



# wwPDB X-ray Structure Validation Summary Report

Jul 22, 2014 – 12:05 PM EDT

PDB ID : 4OCF  
Title : Crystal structure of the disulfide oxidoreductase DsbA (S30XXC33) active site mutant from *Proteus mirabilis*  
Authors : Kurth, F.; Martin, J.L.  
Deposited on : 2014-01-09  
Resolution : 1.98 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>

---

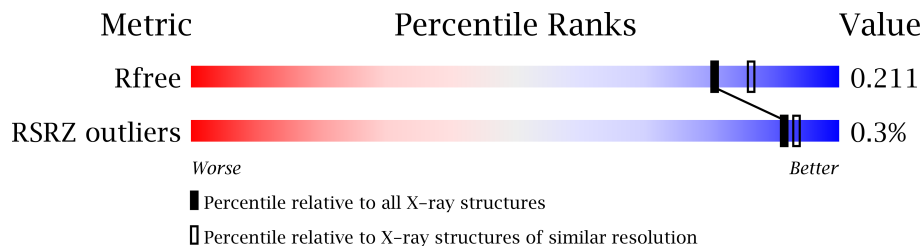
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : **FAILED**  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23489  
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) : stable23489

# 1 Overall quality at a glance

The reported resolution of this entry is 1.98 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	6577 (2.00-1.96)
RSRZ outliers	66119	6578 (2.00-1.96)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

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

Mol	Type	Chain	Res	Geometry	Electron density
2	SCN	A	203	-	X
2	SCN	B	201	-	X
2	SCN	B	204	-	X
2	SCN	D	201	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12410 atoms, of which 5732 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.

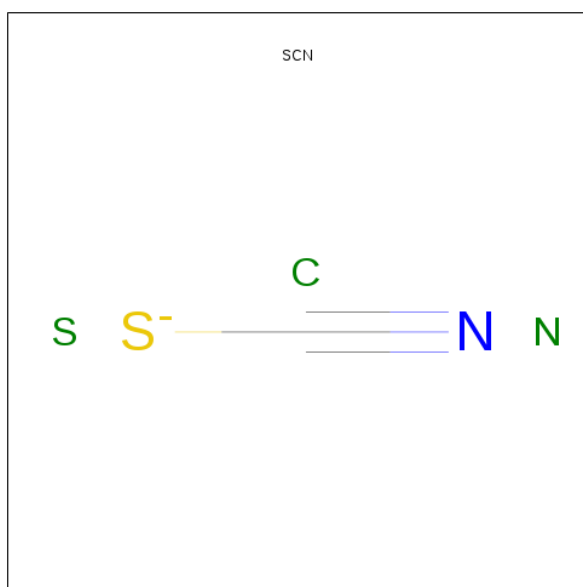
- Molecule 1 is a protein called Thiol:disulfide interchange protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	187	Total	C	H	N	O	S	0	0	0
			2874	923	1420	240	285	6			
1	B	188	Total	C	H	N	O	S	0	1	0
			2923	934	1452	243	287	7			
1	C	187	Total	C	H	N	O	S	0	0	0
			2881	926	1423	241	285	6			
1	D	187	Total	C	H	N	O	S	0	0	0
			2895	926	1437	241	285	6			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	EXPRESSION TAG	UNP B4EZ68
A	0	ASN	-	EXPRESSION TAG	UNP B4EZ68
A	1	ALA	-	EXPRESSION TAG	UNP B4EZ68
A	30	SER	CYS	ENGINEERED MUTATION	UNP B4EZ68
B	-1	SER	-	EXPRESSION TAG	UNP B4EZ68
B	0	ASN	-	EXPRESSION TAG	UNP B4EZ68
B	1	ALA	-	EXPRESSION TAG	UNP B4EZ68
B	30	SER	CYS	ENGINEERED MUTATION	UNP B4EZ68
C	-1	SER	-	EXPRESSION TAG	UNP B4EZ68
C	0	ASN	-	EXPRESSION TAG	UNP B4EZ68
C	1	ALA	-	EXPRESSION TAG	UNP B4EZ68
C	30	SER	CYS	ENGINEERED MUTATION	UNP B4EZ68
D	-1	SER	-	EXPRESSION TAG	UNP B4EZ68
D	0	ASN	-	EXPRESSION TAG	UNP B4EZ68
D	1	ALA	-	EXPRESSION TAG	UNP B4EZ68
D	30	SER	CYS	ENGINEERED MUTATION	UNP B4EZ68

- Molecule 2 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	S	0	0
			3	1	1	1		
2	A	1	Total	C	N	S	0	0
			3	1	1	1		
2	A	1	Total	C	N	S	0	0
			3	1	1	1		
2	B	1	Total	C	N	S	0	0
			3	1	1	1		
2	B	1	Total	C	N	S	0	0
			3	1	1	1		
2	B	1	Total	C	N	S	0	0
			3	1	1	1		
2	C	1	Total	C	N	S	0	0
			3	1	1	1		
2	D	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	195	Total	O	0	0
			195	195		
3	B	227	Total	O	0	0
			227	227		
3	C	203	Total	O	0	0
			203	203		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	185	Total 185	O 185	0	0

