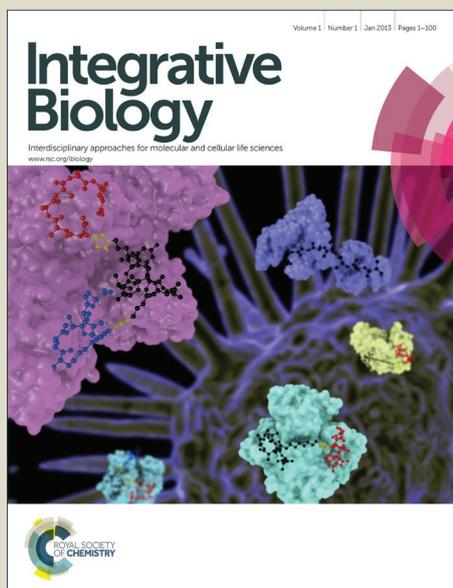


# Integrative Biology

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

# Extracting physical chemistry from mechanics: a new approach to investigate DNA interactions with drugs and proteins in single molecule experiments

M. S. Rocha<sup>\*a</sup>

Received Xth XXXXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXXXX 20XX

First published on the web Xth XXXXXXXXXXXX 200X

DOI: 10.1039/b000000x

In this review we focus on the idea of establishing connections between the mechanical properties of DNA-ligand complexes and the physical chemistry of DNA-ligand interactions. This type of connection is interesting because it opens the possibility of performing a robust characterization of such interactions by using only one experimental technique: single molecule stretching. Furthermore, it also opens new possibilities in comparing results obtained by very different approaches, in special when comparing single molecule techniques to ensemble-averaging techniques. We start the manuscript reviewing important concepts of the DNA mechanics, from the basic mechanical properties to the Worm-Like Chain model. Next we review the basic concepts of the physical chemistry of DNA-ligand interactions, revisiting the most important models used to analyze the binding data and discussing their binding isotherms. Then, we discuss the basic features of the single molecule techniques most used to stretch the DNA-ligand complexes and to obtain “force  $\times$  extension” data, from which the mechanical properties of the complexes can be determined. We also discuss the characteristics of the main types of interactions that can occur between DNA and ligands, from covalent binding to simple electrostatic driven interactions. Finally, we present a historical survey on the attempts to connect mechanics to physical chemistry for DNA-ligand systems, emphasizing a recently developed fitting approach useful to connect the persistence length of the DNA-ligand complexes to the physicochemical properties of the interaction. Such approach in principle can be used for any type of ligand, from drugs to proteins, even if multiple binding modes are present.

## 1 Introduction

The DNA molecule is the biological polymer related to some of the most important vital processes, from the storage and transmission of genetic information to the translation of proteins. Its primary structure is usually described as two parallel strands with a peculiar chemical structure based in complementary base-pairs, which allows the replication of the molecule in an unmistakable way<sup>1,2</sup>. The two DNA strands are arranged forming a double-helix structure that sets important properties to the molecule such as a well-defined negative charge density and a bending stiffness which places DNA in the class of semi-flexible polymers<sup>3–6</sup>.

Since it stores the genetic information of an organism, the DNA molecule may be very long in some cases. In fact, the human genome has approximately 3 billion base pairs, corresponding to a linear contour length of the order of 1 meter. If a DNA molecule with this length is placed disperse in a water-based solution, its radius of gyration will be of the order of 100  $\mu\text{m}$ <sup>7</sup>. How can a molecule with this size

be stored in the nucleus of a cell, which has typical dimensions of the order of a few micrometers<sup>2</sup>? The answer lies, at least partially, in the mechanical properties of the DNA molecule, which must be unique to allow such a condensation. *In vivo*, this process usually occurs mediated by the interaction of the DNA molecule with ligands, especially (but not exclusively) histone proteins. Furthermore, from molecular biology it is known that other important intracellular processes such as cell division and protein binding also depend on the DNA topology, which in turn, depends on the mechanical properties of the DNA molecule<sup>8,9</sup>. DNA topology can be strategically changed during these processes by the action of enzymes such as the topoisomerases, allowing their occurrence efficiently<sup>9,10</sup>.

