



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 6, 2017 – 12:05 PM EST

PDB ID : 4GKR  
Title : Structure of the C-terminal motor domain of Kar3 from *Candida glabrata*  
Authors : Duan, D.; Allingham, J.S.  
Deposited on : unknown  
Resolution : 2.69 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
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-20030345



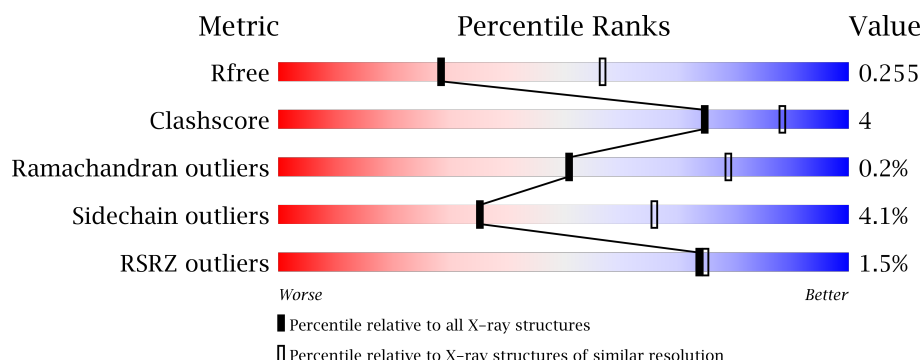
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.69 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	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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	371	 .% 72% 7% 20%
1	B	371	 2% 73% 9% 17%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	A	703	-	-	-	X
4	EDO	B	703	-	-	-	X



## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4486 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 Neck and C-terminal motor domain of Kar3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	0	0
			2156	1373	375	400	8			
1	B	307	Total	C	N	O	S	0	0	0
			2214	1414	383	409	8			

There are 4 discrepancies between the modelled and reference sequences:

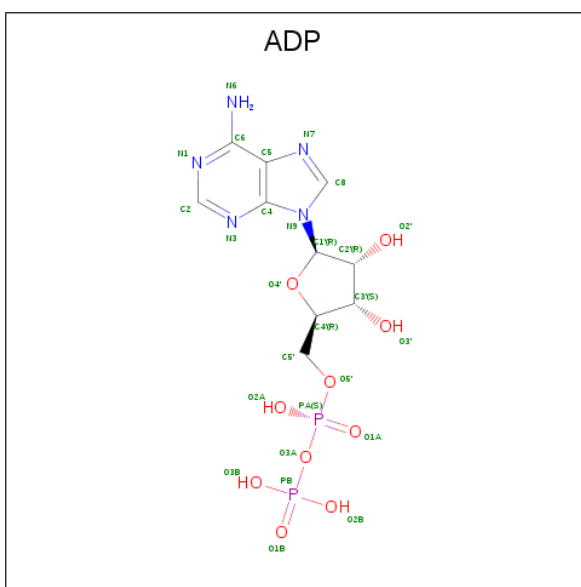
Chain	Residue	Modelled	Actual	Comment	Reference
A	322	MET	-	EXPRESSION TAG	UNP Q6FVW6
A	323	ALA	-	EXPRESSION TAG	UNP Q6FVW6
B	322	MET	-	EXPRESSION TAG	UNP Q6FVW6
B	323	ALA	-	EXPRESSION TAG	UNP Q6FVW6

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		

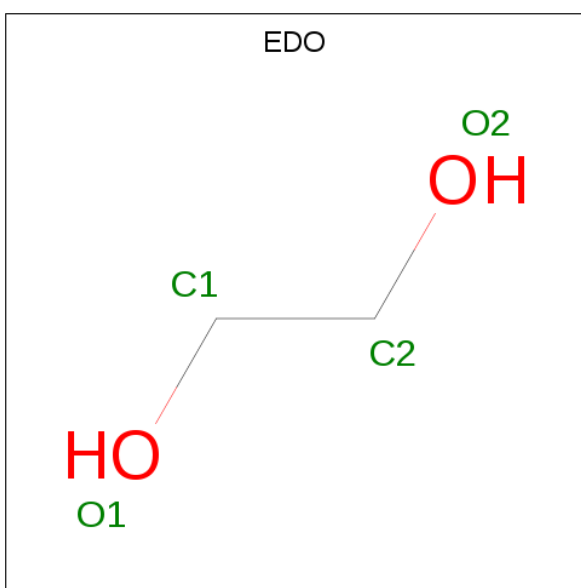
- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>10</sub>P<sub>2</sub>).





