# Chapter 3
# How Proteins Slide on DNA

**Daniel Barsky, Ted A. Laurence, and Česlovas Venclovas**

## 3.1 Introduction

Protein–DNA interactions are required for all the major functions of DNA: transcription and regulation, replication and repair, even the packaging of DNA into chromosomes. Not only are protein–DNA interactions crucial for all these cellular activities, but they are also, in our view, among the most fascinating macromolecular interactions because of their dynamics. In this chapter, we focus on DNA sliding by proteins, particularly diffusive sliding. Such sliding is typically part of the search for a target on the DNA itself or for another protein bound to the DNA. Of particular interest here are the proteins known as DNA sliding clamps that can remain bound to the DNA while diffusing vast distances along the double helix of DNA. We do not yet know the detailed mechanisms of protein sliding on DNA, but we aim to familiarize the reader with what is known observationally and to provide some discussion of potential mechanisms.

*Passive vs. active sliding*. Proteins that interact with DNA can be divided into two groups: those that actively move along DNA ("active sliders") and those that do not. The first group contains proteins such as DNA and RNA polymerases and helicases, and because of the "processivity" of their functions (e.g., incorporating one base after another), it is intuitively obvious that these proteins can slide along DNA, that is, remain in contact with DNA while moving along it. Many proteins in the second group also slide or hop along DNA, either through attachment to active sliders or by diffusion ("passive sliders"). Among the members of the second group are proteins such as DNA replication processivity factors that function to promote the retention of an active slider (such as a DNA polymerase) on the DNA. Because they remain bound to DNA independent of binding to other proteins, the processivity factors have been dubbed "DNA sliding clamps". In isolation on DNA, a DNA sliding clamp becomes a passive slider. DNA sliding clamps exist in all life forms. Examples include PCNA (eukaryotic and archaeal), β clamp (bacterial), UL42/44 (viral), and gp45 (of phages). Other passive

D. Barsky (✉)
Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory,
7000 East Avenue, Livermore, CA 94550, USA
e-mail: barsky@cornell.edu

M.C. Williams and L.J. Maher, III (eds.), *Biophysics of DNA-Protein Interactions*,
Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-0-387-92808-1_3, © Springer Science+Business Media, LLC 2011

sliders are not obviously sliders at all. In a sense, the fact that sequence-specific proteins such as restriction endonucleases and transcription factors can slide along DNA may appear surprising. After all, a biomolecule that binds its specific target sequence with nanomolar specificity might be expected to bind and stay put. The very interactions that stabilize specific binding (i.e., the "binding specificity") would seem to ensure that sliding does not occur. Nevertheless, even proteins that bind their DNA targets very tightly slide or hop along DNA to get there.

*Brief historical perspective*. The concept that proteins passively slide on DNA to locate specific binding sequences is at least four decades old. As with many ideas in science, it is difficult to determine when the idea was first formulated. In the early 1980s, Otto Berg, Peter von Hippel, and colleagues wrote a series of landmark papers. In one of those papers, Berg et al. [1] cited a 1968 book chapter by Adam and Delbrück [2]. Hence, we know the concept is at least four decades old. A defining (and likely independent) moment occurred in 1970 when it was reported that the lacI repressor protein could locate its target site about two orders of magnitude faster than predicted by normal diffusion-collision mechanism (i.e., three-dimensional (3D) diffusion and random collision) [3]. Given the sensitivity of the result to the salt concentration, it was immediately interpreted as evidence that the DNA electrostatically attracts the protein [3]. Later, the search for the specific DNA target was modeled as a two-step process in which the search was "facilitated" by first binding the DNA molecule nonspecifically and then sliding to the target [1]. In the ensuing decades, a combination of biochemical assays and theoretical considerations gave rise to an increasing appreciation of the importance of sliding along DNA for many DNA-binding proteins. The latest chapter in the lacI repressor story is presented by Wang and Austin in this book.

*Sliding, hoping, and jumping*. In their papers, Otto Berg, Peter von Hippel, and colleagues developed a theory of this facilitated diffusion model and followed it with further experiments on lacI [1, 4, 5]. They laid out four possible modes of DNA target searching by proteins – more generally, modes of moving from one site to another on DNA: (1) *sliding* along DNA via continuous one-dimensional (1D) diffusion without dissociation, (2) *hopping*, where the protein effectively diffuses along a single molecule of DNA but does so via a series of dissociation and rebinding events, (3) *jumping*, which, in contrast to facilitated diffusion, amounts to ordinary 3D diffusion between DNA sites, and (4) *intersegmental transfer*, where the protein swaps sites on the DNA via a looped intermediate (Fig. 3.1). To date with some minor differences all models of proteins moving on DNA have defined these same modes [6, 7]. The first, second, and third modes present a continuum transition from 1D diffusion on DNA to 3D diffusion. The fourth mode, intersegmental transfer, does not fit so neatly into this continuum. This mode requires at least two DNA-binding domains in a special arrangement.

While some experimentation and refinement of this theory took place in the ensuing decades, there has been an explosion of activity in just the last few years, particularly through direct observation of protein sliding on DNA. In our review of that activity, we consider only passive sliding. For investigations of active sliders such as polymerases, there are recent reviews [8, 9]. Many proteins have been observed passively sliding on DNA (Table 3.1). These include proteins responsible for detecting DNA damage (hOgg1, MutM, Msh2–Msh6, Ada, Rad51), proteins

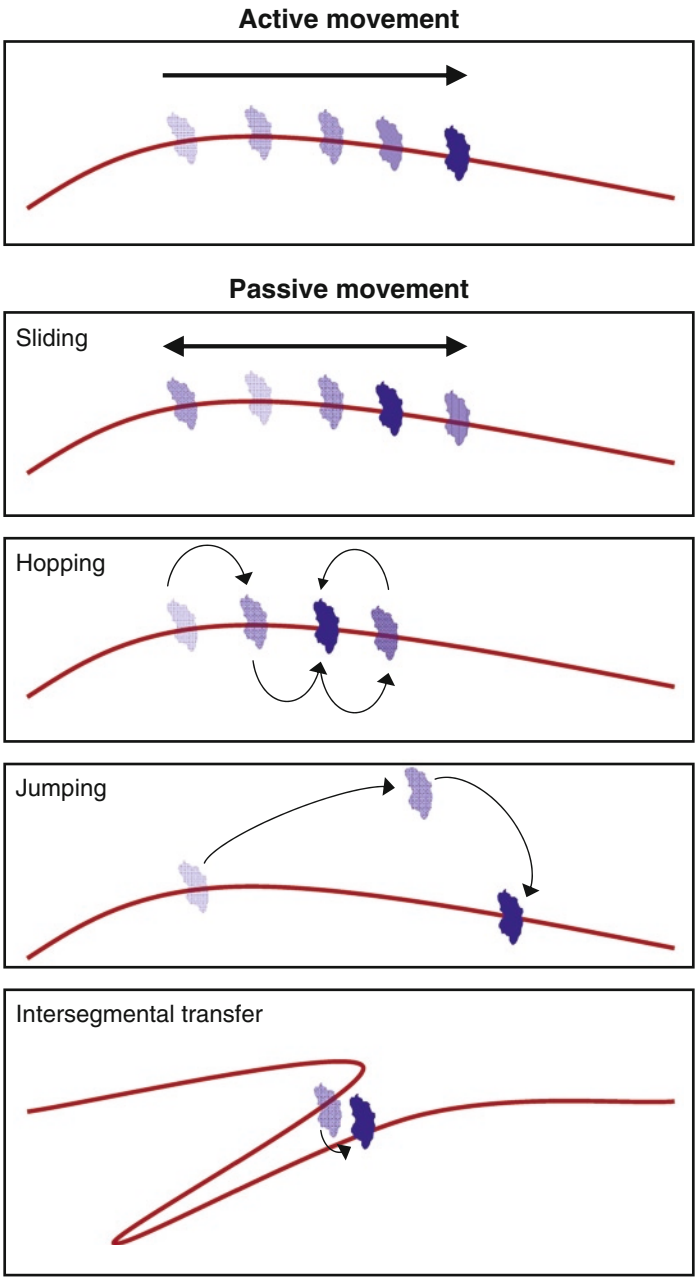


**Fig. 3.1** Modes of proteins searching for DNA targets. This chapter focuses on passive sliding

**Table 3.1** Proteins and their diffusion constants, as measured by single-molecule methods. All of the measurements use single-molecule tracking, except the *Eco*RI [30] and the β clamp [11] measurements which use a biochemical assay and single-molecule FRET, respectively

| Protein | Diffusion constant ($m^2/s$) | References |
|---|---|---|
| *Eco*RV | $3\times10^{-15}$ | Biebricher et al. [20] |
| *Eco*RV | $10^{-14}$ | Bonnet et al. [21] |
| *Eco*RI | $3\times10^{-15}$ | Rau and Sidorova [30] |
| hOgg1 | $6\times10^{-13}$ | Blainey et al. [22] |
| MutM | $4\times10^{-14}$ | Blainey et al. [22] |
| AVP–pVIc | $2\times10^{-12}$ | Blainey et al. [31] |
| *Bam*HI | $6\times10^{-13}$ | Blainey et al. [31] |
| Msh2–Msh6 | $9\times10^{-14}$ to $2\times10^{-16}$ | Gorman et al. [18] |
| Rad51 | $10^{-13}$ to $10^{-14}$ | Graneli et al. [32] |
| RNAP | $10^{-14}$ | Harada et al. [23] |
| LacI | $10^{-13}$ to $2\times10^{-16}$ | Wang et al. [27] |
| C-Ada | $10^{-12}$ | Lin et al. [25] |
| p53 | $3\times10^{-13}$ | Tafvizi et al. [26] |
| PCNA | $10^{-12}$ | Kochaniak et al. [24] |
| β clamp | $10^{-14}$ | Laurence et al. [11] |

involved in cutting DNA in specific places (*Eco*RI, *Eco*RV, *Bam*HI), promoter and repressor DNA transcription factors (lacI, p53), the adenoviral AVP–pVIc complex, the RNA polymerase (RNAP), and the DNA sliding clamps (PCNA, β clamp) that aid critically in DNA replication and are also involved in some forms of DNA repair. Many of these proteins bind to a specific DNA feature; they slide on DNA in the course of searching for that feature. Although a fascinating topic in its own right, in this chapter we are not concerned with the search for DNA targets per se, but rather the mechanism by which proteins actually move along DNA. We thus focus mainly on the first mode of translocation, *bona fide* sliding on DNA, and on its nearest alternative, hopping. From this point of view, the DNA sliding clamps are particularly interesting because these proteins form closed rings around DNA and are therefore topologically constrained to remain bound. As we will see below, the diffusion of DNA sliding clamps remains poorly understood.

## 3.2 Experimental Observations of Sliding

Monitoring protein movement along DNA is an active area of research with many papers reporting evidence for sliding, speeds of target acquisition, and measured diffusion constants. Until recently, biochemical assays commonly involving various pathways and traps were the most frequently used methods, but these have now been largely overtaken by single-molecule assays, especially single-molecule tracking by total internal reflection fluorescence (TIRF) microscopy. Other single-molecule methods include atomic force microscopy (AFM) and fluorescence resonance energy transfer (FRET).