Like the proteins and enzymes exemplified above, many drugs are capable to interact with DNA, modifying its mechanical properties with biological implications *in vivo*. Cancer chemotherapy, for instance, is a field in which the details about DNA interactions with drugs are important. In fact, some classes of drugs such as the anthracyclines and the platinum-based compounds exhibit a strong affinity to interact with the DNA of cancer cells. When these drugs bind to DNA they can inhibit the replication process, thus stopping the tu-

<sup>a</sup> Laboratório de Física Biológica, Departamento de Física, Universidade Federal de Viçosa. Viçosa, Minas Gerais, Brazil. Fax: 55-31 3899-2483 ; Tel: 55-31 3899-3399; E-mail: [marcios.rocha@ufv.br](mailto:marcios.rocha@ufv.br)

mor growth<sup>11,12</sup>. On the other hand, gene therapy is another field of medical sciences in which this kind of knowledge is also important<sup>13,14</sup>. In these therapies, DNA molecules are usually transported from outside to inside living cells in order to replace defective genes, thus correcting cell malfunctions. One approach to accomplish this transport in an easier way, for example, is condensing the DNA molecule by using cationic ligands<sup>15,16</sup>.

In summary, all the examples discussed above show the importance in studying and understanding the details behind DNA interactions with ligands. In fact, many researchers of varied areas such as physics, chemistry, biology, medicine, pharmacy, engineering, etc have paid attention to this topic along the past 20 years, with a fast increase of the number of publications and citations<sup>17</sup>.

In this manuscript we review important topics of the field “DNA-ligand interactions”, emphasizing in how one can connect the changes of the mechanical properties of the DNA induced by the binding ligand to the physicochemical information of such interaction. In particular, we show that if one knows how a mechanical property changes as a function of the ligand concentration in the sample, many insights on the physical chemistry of the interaction can be promptly obtained. This type of connection is interesting because it allows one to perform a robust characterization of the interaction both from the mechanical and physicochemical point of view by using only one experimental technique: single molecule stretching experiments. To discuss such connection, firstly in Section 2 we discuss the basic DNA mechanics, revisiting the main concepts and approaches used in the field. In particular, we revisit the Worm-Like Chain (WLC) model, the standard one used to describe bare DNA mechanics and to investigate the changes of the DNA mechanical properties when interacting with a binding ligand. Then, in Section 3 we discuss the physical chemistry of DNA-ligand interactions, emphasizing the chemical equilibrium states which can usually be described by a binding isotherm. We also revisit the most important models used in the field, discussing their strong points and limitations. In Section 4 we discuss briefly the experimental techniques most used to perform single molecule experiments, emphasizing the key features of each one. In Section 5 we present and discuss the main types of interactions that occur between DNA and ligands: intercalation, covalent binding, electrostatic driven interactions and groove binding. Finally, in Section 6 we present and discuss the main topic of this review: the approaches on how one can connect mechanics to physical chemistry. The final conclusions are presented in Section 7.

## 2 DNA mechanics

During the last decades, DNA mechanics has become a very well studied topic especially due to the advent of single molecule techniques. Such techniques allow one to manipulate and stretch individual DNA molecules, giving access to mechanical information contained in the “force  $\times$  extension” curves. Before single molecule techniques, such type of information was somewhat difficult to be accessed by ensemble-averaging techniques.

From the mid-90s some theoretical models were formulated in order to explain the mechanical behavior of DNA molecules. In particular, most of these models attempt to give a theoretical expression for the “force  $\times$  extension” curve based on key mechanical parameters such as the linear contour length of the polymer chain and its bending stiffness, which can be conveniently represented by the persistence length.

The contour length is the most basic mechanical property of a polymer chain: it is simply the length of the chain measured along its contour, which is proportional to the number of monomers. The persistence length, otherwise, is the correlation length of the polymer chain and thus gives information about the bending stiffness of the polymer. In the case of DNA molecule, the persistence length has basically two components: the intrinsic and the electrostatic one. The first component is related to the bending rigidity due to the molecule composition itself, while the second one is due to the negative charge distribution along the double-helix<sup>18–21</sup>. Since these two components are usually present in most relevant situations, the models in general represent the persistence length by its effective value, which takes into account the two contributions. Following most authors, in this manuscript we will call the effective persistence length only by persistence length.

In water-based solutions under nearly physiological conditions (pH = 7.4, [NaCl] = 150 mM), a disperse bare DNA molecule is in general classified as a semi-flexible (or semi-rigid) polymer due to its intermediate value of the bending stiffness, which corresponds to a persistence length  $A \simeq 50$  nm<sup>6,22–24</sup>. When the contour length of the molecule is sufficiently small (smaller than a persistence length), however, the disperse DNA molecules appears a rigid rod and can be treated approximately as a rigid polymer. In this case the bending rigidity depends somewhat on the size and base-pair sequence of the DNA molecule<sup>25–27</sup>.

