Research Article

Exploring the role of SoxA and SoxX in sulphur oxidation in *Allochromatium vinosum* through Protein-protein docking: An in silico approach

Sujay Ray, Angshuman Bagchi*

Department of Biochemistry and Biophysics University of Kalyani
Kalyani, Nadia India

*Corresponding Author E-mail: angshuman_bagchi@yahoo.com

Abstract:

Thiosulphate ($S_2O_3^{2-}$) is a stable and environmentally abundant sulphur compound of intermediate oxidation state and fulfils an important role in the natural sulphur cycle. There are main two types of thiosulfate-oxidizing Sox enzyme system type one group 1 forms sulfur globules as intermediates *Allochromatium vinosum* (A.vino or A.vinosum)), group 2 which does not form sulphur globules as intermediate form for example *Paracoccus pantotrophus*. Sox genes in A. vinosum, are separated into three gene clusters. Cluster one comprises Alvin_2108 to Alvin_2112. The second gene cluster extends between Alvin_2165 and Alvin_2167. The third gene cluster includes Alvin_2168 to Alvin_2182. Where SoxX, SoxA, are encoded by Alvin_2168 to Alvin_2169. In the present work, homology modeling has been used to build the three dimensional structures of SoxA, and SoxX. With the help of protein-protein docking and Protein Interaction Calculator (PIC) sever the amino acid residues of these proteins involved in the interactions have been identified. The interactions between the SoxA, and SoxX proteins are mediated mainly through hydrogen bonding, hydrophobic interaction, electrostatic interaction. Possible mechanisms of interaction between the SoxA, and SoxX have identified in spite the absence of SoxK.

Key words: Homology modeling, Protein-protein interactions, Docking simulations _ Environmental sulphur balance, Homology modeling, Sox operon, Sulfur oxidation

Introduction:

Fundamental way to maintain the environmental sulphur balance is the Microbial redox reactions of inorganic sulphur. In general these reactions are carried out by different types of microorganism. [1]. Thiosulphate ($S_2O_3^{2-}$) is a rather stable and environmentally abundant sulphur compound of intermediate and is mostly used by chemotrophic and prototrophic sulphur oxidizer bacteria [2]. Sox genes are mainly responsible for carrying out such type of redox reaction. The organism *Allochromatium vinosum* is a Gram-negative proteobacteria that belongs to the family Chromaticeae and is an ideal model organism for studying thiosulphate oxidation [3].

SoxA, & SoxX of A.vino are proteins with 281, and 128 amino acid residues, respectively. However, to date the detailed structural information regarding the interactions between these two proteins have not been properly understood. We here report the further examination of protein-protein interaction of the two Sox encoded proteins SoxA and SoxX on the thiosulfate oxidation in the phototrophic sulfur oxidizing proteobacterium A. Vinosum. In the present study, the three dimensional structures of SoxA & SoxX from A.vino obtained by homology modeling have been described. Molecular docking simulations have been performed in order to find out the possible modes of binding of these proteins. Binding sites of SoxA, SoxX have been predicted and analyzed. These studies provide a detailed and rational structural insight into the plausible molecular mechanism

of the involvements of these proteins in the global sulfur oxidation reaction cycle.

Materials and methods

Sequence analysis and homology modelling:

The amino acid sequences of SoxA, and SoxX proteins of A.vino were obtained from NCBI nucleotide database (Acc. No. NC_013851). These amino acid sequences were used separately to build homology models by Modweb. Modweb is a web server for protein structure modelling using comparative modelling approach. Models were built for each one of the sequence-structure matches using modeller. Modelled residues of SoxX and SoxA were 102 and 260 residue length. The model structures of SoxA and SoxX were based on crystal structures of 1h32A (~30%) and 3oa8B with 30% sequence identity. The structure of SoxX protein were refined by the chiron and followed by koba energy minimization server [4, 5, 6]. The modelled structures were then superimposed separately on each of the crystal templates without altering the coordinate systems of atomic positions in the respective templates. The root mean square deviations (RMSD) for the superimpositions were 1.35 Å for SoxX and 3.708 Å for SoxA.

