# Single-stranded DNA-binding Proteins

**Yousif Shamoo,** *Rice University, Houston, Texas, USA*

Single-stranded DNA binding proteins bind specifically to single-stranded DNA. That specificity relates to *in vivo* functions in DNA replication, recombination and repair.

## Introduction

The emergence of DNA as the predominant hereditary material of life has necessitated the evolution of proteins whose sole function is the maintenance and care of these vital molecules. In a typical cell, there is a wide array of proteins with specialized functions in DNA metabolism that allow the cell to maintain and replicate its genome. In order to copy or make repairs to DNA, the double helix must be unwound to reveal the two complementary strands. The need to manipulate DNA in its single-stranded form has given rise to a specialized group of proteins. This article examines the family of proteins whose primary role is to recognize and bind to single-stranded DNA (ssDNA). By comparing a variety of ssDNA-binding proteins from different species and different metabolic processes, we can discern their common characteristics, as well as the distinctions that make them essential to the cell.

## The OB Fold and ssDNA Recognition

At first glance, the evolution of proteins to bind tightly yet promiscuously to all available sequences of ssDNA would seem an impossible task. In order to understand how ssDNA recognition occurs, it is helpful to consider not only the properties of ssDNA, but also those of the molecules likely to compete with ssDNA for binding, namely double-stranded DNA (dsDNA) and RNA. As the different physical and chemical properties of these nucleic acids are described, we will see how ssDNA-binding proteins have evolved to make use of these differences to bind preferentially to ssDNA. Upon further review, it will be evident that although ssDNA-binding proteins come from a wide range of organisms and metabolic processes, many share a common structural motif called the OB fold (oligosaccharide/oligonucleotide-binding fold) that enables them to bind ssDNA (Murzin, 1993).

### The properties of ssDNA and their relationship to binding

The individual repeating unit of ssDNA is the nucleotide. A nucleotide consists of three fundamental parts: a nucleoside base, a sugar and a phosphate. ssDNA-binding proteins make use of all these features for binding and recognition. As shown in **Figure 1a**, the structure of ssDNA is quite flexible, since there is free rotation about the phosphodiester backbone that links adjacent nucleotides. Each phosphate of the phosphodiester linkage connecting adjacent nucleotides requires four oxygens. Two oxygens make up the diester link that connects the nucleotides (referred to as 'bridging oxygens'), while the other two fulfil the valency requirements of the phosphate ('nonbridging oxygens'). This chemistry leaves one of the nonbridging oxygens with a negative charge. Since each phosphate of the phosphodiester linkage that makes up a DNA strand has one negative charge, ssDNA can be loosely thought of as an electronegative polymer. ssDNA-binding proteins commonly take advantage of this electronegative character and line their DNA-binding surfaces with the electropositive amino acid residues lysine and arginine (**Figure 1b**). In addition, there are examples of charged residues such as arginine forming favourable interactions between its positively charged guanidinium group and the negatively charged pi cloud delocalized above and below the plane of the aromatic nucleic acid base. In addition to electrostatic interactions, hydrogen bonds from the amino acid