Intuitively, it might be expected that it would be difficult to distinguish between sliding and hopping on DNA, and indeed, this is the case [10]. What is known about translocation in terms of distance and time is dictated by what can be observed experimentally. In this respect, the very short time and length scales involved in translocation over a small number of base pairs can make the distinction between sliding and hopping very difficult. Most experiments observe movement over many seconds and across hundreds to tens of thousands of base pairs, although some recent reports probe distances as short as tens of base pairs with time resolution in submicroseconds [11].

Because active sliding tends to be directional, it can be characterized by a simple rate, with units of base pairs per second (or in SI units, m/s). Passive sliding is a diffusional or Brownian process and therefore is not properly characterized by a rate, but rather by a diffusion constant ($D$). The diffusion constant is related to the average distance that a particle moves ($\Delta x$) in a time ($t$) by Einstein's simple relation:

$$\left\langle \Delta x^2 \right\rangle = 2nDt \tag{3.1}$$

where $n$ is the number of dimensions involved in diffusion (i.e., 1, 2, or 3). This has two important consequences. First, the time it takes for a protein to cover a distance by random search grows very rapidly, as the square of the distance. Second, $\Delta x$ is a distance independent of direction, so, for example, after a time interval $t$ a protein sliding on DNA with diffusion constant $D$ is just as likely to be $\Delta x$ "beyond" or "behind" its original position. Despite this distinction between active and passive sliding, it is common nonetheless to use dynamic adjectives such as "fast" or "slow" to characterize passive sliding with relatively high or low diffusion constants.

### *3.2.1 Biochemical Assays*

To date, a great wealth of data has come from biochemical assays. By measuring the rate of product production, biochemical assays that measure the productivity and processivity of active sliders, such as DNA polymerases, provide the sliding rate for these proteins. For example, we know that DNA polymerase III holoenzyme moves along the template strand, reading it and building the opposite strand at the rate of roughly one base every 2 ms [12], a measurement more recently confirmed by single-molecule methods [13]. For the passive sliders, similar strategies can be used, even in cases where no product is produced (nonenzymatic sliders). Although biochemical assays are not ideal for monitoring motion, certain clever experiments can reveal critical information about sliding. In many cases, the biochemical assays make use of DNA topology such as closed loops or even catenated DNA (where one loop runs inside another, as in links of a chain) and careful placement of targets and starting points [14].

One strategy is to load the protein onto DNA at a known location and then measure the time it takes to reach a target, as has been done for a DNA sliding clamp, the *Escherichia coli* β clamp more than a decade ago in the O'Donnell lab. In a control experiment, the β clamp was loaded onto a nicked, circular DNA plasmid (7,200 bp), and the β clamp remained on the circular plasmid with a half-life of over an hour – a long time relative to the life of a bacterial cell [15]. When the circular DNA plasmid was cut open by a site-specific endonuclease, the β clamp disassociated from the DNA within a few minutes by sliding off one of the free ends of the "linearized" plasmid. Although those experiments were intended as qualitative observations of sliding on DNA and not designed to measure the diffusion constant, we can nevertheless calculate at least a lower limit for the diffusion constant. After 1 min of linearizing (endonuclease) reaction, plasmids were analyzed by approximately 15 min of gel filtration, after which little or none of the sliding clamps were present on the linearized DNA [15]. Using this time (~15 min), and assuming that clamps were at random positions on the 7,200-bp plasmid upon linearization, a "survival probability" analysis [16] yields a relation between the time τ, interval length $L$, and diffusion constant $D$, such that

$$\tau = L^2 / \left( D\pi^2 \right), \tag{3.2}$$

and this gives a limit of $D \geq 10^{-15}$ m$^2$/s.

Rau and Sidorova devised a method for measuring sliding on DNA based on the ratio of the dissociation rate of the *Eco*RI restriction endonuclease from DNA fragments containing one specific binding site vs. the dissociation rate of *Eco*RI from DNA fragments containing two specific binding sites [30]. The measurements enabled the authors to determine the sliding rate (Table 3.1), finding it relatively insensitive to salt concentration and osmotic pressure, and indicating that the "protein moves smoothly along the DNA probably following the helical phosphate-sugar backbone of DNA." They estimate that the sliding rate they measure is 2,000-fold slower than the diffusion of free protein in solution. A factor of 40–50 can be accounted for by rotational drag that would result from following a helical path on the DNA. They suggest that the other factor of 40–50 could reflect the requirement for making and breaking salt bridges between the DNA and the protein, or to the disruption of water structure at the protein–DNA interface as the two surfaces move past each other.

Besides biochemical approaches, there have been a variety of spectroscopic approaches, new and old. Older approaches such as fluorescence recovery after photobleaching (FRAP) are poorly suited to such measurements. Extracting information on the diffusion of sliding clamps along DNA would require synchronized loading of clamps on many aligned and uniformly stretched DNA molecules. Even so, some useful information can be obtained from older optical methods. For example, fluorescence anisotropy allows one to create a nonequilibrium situation very simply by exciting only the fluorophores whose excitation dipoles happen to be aligned with excitation light. Although definitive information on diffusive motion is difficult to obtain by this approach, Austin et al. [17] were able to place an upper bound on the diffusion rate of *E. coli* RNA polymerase on DNA using this technique. The newer, single-molecule biophysical methods are much more effective in revealing diffusion-based sliding.

### 3.2.2 Single-Molecule Methods

There has been an explosion of single-molecule research for measuring the sliding of proteins on DNA. The diffusion constants of many passive sliders have now been measured primarily by single-molecule methods. Fourteen examples are listed in Table 3.1. Single-molecule measurements are ideal for measuring diffusion of proteins such as sliding clamps on DNA because single-molecule methods directly observe the Brownian motion involved in diffusion. Diffusion constants are then determined by statistical analysis of the Brownian motion. Two single-molecule approaches have been used to observe diffusion of proteins on DNA. First, *single-molecule tracking* measures the position of a protein with an attached fluorophore moving on stretched DNA [18–27]. Second, *single-molecule FRET* monitors distance changes between a donor-labeled protein moving on acceptor-labeled DNA using FRET efficiency changes [11, 28, 29]. The latter methodology does not require stretched DNA or even surface attachment to reveal diffusion.

*Single-molecule tracking*. This is the most common single-molecule method used to measure protein sliding on DNA. The positions of fluorescently labeled DNA-binding proteins are dynamically monitored as they move on stretched DNA [18–27]. For a recent review, see Gorman and Green [19]. The fluorescent labels for the proteins are small organic fluorophores (such as Alexa 488 or Cy3B) [21–26], fluorescent proteins [27], or quantum dots [18, 20, 24]. These methods directly monitor the position of the fluorescently labeled proteins by measuring their position using a CCD camera (Fig. 3.2a). Excitation is provided by total internal reflection of the fluorescence (TIRF), providing a wide area of excitation that only penetrates ~200 nm into the solution. This is an excellent approach to image wide areas with low background from solution. One exception is provided by Biebricher et al. [20], where the authors take advantage of especially bright quantum dots and use bright field imaging to monitor diffusion away from the glass surface. One current drawback of the TIRF approach is that the CCD cameras used typically have time resolutions in the order of 10–100 ms. This prevents monitoring of faster diffusive motions, which is unfortunate, since the diffusive motion may vary over different length and time scales. For example, there is some evidence of multiple modes of diffusion in the case of PCNA [24]. The time resolution limit is not intrinsic to the method and may be improved with newer technology, although signal-to-noise issues will ultimately limit how fast the position may be monitored. The chapter in this book by Wang and Austin discusses methods to improve the time resolution.

When monitoring the position of the protein directly, stretching of the DNA is necessary (at least to the point of straightening) so that position measured using the camera can be directly linked to position on the DNA. Careful preparation of stretched or elongated DNA is characteristic of this methodology. The mode of stretching is another primary methodological distinction between these studies. The DNA may be attached to the surface at both ends [18, 21, 24], attached at one end and elongated using fluid flow [22, 25, 26], or stretched between beads in optical traps [20, 23]. The proximity of the surface is a possible issue with these studies (except for Biebricher et al. [20] where the DNA is away from the surface),

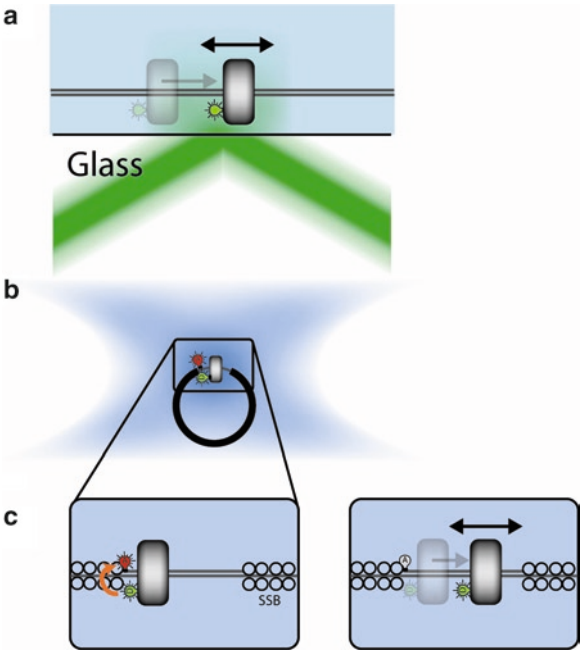


**Fig. 3.2** Single-molecule methods used to track proteins sliding on DNA. (**a**) Single-molecule tracking monitors the position of proteins on stretched DNA using total internal reflection excitation in combination with high sensitivity CCD cameras. The position is obtained by fitting the centroid of the observed fluorescence, obtaining accuracy down to ~50 nm. Typical time resolution is about 30–100 ms. (**b**) Fluorescence resonance energy transfer (FRET) between donor-labeled protein and acceptor-labeled DNA may be used to track position of protein on DNA. Using correlations, fluctuations in distance on faster time scales (<1 ms) may be monitored. (**c**) *Left panel*: When the protein labeled with donor (*green bulb*) is near the acceptor label (*red bulb*), FRET occurs (*orange arrow*). *Right panel*: When the protein diffuses away from the acceptor label, FRET decreases or disappears

but these studies have included several controls that confirm that the biochemical systems operate correctly. In the experiments that stretch the DNA with flow, the flow velocity can affect the diffusion of the proteins, especially in cases where "hops" between sites are expected.

Care must also be taken when interpreting data obtained using large labels such as fluorescent protein domains and quantum dots since they are similar in size to the labeled proteins and can clearly affect diffusive behavior. In fact, Kochaniak et al. [24] used a large quantum dot to detect helical sliding by the logic that if the protein is constrained to follow the helical shape of the DNA by revolving once about every 10 bp, the larger hydrodynamic radius caused by the attached quantum dot would slow down the measured diffusion considerably more if the protein was constrained to rotate while sliding. These issues are explored further in Sect. 3.3.