The PROSA web server was used to calculate Z scores [7]. The result showed that the predicted homology models were well inside the range of typical native structures [8]. PROCHECK analyses were performed in order to assess the stereo chemical qualities of the models and

Ramachandran plots [9,10] were drawn. No residues were found to be present in the disallowed regions of the Ramachandran plots.

Molecular docking simulations:

In order to study the interactions between SoxA and SoxX proteins the models of the SoxA and SoxX proteins were docked using the software ClusPro. ClusPro is fully automated web server for protein-protein docking [11]. The two modelled protein structures of SoxA and SoxX were uploaded through the ClusPro webserver. Using advance option through ClusPro web server unstructured terminus from receptor and ligand were removed. The docked structure of the SoxAX complex that yielded the best clustering size among all the other possible docked structures were selected and the model of the complex protein was then energy minimized using steepest decent technique by fixing the backbone of two proteins in the complex structure to ensure proper interactions. All energy minimizations were done with CHARMM force fields [12] using the program Discovery studio until the structures reached the final RMS gradient of 0.1.

Calculation of protein-protein interactions:

To find out the interactions between the SoxA, SoxX proteins, PIC web server were used. This web server were designed to calculate various kinds of interactions; such as disulphide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic- aromatic interactions, aromatic-sulphur interactions and cation - π interactions within a protein or between proteins in a complex [13].

Results:

Description of the structure of SoxA:

The modelled structure of SoxA is a 260 amino acid residue long protein. The secondary structures of the protein were predicted by PSIPRED. The algorithm uses a protein sequence or multiple alignment of protein sequences as query sequence input and predicts the each amino acid residue into either alpha helix ('H'), β sheet ('E') or random coil ('C') secondary structures with certain confidence for each residue [14]. The protein is made of helix, sheet and coil. There are mainly nine regions (amino acid residues 9-22, 30-31, 40-50, 55-66, 76-79, 103-113, 151-168, 215-225 & 236-249) predicted as helical, one region (amino acid residues 92-94) predicted as β sheet and rest of the structure predicted as coil. The structure is presented in Fig. S1.

Description of the structure of SoxX:

The modelled structure of SoxX is a 102 amino acid residue long protein. The secondary structures of the protein were predicted by PSIPRED [14]. The algorithm

Fig S1: Model structure of SoxA protein from A. Vinosum. With distinct secondary structure showing as alpha helix, beta structure and random coil.

Fig S2: Model structure of SoxX protein from A. Vinosum. With distinct secondary structure showing as alpha helix, beta structure and random coil

Fig 1: Interaction of SoxA (green) and SoxX (red) are shown in the complex.

Table 1: Protein-protein Side chain-Side chain hydrogen bonds where X represents the SoxA protein chain and A represents SoxX protein chain.

POSITION	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	M O	Dd-a	Dh-a	A(d-H-N)	A(a-O=C)
112	X	ARG	NH1	14	A	ASP	OD2	1	2.87	3.58	40.97	999.99
112	X	ARG	NH1	14	A	ASP	OD2	2	2.87	2.57	96.80	999.99
112	X	ARG	NH2	14	A	ASP	OD1	1	2.59	2.79	67.92	999.99
112	X	ARG	NH2	14	A	ASP	OD1	2	2.59	2.00	113.15	999.99
112	X	ARG	NH2	14	A	ASP	OD2	1	2.82	3.49	43.20	999.99
112	X	ARG	NH2	14	A	ASP	OD2	2	2.82	2.53	95.50	999.99
115	X	ARG	NE	14	A	ASP	OD1	-	2.90	2.76	88.26	999.99
115	X	ARG	NH2	24	A	ASP	OD1	1	3.18	2.83	99.79	999.99
115	X	ARG	NH2	24	A	ASP	OD1	2	3.18	2.81	101.59	999.99
165	X	GLU	OE1	29	A	ASN	ND2	1	3.11	3.63	53.18	999.99
165	X	GLU	OE1	29	A	ASN	ND2	2	3.11	3.59	56.48	999.99
165	X	GLU	OE2	29	A	ASN	ND2	1	2.86	3.41	51.06	999.99
165	X	GLU	OE2	29	A	ASN	ND2	2	2.86	3.13	65.83	999.99
19	A	ARG	NE	110	X	TYR	OH	-	2.54	2.46	81.24	999.99
19	A	ARG	NH2	110	X	TYR	OH	1	2.62	2.27	97.41	999.99
29	A	ASN	ND2	165	X	GLU	OE1	1	3.11	2.48	117.23	999.99
29	A	ASN	ND2	165	X	GLU	OE1	2	3.11	3.43	63.46	999.99
29	A	ASN	ND2	165	X	GLU	OE2	1	2.86	2.90	77.15	999.99
29	A	ASN	ND2	165	X	GLU	OE2	2	2.86	2.58	94.93	999.99