(a)
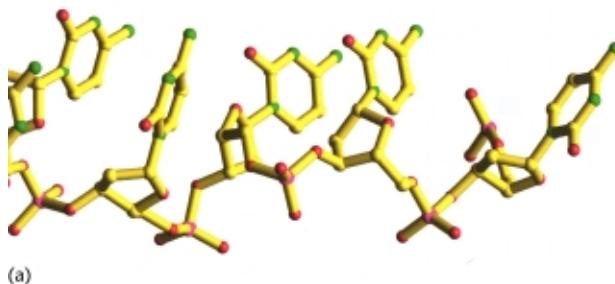
**Figure 1** (a) A segment of ssDNA taken from the human RP-A SSB structure with ssDNA. Note how irregular the structure is. The nucleotide bases are in a variety of orientations and are accessible to stacking interactions with other bases or with side-chains from the protein. The phosphodiester backbone is also highly irregular and phosphate–phosphate (pink) distances are quite variable. Carbon (yellow), nitrogen (green) and oxygen (red) atoms are shown as spheres.
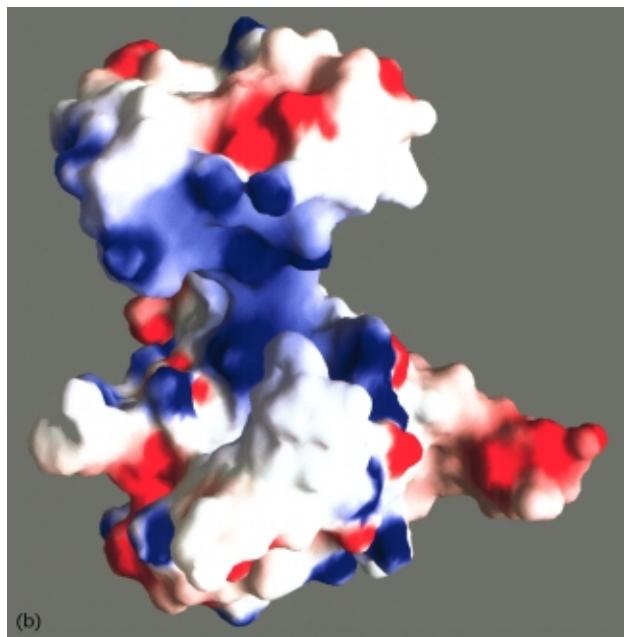
**Figure 1** (b) An electrostatic surface representation of the bacteriophage T4 SSB. The ssDNA-binding cleft of the protein is lined with positively charged amino acids (blue) to make interactions with the negatively charged DNA polymer. Negatively charged regions of the protein surface are shown in red and are well away from the ssDNA-binding cleft at the midsection of the protein.

side-chains or even the amide or carbonyl groups of the polypeptide can be used to interact in a nonsequence-specific manner to the ssDNA. These generalized electrostatic or hydrogen-bonding interactions between amino acid side-chains from the protein to the phosphodiester backbone and bases of the DNA are ubiquitous but make up only a part of the energy and specificity needed for ssDNA binding.

Another component of ssDNA available for recognition and binding is the nucleotide base itself. Although the chemical structures of the purines and pyrimidines are quite different, they share the common property of being aromatic ring structures. Since nucleic acid bases are relatively planar, ssDNA-binding proteins frequently make stacking interactions between the nucleic acid base and its own aromatic and planar side-chains such as the amino acids tyrosine and phenylalanine to bind ssDNA. The generalized hydrophobic nature of stacking interactions provides the protein with an additional strategy for ssDNA binding that does not require overspecialization of the nucleic acid-binding surface into specific nucleotide-binding pockets or hydrogen-bonding patterns that would favour one sequence over another. The stacking interaction also uses a part of the nucleotide that is distinct from the phosphodiester backbone and can therefore be used in conjunction with other interactions.

The sugar moiety of DNA is less accessible to the protein and is therefore used less frequently, but since it is largely planar, stacking interactions appear to be favoured. Since the only chemical difference between RNA and DNA is the presence of a 2′ hydroxyl group on the RNA, the sugar often plays an important role in distinguishing RNA from DNA. Proteins that bind ssRNA can make use of the 2′ hydroxyl group to increase their affinity for ssRNA over ssDNA. In contrast, ssDNA binding proteins can have tightly fitted pockets on their surfaces such that nucleotides with 2′ hydroxyls are excluded from binding. This sort of steric exclusion is a powerful strategy since it is impossible to fit an incorrectly shaped chemical structure into the binding pocket. Typically, enzymes such as DNA or RNA polymerases must have highly selective active sites to prevent misincorporation of incorrect nucleotides into the newly synthesized nucleic acid strand.

