ELSEVIER

Journal of Molecular Structure: THEOCHEM 758 (2006) 113–118

THEO CHEM

www.elsevier.com/locate/theochem

Structural identification of a novel thioredoxin SoxS: Prediction of the function in the process of transport of reductants during sulfur oxidation by the novel global sulfur oxidation reaction cycle

Angshuman Bagchi *, Tapash Chandra Ghosh

Bioinformatics Center, Bose Institute, AJC Bose Centenary Building, P1/12 CIT Scheme VIIM, Kankurgachi, Kolkata, West Bengal 700 054, India

Received 10 August 2005; accepted 10 October 2005
Available online 27 December 2005

Abstract

Microbial redox reactions involving inorganic sulfur compounds, mainly the sulfur anions, are one of the vital reactions responsible for the environmental sulfur balance. These reactions are mediated by phylogenetically diverse prokaryotes, some of which also take part in the extraction of metal ions from their sulfur containing ores. These sulfur oxidizers oxidize inorganic sulfur compounds like sulfide, thiosulfate, etc. to produce reductants that are used for carbon dioxide fixation or in respiratory electron transport chains. The sulfur-oxidizing gene cluster (*sox*) of α-Proteobacteria comprises of at least 15 genes, forming two transcriptional units, viz., *soxSR* and *soxVWXYZABCDEFGH*. SoxS is a periplasmic thioredoxin and an essential component of *sox* operon. It is required for optimal expression of the *sox* gene cluster. All thioredoxins are involved in interaction with DNA polymerase. We have employed homology modeling to construct the three-dimensional structure of the SoxS protein from *Rhodovulum sulfidophilum*. With the help of docking and molecular dynamics studies we have identified the amino acid residues of the protein involved in the interaction with DNA polymerase to structurally classify SoxS as a thioredoxin. The probable biochemical mechanism of the involvement of the protein in sulfur oxidation has also been investigated. Our study provides a rational basis to interpret the structural classification of SoxS as a thioredoxin and thereby to predict the possible molecular mechanism of the regulation of sulfur anion oxidation reactions by these ecologically important organisms.

Keywords: Sulfur oxidation; Homology modeling; Molecular dynamics; Protein–protein interaction

1. Introduction

Microbial redox reactions of sulfur are mainly responsible for cycling of this element in the environment. Sulfur has a unique range of oxidation states that varies from +6 to −2 and as a result of which several important biological processes involving transformations of sulfur from one form to other have been evolved. Sulfur based chemo- or photolithotrophy is one of such processes in which electron transfer from reduced sulfur compounds is used by phylogenetically diverse prokaryotes [1]. These sulfur utilizing bacteria oxidize inorganic sulfur compounds like sulfide, thiosulfate, etc. to produce reductants that are used for carbon dioxide fixation or in respiratory electron transfer chains [2]. The transformation of the sulfur compounds from one form to the other is a major component of the biogeochemical sulfur cycle. Besides some of these bacteria (though they may not contain the *sox* gene cluster) also take part in the extraction of metals from their sulfur containing ores [3,4]. Nevertheless, the biochemical mechanism of the bacterial sulfur oxidation process or its regulation is, in general, still poorly understood.

The sulfur-oxidizing gene cluster (*sox*) of α-Proteobacteria comprises of at least **15** genes, which form two transcriptional units, viz. ***soxSR*** and *soxVWXYZABCDEFGH*. Recent studies with both chemo- and photolithotrophic α-Proteobacteria such as *Paracoccus pantotrophus* (Para) and *Rhodovulum sulfidophilum* (Rsulf) revealed that multiple-gene cluster, *shxVW* (*soxVW*) and *soxXYZABCDEFGH*, is associated with the metabolism sulfur anions [5–8]. SoxXA, SoxYZ, SoxB and SoxCD are required for sulfur-dependent cytochrome *c* reduction. The eight-electron oxidation of a molecule of thiosulfate is governed by cytochrome *c* complex multienzyme system (TOMES) encoded by *soxXYZABCD*. The enzyme system including ShxV and ShxW, which were proposed to be involved in the biogenesis of cytochrome *c*, were shown to be

* Corresponding author. Tel.: +91 33 2334 6626; fax: +91 33 2334 3886.
E-mail address: angshu@bic.boseinst.ernet.in (A. Bagchi).