In the following section we present the most relevant model used to study the mechanics of the DNA molecule: the Worm-Like Chain (WLC) model, which works very well to analyze single molecule stretching experiments if the contour length of the molecule is not too small, as discussed above. In this manuscript we do not intend to review other mechanical models or present an historical survey on this specific subject, since today the WLC model is recognized as the standard one

to study DNA mechanics.

## 2.1 Worm-Like Chain Model (WLC)

The Worm-Like Chain (WLC) is a model derived from polymer physics, and has become in the past years the standard one in analyzing DNA stretching experiments. To introduce this model, let us firstly assume that the polymer itself is a chain formed by rigid rods with lengths  $b$ , connected by freely-rotating vertices. Let us call  $\theta_i$  the angle between the rods  $i$  and  $i + 1$ . The WLC model is then defined by assigning an harmonic bending energy function to the angle formed between the two rods<sup>4,24,25,28–30</sup>,

$$E(\theta_i) = \frac{\kappa}{2b} \theta_i^2, \quad (1)$$

taking the continuum limit with  $b \rightarrow 0$ . The constant  $\kappa$  is the effective elastic bending stiffness of the chain.

In the continuum limit, Eq. 1 can be used to write the total bending energy of the chain<sup>4,24</sup>,

$$E = \frac{\kappa}{2} \int_0^L |C|^2 ds, \quad (2)$$

where  $C$  is the local curvature at each point,  $ds$  is a length element along the polymer, and  $L$  is the contour length of the polymer chain.

The parameter  $\kappa$  is directly related to the polymer persistence length  $A$  by

$$A = \frac{\kappa}{k_B T}, \quad (3)$$

where  $k_B$  is Boltzmann's constant and  $T$  is the absolute temperature.

Equation 2 can be used to deduce the behavior of the force as a function of the polymer extension as one stretches it. This analysis can be performed numerically or analytically using appropriate approximations<sup>24</sup>. In 1995, Marko and Siggia solved the model analytically, obtaining an approximate expression for the force as a function of the polymer extension which has become the most used to analyze DNA stretching experiments in the entropic low-force regime ( $F \leq 5$  pN)<sup>4,23</sup>. Their result is

$$F = \frac{k_B T}{A} \left[ \frac{z}{L} + \frac{1}{4 \left(1 - \frac{z}{L}\right)^2} - \frac{1}{4} \right], \quad (4)$$

where  $F$  is the force and  $z$  is the end-to-end distance (extension) of the DNA molecule.

Despite its renowned utility, this expression is still an approximation, diverging at  $z = L$ . Moreover, Eq. 4 describes well only the entropic regime of the polymer, which is valid for stretching forces typically below  $\sim 5$  pN. In this regime the

applied forces are sufficiently small such that they can change only the polymer conformation in solution, *i. e.*, its entropy.

Also in 1995, Odjik proposed a different approach that accounts for higher forces, in which enthalpic effects start to become relevant for the polymer mechanics<sup>5</sup>. The enthalpic regime is defined as the regime in which the stretching forces became large enough to distort the DNA primary structure and eventually to break chemical bonds. Such effect can be accounted by introducing an enthalpic mechanical parameter to describe the polymer deformation: the stretch modulus  $S$ . The analytical expression proposed by Odjik reads<sup>5</sup>

$$z = L \left[ 1 - \frac{1}{2} \sqrt{\frac{k_B T}{AF} + \frac{F}{S}} \right]. \quad (5)$$

Observe that the stretch modulus  $S$  has units of force. Taking the limit  $S \rightarrow \infty$  and inverting the above equation (isolating  $F$ ), we found an equation similar to the Marko-Siggia expression (Eq. 4) if  $z \sim L$ , *i. e.*, neglecting very small forces. Thus, observe that a polymer in the entropic regime can be interpreted as a polymer that has a stretch modulus  $S$  very high, *i. e.*, that resists deformations on its chemical structure.