## ssDNA specificity and recognition

The combination of electrostatic, hydrogen-bonding and stacking interactions from proteins to ssDNA forms the basis for ssDNA binding and specificity. Unfortunately, dsDNA, dsRNA and single-stranded RNA (ssRNA) share many of these properties. How does an ssDNA-binding protein exclude these competing molecules? The exclusion of double-stranded nucleic acids is relatively simple. Although dsDNA and dsRNA both have an electronegative phosphodiester backbone, they are significantly more inflexible in the spacing and positioning of those negative charges than ssDNA. This inflexibility arises from the fact that nucleotides of double-stranded nucleic acids are engaged in base pairing and form a helix that rigidly determines the spacing and position of each phosphate. For example, in B-form DNA, a single helical turn has 10 base pairs and a rise of 3.4 nm, meaning that each phosphate is separated by 0.34 nm and a rotation of 36° and must be positioned to form a helix. By contrast, ssDNA need not form a helix, and the internucleotide spacing can vary quite widely (**Figure 1a**). Any attempt to place a double-stranded nucleic acid onto an ssDNA-binding surface would almost certainly require a deformation of either the DNA or protein to accommodate the relatively inflexible duplex, an energetically costly bargain. Of course the simplest strategy is to build a ssDNA-binding surface that is too narrow to fit the much wider duplex. Depending on the particular ssDNA-binding protein and its function, either or both of these strategies can be employed successfully to discriminate against double-stranded nucleic acids.

The ability to distinguish ssDNA from ssRNA is perhaps the most vexing problem confronting a ssDNA-binding protein. In fact, ssDNA-binding proteins typically have only modest preferences for ssDNA over ssRNA. As a consequence the *in vivo* concentrations of ssDNA-

binding proteins are usually tightly controlled to prevent their interference in RNA metabolism. The plasticity of the sugar moiety to readily form a wider range of sugar puckers in ssDNA over ssRNA can provide some of the observed discrimination. The stacking of protein side-chains or even adjacent nucleic acid bases with a ribose would be markedly worse than that of deoxyribose since the hydroxyl group would have to be positioned away to allow stacking (a form of steric exclusion). The addition of a hydroxyl group in RNA also has the effect of limiting the available conformations of ssRNA, and may impose a small but meaningful energetic barrier to ssRNA binding over ssDNA in extended irregular structures.

There are examples of proteins that bind ssDNA in a sequence-specific fashion such as the telomere-binding protein (see below) but the principles of ssDNA recognition remain the same. The addition of sequence specificity to a protein's interaction with ssDNA is simply an embellishment of prototypical ssDNA-binding proteins where some of the generalized interactions such as stacking and electrostatic interactions are supplemented with more specific hydrogen-bonding networks or small pockets at the DNA-binding surface that fit only particular nucleotide bases.

