



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:57 PM GMT

PDB ID : 4K0E  
Title : X-ray crystal structure of a heavy metal efflux pump, crystal form II  
Authors : Pak, J.E.; Ngonlong Ekende, E.; Vandebussche, G.; Stroud, R.M.; Center for Structures of Membrane Proteins (CSMP)  
Deposited on : 2013-04-03  
Resolution : 3.71 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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

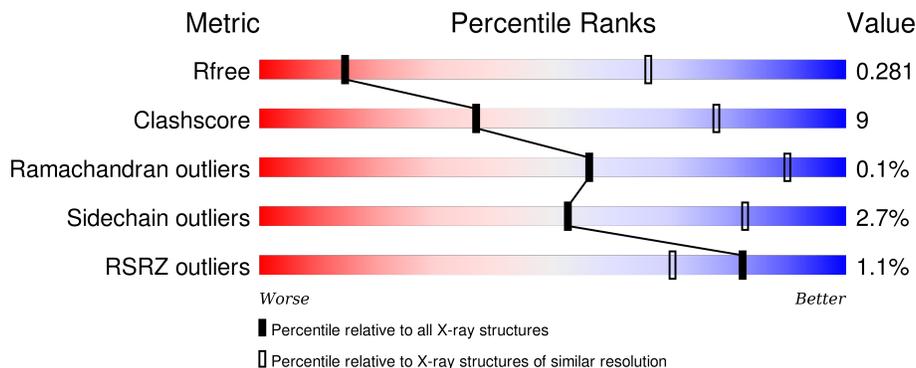
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.71 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	1129 (3.94-3.50)
Clashscore	102246	1252 (3.94-3.50)
Ramachandran outliers	100387	1199 (3.94-3.50)
Sidechain outliers	100360	1197 (3.94-3.50)
RSRZ outliers	91569	1136 (3.94-3.50)

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	1045	 70% 23% • 7%
1	B	1045	 69% 21% • 8%
1	C	1045	 68% 21% • 10%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 22061 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 Heavy metal cation tricomponent efflux pump ZneA(CzcA-like).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	977	Total 7489	C 4809	N 1303	O 1345	S 32	0	0	0
1	B	957	Total 7329	C 4710	N 1271	O 1317	S 31	0	0	0
1	C	942	Total 7239	C 4644	N 1261	O 1302	S 32	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

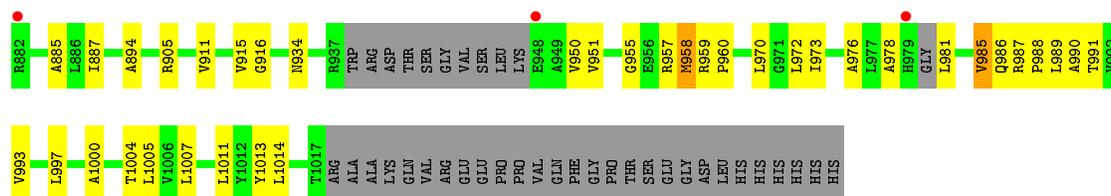
Chain	Residue	Modelled	Actual	Comment	Reference
A	1040	HIS	-	EXPRESSION TAG	UNP Q1LCD8
A	1041	HIS	-	EXPRESSION TAG	UNP Q1LCD8
A	1042	HIS	-	EXPRESSION TAG	UNP Q1LCD8
A	1043	HIS	-	EXPRESSION TAG	UNP Q1LCD8
A	1044	HIS	-	EXPRESSION TAG	UNP Q1LCD8
A	1045	HIS	-	EXPRESSION TAG	UNP Q1LCD8
B	1040	HIS	-	EXPRESSION TAG	UNP Q1LCD8
B	1041	HIS	-	EXPRESSION TAG	UNP Q1LCD8
B	1042	HIS	-	EXPRESSION TAG	UNP Q1LCD8
B	1043	HIS	-	EXPRESSION TAG	UNP Q1LCD8
B	1044	HIS	-	EXPRESSION TAG	UNP Q1LCD8
B	1045	HIS	-	EXPRESSION TAG	UNP Q1LCD8
C	1040	HIS	-	EXPRESSION TAG	UNP Q1LCD8
C	1041	HIS	-	EXPRESSION TAG	UNP Q1LCD8
C	1042	HIS	-	EXPRESSION TAG	UNP Q1LCD8
C	1043	HIS	-	EXPRESSION TAG	UNP Q1LCD8
C	1044	HIS	-	EXPRESSION TAG	UNP Q1LCD8
C	1045	HIS	-	EXPRESSION TAG	UNP Q1LCD8