### 3 Residue-property plots

MolProbity failed to run properly - this section will therefore be empty.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	37.99Å 80.21Å 114.45Å 90.00° 91.06° 90.00°	Depositor
Resolution (Å)	57.22 – 1.98 57.22 – 1.98	Depositor EDS
% Data completeness (in resolution range)	99.7 (57.22-1.98) 98.7 (57.22-1.98)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.05 (at 1.98Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1069)	Depositor
R, $R_{free}$	0.152 , 0.212 0.153 , 0.211	Depositor DCC
$R_{free}$ test set	1989 reflections (4.20%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.1	Xtriage
Anisotropy	0.434	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 25.6	EDS
Estimated twinning fraction	0.070 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 47846 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12410	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.20% 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 ⓘ

MolProbity failed to run properly - this section will therefore be empty.

### 5.2 Close contacts ⓘ

MolProbity failed to run properly - this section will therefore be empty.

### 5.3 Torsion angles

#### 5.3.1 Protein backbone ⓘ

MolProbity failed to run properly - this section will therefore be empty.

#### 5.3.2 Protein sidechains ⓘ

MolProbity failed to run properly - this section will therefore be empty.

#### 5.3.3 RNA ⓘ

MolProbity failed to run properly - this section will therefore be empty.

### 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 ⓘ

9 ligands are modelled in this entry.

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	SCN	A	201	-	2,2,2	1.51	1 (50%)	1,1,1	0.21	0
2	SCN	A	202	-	2,2,2	1.53	1 (50%)	1,1,1	0.37	0
2	SCN	A	203	-	2,2,2	1.98	1 (50%)	1,1,1	0.49	0
2	SCN	B	201	-	2,2,2	1.71	1 (50%)	1,1,1	0.15	0
2	SCN	B	202	-	2,2,2	1.45	1 (50%)	1,1,1	0.15	0
2	SCN	B	203	-	2,2,2	1.48	1 (50%)	1,1,1	0.02	0
2	SCN	B	204	-	2,2,2	1.80	1 (50%)	1,1,1	0.60	0
2	SCN	C	201	-	2,2,2	1.73	1 (50%)	1,1,1	0.32	0
2	SCN	D	201	-	2,2,2	1.80	1 (50%)	1,1,1	0.31	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	SCN	A	201	-	-	0/0/0/0	0/0/0/0
2	SCN	A	202	-	-	0/0/0/0	0/0/0/0
2	SCN	A	203	-	-	0/0/0/0	0/0/0/0
2	SCN	B	201	-	-	0/0/0/0	0/0/0/0
2	SCN	B	202	-	-	0/0/0/0	0/0/0/0
2	SCN	B	203	-	-	0/0/0/0	0/0/0/0
2	SCN	B	204	-	-	0/0/0/0	0/0/0/0
2	SCN	C	201	-	-	0/0/0/0	0/0/0/0
2	SCN	D	201	-	-	0/0/0/0	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	203	SCN	C-S	2.77	1.81	1.63
2	D	201	SCN	C-S	2.52	1.79	1.63
2	B	204	SCN	C-S	2.49	1.79	1.63
2	C	201	SCN	C-S	2.39	1.78	1.63
2	B	201	SCN	C-S	2.38	1.78	1.63

There are no bond angle outliers.

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	187/190 (98%)	-0.49	1 (0%) 88 91	3, 9, 22, 36	0
1	B	188/190 (98%)	-0.49	0 100 100	3, 9, 22, 35	0
1	C	187/190 (98%)	-0.41	0 100 100	3, 9, 24, 35	1 (0%)
1	D	187/190 (98%)	-0.35	1 (0%) 88 91	3, 12, 28, 40	1 (0%)
All	All	749/760 (98%)	-0.44	2 (0%) 91 93	3, 10, 25, 40	2 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	ALA	3.5
1	D	188	LYS	2.6

### 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
2	SCN	A	203	3/3	0.23	14.85	16,16,23,29	0
2	SCN	B	201	3/3	0.17	7.57	18,18,27,38	0
2	SCN	B	204	3/3	0.11	4.69	13,13,18,24	0
2	SCN	D	201	3/3	0.21	2.06	21,21,26,36	0
2	SCN	A	201	3/3	0.10	1.88	9,9,11,13	0
2	SCN	A	202	3/3	0.09	1.36	8,8,11,16	0
2	SCN	B	203	3/3	0.07	0.01	13,13,13,19	0
2	SCN	C	201	3/3	0.07	-0.15	9,9,13,14	0
2	SCN	B	202	3/3	0.07	-0.37	4,4,8,10	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.