Displacement of the canonical single-stranded DNA-binding protein in the Thermoproteales

Sonia Paytubi1,2, Stephen A. McMahon2, Shirley Graham3, Huanting Liu1, Catherine H. Botting1, Kira S. Makarova1, Eugene V. Koonin1, James H. Naismith1,3, and Malcolm F. White1,3

*Biomedical Sciences Research Complex, University of St Andrews, Fife KY16 9ST, United Kingdom; and †National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894

AUTHOR SUMMARY

Proteins are the major structural and operational components of cells. Even the simplest organisms possess hundreds of different proteins, and more complex organisms typically have many thousands. Because all living beings, from microbes to humans, are related by evolution, they share a core set of proteins in common. Proteins perform fundamental roles in key metabolic processes and in the processing of information from DNA via RNA to proteins. A notable example is the ssDNA-binding protein, SSB, which is essential for DNA replication and repair and is widely considered to be one of the few core universal proteins shared by all life forms. Here we demonstrate that one branch of the tree of life has lost this “ubiquitous” protein and replaced it with another, unrelated one. This finding has important implications for our understanding of the plasticity of evolution.

Rapid advances in genome-sequencing technology over the past 15 y yielded a wealth of new information on many divergent parts of the tree of life that can be mined for information to illuminate all aspects of biology. The tree of life consists of three fundamental divisions known as “domains”: Eukarya (organisms with a nucleus where DNA is stored, including plants, fungi, animals) and the prokaryotic Bacteria and Archaea, which lack a nucleus (1). One of the most highly conserved proteins found in all three domains is the SSB protein, which binds and protects ssDNA during replication and repair of damage to the genome. It is an essential protein that is thought to have been present in the last common ancestor of all extant life (2). The defining feature of all known SSBs is the oligonucleotide-binding (OB) fold shown in Fig. P1. Recently we noted that one group of archaeal species, the Thermoproteales, lack a detectable gene for the SSB protein in their genomes (3).

Because a functional SSB is likely to be essential for any organism, we reasoned that the Thermoproteales might have lost the canonical SSB gene and replaced it with another, unrelated gene. To test this hypothesis, we undertook a two-pronged approach comprising a combination of bioinformatics and biochemical route involved direct purification and identification of proteins that could bind to ssDNA in one of the Thermoproteales, Thermoproteus tenax. This approach resulted in the identification of the product of the gene ttX1576 as a candidate for the missing SSB.

We proceeded to characterize the properties of the Ttx1576 protein, (which we renamed “ThermoDBP” for Thermoproteales DNA-binding protein), showing that it has all the biochemical properties consistent with a role as a functional SSB, including a clear preference for ssDNA binding and low sequence specificity. Using crystallographic analysis, we solved the structure of the DNA-binding domain of ThermoDBP, revealing a protein fold with a prominent cleft punctuated with aromatic amino acid residues and lined by positively charged residues. The structure of ThermoDBP immediately suggested a mechanism for the binding of ssDNA along the cleft that is reminiscent of the binding clefts of canonical SSB proteins but is unrelated to them in sequence and detailed structure. The two ssDNA-binding domains are linked by a C-terminal helical coiled-coil domain that allows ThermoDBP to dimerize (Fig. P1).

In conclusion, we used biochemistry and bioinformatics to demonstrate the displacement of an essential, “universal” protein by a completely unrelated one in one branch of the tree of life. The structure of ThermoDBP reveals a unique solution to the problem of ssDNA binding. This result suggests that even the most fundamental, ubiquitous proteins can be replaced during evolution.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank database, www.pdb.org (PDB ID code 3TEK).

1Present address: Departament de Microbiologia, Facultat de Biologia, Universitat de Barcelona, 08028 Barcelona, Spain.

2S.P. and S.A.M. contributed equally to this work.

3To whom correspondence may be addressed. E-mail: naismith@st-and.ac.uk or mfw2@st-and.ac.uk.

See full research article on page E398 of www.pnas.org.

Cite this Author Summary as: PNAS 10.1073/pnas.1113277108.