*Single-molecule FRET*. The second major single-molecule method for monitoring diffusive motion of proteins on DNA is FRET [11, 28, 29]. FRET involves the non-radiative transfer of electronic excitation energy from donor to acceptor fluorescent

molecules via a weak dipole–dipole coupling mechanism. FRET requires a resonance between the emission of a donor molecule and the absorption of an acceptor molecule, and depends critically on the relative orientations of the dipole moments of the donor and acceptor molecules, although for fluorescently labeled biomolecules orientational averaging typically occurs. Most importantly, FRET has a strong dependence on the distance between the two fluorophores:

$$E = 1 \Big/ \left[ 1 + \left( R/R_0 \right)^6 \right], \tag{3.3}$$

where $E$ is the fraction of donor excitations transferred to the acceptor, $R$ is the distance between donor and acceptor, and $R_0$ is the distance at which $E = 0.5$ ($R_0$ is between 4 and 7 nm for typical fluorophores). For experiments monitoring proteins sliding on DNA, the relative motion of a donor-labeled protein and acceptor-labeled DNA (donor and acceptor labels can be swapped) is monitored by measuring $E$ as a function of time for each single protein. The FRET methodology is complementary to the tracking methodology described above because the length scales probed are so different between the methods. FRET allows detection of proximities (i.e., distances between labels) in the range of 2–8 nm; whereas, by using CCD cameras to monitor the position of the protein, the tracking methods approach the limits of their sensitivity well above this range at around 50 nm. In addition, movement of the DNA itself will not affect FRET experiments, but will cause additional errors in tracking experiments. At longer time scales (≥1 ms), there is sufficient signal to monitor distance changes by calculating $E$ as a function of time. At shorter time scales (<1 ms), fluorescence correlations reveal FRET intensity fluctuations caused by distance fluctuations.

Using the FRET method and fluorescence correlation spectroscopy, we determined a diffusion constant of $D \sim 10^{-14}$ m$^2$/s for the *E. coli* β clamp sliding on DNA [11, 33]. Our experiments achieved fine time-resolution measurements using a single detection volume (~2 fl) defined by confocal microscopy, detecting photons with avalanche photodiodes (APD), and obtaining time resolutions down to 100 ns [11]. By loading a single protein (β clamp) onto a small section of each large DNA plasmid, we were able to monitor sliding clamps diffusing on DNA via FRET without any immobilization, that is, while both protein and DNA were freely diffusing in solution (Fig. 3.2b and c). Given that a protein with diffusion constant $D$ moves a mean-square distance given by (3.1), $\Delta x^2 = 2D\Delta t$ within time $\Delta t$, and the fact that FRET efficiency decreases from nearly 1 to nearly 0 within about 5 nm, we can obtain limits on the diffusion constants measured. The time resolution of 100 ns limits one to measuring diffusion constants $D < 10^{-10}$ m$^2$/s. In our detection methodology, proteins and DNA can diffuse in and out of the detection volume, giving rise to "bursts" of fluorescence. Bursts of fluorescence due to FRET are observed when sliding clamps are loaded on DNA, and then the DNA–clamp complexes diffuse in and out of the confocal detection volume. Diffusion of the DNA–sliding clamp complex out of the confocal detection volume in free solution meant that we could not observe the sliding motion for arbitrarily long periods of time and therefore limited our measurements to $D \geq 10^{-14}$ m$^2$/s. The diffusion constant of the β clamp sliding on DNA was very close to this lower limit.

Single-molecule FRET measurements are commonly performed using TIRF microscopy in combination with a CCD camera to monitor several protein/DNA systems simultaneously [28, 29]. Unfortunately, current CCD cameras sensitive enough for single-molecule spectroscopy do not have sufficient time resolution (typically 30–100 ms, down to 1 ms in some cases), limiting measurement to slow diffusive motions $D < 10^{-14}$ m$^2$/s. For many proteins that slide on DNA, it is important to sacrifice the ability to simultaneously monitor many protein–DNA systems for better time resolution.

To date, our measurement of the *E. coli* β clamp has been the only single-molecule FRET measurement of DNA sliding, and it is interesting to compare it with what was more recently reported by single-molecule tracking for the sliding of human PCNA. The proteins are functionally and structurally similar. By single-molecule tracking, Kochaniak et al. determined an overall diffusion constant for PCNA that, at $D = 3 \times 10^{-12}$ m$^2$/s, is two orders of magnitude higher than that for the β clamp [24]. These authors also obtained evidence that 85–99% of the motion reflected the clamp in a slower diffusing mode that likely tracks the groove of DNA. The remainder of the time, the clamp diffuses much more quickly without tracking DNA groove. It is intriguing that for the β clamp, we were unable to detect any evidence for fluctuations that would be expected for $D = 10^{-12}$ m$^2$/s, precisely where our correlation-based detection is most sensitive [11]. Indeed, the lack of fluctuations at the appropriate time scale for such fast diffusion was one evidentiary fact from which we deduced that $D \sim 10^{-14}$ m$^2$/s. Assuming that the dimeric β clamp behaves quite similarly to the trimeric PCNA in terms of sliding behavior, we believe that the large discrepancy between the two measurements likely involves the two modes of diffusion proposed by Kochaniak et al., based on their experimental results [24]. If the sliding clamp spends 2–15% of its time in the fast mode, fluorescence correlation amplitudes for that motion would be correspondingly small (i.e., 2–15% of the total fluctuations, well within the noise level of our previous experiments). In fact, the lack of significant fluctuations at the appropriate time scale is consistent with the multiple diffusion mode interpretation of Kochaniak et al. In the end, additional, improved experiments that measure the diffusion of the sliding clamps over multiple length and time scales will be necessary to resolve the puzzle produced by these seemingly discordant results.

## 3.3 Interpreting the Observations

What factors affect the sliding of proteins on DNA? In a Newtonian sense, sliding should be controlled by entities that can exert a force on proteins. Major components that can induce or retard the motion of the DNA sliding protein include other proteins, solvent, and DNA. In fact, the thermal forces that induce sliding through Brownian motion are mediated by the same objects that can also retard the motion. We can avoid confusing the issue by focusing only on the frictional or retarding forces. The "other proteins" could bind only the DNA sliding protein or they too could bind the DNA. These proteins might be actively sliding proteins, such as polymerases, or other

passively sliding proteins. Since we are interested in how proteins slide on DNA, we will not discuss explicit protein–protein interactions except for the important case of a passively sliding protein affecting the active sliding of another, such as a DNA sliding clamp attached to a DNA polymerase (see below). That leaves just the solvent and the DNA to affect the sliding motion, i.e., to produce retarding forces.

*Three drag terms*. The solvent, through its viscosity, will affect the sliding according to the shape of the protein. This "three-dimensional (3D) solvent drag" is the same drag that occurs in ordinary 3D solvent-phase diffusion, and it is characterized by frictional drag term $\alpha$, which can be calculated by solving the Navier–Stokes equations. For simple shapes, $\alpha$ is given by the product of the viscosity of the medium, the size of the protein, and a geometrical factor (a simple example is given below).

The DNA will affect sliding in two possible ways: (1) frictionless forces that nevertheless create additional solvent drag by compelling the protein to take a longer path (e.g., a helical one) through the solvent and/or to rotate while moving along the DNA, and (2) frictional drag due to a series of energetic barriers that the DNA presents to the sliding protein. For simplicity in terminology, we call (1) "DNA-induced solvent drag" and (2) "DNA surface drag." Together, the 3D solvent drag, the DNA-induced solvent drag, and the DNA surface drag constitute three drag terms that completely describe the retarding forces for proteins sliding on DNA (Fig. 3.3). What can create confusion is that observations of DNA sliding cannot usually detect the actual path (e.g., rotations or revolutions) of the protein and thus assume simple 1D diffusion, sometimes called "effective 1D diffusion". At least one author makes a distinction between 1D and 2D sliding, with 1D/2D indicating that the protein is/is not required to rotate as it slides, respectively [34]. This terminology has not been widely followed; we and others simply describe both of these as "1D" diffusion that either does or does not follow a helical path along the DNA.

In focusing on the passive sliding of a protein, sliding on DNA is most simply characterized by the effective 1D diffusion constant, a quantity that can be measured as previously discussed in Sect. 3.2. To date, over a dozen such diffusion constants have been measured. The most striking feature of the diffusion constants listed in Table 3.1 is that they span a very large range, covering four orders of magnitude, and in a few cases a single protein ranges over three orders of magnitude. So far, the fastest sliding proteins are AVP–pVIc, Ada (C terminus), and PCNA, with diffusion constants of $10^{-12}$ m$^2$/s to $2\times10^{-12}$ m$^2$/s. At the other extreme are the lacI repressor and the Msh2–Msh6 repair protein, whose sliding is characterized by diffusion constants as low as $2\times10^{-16}$ m$^2$/s. (The sliding of SSB on single-stranded DNA might be even slower [29], but it is not directly comparable to sliding on double-stranded DNA because the structure of single-stranded DNA is so different from double-stranded DNA). To put these numbers in perspective, we compare them in two ways. First, we compare the measured diffusion constants with the solvent-phase 3D diffusion constants of small molecules and proteins. Second, we interpret the diffusion constants in terms of distance and time scales in sliding on DNA. For the first, it may be helpful to point out that a "3D" diffusion constant can be equally valid for characterizing 1D diffusion (recall that the Einstein relation always uses the same diffusion constant, but changes through the factor $n$ according

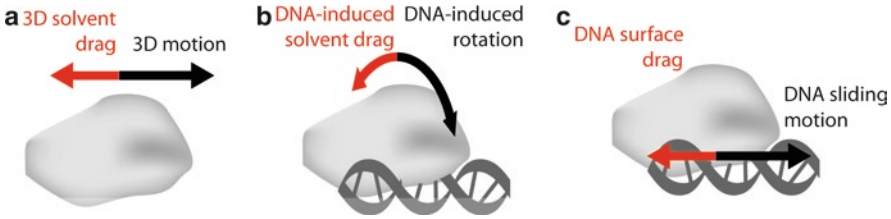


**Fig. 3.3** The three drag terms relevant to DNA sliding. From *left to right*: 3D solvent drag, the drag that determines the "ordinary" diffusion constant for a protein in solution; DNA-induced solvent drag, the "extra" solvent drag that a protein experiences while sliding along DNA due to requirements to rotate or revolve around the DNA; DNA surface drag, the drag that a protein experiences due to the DNA directly inhibiting the sliding of the protein

to the dimensionality). The projection of the 3D diffusion of a protein onto a single coordinate axis is an idealized model of diffusion on a frictionless, rod-like DNA, i.e., with neither DNA-induced solvent drag nor DNA surface drag. Therefore, the 3D diffusion constant can be viewed as the baseline 1D diffusion constant for sliding on DNA. The actual diffusion constant for sliding on DNA will be reduced from that by the DNA-induced solvent drag and DNA surface drag terms.

*The 3D solvent drag*. The diffusion constant for (3D) solvent-phase diffusion decreases as the size of the molecule increases. The diffusion constant of water molecules in pure water is about $3\times10^{-9}$ m$^2$/s. Proteins typically exhibit diffusion constants that are one or two orders of magnitude smaller. For a protein diffusing in solution the diffusion constant can be estimated by Stokes–Einstein relation:

$$D = k_B T/\alpha \tag{3.4}$$

where $k_B$ is the Boltzmann constant, $T$ is the temperature, and $\alpha$ is the frictional drag term introduced above. For the simplest case of a spherical protein, $\alpha$ is given by:

$$\alpha = 6\pi\eta R, \tag{3.5}$$

where $\eta$ is the viscosity of water and $R$ is the radius of the sphere. The physicist's infamous spherical approximation is not as bad as it might appear. For example, we can approximate the donut-shaped β clamp as a sphere, using $a$=4.5 nm, the approximate cylindrical radius of the β clamp; and at temperatures 25–40°C (i.e., 298–313°K) this gives an estimate of $D$ in the range of $5\times10^{-11}$ to $8\times10^{-11}$ m$^2$/s. To avoid the crude spherical approximation, one can also solve the Navier–Stokes equations numerically for the molecular structure of the β clamp [35], using the program, hydropro [36]. At 27°C this calculation gives $D$=$5.7\times10^{-11}$ m$^2$/s, which agrees very well with our previous estimate. These solvent-phase diffusion constants characterize 3D diffusion, not necessarily diffusion on DNA; however, as reasoned above, the 1D diffusion rate for the sliding clamp on DNA cannot be larger than its ordinary (3D) diffusion constant in solution. Hence, we can postulate an upper limit for diffusion on DNA of $D\sim10^{-10}$ m$^2$/s.