In 1999, Bouchiat *et al.* proposed another solution of the WLC model in the entropic regime. Their approach consists in adding six terms to Eq. 4 in order to improve its accuracy<sup>31</sup>. These terms were determined by comparing the results predicted by Eq. 4 to results from an exact numerical solution of the WLC model<sup>31</sup>, which was obtained perturbatively. The resulting expression reads

$$F = \frac{k_B T}{A} \left[ \frac{z}{L} + \frac{1}{4 \left(1 - \frac{z}{L}\right)^2} - \frac{1}{4} + \sum_{i=2}^7 a_i \left(\frac{z}{L}\right)^i \right], \quad (6)$$

where the  $a_i$ 's are constants numerically determined.

In addition to the models discussed above, important contributions to the elucidation of many peculiarities of DNA mechanics were given by the groups of A. Vologodskii, M. D. Frank-Kamenetskii, H. E. Gaub, M. C. Williams, V. Croquette, F. Ritort, C. Bustamante and others, especially concerning the bending of small DNA fragments, strong bending and fluctuations in the double-helix, dependence of DNA rigidity on the temperature and base sequence, DNA twist, overstretching transition, DNA hairpins, etc.<sup>21,22,25–27,32–51</sup>.

## 3 Physical chemistry of DNA-ligand interactions

The study of the physical chemistry of DNA-ligand interactions consists in two different sub-fields: the chemical equilibrium of the interaction and the kinetics of the interaction. Consider the system of interest (DNA + ligand molecules in

solution) as composed by two different partitions where the ligand molecules can stay: the DNA (bound ligand molecules) and the solution (free ligand molecules). The chemical equilibrium is achieved when the average number of molecules in the partitions remains constant in time. The kinetics of the interaction, otherwise, describes the changes that occur between the initial incubation and the final equilibrium state.

In this manuscript we emphasize the physical chemistry of the chemical equilibrium, since the equilibrium states can be represented by a binding isotherm that can be linked to the changes of the mechanical properties of DNA-ligand complexes. Below we discuss the most relevant models that attempt to describe the chemical equilibrium of DNA-ligand interactions. Some studies on the kinetics of such interactions were performed by the groups of M. C. Williams, D. Anselmetti, D. M. Crothers and others<sup>52–60</sup>.

### 3.1 The general problem

Consider two molecules A and B associating in solution to result in a molecule C. This mechanism can be represented by the chemical reaction



where  $K_i$  and  $K_d$  are, respectively, the equilibrium intrinsic binding constants of association and dissociation. They are also known as thermodynamic constants or macroscopic constants. Observe that  $K_i$  represents the association reaction, where the reagents A and B associate to result in the compound C, while  $K_d$  represents the dissociation reaction, *i. e.*, the reverse reaction in which C dissociates in the original reagents A and B.

These constants are defined in term of the molar concentrations of the involved substances,

$$K_i = \frac{[C]}{[A][B]}, \quad (8)$$

and

$$K_d = \frac{[A][B]}{[C]} = K_i^{-1}. \quad (9)$$

Note that in these last two equations,  $[X]$  is the molar concentration (1 M = 1 mol/liter) of the compound X. Also observe that  $K_i$  has units of  $M^{-1}$ , while  $K_d$  has units of M.

### 3.2 Scatchard model

This is the simplest model that can describe the chemical equilibrium of the DNA molecule with ligands in solution. Let us firstly adapt the previous notation for the specific case of DNA-ligand interactions. Call  $[A] \equiv C_f$  the concentration of

free ligands solution and  $[C] \equiv C_b$  the concentration of ligands bound to DNA (result of the reaction). Suppose firstly that each ligand molecule occupies only one base pair of the DNA when bound. Consequently, the concentration of free linkable sites in the DNA molecule can be written as  $[B] \equiv C_{bp} - C_b$ , where  $C_{bp}$  is the concentration of DNA base pairs, which is a constant.

Substituting these definitions in Eq. 8, one has

$$K_i = \frac{C_b}{C_f(C_{bp} - C_b)}. \quad (10)$$

Now we introduce the bound ligand fraction  $r$ ,

$$r = \frac{C_b}{C_{bp}}, \quad (11)$$

such that Eq. 10 can be rewritten as

$$r = \frac{K_i C_f}{1 + K_i C_f}, \quad (12)$$

which is known as the Scatchard binding isotherm, proposed originally in 1949<sup>61</sup>.

Despite its didactic utility, the Scatchard binding isotherm has two important simplifications: (a) It is valid only for very small ligand molecules which occupy only one DNA base-pair when bound, which is not the case for most ligand molecules. (b) It supposes that previous bound ligand molecules do not interfere in the binding mechanism of the subsequent ones, *i. e.*, the interaction is non-cooperative.