## The OB fold: a common single-stranded nucleic acid binding motif

The OB fold is a small domain originally observed in oligosaccharide/oligonucleotide-binding proteins and consists of a five-stranded antiparallel β barrel terminating in an α helix (**Figure 2**; Murzin, 1993). The OB fold presents a narrow ssDNA-binding cleft that is studded with residues in position to make numerous interactions with a small number of nucleotides (typically 2–5) in a relatively compact substructure. β barrels are strongly twisted β sheets and in the OB fold the twist allows amino acid side-
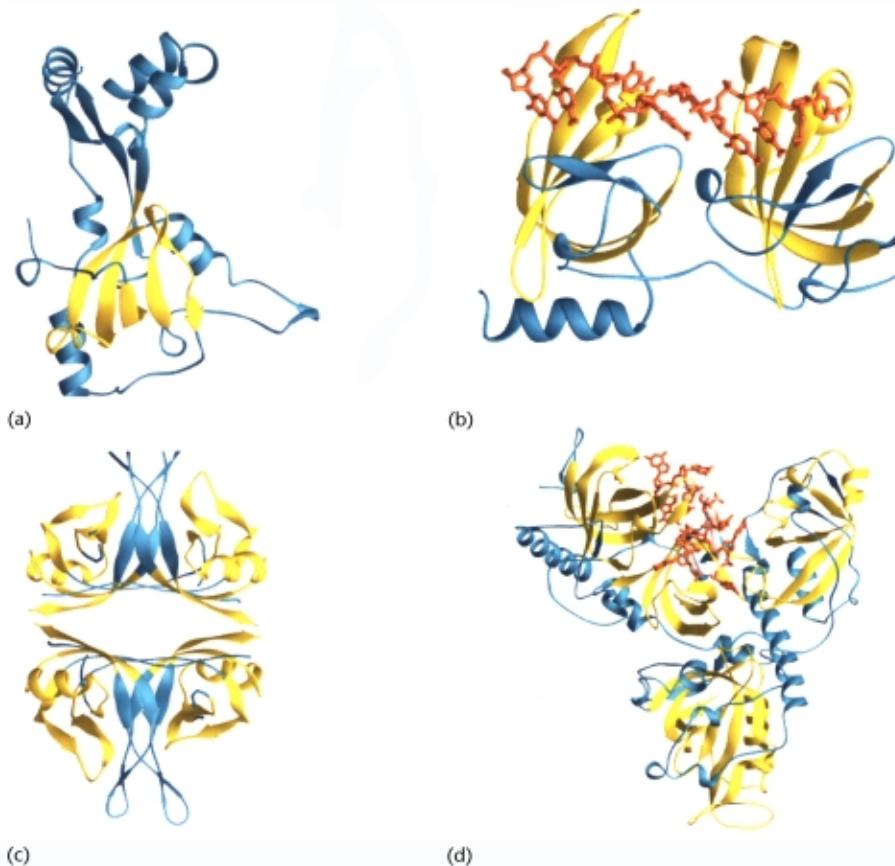


**Figure 2** Ribbon diagram of proteins employing the OB fold (yellow). (a) T4 SSB; (b) RP-A; (c) human mitochondrial SSB; (d) telomere end-binding protein. The OB fold is a conserved ssDNA-binding motif observed in many ssDNA-binding proteins. Several examples of how the OB fold is employed are shown. A ssDNA-binding protein can have one OB fold (T4 SSB) or as many as four in the *E. coli* SSB tetramer. There are also examples of ssDNA-binding proteins that do not have any OB folds (RecA, Figure 3).

chains from the protein surface to interact with all the structural elements that make up the nucleotide (base, sugar and backbone). The association of a single OB fold domain with ssDNA is relatively weak (association constants of $10^5$–$10^6$ L mol$^{-1}$) but because of the modular nature of the OB fold, it can be present in more than one copy to allow a protein to bind longer sequences and raise the affinity. Some proteins, such as *Escherichia coli* and human mitochondrial ssDNA-binding proteins (SSBs), bind as discrete oligomers (tetramers) to ssDNA while others, such as bacteriophage T4 SSB, bind cooperatively to ssDNA as indefinite linear filaments (**Figure 2**).

The OB fold has been found in many ssDNA-binding proteins, including bacteriophage T4 SSB, human RP-A, human mitochondrial SSB, *Escherichia coli* SSB, filamentous phage fd gp5 and *Oxytrichia nova* telomere-binding protein (Suck, 1997). Not surprisingly, the OB fold has also been found in several ssRNA-binding proteins, including *E. coli* aspartyl tRNA synthetase and polynucleotide phosphorylase (Murzin, 1993). The widespread employment of the OB fold in ssDNA-, ssRNA- and oligosaccharide-binding proteins has led to the suggestion that it is related to an ancient ancestral OB fold-containing progenitor and has since diverged into more specialized activities (Bycroft *et al.*, 1997).

## SSBs in DNA Replication, Recombination and Repair

The complex processes of DNA replication, recombination and repair all require that the DNA double helix be at least transiently unwound. In DNA replication, helicases bind to parental DNA and unwind it so that DNA polymerase may read the genetic code to synthesize a new copy or daughter strand. As the duplex DNA is unwound, SSBs are responsible for binding ssDNA until it is utilized by DNA polymerase or other proteins involved in DNA recombination and repair. Historically, SSBs have sometimes been loosely referred to as 'helix-destabilizing proteins' because they can reduce the stability or 'melt' some duplex DNAs. It should be emphasized that SSBs do not unwind dsDNA; rather, they bind and stabilize the ssDNA conformation as it becomes available either enzymatically via helicases or by binding ssDNA 'bubbles' or the transiently frayed 5′ or 3′ ends of an otherwise duplex DNA. Since SSBs must bind all available ssDNA as it becomes accessible, they are highly abundant.