The reported diffusion constants for the DNA sliding clamps are at least two orders of magnitude lower than this.

*The DNA-induced solvent drag*. As described above, the nature of protein–DNA interactions may give rise to an induced rotation and/or revolution of proteins sliding on DNA. In 1979 Schurr wrote a short report in response to early suggestions that low effective diffusion constants for proteins sliding on DNA were due to large energy barriers for proteins sliding on DNA, calculating the extra solvent drag that would occur if the lacI protein were forced to rotate during its translocation along the DNA axis [37]. He introduced the DNA-induced solvent drag term (with other words) to compete with the DNA surface drag term, and estimated for the lacI protein the DNA-induced solvent drag was responsible for a two order of magnitude reduction in the effective diffusion constant. To keep the model simple, and perhaps because of the lack of insights from crystal structures of protein–DNA complexes, Schurr assumed that the protein would rotate on the DNA. He found for the total solvent drag the formula,

$$\alpha = 6\pi\eta R\left[1+\left(4/3\right)\left(2\pi R/b\right)^2\right] \tag{3.6}$$

where $b$ is the distance along the DNA traveled in one rotation, about 3.4 nm for the DNA double helix. In (3.6) the first term is simply the usual Stokes drag (Fig. 3.3a) and the second term is due to the rotational drag (Fig. 3.3b). This model was recently updated to include a term for revolution of a protein on the DNA [38], allowing one to calculate the DNA-induced solvent drag for a nonrotationally symmetric protein as well as for a particle attached to the DNA-binding protein, such as a bead or quantum dot. The observation of rotation/revolution during sliding has been mostly indirect. Using the fact that the rotational/revolutionary part of the DNA-induced solvent drag has a $1/R^3$ dependence on the radius of the protein (or similar for the radius of the revolution) while the 3D solvent drag has a $1/R$ dependence on the radius of the protein, Blainey et al. estimated the contribution of the DNA-induced solvent drag for a handful of proteins based on the effective diffusion constant and concluded that all undergo rotation-coupled sliding [31]. Different arguments for the rotation-coupled sliding of PCNA have been made (see below) [24].

While it may be intuitive that some proteins would follow the helical form of the DNA, by tracking a groove or the phosphate backbone, it is less obvious that the DNA sliding clamps would be required to rotate while sliding – that is, sliding on DNA more like a nut on a bolt than like a washer. To illuminate the subject better, instead of using the above formula for $\alpha$, we can reason as follows. If the β clamp were constrained to follow the groove of the DNA, it would be slowed down by the additional viscous drag as the clamp rotated once around the DNA for every 10 bp or so of linear diffusion – what we termed "DNA-induced solvent drag" (above). We can account for this drag by calculating the rotational diffusion constant of the protein as it rotates on its axis. The hydropro program [36] gives a rotational diffusion constant $D_{\mathrm{rot}} = 2.3\times10^6$ to $3\times10^6$ rad$^2$/s, depending on the axis of the rotation. Using the average value and the rotational analog to the Einstein relation,

$$\left\langle \theta^2 \right\rangle = 2D_\gamma t, \tag{3.7}$$

we find that it takes 7 μs to make a full rotation around the DNA helix. By contrast, using the solvent viscosity-induced diffusion constant $D \sim 10^{-10}$ estimated above, the time it would take to linearly diffuse the distance of one helical turn by simple linear diffusion is just 0.1 μs. Thus, if the clamp were constrained to follow the groove of DNA and therefore rotate as it moved along the DNA, the viscosity of water would slow the clamp down by a factor of about 70 relative to simple linear diffusion. This retardation is nearly as much as the factor of 110, estimated several decades ago by Schurr for the lacI repressor, which turns out to bind to DNA with an altogether different topology [37]. The factor of 70 would give an effective diffusion constant, $D_{\text{eff}} \sim 10^{-12}$ m$^2$/s, that would be measured in our experiments if the diffusion of the sliding clamp were governed only by the viscosity of water and the requirement to follow the groove of DNA (i.e., the combined 3D solvent drag and DNA-induced solvent drag). Because our experiments on the β clamp provided a diffusion constant estimate two orders of magnitude smaller than this value, we have concluded that diffusion of the sliding clamp is retarded well beyond these two drag terms. This implies that there must be significant DNA surface drag (discussed below). It seems logical that the clamp would follow a helical path at least while attached to the polymerase; however, our measurements neither confirmed nor refuted this tendency in the absence of the polymerase.

While it is reasonable to expect essentially the same behavior for the other major DNA sliding clamp, PCNA, a first publication by Kochaniak et al. suggests otherwise [24]. These authors report an apparent diffusion constant of about $10^{-12}$ m$^2$/s, which is what we have just predicted for the combined effects of 3D solvent drag and DNA-induced solvent drag. Yet it may be premature to conclude that for PCNA there is almost no DNA surface drag. First, Kochaniak et al. use two types of viscogens that should differentiate rotational and translational motion, thereby finding evidence for rotationally coupled sliding. Nevertheless, when a large quantum dot (12 nm in radius) was attached to the protein, the protein-dot couple appeared to slide without any rotation, prompting Kochaniak et al. to propose a second, much faster sliding mode that was not coupled to rotation. They proposed that the apparent diffusion constant reflects the sum of a fast linear diffusion and a slow rotationally-coupled diffusion:

$$D_{app} = f\, D_{fast} + (1 - f)\, D_{slow}, \tag{3.8}$$

where $D_{\text{fast}}$ is $10^{-11}$ m$^2$/s or larger and $f$ is the fraction that diffuses at this fast rate. This result leaves $D_{\text{slow}}$ somewhat undefined since it can be close to $10^{-12}$ m$^2$/s for a very small $f$ of ~1–2%, but $D_{\text{slow}}$ can be about $10^{-14}$ m$^2$/s for $f \sim 10\%$. Thus, resolution of this issue awaits improved resolution in the experiments.

*DNA surface drag*. In addition to protein–solvent interactions, protein–DNA interactions can play a major role in limiting the rate of sliding. Most proteins that diffuse on DNA must first adhere, but for the adhesion to cause frictional drag, it must vary as the protein moves. The electrostatic attraction between protein and DNA, for example, does not necessarily create drag. When there is a sufficiently large number of attractive interactions between proteins and DNA, a number of slip–stick interactions can occur, converting kinetic energy of the whole protein moving

into randomized kinetic energy of vibrational motions (i.e., molecular friction). Thus, it is reasonable to postulate a series of local energy minima separated by energetic barriers, where the local minima are due to attraction and adhesion between protein and DNA, while barriers arise as these attractions vary with the changing details of the intermolecular interactions. Therefore, while the presence of attractive forces between the protein and DNA does not necessarily create drag, in practice intermolecular drag is likely. It is interesting to consider that not all sliding proteins must intrinsically adhere to the DNA. The DNA sliding clamps, owing to their ring-like topology, could theoretically remain on the DNA in the absence of any attractive, sticking forces. We will return to these topics in Sect. 3.4 below.

By modeling a protein sliding on DNA as a particle diffusing in a rough potential energy well, it has been shown by Zwanzig that the roughness of the energy landscape determines the diffusion rate [39]. In particular, the author shows that in a rough energy landscape characterized by an average barrier height of $\varepsilon$, the effective diffusion constant $D_{\text{eff}}$ due to the rough energy landscape can be related to the diffusion constant in the absence of the roughness $D$ by the expression,

$$D_{\text{eff}} = D \exp\left[-\left(\varepsilon / k_B T\right)^2\right] \tag{3.9}$$

More precisely, this result is for a Gaussian distribution of barrier heights, where $\varepsilon$ is the standard deviation of the distribution. Although, as Zwanzig points out, this 1D model cannot fully account for the multidimensional dynamics of proteins, it is a useful phenomenological model, allowing us to estimate the energetics that give rise to the drag involved in sliding. Other formulations, based on an Arrhenius model, have been employed [22, 24, 40]. The form of Zwanzig's expression means that large differences in the diffusion constant can reflect small differences in the roughness of the energy landscape. For example, we predicted a 3D diffusion constant for the β clamp sliding on DNA of $D \sim 10^{-10}$ m$^2$/s, and measured a diffusion constant of $D \sim 10^{-14}$ m$^2$/s. If the entire difference in these two diffusion constants was due to DNA surface drag, the average barrier height would be ~3 $k_B T$. Slutsky and Mirny suggest that free 1D diffusion requires energy barriers between positions on DNA to be less than about 2 $k_B T$ [6]. This result implies that the β clamp diffusion constant $D \sim 10^{-14}$ m$^2$/s must include DNA-induced solvent drag. Blainey et al. and Gorman et al. estimated energy barriers from measured diffusion constants and found that the estimated energy barriers were within the 2 $k_B T$ limit [18, 22]. However, Gorman et al. found this to be the case only if the Msh2–Msh6 proteins followed a helical path along DNA, revolving around the DNA with sliding (as we just showed for the β clamp). More recently, Blainey et al. used Zwanzig's formulation to estimate energy barriers near 1 $k_B T$ for several DNA-binding proteins (not all included in Table 3.1) and provided evidence that the proteins track the groove of DNA as they slide [31]. Overall, it appears that DNA surface drag is a dominant factor in sliding. Even the fastest AVP–pVIc protein is retarded by an additional factor of 2–5 by DNA surface drag above retardation due to the combination 3D solvent drag and DNA-induced solvent drag [31].

*Distance and time scales*. Next, we consider what the enormous range of measured diffusion constants (Table 3.1) implies in terms of time and length scales. A natural measure of distance along DNA is the distance from one base pair to the next, ~3.4 Å for B-form DNA, the conformation thought to be most prevalent under biological conditions. Using the Einstein relation we have calculated (Table 3.2) the time it would take a hypothetical protein to diffuse 1, 10 or 100 bp, as well as the mean displacement that occurs in 1 ms. The fastest diffusing proteins ($D \sim 10^{-12}$ m$^2$/s) travel over 100 bp in a millisecond. The slowest proteins scarcely move on the millisecond time scale.

*Sliding vs. hopping*. In addition to the body of experimental observations, there has also been a growing body of theoretical arguments concerning the movement of proteins on DNA. In just the last few years several dozen theoretical papers have been published. A recent paper by Wunderlich and Mirny on DNA target searching gives a good summary of the latest theoretical developments for proteins that diffuse only a certain distance along DNA (sliding mode) before dissociating from the DNA (hopping or jumping mode) [10]. The average such distance is defined to be the protein "sliding length," ($s$) an experimentally determined parameter that is key to the remainder of the analysis. The experimental determination of $s$ can be, however, thwarted by the presence of very small hops, which, the authors point out, would be missed by single-molecule observations. For some proteins, the theory predicts very small hops, with a median distance of about 1 bp, but such small hops could arguably be considered simply part of the sliding mechanism with one important difference: hops are more likely to lead to jumps. Because such small hops do not affect the diffusion constant, the authors conclude "the major contribution of hops is to the duration of sliding rather than to its rate" [10].

