



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 03:14 PM GMT

PDB ID : 3PUQ  
Title : CEKDM7A from C.Elegans, complex with alpha-KG  
Authors : Yang, Y.; Wang, P.; Xu, W.; Xu, Y.  
Deposited on : 2010-12-06  
Resolution : 2.25 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

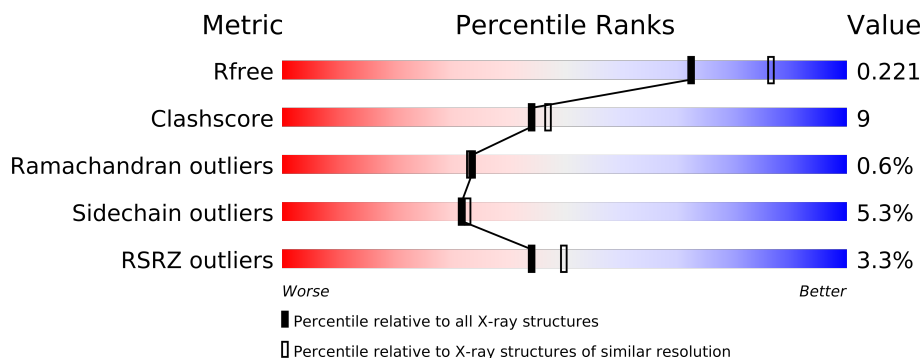
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.25 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))
$R_{free}$	66092	1108 (2.28-2.24)
Clashscore	79885	1326 (2.28-2.24)
Ramachandran outliers	78287	1291 (2.28-2.24)
Sidechain outliers	78261	1291 (2.28-2.24)
RSRZ outliers	66119	1110 (2.28-2.24)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	528	
1	C	528	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	GOL	A	5	-	X
5	GOL	C	5	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8612 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Lysine-specific demethylase 7 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	472	Total	C	N	O	S	0	0	0
			3879	2480	651	721	27			
1	C	471	Total	C	N	O	S	0	0	0
			3872	2476	650	719	27			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	184	GLU	-	EXPRESSION TAG	UNP Q9GYI0
A	185	PHE	-	EXPRESSION TAG	UNP Q9GYI0
A	186	HIS	-	EXPRESSION TAG	UNP Q9GYI0
A	187	MET	-	EXPRESSION TAG	UNP Q9GYI0
C	184	GLU	-	EXPRESSION TAG	UNP Q9GYI0
C	185	PHE	-	EXPRESSION TAG	UNP Q9GYI0
C	186	HIS	-	EXPRESSION TAG	UNP Q9GYI0
C	187	MET	-	EXPRESSION TAG	UNP Q9GYI0

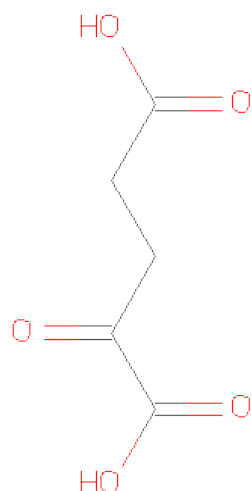
- Molecule 2 is FE (II) ION (three-letter code: FE2) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Fe	0	0
			1	1		
2	C	1	Total	Fe	0	0
			1	1		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

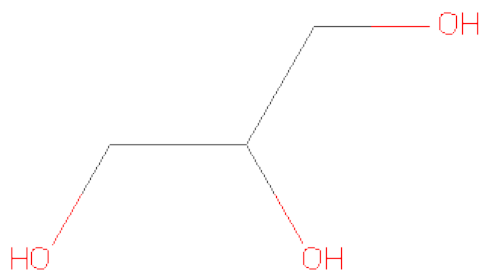
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Zn	0	0
			2	2		
3	C	2	Total	Zn	0	0
			2	2		

- Molecule 4 is 2-OXOGLUTARIC ACID (three-letter code: AKG) (formula:  $C_5H_6O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	5	5		
4	C	1	Total	C	O	0	0
			10	5	5		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	402	Total	O	0	0
			402	402		
6	C	421	Total	O	0	0
			421	421		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.59Å 78.14Å 102.93Å 90.00° 92.07° 90.00°	Depositor
Resolution (Å)	50.00 – 2.25 38.60 – 2.24	Depositor EDS
% Data completeness (in resolution range)	98.1 (50.00-2.25) 99.1 (38.60-2.24)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.09 (at 2.24Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.188 , 0.224 0.183 , 0.221	Depositor DCC
$R_{free}$ test set	3564 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtriage
Anisotropy	0.212	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 42.4	EDS
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 70564 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8612	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.24% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ZN, AKG, FE2