Several members of this important class have been structurally characterized, including bacteriophage T4 SSB (Shamoo *et al.*, 1995), *E. coli* SSB (Raghunathan *et al.*, 1997), human RP-A (Bochkarev *et al.*, 1997), human mitochondria (mt) SSB (Yang *et al.*, 1997) and adenovirus (DBP) (Tucker *et al.*, 1994; **Figure 2**). As shown in **Figure 2**, an SSB can contain from one (T4 SSB) to as many as four

copies (*E. coli* and human mitochondrial SSB) of the OB fold in its active state.

The SSBs are responsible for several important functions in DNA replication, including:

1. configuring the ssDNA for use by DNA polymerase – this entails arraying the template DNA in an extended single-stranded conformation for efficient use by the polymerase as it passes from the SSB into the polymerase active site;
2. removal of adventitious secondary structures from replication fork ssDNA – as DNA is unwound there is a tendency for the newly unpaired DNA bases to form into secondary structures, such as hairpins, that interfere with replication, SSBs denature these structures;
3. protecting ssDNA from nucleases – *in vivo* the cell has a large endogenous population of nucleases that are a part of DNA turnover and defence against invasion by foreign organisms. ssDNA at the replication fork is vulnerable to the action of nucleases and as a consequence SSBs must protect DNA.

In addition to their role in replication, SSBs are typically involved in DNA recombination and repair. In recombination, they stimulate renaturation of complementary sequences and interact directly with other ssDNA-binding proteins such as RecA and its homologues to facilitate homologous recombination. The mechanism by which SSBs facilitate renaturation of complementary ssDNA is essentially the converse of their role in replication. By removing secondary structures and promoting an extended structure in ssDNA, the rate at which complementary strands can properly anneal and then zipper back up into dsDNA is significantly enhanced. Recombination proteins such as RecA recognize ssDNA coated with SSB and are able to displace the SSB. In many ways, an SSB fulfils the same role in recombination as replication, that is, by binding to ssDNA and passing it off to the recombination proteins such that the substrate (ssDNA) is free of secondary structures and is in a form that facilitates subsequent enzymatic events.

In repair, damage to DNA from ultraviolet light or other agents often produces lesions that disrupt base pairing of the DNA and thereby gives the appearance of partially denatured DNA. The binding of the SSB to the damage serves two purposes: first, to protect the now vulnerable DNA from nucleases, and to recruit repair enzymes to the damage. Again, through a series of highly specific protein–protein interactions, the SSB is displaced such that repairs to the DNA can begin.

The interaction of SSBs with proteins involved in DNA metabolism are typically quite specific, such that T4 SSB works only with members of the T4 DNA replication, recombination and repair system, while *E. coli* SSB works

only with the *E. coli* proteins. The interactions of bacteriophage T4 SSB are typical for the SSB family. Not only can T4 SSB interact with its cognate DNA polymerase and helicase at the replication fork, it also interacts with the DNA repair enzyme UvsX, and makes cooperative interactions with itself to bind long stretches of ssDNA. These specific interactions are mediated by subdomains of the protein that are widely diverse in structure, unlike the generally well-conserved OB fold.

## RecA: an ssDNA-binding Protein that can Bind More Than One Strand of DNA Simultaneously

*E. coli* RecA is the prototype for proteins whose primary role is to enzymatically catalyse homologous recombination for DNA repair. RecA homologues are ubiquitous and have been found in minimalist organisms, such as bacteriophage T4 (UvsX) and *Mycoplasma*, to eukaryotes including yeast (RAD51) and humans (hRad51), as well as over 60 bacterial species (Roca and Cox, 1997). The presence of RecA-like proteins in so many organisms is a testament to their utility in evolutionary fitness.