Within the observational limitations of the various experiments it has been difficult to separate true sliding from hopping. An approach to remedy this is presented in the chapter of this volume on lacI diffusion by Wang and Austin. They note that with millisecond dissociation times estimated for lacI, there is expected to be a significant amount of hopping. In general, fast dissociation kinetics (short dissociation times) can implicate hopping when compared to mean sliding times. On the other hand, long dissociation times would indicate little hopping. At the extreme are the DNA sliding clamps that not only display very long dissociation

**Table 3.2** Effect of diffusion constant on mean time for three log displacement as well as the mean displacement that occurs in 1 ms, based on the Einstein relation for one-dimensional diffusion given by (3.1)

| $D$ (m$^2$/s) | Time (μs) (for 1 bp) | Time (ms) (for 10 bp) | Time (ms) (for 100 bp) | Displacement (bp) (in 1 ms) |
|---|---|---|---|---|
| $10^{-16}$ | 578 | 57.8 | 5780 | 1.3 |
| $10^{-15}$ | 57.8 | 5.78 | 578 | 4.2 |
| $10^{-14}$ | 5.78 | 0.578 | 57.8 | 13.2 |
| $10^{-13}$ | 0.578 | 57.8 (μs) | 5.78 | 41.6 |
| $10^{-12}$ | 57.8 (ns) | 5.78 (μs) | 0.578 | 131.5 |
| $10^{-11}$ | 5.78 (ns) | 0.578 (μs) | 0.058 | 415.9 |

times but must be actively reloaded onto DNA. Such proteins are not expected to display hopping. Nevertheless, in a sort of gray area in the definition of hopping, it is possible that hop-like diffusion can occur where the clamp protein is not in contact with DNA. This case would pose an important exception to the above theory since the sliding clamp topology (closed ring around DNA) implies that hops never lead to jumps (zero probability), and moreover, the clamp remains bound to the DNA without the *prima facie* requirement for any overall protein–DNA attraction.

In the earliest report of lacI sliding on DNA, it was noted that the protein–DNA interactions are dominated by electrostatic forces, a result gleaned from the sensitivity of the sliding to the ionic strength of the solution [3]. In fact, this has become a standard method for distinguishing hopping from sliding. In order to experimentally increase hopping, one can increase the ionic strength of the solution. As mentioned above, DNA sliding clamps cannot hop in the ordinary sense, but they can potentially undergo hop-like diffusion ("ice skate" [41]). Kochaniak et al. found that the diffusion coefficient of PCNA along DNA changed by only a factor of 2 upon a 13-fold change in ionic strength (from 41 to 541 mM). While this is a very small effect, it may indicate some smoothening of the energy landscape due to interference in transient ionic bonds between protein and DNA.

Other proteins may share at least some of the topological character of the DNA sliding clamps. There is evidence, for example, that Rad51 and Msh2–Msh6 complexes also form rings or "clamp-like" structures on DNA. In the case of Rad51, this inference is made on the basis of (1) previous structural considerations, (2) lack of helical extension of the DNA templates, (3) no observations of proteins moving past one another, (4) implausibility of a helical form sliding so well along DNA [32]. In the case of Msh2–Msh6, the sliding clamp model is consistent with both crystal structures and electron micrographs that show the Msh2–Msh6 complex completely encircling DNA. Gorman et al. suppose that the complex would have to "at least partially unfold to allow dissociation, and this requirement for a large-scale structural reorganization may account for the long time periods that Msh2–Msh6 is able to slide on DNA" [18]. "Hopping [of the complex] is inconsistent with the finding that the diffusion coefficients were insensitive to salt concentration, different proteins bound to the same DNA were unable to bypass one another, and they were resistant to challenge with excess competitor oligonucleotide" [18].

*Diffusion constants and biological function*. While most of the research in the area of passive sliding is aimed at proteins that slide (and hop, jump, etc.) in search of targets on DNA, DNA sliding clamps must slide to serve their function. For proteins that search for a target, the diffusion constant conveys a key aspect of the speed of the survey search from a given starting point. For proteins that slide as their primary function, the diffusion constant describes the drag imposed on that function. The β clamp again provides an excellent example. The reported value of $D \sim 10^{-14}$ m$^2$/s implies significant friction (attractive or sticky interactions) between the protein and DNA, since this value is at least two orders of magnitude slower than would be predicted for a protein of this size diffusing through water alone. This may appear surprising since the polymerase III core ($\alpha$, $\theta$ and $\varepsilon$ subunits) replicates DNA approximately 50-fold faster when the β clamp is included in the

reaction [42]. Nevertheless, we can reason that even the "slow" diffusion constant is large enough to imply almost no friction for the DNA polymerase to work against. In isolation the β clamp moves from one base pair to the next (0.34 nm) in 1/100th the time the polymerase takes to insert a single base (~1 ms). Note that the polymerase, a Brownian motor, presumably also moves forward by random diffusion on a submillisecond time scale, but each freshly inserted DNA base becomes the ratchet tooth that prevents backward motion.

Even though the β clamp easily moves from one base to the next in just 6 μs, longer distances require much more time due to the nature of diffusion (Table 3.2). Since the β clamp spans a distance of about 12 bp on dsDNA (~4 nm), the time to diffuse just one clamp width away is about a millisecond. We have suggested that this relatively slow sliding of the clamp may play a role in preventing the polymerase–clamp complex from excessive drift during the polymerization process. It is noteworthy that recent single-molecule observations of HIV reverse transcriptase (an RNA polymerase that does not attach to a sliding clamp) recorded frequent events involving polymerase sliding away from the polymerization site [28]. The relatively slow sliding of the β clamp on DNA could mean that the attached polymerase need not maintain contact with the DNA to keep its place on the template. Indeed, such a feature may be required for efficient switching between polymerases and other DNA replication and repair factors. It is known, for example, that in the absence of the β clamp, the main bacterial polymerase – the $\alpha$ subunit – loses its grip on the template and rapidly dissociates from the DNA. It is therefore plausible that the polymerase often loses direct contact with the DNA during synthesis, especially during polymerase swapping events [43, 44]. During these transitions, it may be the β clamp that maintains the correct position on the DNA. Thus, by sticking significantly as it passes along the DNA, the β clamp may help prevent rapid polymerase drift.

Ultimately, do all DNA-binding proteins slide on DNA? Possibly not, but it may be difficult to rule out any sliding. One can imagine proteins that cannot slide, for example, because they are too stiff (see below). In some cases there is evidence that sliding does not occur over long distances. One particularly interesting example is RecA which facilitates homologous recombination by matching single-stranded DNA to a complementary region in double-stranded DNA. In complex with single-stranded DNA, RecA was shown to reach its target on double-stranded DNA independent of its position, indicating that it did not reach its target by sliding along DNA [45]. Nevertheless, this work falls short of proving that RecA does not slide on DNA at all.

## 3.4 Models and Mechanisms of Sliding

### *3.4.1 Implications of Search*

Results from theoretical treatments of DNA target search have implications for the nature of sliding and hopping on DNA. Several reports, based on theoretical considerations, suggest that many proteins interact with DNA in at least two modes.

In one mode the protein slides and/or hops along DNA with high diffusion rate. In another mode the protein binds tightly to a particular site and remains there. Two DNA interaction modes provide an obvious solution to the apparent paradox of protein sliding over great distances while also tightly binding to the specific target sequence. The mechanism for switching between these modes invokes the general idea that protein–DNA interactions are nonspecific during sliding/hopping and specific during tight binding. One diffusion rate for each protein is probably inadequate to describe target acquisition. During fast diffusion there is too little interaction to detect the specific target. Therefore, a slower diffusion mode may be required by the biology (and implicated by kinetics experiments). In the language of proteins searching for a target, the two modes would be a fast diffusion over nonspecific DNA and a slow diffusion close to the target site. This also seems to fit with the crystal structures of proteins bound "specifically" and "nonspecifically" to DNA (as discussed below). It should be mentioned that there has been some controversy about the thermodynamics of such a switch. These considerations place limits on the frequency of conformational changes that control the modes [46].

Kinetics studies over the past decade, such as those performed with *Eco*RI, suggested that while proteins are sliding along DNA, they pause at sites that resemble their recognition sites [47]. This provided evidence for a fast and slow mode of sliding, fitting well with the aforementioned theories of DNA target search. From X-ray crystallography, we have structures for a handful of proteins bound "specifically" to a DNA target or bound "nonspecifically" to generic DNA. These two states are thought to be indicative of target binding vs. target searching, perhaps revealing intermolecular contacts in sliding quickly vs. sliding slowly, and the large differences in the intermolecular conformations have been interpreted in this way. We will now discuss the theoretical arguments for multiple DNA interaction modes and then the experimental evidence supporting this view. Work has focused on analyzing crystal structures of proteins bound to DNA, as well as simplified simulations and calculations. Although mechanisms largely remain unclear, some clues are available from structural analyses and simulations, as discussed below.

### *3.4.2 Analysis and Implications of Crystal Structures*

The ways in which proteins interact with DNA are varied. The protein may wrap around the DNA partly or completely, may grab the DNA deeply with many intermolecular contacts or barely touch the DNA, and may twist, splay, or sharply bend the DNA. The nature of the protein–DNA interactions can strongly affect the structure–function relationship of the complex, especially for enzymatic proteins such as the restriction endonucleases. A decade ago Jones et al. published a summary of protein–DNA interactions, analyzed from 26 protein–DNA complexes and compared with some 36 protein–protein interactions [48]. The main findings were that the protein–DNA interface is largely polar, the protein–DNA interactions are often mediated by water molecules, and most of the protein–DNA contacts are to the sugar-phosphate

backbone, not the bases. Positively charged arginine residues were the most likely to be involved in the protein–DNA interface, followed by polar threonine, asparagine, and positively charged lysine residues. Among the least observed residues to contact the DNA were the negatively charged aspartic and glutamic acid residues. Traditionally the emphasis in the crystallography of protein–DNA complexes was to understand how specific sequence recognition occurs. Due to the high stability of the sequence-specific complex, this turns out to be easier than obtaining structures for complexes with non-specific DNA sequences. At that time the only structure of a protein bound to DNA "nonspecifically" was the structure of *Eco*RV bound to noncognate DNA. This remains an illuminating example when compared with the structure of the protein bound specifically to the cognate DNA (Fig. 3.4). The gist of the comparison is that there are many fewer contacts overall and no contacts to bases in the nonspecific complex.