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	0.39	1/3980 (0.0%)	0.54	0/5376
1	C	0.38	0/3973	0.53	0/5365
All	All	0.38	1/7953 (0.0%)	0.54	0/10741

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	247	CYS	CB-SG	-5.94	1.72	1.81

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3879	0	3745	65	0
1	C	3872	0	3738	68	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
3	A	2	0	0	0	0
3	C	2	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	10	0	4	0	0
4	C	10	0	4	2	0
5	A	6	0	8	0	0
5	C	6	0	8	0	0
6	A	402	0	0	5	0
6	C	421	0	0	4	0
All	All	8612	0	7507	132	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

The worst 5 of 132 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:247:CYS:SG	1:C:249:THR:HG23	1.88	1.14
1:A:191:THR:HB	1:A:192:PRO:HD3	1.31	1.09
1:C:513:ILE:HD13	1:C:550:ARG:HH12	1.24	1.00
1:C:369:PRO:HG3	1:C:436:ILE:HG22	1.47	0.94
1:A:247:CYS:SG	1:A:249:THR:HG23	2.14	0.87

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	466/528 (88%)	454 (97%)	9 (2%)	3 (1%)	33	33
1	C	465/528 (88%)	451 (97%)	11 (2%)	3 (1%)	33	33
All	All	931/1056 (88%)	905 (97%)	20 (2%)	6 (1%)	33	33

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	194	GLU
1	C	455	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	455	ASP
1	A	194	GLU
1	A	471	LEU

### 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	423/475 (89%)	402 (95%)	21 (5%)	34	36
1	C	422/475 (89%)	398 (94%)	24 (6%)	29	29
All	All	845/950 (89%)	800 (95%)	45 (5%)	32	33

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	664	MET
1	C	249	THR
1	C	648	ASN
1	C	194	GLU
1	C	258	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	705	ASN
1	C	252	HIS
1	C	444	GLN
1	A	596	HIS
1	A	648	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 10 ligands modelled in this entry, 6 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	$\# Z  > 2$
4	AKG	A	4	2	9,9,9	1.01	0	11,11,11	1.55	3 (27%)
5	GOL	A	5	-	5,5,5	0.40	0	5,5,5	0.38	0
4	AKG	C	4	2	9,9,9	1.05	0	11,11,11	1.91	6 (54%)
5	GOL	C	5	-	5,5,5	0.34	0	5,5,5	0.44	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
4	AKG	A	4	2	-	0/9/9/9	0/0/0/0
5	GOL	A	5	-	-	0/4/4/4	0/0/0/0
4	AKG	C	4	2	-	0/9/9/9	0/0/0/0
5	GOL	C	5	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	4	AKG	C4-C3-C2	3.17	119.82	112.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	4	AKG	O2-C1-C2	2.76	120.44	114.37
4	A	4	AKG	C3-C4-C5	2.65	118.45	113.53
4	C	4	AKG	C3-C4-C5	2.62	118.41	113.53
4	A	4	AKG	O2-C1-O1	-2.43	117.87	123.62

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	472/528 (89%)	-0.02	13 (2%)	50	57	16, 33, 60, 100	0
1	C	471/528 (89%)	-0.04	18 (3%)	38	44	18, 33, 58, 94	0
All	All	943/1056 (89%)	-0.03	31 (3%)	44	50	16, 33, 60, 100	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	606	ARG	9.1
1	C	456	VAL	8.7
1	A	456	VAL	7.9
1	C	606	ARG	7.5
1	C	470	TYR	4.4

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GOL	A	5	6/6	0.29	8.32	29,40,46,49	0
5	GOL	C	5	6/6	0.30	2.93	28,44,47,52	0
4	AKG	A	4	10/10	0.16	0.09	21,26,31,32	0
2	FE2	A	1	1/1	0.16	-0.58	17,17,17,17	0
4	AKG	C	4	10/10	0.13	-0.80	22,26,33,36	0
3	ZN	A	3	1/1	0.07	-0.91	34,34,34,34	0
3	ZN	C	3	1/1	0.06	-1.02	38,38,38,38	0
2	FE2	C	1	1/1	0.13	-1.68	17,17,17,17	0
3	ZN	C	2	1/1	0.06	-1.70	46,46,46,46	0
3	ZN	A	2	1/1	0.05	-1.89	45,45,45,45	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.