Novel Mechanism for DNA Duplex Unwinding by a Single Protein

Duplex DNA is generally unwound by protein oligomers prior to replication. The Rep protein of plasmid ColE2-P9 is an essential initiator for plasmid DNA replication. This protein binds the replication origin (Ori) in a sequence-specific manner as a monomer and unwinds DNA. The crystal structure of the DNA-binding domain of ColE2-P9 Rep in complex with Ori DNA unveiled the basis for Ori-specific recognition by the protein and also revealed that it unwinds DNA by the concerted actions of its three contiguous structural modules. A novel mechanism for DNA unwinding by a single protein has been proposed here.

Since a DNA molecule in cells usually forms a stable double-helix structure, the protein machineries acting on single-stranded DNA must separate the paired DNA strands before their actions take an effect. Many DNA-binding proteins enlarge the contact areas, contributing to the interaction with DNA, by increasing their molecular size, including multimer formation and/or complex formation with other cooperative proteins, so that binding specificity and affinity are strong enough to disrupt the stable double-helix structure of DNA. One of those machineries, which initiate double-stranded DNA replication, is protein-based.

DNA replication is the universal process for transmitting genetic information. The process is initiated from a specific region called replication origin. To initiate DNA replication, initiator proteins specifically bind to the origins in the chromosomal DNA. A number of eukaryotic and prokaryotic initiator proteins bind to the origins and form higher order structures consisting of multiple proteins to unwind locally the double-helix structure of DNA. Subsequently, the replication proteins, such as helicase, primase, and DNA polymerase, are loaded onto the unwound region.

The replication initiation of most plasmids, a small DNA molecule in a cell that is physically separated from the host chromosomal DNA and can replicate independently, requires a specific initiator protein, Rep, encoded by each plasmid. Multiple Rep molecules specifically bind to the cognate replication origin (Ori) and recruit host replication proteins. DNA unwinding often relies on the host initiation protein, and DNA replication initiates more or less similarly to the host chromosomal DNA replication.

The Rep of the *Escherichia coli* plasmid ColE2-P9 is unique, because this initiator binds to Ori of the plasmid in a sequence-specific manner as a monomer and unwinds locally its double-helix structure. Furthermore, the Rep exhibits Ori-specific primase activity; synthesizes the primer RNA necessary for the DNA chain synthesis by DNA polymerase. Thus, the replication of the plasmid ColE2-P9 initiates without host replication proteins, except for DNA polymerase I. Although the Rep of ColE2-P9 exhibits multiple functions necessary to initiate DNA replication, this protein is relatively small (34 kDa), and its C-terminal DNA-binding domain (E2Rep-DBD, 14 kDa) is sufficient only for specific binding to Ori of ColE2-P9 and its unwinding. How does this small molecule ensure binding specificity and affinity to DNA and how does it disrupt its double-helix structure?

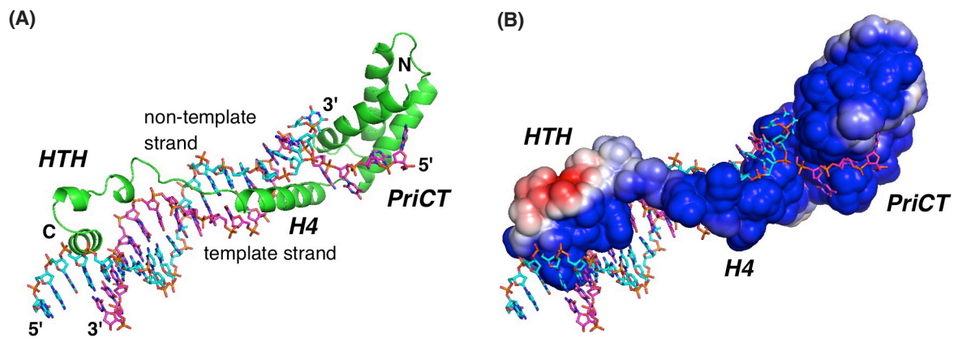

Figure 1: The representation of the crystal structure of the E2Rep-DBD in complex with Ori DNA. Combined with the results of the biochemical and genetic analyses, the structure is shown to represent the snapshot during the DNA unwinding process uniquely facilitated by a single protein molecule.(A) Overall view of the complex. The green ribbon shows E2Rep-DBD, and the magenta and cyan stick models show the template strand for the primer RNA synthesis and the non-template strand of DNA, respectively. (B) Surface charge representation of E2Rep-DBD. The viewpoint is the same as that of panel A. The blue and red colors on the molecular surface show the positively and negatively charged regions, respectively.