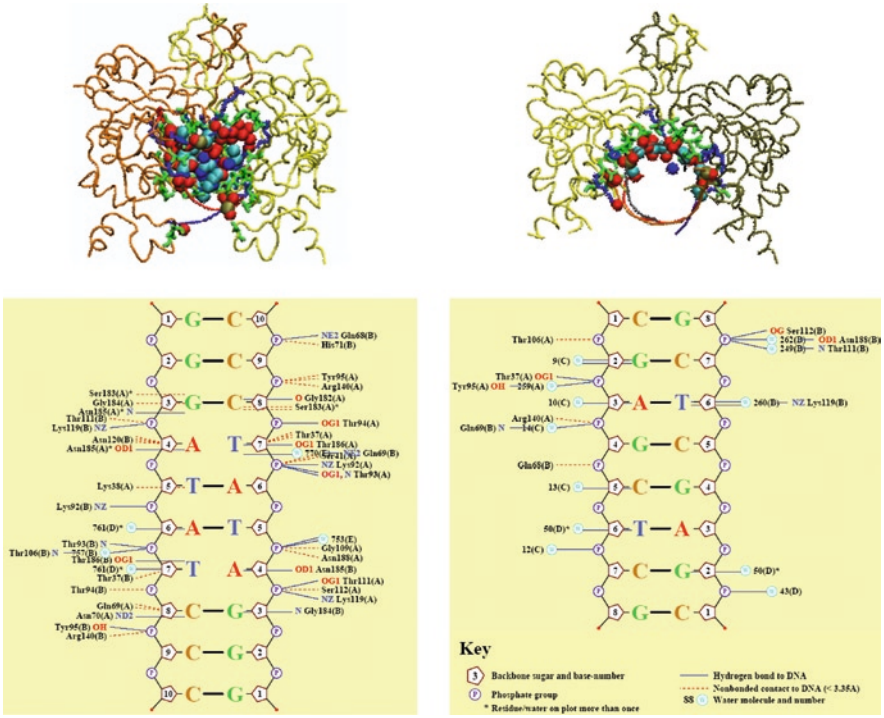


**Fig. 3.4** Comparison of protein–DNA interactions for specific (*left*) and nonspecific (*right*) binding of *Eco*RV. *Top*: Axial views of the interactions. The proteins and DNA are represented by smooth tubes of the respective backbones with each chain colored differently. The protein residues that contact the DNA (i.e., with nonhydrogen atom centers that are within 3.8 Å of any DNA nonhydrogen atom centers) are shown in licorice, with polar residues in *green*, positively charged (basic) residues in *blue*. *Bottom*: Output of nucplot for the two crystal structures, detailing the protein–DNA interactions [62]

Moreover, while the dimeric protein completely wraps around the DNA in the specific complex, the dimer is much more open in the nonspecific complex. There are more than twice as many water molecules bound to the DNA in the nonspecific complex. More specific/nonspecific pairs of structures of protein–DNA have been determined since then, but the initial summary has remained valid. Nonspecific protein–DNA complexes are looser.

An important starting point for mechanistic knowledge of the sliding state is therefore provided by the very few crystal structures of protein–DNA complexes that appear to capture nonspecific interactions. But how well do crystal structures of "nonspecific" protein–DNA complexes reveal the freely diffusing/sliding state? On one hand, the structures are clearly different from those that form when the specific target is present. On the other hand, there is almost a contradiction between the concept of a crystal structure and that of a nonspecifically bound sliding protein. In fact, various tricks are employed to promote uniformity of complexes, as required for crystallization. For example, the DNA sequence is often very similar to the target sequence, leading to structures that are probably not completely representative of nonspecificity. It is nevertheless hoped that at least some features of the nonspecific state are revealed. In fact, assuming that there are two modes of sliding (i.e., fast and slow search), near-specific structures may turn out to be more representative of the slow mode. Perhaps then, the most relevant structures for understanding of protein sliding are structures of proteins that do not specifically bind anywhere on DNA. A prominent example is the DNA sliding clamps that, not surprisingly, resisted cocrystallization with DNA until very recently. With all the inherent limitations of static structures in the illumination of motion, protein–DNA cocrystal structures may give promising leads in understanding the mechanics of sliding.

*Focus on DNA sliding clamps*. As mentioned previously, DNA sliding clamps are ring-shaped proteins that encircle DNA helix, topologically trapping DNA inside the ring and providing a mobile platform to which DNA polymerases and other proteins can attach. The fact that sliding clamps are present in all life forms underlies the evolutionary success of this way of keeping other proteins in close association with DNA. The oligomeric state varies among the kingdoms: the bacterial sliding clamp (polIII $\beta$ subunit or "β clamp") is a dimer, while the eukaryotic and archaeal sliding clamp (PCNA) is trimer. Despite these differences, all sliding clamps have remarkably similar 3D structures featuring pseudo sixfold symmetry. The central channel of sliding clamps is ~35 Å in diameter, significantly larger than the diameter of the DNA double helix in the canonical B-form (~24 Å). Since the first determination of the structure of a DNA sliding clamp, the β clamp, in 1992 [35], these features have raised questions regarding the relative arrangement of the clamp and the DNA. Does the clamp bind directly to DNA or can it "levitate" without making direct contacts with DNA? What is the mechanism of clamp movement/diffusion along the DNA?

Both biochemical and crystallographic studies have shown that the *E. coli* β clamp can bind DNA directly [49]. The crystal structure of the β clamp–DNA complex revealed that the plane of the clamp is not perpendicular to the DNA axis. Instead it

is tilted by an angle of about 22° (Fig. 3.5). Since the central channel of the β clamp is wider than the DNA double helix, the tilt allows the DNA to contact both sides of the clamp channel (indeed both protomers of the dimer). The contacts to the DNA are to the phosphate backbone. As we discussed above, the DNA sliding clamp makes an unlikely target for crystallization. Indeed, several aspects of the crystal structure are worth mentioning. First, the DNA is not completely double stranded. It is a "primed site," meaning that it has a four-base single-stranded 5′ overhang, and there are half a dozen contacts of the clamp to the bases (not just the backbone) of the 5′ overhang. Furthermore, due to the crystallographic arrangement, these ssDNA contacts are with neighboring β clamps, not the β clamp penetrated by the dsDNA. Thus, given these "crystallization artifacts," it is fair to ask if β clamp binding to DNA in solution shows similarity to what is observed in the crystal structure.

There are several lines of evidence suggesting that the β clamp-DNA arrangement seen in the crystal does closely reflect the situation in solution. First, a series of molecular dynamics (MD) simulations have been performed in our own lab for the β clamp bound to 18 bp of DNA. This study starts with the DNA centered and aligned to the central axis of the β clamp on DNA. We have found that the clamp quickly adopts a tilted orientation in respect to the DNA, and the tilt angle relative to the central 12 bp is actually slightly larger than that observed in the crystal structure of β–DNA complex. In the simulations, the angle vacillates mostly between 20° and 30°. (A snapshot created at 10 ns is shown in Fig. 3.5). Many of the same residues seen to interact with the DNA in simulations were also identified as contacts by crystallography (Barsky, unpublished results). In addition, our analysis of conserved residues in bacterial $\beta$ subunits (Venclovas, unpublished data) revealed that the majority of the positively charged residues of the β clamp (Arg, Lys, His) interacting with DNA in the simulations appear to be strongly conserved by evolution. Altogether these data are consistent with the β clamp/DNA interaction in the crystal structure.

Having "verified" the plausibility of the crystal structure of the β clamp–DNA complex, we might now ask if there is anything special about how a sliding clamp binds DNA (other than the enclosing topology) that might give a clue to the

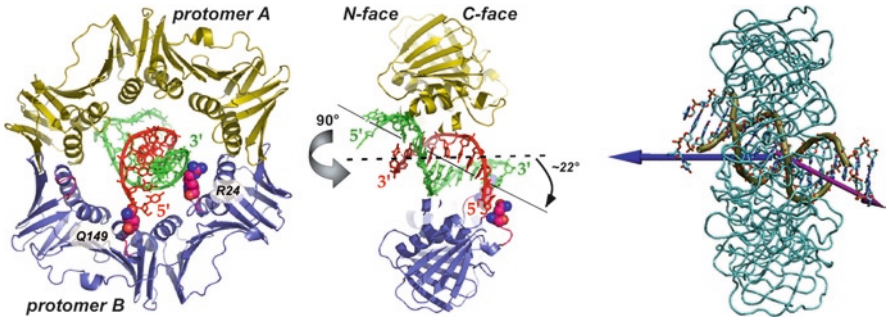


**Fig. 3.5** *Front* and *side views* of the β clamp–DNA complex from X-ray crystallography (*left*) [49] and a snapshot at 10 ns from an MD simulation (*right*) showing a 30° angle between the clamp axis and the DNA axis (Barsky, unpublished results)

sliding mechanism. From the point of view of evolutionary conservation, there are several other highly conserved charged residues that either might provide alternative contacts (R176, H175) during β-clamp diffusion on DNA, or (as in the case of K235) might participate in binding in the same mode, but on a longer DNA helix than the one in the crystal structure. The strong evolutionary conservation of positively charged residues inside the inner channel is not surprising considering their anticipated role in binding DNA backbone. However, the surprising finding is that a number of negatively charged residues (Asp, Glu) lining the central channel that are as (or even more) conserved as those with positive charges. Interestingly, these conserved acidic residues appear to be mostly interspersed between positively charged residues that either contact DNA (as seen in the crystal structure) or could potentially contact DNA (Fig. 3.6). Strong conservation suggests a potential role in clamp diffusion, perhaps allowing Arg or Lys side chains to alternate between binding DNA phosphates and forming salt bridges with neighboring Asp and Glu side chains on the clamp. This might significantly lower the activation energy required for breaking charged interactions with DNA phosphates during clamp movement on DNA. Evidence that protein–DNA interactions can be weakened by the "masking" of positively charged residues by forming salt bridges with nearby negatively charged residues has been reported [50].

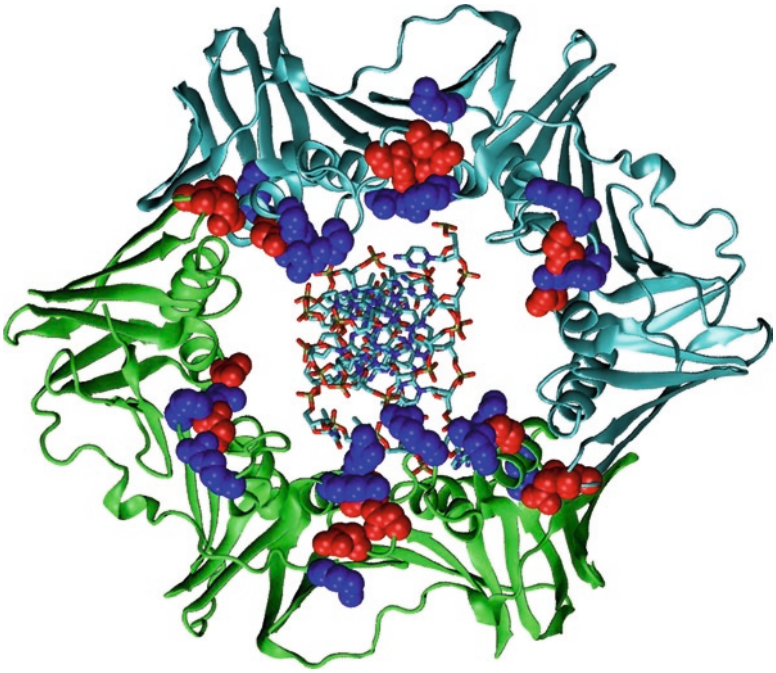

**Fig. 3.6** Crystal structure of β clamp/DNA complex (PDB ID: 3bep). The β clamp monomers are shown in differently colored ribbons. The side chains of the most highly conserved positively (Arg, Lys, His) and negatively charged (Asp, Glu) residues lining the inner channel are shown in a space-filled representation in *blue* and *red colors*, respectively

Our MD simulations of the β clamp show that only a small subset of the 24 positively charged residues interact for substantial times with the DNA. For example, in the 18 bp simulation, only nine residues had contact times (ionic bonds or salt bridges of 3 Å or less between the protein and DNA) totaling more than 10% of the simulation, while two pairs of residues (from each side of the dimer) contacted the DNA more than 80% of the time. This supports the idea that key residues dominate the protein–DNA interactions. This also lends some support to the "ice-skating" model of interaction and sliding.