The first simplification can be bypassed by introducing the parameter  $r_{max}$ , the bound ligand fraction at saturation, *i. e.* the maximum value of the bound ligand fraction  $r$ . Observe that the inverse of  $r_{max}$  is the mean number of base pairs occupied by each bound ligand molecule  $N = 1/r_{max}$ . The corrected binding isotherm then reads

$$r = \frac{r_{max} K_i C_f}{1 + K_i C_f}, \quad (13)$$

### 3.3 Hill model

The Hill binding isotherm was originally proposed by A. V. Hill in 1910 to describe the binding of oxygen to hemoglobin inside red blood cells<sup>62</sup>.

Basically the model introduces the Hill exponent  $n$ , a cooperativity parameter which is a lower bound for the number of cooperating ligand molecules involved in the reaction<sup>63,64</sup>. The binding isotherm reads

$$r = \frac{r_{max} (K_i C_f)^n}{1 + (K_i C_f)^n}. \quad (14)$$

The apparent binding association constant of the reaction is defined as  $K_A = K_i^n$ . Observe that if  $n > 1$ , the interaction is

positively cooperative, *i. e.*, a bound ligand molecule increases the apparent affinity of DNA for subsequent ligand binding. If  $n < 1$ , otherwise, the interaction is negatively cooperative and a bound ligand molecule decreases the apparent affinity of DNA for subsequent ligand binding. If  $n = 1$ , the interaction is non-cooperative and the affinity is independent of the number of previously bound ligand molecules.

The Hill binding isotherm has achieved a particular success to describe positively cooperative “none-or-all” processes ( $n > 1$ ), in which the cooperating ligand molecules bound practically simultaneously to the binding site forming a bound cluster<sup>63,65</sup>. On the other hand, when  $n = 1$  the Hill isotherm reduces to the Scatchard one and is therefore able to describe individual binding of ligand molecules<sup>64</sup>. Finally, to the best of our knowledge there is no report in the literature of a negatively cooperative DNA-ligand interaction described by a Hill binding isotherm.

### 3.4 Neighbor exclusion model (NEM)

This model was proposed in 1974 by McGhee and von Hippel with the purpose of analyzing in detail the neighbor exclusion effects due to large ligand molecules that occupy more than one DNA base-pair<sup>66,67</sup>. The authors have accounted for the ligand size by introducing the exclusion parameter  $N$ , the number of base-pairs that a ligand molecule effectively occupies when binding to DNA. This parameter was cited earlier in connection to the saturated bound ligand fraction,  $N = 1/r_{max}$ .

The model has a non-cooperative and a cooperative version, but the last one has been used only in a few works<sup>68,69</sup> to analyze experimental data because the binding isotherm is somewhat intricate.

The non-cooperative binding isotherm reads

$$\frac{r}{C_f} = K_i(1 - Nr) \left[ \frac{1 - Nr}{1 - (N-1)r} \right]^{N-1}, \quad (15)$$

and the cooperative binding isotherm reads

$$\frac{r}{C_f} = K_i(1 - Nr) \left[ \frac{(2\omega - 1)(1 - Nr) + r - R}{2(\omega - 1)(1 - Nr)} \right]^{N-1} \times \left[ \frac{1 - (N+1)r + R}{2(1 - Nr)} \right]^2, \quad (16)$$

with

$$R = \sqrt{[1 - (N+1)r]^2 + 4\omega r(1 - Nr)}. \quad (17)$$

Here  $\omega$  is the cooperativity parameter. For  $\omega$  smaller, equal, or larger than unity, one has negative, non-cooperative, or positive cooperativity, respectively.

The major advantage of this model is to treat in more detail the effects related to the ligand size. This feature is particularly important in the analysis of DNA interactions with intercalators, a class of ligands in which neighbor-exclusion effects is extremely important<sup>70–72</sup>. In fact, NEM has become in the past years the standard binding isotherm used to analyze DNA interactions with intercalators<sup>73–77</sup>.

## 4 Single molecule experimental methods

In this section we briefly discuss the single molecule experimental techniques commonly used to measure the mechanical properties of the DNA-ligand complexes.

The main advantage of single molecule techniques is the possibility to study a particular DNA molecule free from the influence of other molecules in the sample. Single molecule stretching experiments such as those performed with optical or magnetic tweezers usually give insights on the global (long length scale) mechanical properties of individual DNA molecules. In fact, mechanical parameters such as the persistence and contour lengths and the stretch modulus can be extracted by analyzing the “force  $\times$  extension” curves of the complexes, which can be obtained in single molecule approaches.

