



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 07:38 PM GMT

PDB ID : 4MCU  
Title : Crystal structure of disulfide oxidoreductase from *Klebsiella pneumoniae* in reduced state  
Authors : Kurth, F.; Premkumar, L.; Martin, J.L.  
Deposited on : 2013-08-21  
Resolution : 1.99 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>

---

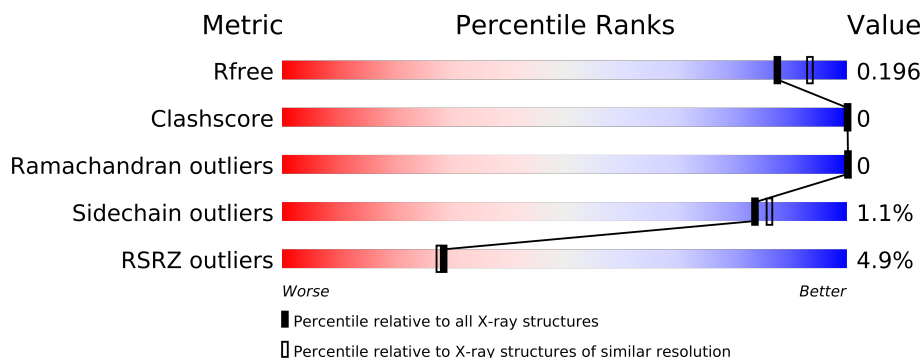
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.99 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	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. 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	190	
1	B	190	
1	C	190	
1	D	190	
1	E	190	
1	F	190	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 17730 atoms, of which 8546 are hydrogens 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 Thiol:disulfide interchange protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	188	Total	C	H	N	O	S	0	0	0
			2881	938	1416	238	282	7			
1	B	189	Total	C	H	N	O	S	0	0	0
			2912	944	1437	240	284	7			
1	C	186	Total	C	H	N	O	S	0	2	0
			2890	943	1420	238	282	7			
1	D	185	Total	C	H	N	O	S	0	0	0
			2850	928	1403	234	278	7			
1	E	190	Total	C	H	N	O	S	5	2	0
			2969	962	1465	247	288	7			
1	F	186	Total	C	H	N	O	S	0	0	0
			2857	931	1405	235	279	7			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	EXPRESSION TAG	UNP B5XZJ6
A	0	ASN	-	EXPRESSION TAG	UNP B5XZJ6
B	-1	SER	-	EXPRESSION TAG	UNP B5XZJ6
B	0	ASN	-	EXPRESSION TAG	UNP B5XZJ6
C	-1	SER	-	EXPRESSION TAG	UNP B5XZJ6
C	0	ASN	-	EXPRESSION TAG	UNP B5XZJ6
D	-1	SER	-	EXPRESSION TAG	UNP B5XZJ6
D	0	ASN	-	EXPRESSION TAG	UNP B5XZJ6
E	-1	SER	-	EXPRESSION TAG	UNP B5XZJ6
E	0	ASN	-	EXPRESSION TAG	UNP B5XZJ6
F	-1	SER	-	EXPRESSION TAG	UNP B5XZJ6
F	0	ASN	-	EXPRESSION TAG	UNP B5XZJ6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	87	Total 87	O 87	0	0
2	B	67	Total 67	O 67	0	0
2	C	83	Total 83	O 83	0	0
2	D	52	Total 52	O 52	0	0
2	E	55	Total 55	O 55	0	0
2	F	27	Total 27	O 27	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.54Å 91.54Å 147.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	53.90 – 1.99 53.94 – 1.99	Depositor EDS
% Data completeness (in resolution range)	99.9 (53.90-1.99) 99.9 (53.94-1.99)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.51 (at 1.98Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.162 , 0.196 0.161 , 0.196	Depositor DCC
$R_{free}$ test set	4810 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.6	Xtriage
Anisotropy	0.298	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 30.6	EDS
Estimated twinning fraction	0.069 for -h,-k,l 0.419 for h,-h-k,-l 0.069 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 94693 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	17730	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.44% 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

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	0.28	0/1496	0.43	0/2032
1	B	0.28	0/1506	0.44	0/2044
1	C	0.28	0/1503	0.43	0/2037
1	D	0.28	0/1478	0.46	0/2007
1	E	0.27	0/1539	0.43	0/2087
1	F	0.26	0/1483	0.44	0/2014
All	All	0.28	0/9005	0.44	0/12221

There are no bond length outliers.

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	1465	1416	0	0	0
1	B	1475	1437	0	0	0
1	C	1470	1420	0	0	0
1	D	1447	1403	0	0	0
1	E	1504	1465	0	0	0
1	F	1452	1405	0	0	0
2	A	87	0	0	0	0
2	B	67	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	83	0	0	0	0
2	D	52	0	0	0	0
2	E	55	0	0	0	0
2	F	27	0	0	0	0
All	All	9184	8546	0	0	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 0.

There are no clashes within the asymmetric unit.

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	186/190 (98%)	182 (98%)	4 (2%)	0	100	100
1	B	187/190 (98%)	184 (98%)	3 (2%)	0	100	100
1	C	185/190 (97%)	182 (98%)	3 (2%)	0	100	100
1	D	183/190 (96%)	172 (94%)	11 (6%)	0	100	100
1	E	190/190 (100%)	189 (100%)	1 (0%)	0	100	100
1	F	184/190 (97%)	177 (96%)	7 (4%)	0	100	100
All	All	1115/1140 (98%)	1086 (97%)	29 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	156/160 (98%)	156 (100%)	0	100	100
1	B	158/160 (99%)	156 (99%)	2 (1%)	80	82
1	C	158/160 (99%)	156 (99%)	2 (1%)	80	82
1	D	155/160 (97%)	153 (99%)	2 (1%)	80	82
1	E	162/160 (101%)	159 (98%)	3 (2%)	69	71
1	F	155/160 (97%)	154 (99%)	1 (1%)	92	94
All	All	944/960 (98%)	934 (99%)	10 (1%)	84	86

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	44	ASP
1	D	101	VAL
1	E	92	PHE
1	C	163	GLN
1	E	87	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	126	ASN
1	D	45	ASN
1	F	126	ASN
1	C	161	ASN
1	D	96	GLN

### 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 ⓘ

There are no ligands in this entry.

## 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	188/190 (98%)	0.09	8 (4%) 34 32	17, 28, 47, 110	0
1	B	189/190 (99%)	0.06	8 (4%) 35 34	21, 32, 56, 82	0
1	C	186/190 (97%)	0.04	3 (1%) 68 69	20, 32, 51, 70	0
1	D	185/190 (97%)	0.16	8 (4%) 34 32	23, 42, 77, 103	0
1	E	190/190 (100%)	0.12	11 (5%) 22 22	28, 39, 62, 89	1 (0%)
1	F	186/190 (97%)	0.27	17 (9%) 9 9	29, 48, 71, 112	0
All	All	1124/1140 (98%)	0.12	55 (4%) 28 28	17, 37, 66, 112	1 (0%)

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	116	VAL	3.8
1	A	185	VAL	3.5
1	F	107	ILE	3.5
1	D	3	ILE	3.3
1	F	87	ILE	3.2

### 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 ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.