These observations for the β clamp on DNA may be relevant for other DNA sliding clamps. In particular, binding at an angle has been observed for an X-ray crystallographic structure of the yeast PCNA–DNA complex [51] and the structure obtained by electron microscopy for the archaeal DNA ligase–PCNA–DNA complex [52]. Again, while there are reasonable concerns about the relevance of the crystallographic structure, MD simulations have shown that similar to the β clamp, PCNA quickly tilts with respect to the DNA axis [53]. Still, some concerns remain. In the case of PCNA, the angle between clamp and DNA in the crystal structure is substantially larger (40°) than what was reported from the MD simulations (20°). In fact, the DNA is partly disordered and was by necessity modeled as a rigid body, leading the authors to conclude that the orientation of DNA is not definitive [51]. The final structure was termed "an X-ray derived model" (PDB code: 3K4X). This may indicate a highly dynamic complex – exactly what might be expected for a sliding clamp on DNA. The differences in the crystallographic results for PCNA and the β clamp are not easy to explain. Differences might be related to biologically unimportant differences in how the crystals developed, or might be caused by fundamental differences between the two sliding clamps.

From data related to crystal structures, sequence conservation, and modeling, a picture emerges wherein DNA sliding clamps make specific contacts only to the DNA backbone in keeping with the "nonspecific" nature of the interaction. Contacts are transient, and may be supplemented by additional possible contacts and moderated by possible repulsive interactions. Despite the apparent importance of these electrostatic interactions, the sliding of PCNA was only weakly sensitive to the ionic strength – a tenfold increase in the salt concentration led to only a twofold increase in the diffusion constant. The change was in the expected direction. Two recent papers [54, 55] describing crystal structures of PCNA from halophilic archaea (HvPCNA) may be informative in this respect. The inner surface of the archeal PCNA is less positively charged compared to eukaryotic PCNA. The authors note that "strikingly, the positive surface charge considered key to PCNA's role as a sliding clamp is dramatically reduced in the halophilic protein. Instead, bound cations within the solvation shell of HvPCNA may permit sliding along negatively charged DNA by reducing electrostatic repulsion effects" [54]. Perhaps future simulations and experimental observations of diffusion will take up this suggestion.

Finally, we consider the two sliding clamp diffusion modes inferred by Kochaniak et al. who suggested that in the slow mode the proteins follow the DNA helix (undergo rotationally coupled sliding) but not in the fast mode [24]. That the two diffusion modes are due to changes in the protein–DNA interactions is

made especially plausible given the stability of the sliding clamp proteins. It is unlikely that sliding clamps undergo the types of conformational changes that are ascribed to other dual-mode sliding proteins. Another idea is that the tilt of the clamp relative to the DNA could affect the diffusion rate. In fact, it may be argued that since the sliding clamp has been observed to slide over 13-nucleotide ssDNA loops and other secondary structures of DNA [56], it need not always remain oriented in a particular way or follow the grooves of the DNA. Nevertheless, whether the change is in helix tracking, in tilt, or both, it is unknown how a particular state would persist while sliding over many base pairs. Perhaps the solution lies in the details of the electrostatics – the particular placement of charges – rearranged so that phosphate backbone would no longer be tracked. Ionic bonds between oppositely charged side chains or long-lived bound ions might accomplish this. Before the structures of DNA bound to the sliding clamps were determined, it had been suggested that the clamps can "ice skate" along the DNA [41]. This still might be true.

### *3.4.3 Modeling and Simulation*

*Monte Carlo simulations*. Interesting Monte Carlo studies have predicted nonintuitive behavior for proteins sliding on DNA. Dahirel et al. [57] showed by means of Monte Carlo simulations and analytical calculations that there is a counterintuitive repulsion between two oppositely charged macromolecules at a nanometer range. This predicted force is thought to arise from a local increase of the osmotic pressure exerted by ions trapped at the interface. For DNA-binding proteins with concave surfaces, and for realistic protein charge densities, the authors predicted that the DNA–protein interaction free energy has a minimum at a finite surface-to-surface separation. This defines a separation distance at which proteins might easily slide on DNA. When a protein encounters its specific target sequence, the proposed free energy barrier would be overcome by favorable hydrogen bonds, thus enabling sequence recognition. In other Monte Carlo simulations, Barbi et al. show that the consequences of specific site recognition during the process of sliding over nonspecific sites produces "nontrivial" sequence-dependent dynamics that do not necessarily resolve neatly into two modes (or two effective diffusion constants) [40].

*Electrostatics calculations and Brownian dynamics*. Sun et al. [58] used electrostatic free energy calculations (based on numerical solutions of the Poisson–Boltzmann equation) to take into account both energetic and structural considerations for *Bam*HI restriction endonuclease sliding along DNA. The authors equate nonspecific binding with the sliding state, and they argue from experiments measuring the change in heat capacity and the salt-dependence of binding that the nature of the protein–DNA interactions is dominated by through-solvent electrostatics (no protein–DNA contacts). There are a number of reports that nonspecific binding gives rise to negligible heat capacity change, which Sun et al. and others ascribe mainly to a small change in the hydrophobic contribution to the free energy. This, in turn, implies that there is little

change in the hydration of the protein upon binding. Thus, nonspecific binding may imply little contact with DNA, an inference that is at least partly supported by crystal structures. Nonspecific protein binding is known to have a stronger salt-concentration dependence, indicating that sequence nonspecific protein–DNA interactions are dominated by electrostatic forces. Sun et al. apply their modeling to the *Bam*HI–DNA complex. The authors argue that the nonspecific structure of the cocomplex, where just 1 bp was changed from the cognate target sequence, does not represent an electrostatic free energy minimum. To achieve the minimum energy state, the protein would need to be tilted relative to the DNA axis, and moved away from the DNA by ~15 Å. The authors calculated the effect of protein revolution on sliding and found that the protein would need to overcome a barrier of ~2.5 kcal/mol if it did not follow the helical pitch. This barrier would arise from the requirement of crossing from one major groove to the next. The authors also calculated Coulombic interaction energies for charged residues interacting with DNA. Such interactions were predicted to contribute significant favorable free energies (7–11 kcal/mol), with some of the interactions over surprisingly large distances (up to 24 Å). The *Bam*HI diffusion constant has recently been measured (Table 3.1), and it has been inferred that the protein rotates while sliding [31]. It would be interesting to test the prediction of Sun et al. that the specific mutation E51K would affect the *Bam*HI sliding rate [58].

*MD simulations*. Of all the levels of modeling, by far the most complete and detailed simulations are atomistic molecular dynamics (MD) simulations. We have already discussed some results from our simulations of the β clamp in the section on crystal structures. It is worth considering the feasibility of simulating the actual sliding of proteins on DNA using molecular dynamics (MD) simulations. Currently, state-of-the-art for MD simulations of roughly 100,000 atoms is on the order of 100 ns. Unfortunately, for even the fastest reported diffusion rates ($D = 3 \times 10^{-12}$ m$^2$/s), one could expect to observe only about 2 bp of displacement in 100 ns (Table 3.2). Given the flexibility of the DNA and the proteins, 2 bp of overall movement would be difficult to observe unambiguously. Nevertheless, in 500 ns simulations, clear displacements of roughly 5 bp should occur for this diffusion rate. We conclude that the direct observation of sliding on DNA by MD simulation should be feasible in the not-too-distant future.

### *3.4.4 Proposed Molecular Models of Sliding*

Here we briefly speculate on the molecular mechanisms of sliding. In addition to the previously mentioned "ice-skating" model of sliding, we propose two new models: the "inchworm" and "centipede" models. In the inchworm model the protein has two or more binding sites that independently bind the DNA and release it. If the position on the DNA can change between binding events, sliding can occur. In the centipede model the protein has many residues that can form attractive contacts to the DNA. The exchange of these residues can lead to overall movement on the DNA.

Halford postulates that what looks like sliding occurs by the same mechanism as dissociation. “The transfer of the protein from one site to another 1 bp along the DNA is likely to be similar to that for the dissociation of the protein into free solution: both processes involve the same number of bond-breaking events between protein and DNA backbone. For example, a hydrogen bond between a particular functional group on the protein and another on the DNA cannot be maintained while the protein moves 1 bp along the DNA, as this movement, a 3.4 Å translocation and a 34 degree rotation, must extend the length of the hydrogen bond to beyond its breaking point” [59]. By this picture, all true sliding – as opposed to hopping – would be some form of “ice-skating” where no bonds persist, and any exchanging bonds are likely mediated by water molecules. This might be one mechanism of sliding. However, evidence from X-ray crystallography suggests that specific protein–DNA contacts may occur even when the two molecules interact nonspecifically. Thus, there are likely other mechanisms of sliding.

In fact, the picture of sliding as a series of dissociation events inadvertently suggests that true sliding almost certainly requires flexibility, either as global conformational changes (e.g., an “inchworm” model), or as local side chain rearrangements (e.g., a “centipede” model). Even hopping need not happen in a rigid, all-or-none fashion. Protein flexibility could allow the protein to peel off the DNA, much the way dsDNA can fray at the ends. The inchworm and centipede models both involve partial binding and unbinding events in the sliding mechanism.

In the inchworm model in addition to the requirement for separate binding sites, the protein must have internal flexibility to contract and extend so that binding and release events effect translation on the DNA. Note that the requirement for more than one binding sites is similar to what is required for intersegmental transfer (Fig. 3.1). Because intersegmental transfer may involve multimeric protein complexes, each monomer of which can bind DNA, it is not clear if any of the known or suspected intersegmental transfer proteins can slide by the inchworm model. Two known examples are Oct-1 [60] and APOBEC3G, an antiviral restriction factor that acts on foreign single-stranded DNA [61]. Other candidates are the RecA protein and the single-strand binding protein, SSB. It would be interesting to investigate whether any proteins that move by intersegmental transfer can also slide via the inchworm model. Perhaps this can be tested by attaching a pair of FRET labels to the protein and observing it move on stretched DNA where intersegmental transfer is not possible.

In the centipede model the protein has many potential residue contacts with the DNA. As contacts randomly break and reform, the collective effect is a random walk of the protein on the DNA. Many attractive contacts would tend to promote cooperativity and should slow sliding. If, however, many residues are available to bind sites on the DNA (e.g., Lys and Arg residues that can bind phosphates on the DNA backbone), the competition may actually accelerate sliding. Thus, we call this variation of the centipede model a “cooperative/competitive” model of sliding. We propose this model for the DNA sliding clamps in the slow mode. Residues that may inhibit protein–DNA binding, such as the conserved, negatively charged residues of the β clamp, can also accelerate sliding.

## 3.5 Conclusion

Protein sliding on DNA has been shown to be of intense interest. Experimental methods and theoretical treatments are progressing rapidly toward a mechanistic understanding. The dynamics of protein sliding on DNA are challenging, yet clues are appearing as experimental and simulation methods and resources improve. Given this auspicious situation, it is anticipated that a bumper crop of testable hypotheses will soon emerge regarding all aspects of protein sliding on DNA.

