# Structure of the Chromatin Protein Alba from *Archaeoglobus fulgidus*

K. Zhao, X. Chai, R. Marmorstein
The Wistar Institute and Department of Chemistry,
University of Pennsylvania, Philadelphia, PA, U.S.A.

## Introduction

A homologue of the yeast Sir2 histone deacetylase from the archaea *Sulfolobus solfataricus* deacetylates a chromatin, non-sequence-specific DNA-binding protein, Alba (formally known as Sso10b) to promote Alba/DNA association and transcriptional repression [1]. The Alba protein provides a particularly interesting model for understanding how the acetylation status of eukaryotic histones may affect its biological activity. Like the eukaryotic histones, Alba binds DNA nonspecifically, it is acetylated in its N-terminal terminus to affect its DNA binding and transcriptional regulatory properties, and it also forms higher-order structures to bind DNA [2-4]. Interestingly, Alba homologues are also found in eukaryotes [4], although their sequence divergence, which is greater than that of the archaeal proteins, suggests that they may have evolved a new function in eukaryotes. To understand the mechanistic basis for how the acetylation status of Alba may affect its DNA-binding properties, we have determined the high-resolution crystal structure of Alba from the archaea *Archaeoglobus fulgidus* (Af-Alba) [5].

## Methods and Materials

The full-length Alba gene from *Archaeoglobus fulgidus* (Af-Alba) was polymerase chain reaction (PCR)-amplified from *Archaeoglobus fulgidus* genomic DNA and cloned into the pGEX4T-1 vector to express the N-terminal glutathione-S-transferase (GST)-fusion protein. The protein was purified by using a combination of glutathione affinity, thrombin protease cleavage to remove glutathione, and gel filtration chromatography. Selenomethionine-derivatized Af-Alba protein was overexpressed from pGEX4T-1/GST-Alba-transformed bacterial strain B834 (DE3) (Novagen) and grown in a MOPS-based minimal medium as described elsewhere [6]. Selenomethionine-derivatized Af-Alba protein was purified as described for the unmodified protein.

Crystals of native Af-Alba and selenomethionine-derivatized Af-Alba were grown at room temperature by using the hanging-drop vapor-diffusion method. Crystals were obtained in two different forms, in space group $P4_32_12$ with two molecules per asymmetric unit cell and in space group $I2_12_12_1$ with one molecule per asymmetric unit cell. The $P4_32_12$ crystal form was obtained by mixing 10 mg/mL protein, 1 μL:1 μL, with a reservoir solution containing 15% isopropanol, 50 mM Na Cacodylate (pH 6.0), 100 mM KCl, and 25 mM $MgCl_2$, then equilibrating over 0.5 mL of reservoir solution. The $I2_12_12_1$ crystal form was obtained by mixing 6 mg/ml protein, 1 μL:1 μL, with a reservoir solution containing 30% PEG 400, 100 mM Tris-HCl (pH 8.5), and 200 mM Na citrate. Both crystal forms grew to a typical size of 100 × 100 × 30 mm over 6 days. The $P4_32_12$ and $I2_12_12_1$ crystal forms were transferred to a reservoir solution supplemented with 15% MPD and 40% PEG400 cryoprotecting reagent, respectively, prior to flash-freezing and storage in solid propane before data collection.

Native data from both crystal forms and a three-wavelength (peak, inflection, and remote) multiwavelength anomalous diffraction (MAD) data set from selenomethionine-derivatized $P4_32_12$ crystals were collected at beamline station BM-19-B at the APS by using an ADSC Quantum-4 charge-coupled device (CCD) detector at 100K. All data were processed with the HKL 2000 suite (HKL Research Inc.). Four selenium sites were identified in the $P4_32_12$ crystal form by using CNS [7] and SOLVE [8] and confirmed with cross difference Fourier maps. Those sites were further refined by CNS and SOLVE. The map was improved by solvent flattening, and the program O [9] was used to build residues 4-89 of the protein by using the selenomethionine sites as guides. Refinement and model building was carried out by using simulated annealing [10] and torsion angle dynamic [11] refinement protocols in CNS, and iterative manual adjustments to the model were made by using the program O, with reference to $2F_o$-$F_c$ and $F_o$-$F_c$ electron density maps. Both structures were refined with excellent residuals and stereochemistry. The $P4_32_12$ crystal form was refined to 2.65-Å resolution with an R-working of 23.7% and an R-free of 28.5%, and the $I2_12_12_1$ form was refined to 2.0-Å resolution with an R-working of 22.6% and an R-free of 25.6%.

## Results and Discussion

The Af-Alba protein monomer adopts an elongated shape with dimensions of roughly 23 × 26 × 50 Å with an αβ fold (Fig. 1). Both $P4_32_12$ and $I2_12_12_1$ crystal forms reveal a twofold symmetric Af-Alba dimer burying about 1500 Å$^2$ of solvent-excluded surface, in which the molecules interact side by side along the long dimension of the monomer and with the twofold axis

aligned perpendicular to the long dimension of the dimer (Fig. 1). Not surprisingly, the dimers in both crystal forms reported on here are essentially superimpossable, with an rms deviation of 1.3 Å for Cα atoms. Notably, nearly all of the residues involved in dimer interactions are conserved within the Alba proteins, including the Alba homologues from eukaryotic organisms, suggesting that this dimer is a conserved feature of these proteins.