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
2	B	1	Total 1	Zn 1	0	0
2	A	2	Total 2	Zn 2	0	0
2	C	1	Total 1	Zn 1	0	0







## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	215.34Å 127.07Å 163.33Å 90.00° 93.12° 90.00°	Depositor
Resolution (Å)	19.94 – 3.71 20.02 – 3.71	Depositor EDS
% Data completeness (in resolution range)	(Not available) (19.94-3.71) 85.5 (20.02-3.71)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 3.71Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.229 , 0.283 0.229 , 0.281	Depositor DCC
$R_{free}$ test set	1927 reflections (4.87%)	DCC
Wilson B-factor (Å <sup>2</sup> )	96.7	Xtriage
Anisotropy	0.269	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 48.9	EDS
Estimated twinning fraction	0.017 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.023 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 44707 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	22061	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	122.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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, 0.375 respectively for untwinned datasets, and 0.333, 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:  
ZN

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	0.24	0/7618	0.45	0/10362
1	B	0.23	0/7457	0.44	1/10146 (0.0%)
1	C	0.22	0/7358	0.43	1/9996 (0.0%)
All	All	0.23	0/22433	0.44	2/30504 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	517	LEU	CA-CB-CG	5.41	127.74	115.30
1	B	517	LEU	CA-CB-CG	5.34	127.59	115.30

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	7489	0	7783	153	0
1	B	7329	0	7597	146	0
1	C	7239	0	7499	137	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	22061	0	22879	423	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 9.

The worst 5 of 423 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:70:GLU:HA	1:A:82:MET:HE1	1.62	0.81
1:A:984:ASP:HB3	1:A:987:ARG:HE	1.51	0.75
1:B:623:SER:OG	1:B:624:GLY:N	2.15	0.75
1:C:735:THR:HA	1:C:740:SER:HB3	1.71	0.73
1:B:978:ALA:HA	1:B:987:ARG:HE	1.55	0.71

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	967/1045 (92%)	909 (94%)	57 (6%)	1 (0%)	56 90
1	B	945/1045 (90%)	882 (93%)	61 (6%)	2 (0%)	52 87
1	C	920/1045 (88%)	855 (93%)	64 (7%)	1 (0%)	56 90
All	All	2832/3135 (90%)	2646 (93%)	182 (6%)	4 (0%)	56 90

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	243	VAL
1	B	542	ILE
1	A	841	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	171	VAL

### 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	805/864 (93%)	785 (98%)	20 (2%)	55	84
1	B	785/864 (91%)	765 (98%)	20 (2%)	55	84
1	C	778/864 (90%)	754 (97%)	24 (3%)	47	80
All	All	2368/2592 (91%)	2304 (97%)	64 (3%)	52	83

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	549	TYR
1	B	710	VAL
1	C	745	VAL
1	B	551	ASP
1	B	676	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	558	GLN
1	B	764	ASN
1	C	753	ASN
1	B	597	ASN
1	B	674	ASN

### 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, 4 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	977/1045 (93%)	-0.52	10 (1%) 84 73	59, 104, 178, 228	0
1	B	957/1045 (91%)	-0.48	8 (0%) 87 78	61, 109, 168, 213	0
1	C	942/1045 (90%)	-0.27	15 (1%) 74 61	78, 140, 198, 251	0
All	All	2876/3135 (91%)	-0.43	33 (1%) 82 70	59, 118, 187, 251	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	11	ASN	4.6
1	A	427	GLN	4.0
1	A	502	ALA	3.6
1	C	509	ARG	3.4
1	C	764	ASN	3.4

### 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
2	ZN	C	1101	1/1	0.83	0.17	-0.82	151,151,151,151	0
2	ZN	A	1101	1/1	0.97	0.20	-	107,107,107,107	0
2	ZN	A	1102	1/1	0.90	0.07	-	138,138,138,138	0
2	ZN	B	1101	1/1	0.97	0.08	-	99,99,99,99	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.