0166-1280/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2005.10.023

inducible by thiosulfate [5–8]. Two ORFs, ORF2 and ORF1, were identified immediate upstream of *shxV* but were located in the opposite direction of *sox* structural genes. These two ORFs predict gene products (SoxS and SoxR, respectively) to be a periplasmic thioredoxin (SoxS) and a transcriptional regulator of ArsR family (SoxR), respectively [9,10]. SoxS of Rsulf is a protein with 135 amino acid residues [9]. It has been observed that SoxS is an essential periplasmic component required for the optimal expression *sox* gene cluster [11]. Though a considerable progress in the genetics of sulfur lithotrophy is noted, molecular mechanism of regulation of *sox* gene expression has not yet been addressed. In the present study, we seek to understand the potential of SoxS protein as the mediator of transport of reductants in the *sox* operon. We describe the three-dimensional structure of SoxS obtained by homology modeling. Since SoxS is a thioredoxin and generally thioredoxins interact strongly with DNA polymerase [12], we have used molecular docking and molecular dynamics in order to structurally classify SoxS as a thioredoxin. The redox active binding site of SoxS has been predicted, analyzed and compared to that of previously reported experimental results. These studies provide detailed structural information on the binding of SoxS as well as its mode of action. As this is the first report regarding the structural basis of the involvements of SoxS in the process of biochemical oxidation of sulfur anions, our studies may contribute towards the understanding of the molecular mechanism of sulfur anion oxidation by these ecologically important microbial species.

2. Material and methods

2.1. Sequence analysis and homology modeling of monomeric SoxS

The amino acid sequence of SoxS of Rsulf was obtained from Entrez database (Accession No. AAO11779). The amino acid sequence was used to search Brookhaven protein data bank (PDB) [13] using the software BLAST [14] for finding suitable template for homology modeling. Structures having more than 25% sequence identity with the target sequence of SoxS were selected. A structural alignment was performed among these structures to identify the core region and the thioredoxin motifs and the top three structures with low Cα root mean square deviation (rmsd) values for the core thioredoxin motif were selected as templates. The sequence of SoxS was also used to search structural databases using FUGUE server [15]. Similar results as obtained from BLAST were again found. The templates selected in this way were thioredoxin-2 (pdb code: 1THX) from Anabaena, human thioredoxins (pdb code: 3TRX and 1ERV) and thioredoxin from *Escherichia coli* (pdb code: 2TRX). The protein was modeled using the corresponding crystal structures as templates. Homology modeling was performed using the program Homology of Insight II [Accelrys, San Diego, CA, USA] on a Silcon Graphics Indigo II workstation.

Modeled structure was then superimposed separately on each of the crystal templates without altering the coordinate system of atomic positions in the respective templates. The mean rms deviations for the superimpositions were 0.2, 0.7, 0.9, and 0.7 Å on 1THX, 3TRX, 1ERV and 2TRX, respectively.

The models were energy minimized fixing the backbones to ensure proper interactions. Conjugate gradient (CG) method was employed for minimization with the consistent valence force field (CVFF) [16] using the program DISCOVER until all the structures reached the final derivative of 0.001 kcal/mol.

2.2. Validation of the model

Regarding the main chain properties of the modeled protein, no considerable bad contacts neither Cα tetrahedron distortion nor hydrogen bond energy problems were found. Moreover, the average *G* factor, the measure of the normality degree of the protein properties, was of −1.04, which is inside the permitted values for homology models. Furthermore, there were no distortions of the side chain torsion angles found. The *Z*-scores calculated using the software PROSA 2003 [17] showed that the predicted model of the protein were well inside the range of a typical native structure [18]. The residue profiles of the three-dimensional model were further checked by VERIFY3D [19]. PROCHECK [20] analysis was performed in order to assess the stereo-chemical qualities of the three-dimensional model and Ramachandran plots [21] were drawn.

2.3. Docking and molecular dynamics calculations

It is well known that thioredoxin proteins interact with DNA polymerase [12]. To study the interaction between SoxS and DNA polymerase, the coordinates of the DNA polymerase bound to thioredoxin (pdb code: 1TKD) [22] protein were extracted from PDB. The coordinates of the thioredoxin motif of modeled SoxS were superimposed on the corresponding motif of the thioredoxin protein bound to DNA polymerase in 1TKD. The rmsd of the superimposition was 0.5 Å. The model of SoxS was subsequently merged with the crystal structure of 1TKD and the thioredoxin protein bound to the DNA polymerase was then removed to form a SoxS–DNA polymerase complex. The model of SoxS was also docked to DNA polymerase using the software GRAMM [23], DOT [24] and ZDOCK [25,26], using the ClusPro server [27,28] and also with Patchdock server [29] in order to get a comprehensive result. The docked structures, that yielded the best score were selected and analyzed visually using the software Insight II. Molecular dynamics (MD) simulations were performed on the docked structures to predict the favorable binding interactions. The docked structures were solvated with an average of 2000 simple point charges [30] water molecules. The system was minimized initially keeping the water and the backbones of the proteins fixed. In the next step of minimization, the protein complex was kept fixed and the water molecules were allowed to move. The first few rounds of minimizations were performed by steepest descent (SD) method and then CG method was employed. The minimized system was equilibrated for a period of 10 ps with positional restraints. Then a 120 ps MD run was