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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $\text{C}_2\text{H}_6\text{O}_2$ ).



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



- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	27	Total 27	O 27	0	0
5	B	25	Total 25	O 25	0	0



i

- Molecule 1: Neck and C-terminal motor domain of Kar3





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.71Å 84.07Å 151.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.69 40.60 – 2.69	Depositor EDS
% Data completeness (in resolution range)	95.8 (20.00-2.69) 95.8 (40.60-2.69)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.98 (at 2.69Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.210 , 0.257 0.206 , 0.255	Depositor DCC
$R_{free}$ test set	1057 reflections (5.42%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.5	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 54.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4486	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, EDO, ADP

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	0.57	0/2196	0.64	0/2993
1	B	0.55	1/2254 (0.0%)	0.61	0/3071
All	All	0.56	1/4450 (0.0%)	0.63	0/6064

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	463	TRP	CD2-CE2	5.05	1.47	1.41

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	2156	0	1980	17	0
1	B	2214	0	2014	14	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	27	0	12	0	0
3	B	27	0	12	0	0
4	A	4	0	6	3	0
4	B	4	0	6	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	27	0	0	0	0
5	B	25	0	0	0	0
All	All	4486	0	4030	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 4.

All (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:B:661:VAL:HG21	1:B:680:ALA:HB2	1.51	0.92
1:B:386:THR:HG23	1:B:394:VAL:HG13	1.58	0.84
1:B:371:ILE:HG12	1:B:387:ILE:HD13	1.73	0.71
1:B:482:GLU:OE1	1:B:548:ARG:NH2	2.22	0.71
1:A:661:VAL:HG21	1:A:680:ALA:HB2	1.73	0.70
1:A:661:VAL:HG21	1:A:680:ALA:CB	2.24	0.67
1:A:426:GLY:HA2	1:A:582:GLN:O	2.00	0.60
1:A:486:GLU:HG2	4:A:703:EDO:H11	1.85	0.57
1:A:409:ASN:HD22	1:A:450:ASP:HB3	1.70	0.56
1:B:353:VAL:HG21	1:B:684:ASN:HB3	1.87	0.56
1:A:483:ILE:HG21	4:A:703:EDO:H22	1.87	0.56
1:A:483:ILE:CG2	4:A:703:EDO:H22	2.38	0.54
1:A:460:ILE:O	1:A:464:ILE:HD13	2.10	0.51
1:B:518:LEU:O	1:B:520:THR:HG23	2.10	0.51
1:A:376:PHE:HA	1:A:383:GLN:HG3	1.94	0.50
1:A:422:SER:HB3	1:A:427:TYR:HB2	1.93	0.50
1:B:356:ARG:NH2	1:B:409:ASN:OD1	2.46	0.48
1:B:386:THR:CG2	1:B:394:VAL:HG13	2.37	0.48
1:A:409:ASN:CB	1:A:450:ASP:HB3	2.46	0.46
1:B:424:LEU:HD23	1:B:570:LEU:HD13	1.98	0.46
1:B:387:ILE:HG21	1:B:674:ILE:HD11	1.98	0.43
1:A:564:SER:HB2	1:A:589:ASP:HB3	2.00	0.42
1:B:624:HIS:HD2	1:B:682:LYS:O	2.02	0.42
1:B:564:SER:OG	1:B:589:ASP:HB3	2.20	0.42
1:A:623:ILE:HG21	1:A:683:VAL:HG13	2.01	0.41
1:A:661:VAL:CG2	1:A:680:ALA:HB2	2.46	0.41
1:A:351:ILE:HD12	1:A:627:ASN:OD1	2.20	0.41
1:A:387:ILE:HD11	1:A:677:LEU:HD12	2.02	0.41
1:B:377:ASP:OD2	1:B:380:ASN:HB2	2.21	0.41
1:B:511:GLU:OE1	1:B:513:ARG:NH1	2.54	0.41
1:A:638:PHE:CZ	1:A:652:ILE:HG23	2.57	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	286/371 (77%)	280 (98%)	6 (2%)	0	100	100
1	B	297/371 (80%)	284 (96%)	12 (4%)	1 (0%)	44	73
All	All	583/742 (79%)	564 (97%)	18 (3%)	1 (0%)	51	79