Useful reviews which discuss and compare single molecule techniques can be found in the literature<sup>78–80</sup>.

### 4.1 Optical tweezers

Since the seminal works of Ashkin and collaborators<sup>81,82</sup>, optical trapping and manipulation have found various applications in many areas of science such as physics, biology and chemistry. Today, the most common optical tweezers are mounted by focusing a laser beam with a microscope objective of large numerical aperture. This apparatus can trap small dielectric objects near the lens focus, being a powerful tool to manipulate beads, particles and biological systems with typical sizes in the micrometer range<sup>82,83</sup>. The typical forces obtained with this apparatus are between 0.1 - 400 piconewtons, which are in the range of many biological forces such as the entropic and enthalpic forces on biopolymers and molecular motors. For an introductory review about the basic theory and features of optical tweezers, see ref.<sup>84</sup>. Other useful reviews on instrumentation and recent advances on the technique can also be found in the literature<sup>85–88</sup>.

To perform precise quantitative measurements with optical tweezers, size-calibrated dielectric beads have become the standard objects to be captured because of their perfect symmetry which facilitates trap calibration and position detection. A dielectric bead trapped in an optical tweezers is an overdamped Brownian harmonic oscillator, such that the optical

trap can be characterized by its trap stiffness  $\kappa$  which depends on the bead size and refractive index<sup>84</sup>.

In the last decades, optical tweezers have been largely used to study the mechanical properties of DNA/RNA molecules and their complexes formed with drugs or proteins. Useful reviews on this subject can be found in the literature<sup>89–92</sup>. Basically, the classic experiment consists in attaching one end of the DNA molecule to a polystyrene or silica bead and the other end to a substrate (a microscope coverslip or a second bead attached to a micropipette, for example). The optical tweezers is then used to trap the bead and so the DNA molecule can be manipulated and stretched by moving the laser beam or the microscope stage. The force as a function of extension can be measured as one stretches the DNA molecule. To perform this task, one needs to detect the bead position and to calibrate the tweezers (determine the trap stiffness  $\kappa$ ). There are many techniques which can be used to perform this kind of measurement, such as dynamic light scattering<sup>93,94</sup>, back-focal plane interferometry<sup>95</sup>, statistics of thermal fluctuations<sup>96</sup>, simple videomicroscopy<sup>64,97,98</sup>, calibration using hydrodynamic drag forces<sup>56</sup> or by using other types of detectors<sup>91</sup>. For a recent review on measuring with optical tweezers, see ref.<sup>85</sup>.

Figure 1 shows a typical “force  $\times$  extension” curve of a single bare  $\lambda$ -DNA molecule ( $\sim 48,500$  base-pairs) obtained by performing a DNA stretching experiment in the entropic regime with optical tweezers. The trap calibration and the bead position detection were performed in this case by using videomicroscopy<sup>97</sup>, and the solid line corresponds to a fitting using the Marko-Siggia WLC model (Eq. 4). From the fitting one can promptly determine the persistence and contour lengths of the DNA molecule, obtaining for this particular curve  $A = (50 \pm 2)$  nm and  $L = (15.6 \pm 0.1)$   $\mu\text{m}$ .

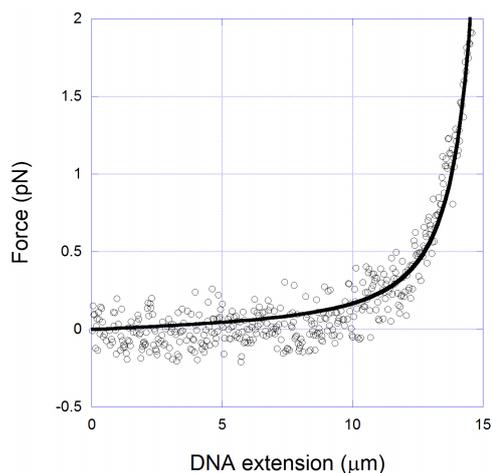
## 4.2 Magnetic tweezers

The idea behind magnetic tweezers is very similar to its optical analogue, the main difference is that in this case the forces are exerted by an external magnetic field applied around the sample. Paramagnetic beads are used instead of dielectric ones in order to be manipulated with the magnetic field. The typical forces obtained are of the order of hundredths of picoNewtons to hundreds of picoNewtons.