Like DNA replication, DNA repair is central to survival. Ultraviolet irradiation or oxidation can cause a myriad of damage to DNA, including double-strand breaks, point lesions, thymine–thymine dimers and single-stranded gaps. The process of homologous recombinational repair allows an organism to swap undamaged segments from one strand to another to produce a complete uncompromised genome. To do homologous repair, RecA must bind in a generally nonsequence-specific manner to the damaged ssDNA, then bind a second strand (the strand with the homologous but undamaged segment), and the complementary strand that will be required to produce the repaired dsDNA (**Figure 3**). Since recombination requires that relatively long stretches of DNA be paired and unpaired, RecA monomers polymerize to form a right-handed helical filament (see **Figure 3**) on DNA. Although the RecA monomer binds only three nucleotides of ssDNA, cooperative RecA–RecA interactions allow RecA filaments to span thousands of nucleotides. The formation of these RecA filaments is essential to homologous recombination.

Electron microscopy studies of RecA bound to ssDNA have demonstrated that RecA can form a 'collapsed' filament with a helical pitch of 6.4 nm. Although these 'collapsed' filaments are observed in the RecA complex with ssDNA, filaments containing ssDNA, dsDNA and RecA require the presence of ATP. In the presence of ATP or a nonhydrolysable analogue, the RecA filament is substantially extended to a helical pitch of 9.5 nm. The result of this helical extension is a significant underwinding
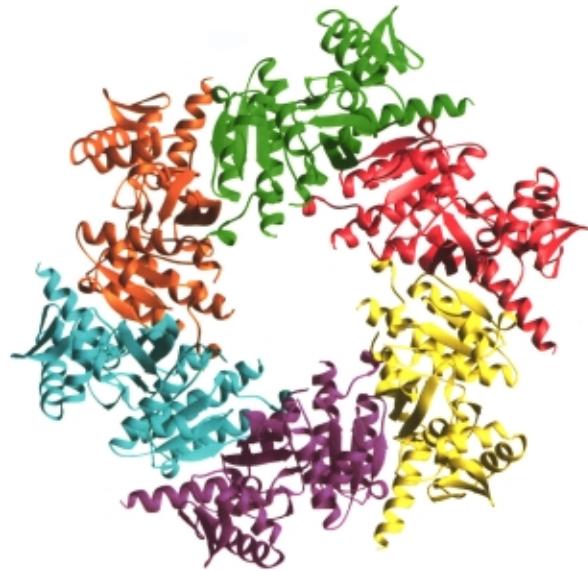


**Figure 3** A ribbon view of one turn of the RecA filament as seen along the helical axis in the crystal structure of RecA. The structure shown here corresponds to the 'collapsed' or ADP filament and has a helical pitch of 6.4 nm and a rise of 0.21 nm per nucleotide. The extended or ATP filament is extended significantly and has a helical pitch of 9.5 nm (six RecA molecules per turn). The extended filament is responsible for homologous strand exchange.

of the DNA (0.51 nm axial rise per base pair) compared to B-form dsDNA (0.34 nm axial rise per base pair). Although the underwinding of the dsDNA does not result in strand separation, there is significant evidence that all three strands within the RecA filament may interact to facilitate strand exchange. Since nonhydolysable ATP analogues support formation of the helical RecA filament, it is clear that ATP hydrolysis is not a prerequisite for filament formation. The ability of RecA to act as a DNA-dependent ATPase has been the subject of intensive investigation but remains poorly understood. Recent studies have suggested that ATP hydrolysis may be related to maintenance of the extended filament by recycling RecA monomers, or perhaps to movement of the RecA filament along DNA (branch migration). In addition, ATP hydrolysis is required to displace ssDNA from the RecA filament as homologous pairing proceeds.