uses a protein sequence or multiple alignment of protein sequences as query sequence input and predicts the each amino acid residue into either alpha helix ('H'), beta sheet ('E') or random coil ('C') secondary structures with certain confidence for each residue [14]. The protein is made up helix and coil regions. There are mainly four regions (amino acid residues 15-22, 52-54, 59-68, 89-101) predicted as helical and rest of the structure predicted as coil. The structure is presented in Fig. S2.

Interaction of SoxA with SoxX:

In order to find the interactions between the proteins the three dimensional coordinates of the proteins were docked by the software tool ClusPro. SoxA and SoxX are found to interact strongly with each other. The protein-protein interface is found to mainly contain the polar amino acid residues. There are also hydrophobic interactions between two proteins.

There are extensive H-bonding interactions involving both the main and the side chains of the two protein molecules. Apart from this there are protein-protein ionic and cation - π interaction

Discussion:

In this study, an attempt has been made to elucidate the structural basis of the involvements of SoxA, SoxX, in binding. For that matter the three-dimensional structures of the proteins SoxA, SoxX, have been built and analysed. Since there have been no previous reports regarding the structural biology of these proteins, results from this study may give a new way to understand the three dimensional structures of SoxA, SoxX as well as to elucidate the structural basis of the molecular functions of these proteins. This model provides a rational framework for designing experiments to determine the

contribution of the various amino acid residues in these proteins to predict the molecular basis of their interactions.

Acknowledgement

The authors would like to thank the DST -PURSE programme 2012-2015 going on in the department of Biochemistry and Biophysics, University of Kalyani for providing different instrumental and infrastructural support. The support from the Bioinformatics Infrastructural Facility, Department of Biochemistry and Biophysics, University of Kalyani is duly acknowledged.

References:

1. Freidrich CG, Bardischewsky F, Rother D, Quentmeier A, Fischer J (2005) Prokaryotic sulfur oxidation. *Curr Opin Micro-biol* 8:253–259
2. Bagchi A (2012) Structural insight into the mode of interactions of SoxL from *Allochrochromatium vinosum* in the global sulfur oxidation cycle. *Mol Biol Rep* 39:10243–10248
3. Hensen D, Sperling D, Trüper HG, Brune DC, Dahl C (2006) Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochrochromatium vinosum*. *Mol Microbiol* 62:794–810
4. Ramachandran, S., Kota, P., Ding, F. and Dokholyan (2011), N. V., *PROTEINS: Structure, Function and Bioinformatics*, 79: 261-270
5. Chopra et al. (2010) Consistent refinement of submitted models at CASP using a knowledge-based potential. *Proteins* vol. 78 (12) pp. 2668-78
6. Chopra et al. (2008) Solvent dramatically affects protein structure refinement. *Proc Natl Acad Sci USA* vol. 105 (51) pp. 20239-44
7. Sippl MJ (1993) Recognition of errors in three-dimensional structures in proteins. *Proteins* 17:355–362
8. Wiederstein M (2004) *Evolutionary methods in biotechnology*. Wiley-VCH, Weinheim
9. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemistry of protein structures. *J Appl Crystallography* 26:283–291
10. Ramachandran GN, Sashisekharan V (1968) Conformation of polypeptides and proteins. *Adv Protein Chem* 23:283–438
11. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: a fully automated algorithm for protein-protein docking *Nucleic Acids Research*. 2004 Jul 1
12. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comp Chem* 4:187–217
13. K. G. Tina, R. Bhadra and N. Srinivasan, PIC: Protein Interactions Calculator, *Nucleic Acids Research*, 2007, Vol. 35, Web Server issue W473–W476. Download
14. Jones DT. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292: 195-202