Basically, the force applied on the paramagnetic beads can be written as

$$\vec{F} = -\frac{1}{2} \nabla (\vec{\mu} \cdot \vec{B}), \quad (18)$$

where  $\vec{\mu}$  is the magnetic dipole moment induced in the bead and  $\vec{B}$  is the applied magnetic field. Observe that for moderate magnetic fields one has  $\vec{\mu} \propto \vec{B}$  and the resulting force is proportional to the gradient of the field intensity.



**Fig. 1** “Force  $\times$  extension” curve of a bare  $\lambda$ -DNA molecule in the entropic regime. *Circles*: experimental data obtained with optical tweezers; *Solid line*: a fitting to the Marko-Siggia Worm-Like Chain (WLC) model (Eq. 4). For this particular DNA molecule we have found from the fitting  $A = (50 \pm 2)$  nm and  $L = (15.6 \pm 0.1)$   $\mu\text{m}$ .

Reviews on the basic and advanced features of magnetic tweezers can be found in the literature<sup>99–101</sup>.

An advantage of this technique in relation to optical tweezers is its convenience to apply torques on the magnetic beads by rotating the external magnetic field, which allows one to rotate the tethered DNA molecules and therefore to study quantities such as the torsional rigidity and the degree of supercoiling<sup>22,39,50,102–105</sup>. These quantities are also mechanical properties important to some biological processes in which the double-helix must be unwound, such as in DNA replication. Another advantage of the magnetic tweezers is its convenience to perform constant-force experiments, working as a force-clamp trap (it is just a matter of choosing the adequate magnetic field - see Eq. 18). Constant-force experiments can be performed with optical tweezers only using non-conventional (and more intricate) approaches such as by using a force-feedback electronics or working in anharmonic regions of the optical potential<sup>106,107</sup>. Among the disadvantages of using magnetic tweezers, one can cite the hysteresis of the magnetic field and heat generation around the sample if the field is produced by current distributions, aside the more intricate calibration of the apparatus and its restriction to applications with magnetic materials. A recent work by Neuman and Nagy provides a detailed comparison between optical and magnetic tweezers, and also atomic force microscopy<sup>79</sup>.

## 4.3 Atomic force microscopy (AFM)

Atomic force microscopy (AFM) is another important tool in single-molecule studies of DNA-ligand interactions<sup>108,109</sup>. In

the last 20 years, a number of different protocols have been developed in order to deposit DNA molecules on a flat surface and to image them reliably and reproducibly. Today, the standard surfaces used to deposit the DNA molecules are mica substrates and less often silicon substrates, because of their low rugosity. A number of buffer solutions containing divalent cations (such as  $\text{Mg}^{2+}$ ,  $\text{Ni}^{2+}$ , etc) have been used to enhance the DNA adsorption onto the substrate, which is otherwise poor. Moreover, divalent cations also allow the polymer chain to equilibrate on the flat 2D adsorbing surface, preventing chain kinetic trapping, which must be avoided in order to study the equilibrium properties of the adsorbed DNA or DNA-ligand complex<sup>108</sup>. Once adsorbed, DNA molecules and DNA-ligand complexes can be imaged using the AFM usually operating in the tapping mode, which minimizes possible damages to the sample due to the tip-surface interactions during the scanning. The images obtained are topographical maps which associate a certain height to each point on the sample. By analyzing these images, several DNA statistical parameters such as the mean contour length, the persistence length and bending angles can be estimated. As discussed by Rivetti *et al.*<sup>110</sup>, this analysis can be performed, for example, by measuring the mean-squared end-to-end distance  $\langle R^2 \rangle$  of the polymer. In fact, statistical mechanics of polymers predicts that for 2D worm-like chains (which is the case of deposited DNA molecules),  $\langle R^2 \rangle$  is given by

$$\langle R^2 \rangle = 4AL \left[ 1 - \frac{2A}{L} \left( 1 - e^{-\frac{L}{2A}} \right) \right]. \quad (19)$$

By using this equation it is possible to determine the persistence length  $A$  by measuring  $\langle R^2 \rangle$  and the contour length  $L$  for the deposited DNA molecules.

On the other hand, the visualization of DNA condensates formed with polycations with the AFM technique is nowadays a routine in many laboratories. The morphology of these condensed DNA complexes seen in the AFM images are as well as clear and well-defined as the images produced by other kinds of microscopy techniques such as electron microscopy (EM). The structure of DNA-protein complexes has also been a target of a number of studies. Within the limits of the AFM technique are, for instance, the visualization of sharp kinks and cross-links introduced in the DNA molecule by histone-like proteins, the structure of nucleosome particles, the determination of protein binding-sites and more recently the determination of protein association constants to DNA<sup>111</sup>.

