doi:10.1016/j.jmb.2007.07.051 J. Mol. Biol. (2007) 372, 1293–1304

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Putative DNA-Binding Protein Sto12a from the Thermoacidophilic Archaeon *Sulfolobus tokodaii* Contains Intrachain and Interchain Disulfide Bonds

Akeo Shinkai[1]†, Shun-ichi Sekine[1,2]†, Akiko Urushibata[3], Takaho Terada[3], Mikako Shirouzu[3] and Shigeyuki Yokoyama[1,2,3]*

[1]*RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan*

[2]*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[3]*Genomic Sciences Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan*

Received 11 May 2007; received in revised form 20 July 2007; accepted 23 July 2007
Available online
2 August 2007

The Sto12a protein, from the thermoacidophilic archaeon *Sulfolobus tokodaii*, has been identified as a small putative DNA-binding protein. Most of the proteins with a high level of amino acid sequence homology to this protein are derived from members of the Sulfolobaceae family, including a transcriptional regulator. We determined the crystal structure of Sto12a at 2.05 Å resolution by multiple-wavelength anomalous dispersion phasing from the selenomethionine-containing protein crystal. This is the first structure of a member of this family of DNA-binding proteins. The Sto12a protein forms a homodimer, and the structure is composed of an N-terminal α-helix, a winged-helix–turn–helix domain, and a C-terminal α-helix that forms an interchain antiparallel coiled coil. The two winged-helix domains are located at both ends of the coiled coil, with putative DNA-recognition helices separated by approximately 34 Å. A structural homology search indicated that the winged-helix domain shared a high level of homology with those found in B-DNA- or Z-DNA-binding proteins from various species, including archaea, bacteria, and human, despite a low level of sequence similarity. The unique structural features of the Sto12a protein include intrachain and interchain disulfide bonds, which stabilize the chain and homodimer structures. There are three cysteine residues: Cys15 and Cys16 in the N-terminal α-helix, and Cys100 in the C-terminal α-helix. Cys15 is involved in an interchain disulfide bridge with the other Cys15, and Cys16 forms an intrachain disulfide bridge with Cys100. This is a novel fold among winged-helix DNA-binding proteins. Possible DNA-binding interactions of the Sto12a protein are discussed based on the crystal structure of Sto12a and comparisons to other winged-helix DNA-binding proteins.

Keywords: disulfide bond; DNA-binding protein; *Sulfolobus tokodaii*; thermophile; winged-helix–turn–helix domain

Edited by K. Morikawa

Introduction

The genomes of *Sulfolobus* species (*S. tokodaii*, *S. solfataricus*, and *S. acidocaldarius*), which are aerobic thermoacidophilic members of the Crenarchaeota, have a number of open reading frames encoding small putative DNA-binding proteins.[1–3] Based on the properties of these DNA-binding proteins, they have been loosely classified into two categories: transcriptional regulators and chromatin proteins.

Most of the transcriptional regulators that have been studied so far are leucine-responsive regulatory protein (Lrp)/AsnC family members,[4] also termed feast/famine regulatory proteins,[5] which are broadly distributed in archaea and bacteria, but not in eukarya. The *S. solfataricus* Lrs14,[6] Ss-Lrp,[7] and Ss-LrpB proteins,[8] and the *S. acidocaldarius* Sa-Lrp protein[9] bind their own promoter regions and may autoregulate their expression. The Lrs14

*Corresponding author. E-mail address: yokoyama@biochem.s.u-tokyo.ac.jp.

†A.S. and S.S. contributed equally to this work.

Abbreviations used: Lrp, leucine-responsive regulatory protein; SH, sulfhydryl; RFX, regulatory factor X.

protein negatively regulates its expression by inhibiting the binding of TATA-box-binding protein and transcription factor B to the TATA box and transcription factor B recognition elements, respectively.[6,7] In contrast, *S. solfataricus* LysM has been considered to be an activator of the *lysWXJK* operon because it binds upstream of the transcription factor B recognition element of the promoter when the operon is highly transcribed in the absence of lysine.[10] In addition to these Lrp/AsnC family proteins, a different type of regulatory protein, *S. solfataricus* Sta1, is involved in transcription from the promoters of the crenarchaeal virus SIRV1.[11]