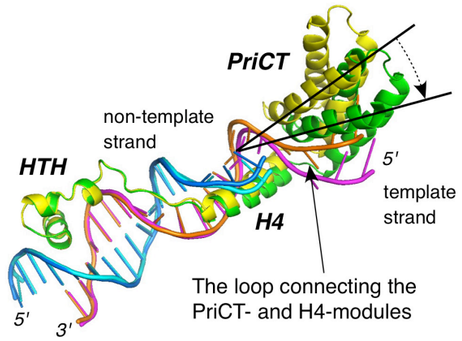

Figure 2: Comparison of the two complex molecules (the proteins are drawn in green and yellow) in the crystallographic asymmetric unit and the result of mutational analysis suggested that the structural flexibility between the H4 and PriCT-modules and the shift in position of the PriCT-module relative to the other modules is important for the unwinding of DNA.

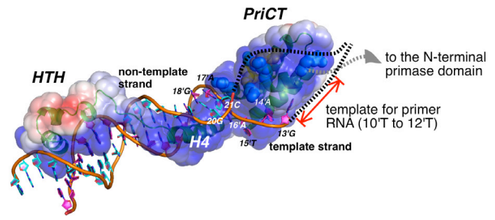

Figure 3: Model for unwinding of the complete Ori by Rep. Black dashed lines indicate the supposed trajectory of each of the DNA strands. The PriCT-module works as a wedge separating the DNA strands, and presumably also as a platform during the RNA primer synthesis by the primase domain. The N-terminal domain of Rep has been supposed to be the primase domain, and the trajectory of the linker connecting the primase domain to E2Rep-DBD is indicated with gray dashed lines.

A research group at the National Institute of Genetics determined the crystal structure of E2Rep-DBD in complex with Ori DNA of ColE2-P9 and revealed the molecular mechanism, using the data collected at the structural biology beam lines in the Photon Factory, KEK [1]. The crystal structure showed that E2Rep-DBD comprises three structural modules (PriCT, H4, and HTH), connected with flexible linker regions, and forms an elongated fold (Fig. 1). As E2Rep-DBD binds to DNA along with the molecule, this protein is capable of interacting with a long region of DNA. One third of the surface area of the molecule is used for interaction with DNA, and the area is comparable with that of the DNA-binding protein complex that is more than twice larger than the E2Rep-DBD molecule. The tight binding of the C-terminal HTH-module and the H4-module of E2Rep-DBD induces distortion of the double-helix structure of Ori DNA. The N-terminal PriCT-module separates the DNA strands, and the module holds the single-stranded DNA to stabilize the unwounded conformation (Fig. 2). The structural study showed that the E2Rep-DBD binds specifically to Ori and unwinds its double-helix structure by the concerted actions of its three contiguous structural modules. The study also suggested that the PriCT-module works as a platform to hold the template strand during the subsequent RNA primer synthesis by the primase domain, because the template region for the RNA primer locates immediate downstream of the bases specifically bound by the PriCT-module (Fig. 3). Although the PriCT domain is conserved in the C-termini of some archaeo-eukaryotic primases, its function has been unknown. The structural study showed that the PriCT domain of the ColE2-P9 Rep plays a central role in DNA unwinding and the stabilization of the unwound structure. The conservation of the domain indicates that it likely plays a similar role in the other proteins. The study is the first report providing the structural basis for the functional importance of the conserved PriCT domain and also reveals a novel mechanism for DNA unwinding by a single protein.

REFERENCE

[1] H. Itou, M. Yagura, Y. Shirakihara, and T. Itoh, *J. Biol. Chem.* **290**, 3601 (2015).

BEAMLINE

BL-17A and AR-NE3A

H. Itou (National Inst. of Genetics)