADP	A	703	27/27	0.95	0.15	26,37,48,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.