Biological Science

NW12A, 5A/2006G169

Recognition of the MAR sequence of DNA by transcription factor SATB1

Kazuhiko Yamasaki*[1], Toshihiko Akiba[1], Kazuaki Harata[1]
[1]AIST, Tsukuba, Ibaraki 305-8566, Japan

Introduction

Special AT-rich sequence binding protein 1 (SATB1) is a transcription factor that specifically binds to a sequence with a special physical property appearing in the nuclear matrix attachment region (MAR) of genomic DNA [1]. SATB1, predominantly expressed in thymocytes, recruits histone deacetylase complex to the MAR site inside the interleukin-2 receptor α gene, in order to repress its expression during the maturation stage of the T-cell development [2]. Previously we have determined the structure of the MAR-binding domain, which is classified to a CUT domain, of SATB1 by NMR spectroscopy and suggested by surface plasmon resonance experiments that it binds to the DNA from the major groove side [3], although the minor-groove binding had been predicted before [1]. In the present study, we determined the structure of the complex of SATB1 MAR-binding domain and MAR-DNA by crystallography and revealed the structural basis for the specific recognition of the MAR sequence [4]. This is the first reported structure of a CUT domain in the complex with DNA.

Method and Results

Crystals of the protein-DNA complex were produced by sitting-drop vapor diffusion method at 293K in 50 mM Tris-HCl buffer (pH 8.0) containing 20% polyethylene glycol 20,000, 10 mM $MgCl_2$ or $MgSO_4$, and 20% ethylene glycol. Diffraction data up to 1.75 Å resolution were obtained under a nitrogen gas stream at 95 K on beam lines NW12A and BL5A. Those up to 2.0 Å were also obtained by on a laboratory diffractometor. Structures were determined from both the data sets by a molecular replacement method using the NMR structure of the MAR-binding domain [3].

The structure revealed that SATB1 MAR-binding domain binds to the standard B-DNA from the major groove side (Fig. 1A). Bases of 5'-CTAATA-3' sequence were contacted by direct or water-mediated hydrogen bonds and apolar or van der Waals contacts. Among them a hydrogen-bonding network involving two Gln residues and an adenine nucleotide (Fig. 1B) is of special note in that it includes the only direct hydrogen bonds contributing to the base recognition. The importance of this interaction was shown by a replacement of Gln402 by Ala, resulting in a ~50-fold decrease of the DNA-binding activity.

The revealed binding mode is very similar to that of the POU-specific domains of POU-family transcription factors, such as Pit-1 (Fig. 1C), OCT-1, and HNF-1α. Although the POU-specific domains possess four helices in common, which is different from the five-helix SATB1 MAR-binding domain, they commonly use third helix of the four as the recognition helix. Also the hydrogen-binding network involving two Gln residues and an adenine nucleotide is observed (Fig. 1D). Together with the fact that CUT-domain-containing proteins and POU-family transcription factors possess homeodomain in the region C-terminal of the CUT or POU-specific domain, the evolutionary relationship between the two groups of transcription factors was strongly suggested.

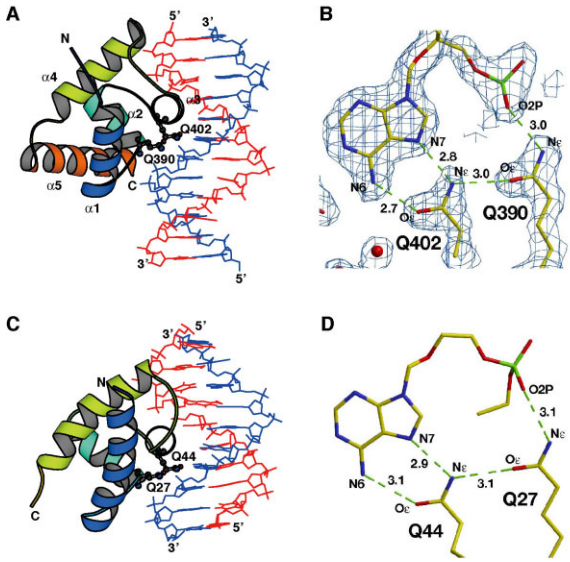

Fig. 1: (A) Structure of the complex of SATB1 MAR-binding domain and DNA. (B) A hydrogen-bonding network observed among two Gln residues and an adenine base in the complex. The $2F_o$-F_c electron densities contoured at 1.5 σ are shown. (C) Structure of the complex of Pit-1 POU-specific domain and DNA (PDB entry 1AU7). (D) A similar hydrogen-bonding network observed in the Pit-1 complex.

References

[1] L. A. Dickinson et al., *Cell. 70*, 631 (1992).
[2] D. Yasui et al., *Nature 419*, 641 (2002).
[3] H. Yamaguchi, M. Tateno, and K. Yamasaki, *J. Biol. Chem. 281*, 5319 (2006).
[4] K. Yamasaki, T. Akiba, T. Yamasaki, and K. Harata, *Nucl. Acids Res.*, in press (2007).

*k-yamasaki@aist.go.jp