As for chromatin proteins, the Sul7d, Alba, and Sul10a families are thought to be involved in chromatin structure, regulation, and DNA packaging.[12–16] The Sul7d protein, which has been found only in *Sulfolobus* species, reportedly compacts relaxed or positively supercoiled DNA and plays a role in DNA packaging.[17] Furthermore, it may also be involved in DNA repair.[18] Interestingly, this protein has an intrinsic ATPase activity and promotes ATPase-hydrolysis-dependent renaturation of protein aggregates.[19,20] Unlike Sul7d, Alba is found in many other archaea and some eukaryotic species and has been observed to induce negative supercoiling of DNA.[21] The crystal structures of these proteins have been determined, and DNA-binding modes have been proposed.[12,13] *S. solfataricus* Smj12 is homologous in its amino acid sequence to the transcriptional regulator Lrs14; however, unlike Lrs14, Smj12 is a nonspecific DNA-binding protein, and it induces positive supercoiling of DNA, in contrast to Sul7d and Alba.[22] The structural basis of the broad functional range of this class of small DNA-binding proteins remains unknown.

The open reading frame ST1889 from *S. tokodaii* strain 7, which we refer to as Sto12a in this study, has been identified as a putative DNA-binding protein.[1] Most of the homologs of Sto12a are derived from members of the Sulfolobaceae family; therefore, Sto12a is a typical small DNA-binding protein of Sulfolobaceae members. However, no three-dimensional structure is available for any member of this family of small DNA-binding proteins. Here, we report the crystal structure of recombinant Sto12a, which revealed that this protein has two winged-helix–turn–helix motifs separated by a two-stranded, antiparallel coiled coil. Interestingly, the Sto12a protein contains interchain and intrachain disulfide bonds. The existence of these disulfide bonds distinguishes this Sulfolobaceae protein family from other winged-helix DNA-binding protein families.

Results

Primary structure of Sto12a

Sto12a (ST1889) is a 109-amino-acid protein with a predicted molecular mass of 12.5 kDa and a potential p*I* of 9.35.[1] It belongs to COG3355 and is classified as a predicted transcriptional regulator in the Clusters of Orthologous Groups of proteins classification system.[23] A BLAST search revealed that the Sto12a protein shared a significant sequence similarity to putative archaeal DNA-binding proteins. Among the 14 closest homologs, 13 were from *S. tokodaii*, *S. acidocaldarius*, *S. solfataricus*, and *Metallosphaera sedula* DSM 5348, which are members of the Sulfolobaceae family, with *e* values of 1×10^{-35} to 5×10^{-9} (Fig. 1). The other homolog was *Hyperthermus butylicus* (strain DSM 5456/JCM 9403) Hbut_1261 (6×10^{-10}), belonging to Pyrodictiaceae family. *S. solfataricus* Lrs14, a transcriptional regulator of the Lrp/AsnC family,[6] and *S. solfataricus* Sta1, which is involved in the transcriptional activation of viral promoters,[11] were among the 14 homologs (Fig. 1).

Initial characterization of recombinant Sto12a

The recombinant selenomethionyl Sto12a protein was purified from *Escherichia coli* cells in the presence of 1 mM DTT. An SDS-PAGE analysis revealed two bands corresponding to about 25 and 12.5 kDa (Fig. 2). Under nonreduced conditions, the ~25-kDa band was mainly detected, regardless of whether the sample had undergone heat treatment. It was converted to the ~12.5-kDa band in the presence of 640 mM 2-mercaptoethanol. The Sto12a protein has three cysteine residues, Cys15, Cys16, and Cys100 (Fig. 1), and an N-terminal sequencing analysis revealed that both the ~25-kDa and the ~12.5-kDa bands have the same sequence of five N-terminal amino acids M-M-K-E-K (data not shown). These results indicate that the ~25-kDa band corresponds to the Sto12a homodimer, formed by disulfide bonding at the interface, while the ~12.5-kDa band corresponds to the monomer dissociated from the dimer by the addition of an extremely high concentration of a reducing agent. It is unusual that the Sto12a protein can dimerize through disulfide bond(s) even in the presence of 1 mM DTT.

In order to confirm the existence of disulfide bridges in Sto12a, the ~25-kDa band detected under nonreduced conditions (Fig. 2), corresponding to the Sto12a dimer, was excised from a gel and mixed with 5,5′-dithiobis(2-nitrobenzoic acid), which reacts with free sulfhydryl (SH) groups and produces 2-nitro-5-mercaptobenzoic acid. After the reaction, we measured absorbance at 412 nm (A_{412}) to detect reaction products; however, the A_{412} value was almost the same as that of the background (data not shown). These results suggest that the Sto12a dimer does not contain free SH groups and that the three cysteine residues in the monomer all form disulfide bonds.

Crystal structure of Sto12a

We used the sample containing both the monomer and the dimer (Fig. 2) for crystallization because it was difficult to separate the two forms. The three-