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	607	ARG

### 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	209/332 (63%)	204 (98%)	5 (2%)	54	83
1	B	210/332 (63%)	198 (94%)	12 (6%)	24	51
All	All	419/664 (63%)	402 (96%)	17 (4%)	35	66

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	385	MET
1	A	513	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	518	LEU
1	A	528	THR
1	A	627	ASN
1	B	356	ARG
1	B	366	ASP
1	B	392	SER
1	B	464	ILE
1	B	473	SER
1	B	478	CYS
1	B	513	ARG
1	B	528	THR
1	B	549	SER
1	B	570	LEU
1	B	584	ILE
1	B	658	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	B	624	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 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	ADP	A	702	2	25,29,29	1.13	2 (8%)	24,45,45	1.82	6 (25%)
4	EDO	A	703	-	3,3,3	0.45	0	2,2,2	0.28	0
3	ADP	B	702	2	25,29,29	1.08	1 (4%)	24,45,45	1.82	3 (12%)
4	EDO	B	703	-	3,3,3	0.55	0	2,2,2	0.18	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
3	ADP	A	702	2	-	0/12/32/32	0/3/3/3
4	EDO	A	703	-	-	0/1/1/1	0/0/0/0
3	ADP	B	702	2	-	0/12/32/32	0/3/3/3
4	EDO	B	703	-	-	0/1/1/1	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	702	ADP	O4'-C1'	2.09	1.44	1.41
3	B	702	ADP	C5-C4	2.95	1.47	1.40
3	A	702	ADP	C5-C4	3.40	1.48	1.40

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	702	ADP	N3-C2-N1	-6.68	123.04	128.86
3	A	702	ADP	N3-C2-N1	-6.14	123.51	128.86
3	A	702	ADP	C4-C5-N7	-2.68	106.82	109.41
3	B	702	ADP	C4-C5-N7	-2.47	107.02	109.41
3	A	702	ADP	O3'-C3'-C4'	-2.01	105.20	111.09
3	A	702	ADP	O5'-PA-O1A	-2.01	101.13	109.25
3	A	702	ADP	O3B-PB-O2B	2.14	116.27	107.61
3	A	702	ADP	C4'-O4'-C1'	2.18	112.08	109.77
3	B	702	ADP	C2-N1-C6	2.20	122.61	118.77



There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	703	EDO	3	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	296/371 (79%)	-0.14	2 (0%) 87 88	25, 45, 82, 103	1 (0%)
1	B	307/371 (82%)	-0.19	7 (2%) 61 61	24, 52, 90, 115	0
All	All	603/742 (81%)	-0.17	9 (1%) 74 75	24, 48, 86, 115	1 (0%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	690	THR	3.7
1	B	576	GLY	3.4
1	B	555	ALA	2.9
1	B	609	THR	2.5
1	B	532	ASP	2.5
1	B	608	GLU	2.2
1	A	361	LEU	2.2
1	A	362	PRO	2.2
1	B	464	ILE	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. 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(Å <sup>2</sup> )	Q<0.9
4	EDO	B	703	4/4	0.86	0.30	6.33	52,53,54,59	0
4	EDO	A	703	4/4	0.94	0.24	2.21	45,45,46,47	0
3	ADP	B	702	27/27	0.99	0.15	-0.19	22,33,40,45	0
3	ADP	A	702	27/27	0.98	0.15	-0.58	27,35,41,46	0
2	MG	A	701	1/1	1.00	0.08	-	33,33,33,33	0
2	MG	B	701	1/1	0.98	0.11	-	26,26,26,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.