Both crystal forms of Af-Alba show very similar dimer-dimer contacts in the crystal lattice, with an rms deviation of 4.3 Å for the Cα atoms between both tetramers. Like the dimerization interaction, the dimer-dimer interactions are twofold symmetric, and, in this case, the interface is formed primarily by antiparallel interactions between hydrophobic residues in the a1 helices of one subunit of each of the dimers. However, the interface also involves the C-terminal tip of the α2 helix (Fig. 1). The dimer-dimer interface involved archaeal conserved hydorphobic residues as well as the lysine 11 deacetylation target of Sir2 that serves to disrupt Alba binding to DNA. On the basis of these structural observations, we propose that the *in vivo* deacetylation of lysine 11 of archaeal Alba by Sir2 promotes dimer-dimer formation for optimal DNA binding. Subsequent solution studies confirm this hypothesis [5].

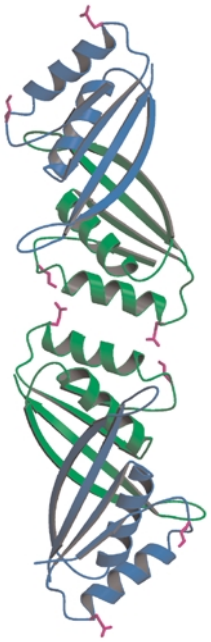

*FIG. 1. Structure of the Alba tetramer. The lysine 11 deactyaltion target of Sir2 and its interacting glutamine residue are highlighted in purple. © 2003 Journal of Biological Chemistry.*

Like the archaeal Alba protein, eukaryotic histone proteins are also acetylated to modulate gene expression. In the case of the histone proteins, hyperacetylation is generally correlated with gene activation, while hypoacetylation is correlated with repression. The mechanism by which the acetylation status of histones regulates gene expression is still unclear. In one model, the unacetylated histone tails interact with neighboring nucleosomes in chromatin, and acetylation is proposed to relieve these interactions to destabilize repressive higher-order chromatin. The studies we present here suggest that the acetylation-induced regulation of archaeal Alba proteins involves the modulation of the higher-order Alba structure. This would be consistent with a model in which histone tail acetylation would disrupt the higher-order chromatin structure. Confirmation of this model would certainly require further direct studies involving histone proteins. Nonetheless, the structural and functional studies presented here suggest that the acetylation status of archaeal Alba affects its oligomerization status and that the evolution to eukaryotic systems and histone proteins might have maintained this mechanism of transcriptional regulation.

## Acknowledgments

We thank A. Clements Egan, M. Fitzgerald, H. Peng, G. Da, K. Li, L. Xu, and F. Xue for useful discussions and A. Joachimiak, R. Zhang, N. Duke, and the SBC-CAT staff for access to and assistance with beamline BM-19 at the APS. This work was supported by grants from the National Institutes of Health (GM52880 and GM60293) to R. Marmorstein and from the Commonwealth Universal Research Enhancement Program, Pennsylvania Department of Health, to The Wistar Institute. The data on atomic coordinates were deposited in the Protein Data Bank (PDB), Research Collaboratory for Structural Bioinformatics (RCSB, Rutgers University, New Brunswick, NJ, http://www.rcsb.org) under accession 1NFH and 1NFJ. Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38. This report came from K. Zhao, C. Xiaomei, and R. Marmorstein, "Structure of a Sir2 substrate, Alba, reveals a mechanism for deactylation-induced enhancement of DNA binding," J. Biol. Chem. **278**, 26071-26077 (2003).

## References

[1] S.D. Bell, C.H. Botting, B.N. Wardleworth, S.P. Jackson, and M.F. White, Science **296**, 148-151 (2002).
[2] R. Lurz, M. Grote, J. Dijk, R. Reinhardt, and B. Dobrinski, EMBO J. **5**, 3715-3721 (1986).
[3] B.N. Wardleworth, R.J. Russell, S.D. Bell, G.L. Taylor, and M.F. White, EMBO J. **21**, 4654-4662 (2002).
[4] M.F. White and S.D. Bell, Trends Genetics **18**, 621-626 (2002).
[5] K. Zhao, X. Chai, and R. Marmorstein, J. Biol. Chem. **278**, 26071-26077 (2003).
[6] J.R. Rojas, R.C. Trievel, J. Zhou, Y. Mo, X. Li, S.L. Berger, D. Allis, and R. Marmorstein, Nature **401**, 93-98 (1999).
[7] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren, Acta Crystallogr. D **54**, 905-921 (1998).
[8] T.C. Terwilliger and J. Berendzen, Acta Crystallogr. D **55**, 849-861 (1999).
[9] T.A. Jones, J. Appl. Crystallogr. **11**, 268-272 (1978).
[10] A.T. Brunger, and A. Krukowski, Acta Crystallogr. A **46**, 585-593 (1990).
[11] L.M. Rice and A.T. Brunger, *Proteins* **19**, 277-290 (1994).