The role of RecA in DNA repair is not limited to facilitating homologous recombination of DNA, but also includes an important role in the SOS response to widespread DNA damage. The SOS repair mechanism is a broadly based response by the cell to massive DNA damage that allows DNA replication to take place under highly adverse conditions. Although this pathway of replication is highly mutagenic, the possibility that some small fraction of the damaged cells will survive is more evolutionarily advantageous than losing the entire population. The SOS response relies principally upon RecA and

LexA proteins. Upon damage, DNA replication stops and as ssDNA is incorporated into the extended RecA filament, the LexA repressor protein binds deeply within the RecA filament. The binding of LexA to the ssDNA:RecA filament stimulates an autocatalytic protease action within LexA to start the SOS response. By stimulating the autocatalytic cleavage of LexA, RecA can be classified as a coprotease as well as a DNA recombination protein.

The three-dimensional structure of RecA was determined in the absence of DNA by X-ray crystallography (Story *et al.*, 1992). Despite the absence of DNA, the RecA crystal structure showed that the RecA molecules formed a helical structure consistent with the helical dimensions measured for the 'collapsed' filament. An examination of the RecA structure showed a central domain and that the most conserved region corresponded to the ATP-binding site called a 'P loop'. In addition, the central domain is thought to be responsible for DNA binding, while the residues at the N-terminus of the protein are essential to formation of the RecA filament via RecA–RecA interactions. RecA does not use the OB fold for ssDNA binding, presumably because it must bind at least three strands of DNA in order to facilitate homologous recombination. The need to bind three DNA strands in a proposed triplex DNA structure probably demands an overall architecture for binding different from that seen in the OB fold-containing family of proteins. Sequence and structure analysis of RecA and its homologues has led to the proposal that RecA may have diverged from a common ancestor that was present prior to the divergence of prokaryotes and eukaryotes.

## Telomere End-binding Protein (TEBP) – a Sequence-specific ssDNA-binding Protein

Eukaryotic chromosomal stability relies on proper maintenance of the DNA sequences at the very ends of the chromosome, called telomeres (Gottschling and Stoddard, 1999). Faithful replication of chromosomal DNA requires intact telomeric sequences. The telomere sequence itself varies with the eukaryote, but is typically a short, repeated, TG-rich DNA sequence at the $3'$ terminus of the chromosome. Failure to maintain telomeres causes cells to lose their ability to divide, resulting in senescence. In the ciliated protozoa *Oxytrichia nova*, telomeres are only 36 nucleotides long and the $3'$ end 16 nucleotides are single-stranded and made up of two repeats of the sequence TTTTGGGG ($T_4G_4$). As described previously in this chapter, exposed ssDNA is vulnerable to the action of nucleases and other processes that might damage it. To safeguard the vulnerable telomere $3'$ end, telomere end-binding protein (TEBP) acts as a cap to sequester the ssDNA. Unlike the SSBs or RecA, TEBP binds tightly and in a sequence-dependent manner to the $(T_4G_4)_2$ sequence. TEBP is made up of two polypeptides, $\alpha$ and $\beta$ that dimerize in the presence of $(T_4G_4)_2$ to form an exceptionally tight 1:1:1 complex.

The structure of TEBP bound to a telomeric DNA sequence has been determined and shows that TEBP uses many of the same strategies for ssDNA binding as those employed by nonsequence-specific ssDNA-binding proteins (Horvath *et al.*, 1998). An examination of the structure shows that $\alpha$-TEBP has three copies of the OB fold motif while $\beta$-TEBP has one (**Figure 2**). The telomeric ssDNA sequence is bound by two OB folds of the $\alpha$-TEBP and one from $\beta$-TEBP. The C-terminal OB fold of $\alpha$-TEBP appears to be involved in making protein–protein interactions with an oligopeptide loop from $\beta$-TEBP. To date, the use of an OB fold to make protein:protein contacts appears to be unique to $\alpha$-TEBP. The OB folds of TEBP form a deep V-shaped cleft with the ssDNA in a highly irregular conformation with the $5'$ and $3'$ ends nearly meeting to form a semicircle. As expected, TEBP uses a combination of nonsequence-specific electrostatic, hydrogen-bonding and stacking interactions to achieve a basal level of ssDNA affinity and then further increases its affinity for the $(T_4G_4)_2$ sequence by making highly specific interactions that are tailored to match the shape and hydrogen bonding characteristics of each nucleotide. The ssDNA-binding surface of TEBP has deep pockets whose surfaces are complementary to specific nucleotides. In addition, several of these pockets are formed from different OB folds, consistent with the notion that the $\alpha$ and $\beta$ subunits of TEBP dimerize only in the presence of the $(T_4G_4)_2$ sequence.