## References

1. Berg OG, Winter RB, Vonhippel PH (1981) Diffusion-driven mechanisms of protein translocation on nucleic-acids. 1. Models and theory. Biochemistry 20:6929–6948
2. Adam G, Delbrück M (1968) Reduction of dimensionality in biological diffusion processes. In: Rich A, Davidson N (eds) Structural chemistry and molecular biology. W.H. Freeman and Company, San Francisco, pp 198–215
3. Riggs AD, Bourgeoi S, Cohn M (1970) The lac repressor–operator interaction. 3. Kinetic studies. J Mol Biol 53:401–417
4. Winter RB, Vonhippel PH (1981) Diffusion-driven mechanisms of protein translocation on nucleic-acids. 2. The *Escherichia coli* repressor–operator interaction – equilibrium measurements. Biochemistry 20:6948–6960
5. Winter RB, Berg OG, Vonhippel PH (1981) Diffusion-driven mechanisms of protein translocation on nucleic-acids. 3. The *Escherichia coli* lac repressor–operator interaction – kinetic measurements and conclusions. Biochemistry 20:6961–6977
6. Slutsky M, Mirny LA (2004) Kinetics of protein–DNA interaction: facilitated target location in sequence-dependent potential. Biophys J 87:4021–4035
7. Halford SE, Marko JF (2004) How do site-specific DNA-binding proteins find their targets? Nucleic Acids Res 32:3040–3052
8. Greenleaf WJ, Woodside MT, Block SM (2007) High-resolution, single-molecule measurements of biomolecular motion. Annu Rev Biophys Biomol Struct 36:171–190
9. Herbert KM, Greenleaf WJ, Block SM (2008) Single-molecule studies of RNA polymerase: motoring along. Annu Rev Biochem 77:149–176
10. Wunderlich Z, Mirny LA (2008) Spatial effects on the speed and reliability of protein–DNA search. Nucleic Acids Res 36:3570–3578
11. Laurence TA, Kwon Y, Johnson A, Hollars CW, O'Donnell M, Camarero JA, Barsky D (2008) Motion of a DNA sliding clamp observed by single molecule fluorescence spectroscopy. J Biol Chem 283:22895–22906
12. Maki S, Kornberg A (1988) DNA polymerase III holoenzyme of *Escherichia coli*. 3. Distinctive processive polymerases reconstituted from purified subunits. J Biol Chem 263:6561–6569
13. Yao NY, O'Donnell M (2009) Replisome structure and conformational dynamics underlie fork progression past obstacles. Curr Opin Cell Biol 21:336–343
14. Gowers DM, Halford SE (2003) Protein motion from non-specific to specific DNA by three-dimensional routes aided by supercoiling. EMBO J 22:1410–1418
15. Yao N, Turner J, Kelman Z, Stukenberg PT, Dean F, Shechter D, Pan ZQ, Hurwitz J, Odonnell M (1996) Clamp loading, unloading and intrinsic stability of the PCNA, beta and gp45 sliding clamps of human, *E. coli* and T4 replicases. Genes Cells 1:101–113
16. Redner S (2001) A guide to first-passage processes. Cambridge University Press, Cambridge
17. Austin RH, Karohl J, Jovin TM (1983) Rotational diffusion of *Escherichia coli* RNA polymerase free and bound to deoxyribonucleic acid in nonspecific complexes. Biochemistry 22:3082–3090

18. Gorman J, Chowdhury A, Surtees JA, Shimada J, Reichman DR, Alani E, Greene EC (2007) Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2–Msh6. Mol Cell 28:359–370
19. Gorman J, Greene EC (2008) Visualizing one-dimensional diffusion of proteins along DNA. Nat Struct Mol Biol 15:768–774
20. Biebricher A, Wende W, Escude C, Pingoud A, Desbiolles P (2009) Tracking of single quantum dot labeled EcoRV sliding along DNA manipulated by double optical tweezers. Biophys J 96:L50–L52
21. Bonnet I, Biebricher A, Porte PL, Loverdo C, Benichou O, Voituriez R, Escude C, Wende W, Pingoud A, Desbiolles P (2008) Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. Nucleic Acids Res 36:4118–4127
22. Blainey PC, van Oijent AM, Banerjee A, Verdine GL, Xie XS (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. Proc Natl Acad Sci USA 103:5752–5757
23. Harada Y, Funatsu T, Murakami K, Nonoyama Y, Ishihama A, Yanagida T (1999) Single-molecule imaging of RNA polymerase–DNA interactions in real time. Biophys J 76:709–715
24. Kochaniak AB, Habuchi S, Loparo JJ, Chang DJ, Cimprich KA, Walter JC, van Oijen AM (2009) Proliferating cell nuclear antigen uses two distinct modes to move along DNA. J Biol Chem 284:17700–17710
25. Lin Y, Zhao T, Jian X, Farooqui Z, Qu X, He C, Dinner AR, Scherer NF (2009) Using the bias from flow to elucidate single DNA repair protein sliding and interactions with DNA. Biophys J 96:1911–1917
26. Tafvizi A, Huang F, Leith JS, Fersht AR, Mirny LA, van Oijen AM (2008) Tumor suppressor p53 slides on DNA with low friction and high stability. Biophys J 95:L1–L3
27. Wang YM, Austin RH, Cox EC (2006) Single molecule measurements of repressor protein 1D diffusion on DNA. Phys Rev Lett 97:4
28. Liu SX, Abbondanzieri EA, Rausch JW, Le Grice SFJ, Zhuang XW (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. Science 322:1092–1097
29. Roy R, Kozlov AG, Lohman TM, Ha T (2009) SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. Nature 461:1092–1097
30. Rau DC, Sidorova NY (2010) Diffusion of the restriction nuclease EcoRI along DNA. J Mol Biol 395:408–416
31. Blainey PC, Luo GB, Kou SC, Mangel WF, Verdine GL, Bagchi B, Xie XS (2009) Nonspecifically bound proteins spin while diffusing along DNA. Nat Struct Mol Biol 16:1224–1229
32. Graneli A, Yeykal CC, Robertson RB, Greene EC (2006) Long-distance lateral diffusion of human Rad51 on double-stranded DNA. Proc Natl Acad Sci USA 103:1221–1226
33. Laurence TA, Kwon Y, Yin E, Hollars CW, Camarero JA, Barsky D (2007) Correlation spectroscopy of minor fluorescent species: signal purification and distribution analysis. Biophys J 92:2184–2198
34. Kampmann M (2004) Obstacle bypass in protein motion along DNA by two-dimensional rather than one-dimensional sliding. J Biol Chem 279:38715–38720
35. Kong XP, Onrust R, Odonnell M, Kuriyan J (1992) Three-dimensional structure of the beta-subunit of *Escherichia coli* DNA polymerase III holoenzyme – a sliding DNA clamp. Cell 69:425–437
36. García De La Torre J, Huertas M, Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys J 78:719–730
37. Schurr J (1979) The one-dimensional diffusion coefficient of proteins absorbed on DNA hydrodynamic considerations. Biophys Chem 9:413–414
38. Bagchi B, Blainey PC, Xie XS (2008) Diffusion constant of a nonspecifically bound protein undergoing curvilinear motion along DNA. J Phys Chem B 112:6282–6284
39. Zwanzig R (1988) Diffusion in a rough potential. Proc Natl Acad Sci USA 85:2029–2030

40. Barbi M, Place C, Popkov V, Salerno M (2004) A model of sequence-dependent protein diffusion along DNA. J Biol Phys 30:203–226
41. Johnson A, O'Donnell M (2005) Cellular DNA replicases: components and dynamics at the replication fork. Annu Rev Biochem 74:283–315
42. Fay P, Johanson K, McHenry C, Bambara R (1981) Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of *Escherichia coli*. J Biol Chem 256:976–983
43. Indiani C, McInerney P, Georgescu R, Goodman M, O'Donnell M (2005) A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. Mol Cell 19:805–815
44. Indiani C, Langston LD, Yurieva O, Goodman MF, O'Donnell M (2009) Translesion DNA polymerases remodel the replisome and alter the speed of the replicative helicase. Proc Natl Acad Sci USA 106:6031–6038
45. Adzuma K (1998) No sliding during homology search by RecA protein. J Biol Chem 273:31565–31573
46. Hu LH, Grosberg AY, Bruinsma R (2008) Are DNA transcription factor proteins Maxwellian Demons? Biophys J 95:1151–1156
47. Jeltsch A, Alves J, Wolfes H, Maass G, Pingoud A (1994) Pausing of the restriction-endonuclease EcoRI during linear diffusion on DNA. Biochemistry 33:10215–10219
48. Jones S, van Heyningen P, Berman HM, Thornton JM (1999) Protein–DNA interactions: a structural analysis. J Mol Biol 287:877–896
49. Georgescu RE, Kim SS, Yurieva O, Kuriyan J, Kong XP, O'Donnell M (2008) Structure of a sliding clamp on DNA. Cell 132:43–54
50. Holbrook JA, Tsodikov OV, Saecker RM, Record MT (2001) Specific and non-specific interactions of integration host factor with DNA: thermodynamic evidence for disruption of multiple IHF surface salt-bridges coupled to DNA binding. J Mol Biol 310:379–401
51. McNally R, Bowman GD, Goedken ER, O'Donnell M, Kuriyan J (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. BMC Struct Biol 10:3
52. Mayanagia K, Kiyonari S, Saito M, Shirai T, Ishino Y, Morikawa K (2009) Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA–DNA complex architecture. Proc Natl Acad Sci USA 106:4647–4652
53. Ivanov I, Chapados BR, McCammon JA, Tainer JA (2006) Proliferating cell nuclear antigen loaded onto double-stranded DNA: dynamics, minor groove interactions and functional implications. Nucleic Acids Res 34:6023–6033
54. Winter JA, Christofi P, Morroll S, Bunting KA (2009) The crystal structure of *Haloferax volcanii* proliferating cell nuclear antigen reveals unique surface charge characteristics due to halophilic adaptation. BMC Struct Biol 9:55
55. Morgunova E, Gray FC, MacNeill SA, Ladenstein R (2009) Structural insights into the adaptation of proliferating cell nuclear antigen (PCNA) from *Haloferax volcanii* to a high-salt environment. Acta Crystallogr D Biol Crystallogr 65:1081–1088
56. Yao N, Hurwitz J, O'Donnell M (2000) Dynamics of beta and proliferating cell nuclear antigen sliding clamps in traversing DNA secondary structure. J Biol Chem 275:1421–1432
57. Dahirel V, Paillusson F, Jardat M, Barbi M, Victor JM (2009) Nonspecific DNA–protein interaction: why proteins can diffuse along DNA. Phys Rev Lett 102:4
58. Sun J, Viadiu H, Aggarwal AK, Weinstein H (2003) Energetic and structural considerations for the mechanism of protein sliding along DNA in the nonspecific BamHI–DNA complex. Biophys J 84:3317–3325
59. Halford SE (2009) An end to 40 years of mistakes in DNA-protein association kinetics? Biochem Soc Trans 37:343–348
60. Doucleff M, Clore GM (2008) Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1. Proc Natl Acad Sci USA 105:13871–13876
61. Nowarski R, Britan-Rosich E, Shiloach T, Kotler M (2008) Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. Nat Struct Mol Biol 15:1059–1066
62. Luscombe NM, Laskowski RA, Thornton JM (1997) NUCPLOT: a program to generate schematic diagrams of protein–nucleic acid interactions. Nucleic Acids Res 25:4940–4945