The three-dimensional structure of the TEBP:$(T_4G_4)_2$ complex also suggests that complex formation may proceed through a cofolding mechanism rather than a more static model where a rigid TEBP molecule would bind to ssDNA (Horvath *et al.*, 1998). In a cofolding model, TEBP would be flexible in the absence of ssDNA, but once the telomeric ssDNA is present, the protein and DNA bind synergistically to induce formation of the nucleotide-binding pockets that are made up from different OB folds. This model is supported by the observation that the $3'$ end and much of the $(T_4G_4)_2$ sequence is deeply buried, and it is therefore difficult to see how the TEBP:$(T_4G_4)_2$ complex could form if either component were rigid. A comparison of the interaction of TEBP with telomeric ssDNA and the interaction of SSBs with replication fork ssDNA suggests that despite their different physiological roles, the predominant function of these proteins is to protect exposed ssDNA from damage.

## Summary

The accurate and faithful replication of DNA has necessitated the evolution of proteins specialized in the

handling of ssDNA. ssDNA-binding proteins are ubiquitous in DNA metabolism and have essential roles in DNA replication, recombination and repair. Many of these proteins employ a common structural motif called the OB fold to bind tightly and preferentially to ssDNA. Depending on the protein, more than one OB fold can be used for binding to ssDNA. Although OB folds are quite common, ssDNA-binding proteins such as RecA and adenovirus DBP (SSB) have none. This illustrates the principle that evolution often devises more than one solution to a problem (Tucker *et al*., 1994). The fact that some ssDNA-binding proteins do not have OB folds but still bind tightly to ssDNA via nonsequence-specific electrostatic, hydrogen-bonding and stacking interactions demonstrates that although protein folds may be quite variable, the chemical basis of ssDNA recognition remains a conserved theme in all ssDNA-binding proteins.

# References

Bochkarev A, Pfuetzner RA, Edwards AM and Frappier L (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**: 176–181.

Bycroft M, Hubbard TJP, Proctor M, Freund SMV and Murzin AG (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**: 235–242.

Gottschling DE and Stoddard B (1999) Telomeres: structure of a chromosome's aglet. *Current Biology* **9**: R164–R167.

Horvath MP, Schweiker VL, Bevilacqua JM, Ruggles JA and Schultz SC (1998) Crystal structure of the *Oxytrichia nova* telomere end binding protein complexed with single strand DNA. *Cell* **95**: 963–974.

Murzin AG (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO Journal* **12**: 861–867.

Raghunathan S, Ricard CS, Lohman TM and Waksman G (1997) Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9 Å resolution. *Proceedings of the National Academy of Sciences of the USA* **94**: 6652–6657.

Roca AI and Cox MM (1997) RecA protein: structure, function, and role in recombinational DNA repair. *Progress in Nucleic Acids Research* **56**: 129–223.

Shamoo Y, Friedman AM, Parsons MR, Konigsberg WH and Steitz TA (1995) Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature* **376**: 362–366.

Story RM, Weber IT and Steitz TA (1992) The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**: 318–325.

Suck D (1997) Common fold, common function, common origin? *Nature Structural Biology* **4**: 161–165.

Tucker PA, Tsernoglou D, Tucker AD *et al*. (1994) Crystal structure of the adenovirus DNA binding protein reveals a hook-on model for cooperative DNA binding. *EMBO Journal* **13**: 2994–3002.

Yang C, Curth U and Kang CH (1997) Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nature Structural Biology* **4**: 153–157.