Finally, besides being a powerful tool for visualizing single molecules, the AFM apparatus can also be used to perform force spectroscopy in solution like optical or magnetic tweezers, allowing one to determine the “force  $\times$  extension” curves of the DNA-ligand complexes<sup>36,112,113</sup>.

Specific reviews on the application of the AFM technique to single molecule studies can be found in the literature<sup>113,114</sup>.

As a final remark, with the improvement of fluorescent-based optical technology along the last decades, fluorescence microscopy has also become another important tool to visualize DNA structure, conformation changes and interactions with ligands at single molecule level<sup>115–123</sup>. The technique can be used as complementary to AFM, with the advantage that one does not need to deposit the molecules in a substrate.

## 5 DNA-ligand interactions

DNA can interact with ligands in many different ways, from covalent binding to simple electrostatic driven interactions. Here we describe briefly the most relevant types of interactions, discussing the main features of each one.

### 5.1 Covalent ligands

The covalent binding of drugs to DNA is usually irreversible and completely inhibits DNA processes. The platinum-based compounds are examples of drugs which can interact with DNA by covalent binding<sup>124,125</sup>. Cisplatin and its related compounds carboplatin and oxaliplatin are antitumor platinum-based molecules usually used in cancer chemotherapies. The action of these complexes as anticancer drugs consists in damaging the DNA molecule with adducts that form various types of crosslinks, which introduce strong structural perturbations and impede DNA replication<sup>126,127</sup>. The clinical use of these complexes, however, is limited due to their several side effects and the development of drug resistance.

Another example of covalent binding is found in the interaction of drugs from the class of furocoumarins (psoralen, angelicin, etc) with DNA when one illuminates the complex with ultraviolet-A (UVA) light<sup>128</sup>. Psoralen is a well-known drug used in the treatment of skin diseases like psoriasis, vitiligo, and some other kinds of dermatitis<sup>129</sup>. The most common therapy is called PUVA (psoralen followed by UVA light), which consists in taking a medicine containing psoralen and exposing the patient to UVA light. The drug effectively increases the skin sensitivity to UVA and the skin melanin level<sup>130,131</sup>. It is well established in the literature that when a DNA-psoralen complex is illuminated with UVA light, the drug molecules absorb photons and form covalent bonds preferentially with the thymines<sup>128,132</sup>. When there is no illumination at the sample, however, psoralen interacts with DNA by intercalative binding - see next section. The effects of covalent binding on the mechanical properties of the DNA-psoralen complexes were recently studied<sup>94,133</sup>. In particular, it was shown that the contour and persistence lengths of the complexes depend on the psoralen concentration and on the exposure time to UVA light<sup>133</sup>.

## 5.2 Intercalators

Intercalative binding is one of the most common interactions between DNA and drugs, and was firstly described by L. S. Lerman in 1961<sup>134,135</sup>. It is characterized by the insertion of a flat aromatic molecule between two adjacent DNA base pairs. The complex is thought to be stabilized by the stacking interactions between the ligand and the DNA bases<sup>136</sup>. Intercalators also introduce strong structural perturbations on the double-helix structure. To accommodate the intercalated molecules, there is an increase in the DNA contour length, which is accompanied by an unwinding of the double-helix by a certain angle per intercalated molecule<sup>56,70–73,137</sup>. Daunomycin, doxorubicin and ethidium bromide (EtBr) are classic examples of drugs which intercalate in the DNA molecule and can modify its elasticity depending on the drug concentration. Daunomycin and doxorubicin are anthracycline antibiotics used in the treatment of various cancers such as some types of leukemias, sarcomas, lymphomas, myelomas, neuroblastomas, as well as cancers in the breast, head, ovary, pancreas, prostate, stomach, liver, lung and others. They inhibit DNA replication and transcription when intercalating, impeding cell duplication<sup>73</sup>. Ethidium bromide (EtBr) is commonly used as a fluorescent stain for identifying and visualizing nucleic acid bands in electrophoresis and in other methods of nucleic acid separation. Other known intercalators are the DNA fluorescent stains acridine orange, methylene blue<sup>138</sup> and diaminobenzidine<sup>76</sup>. More examples and specific reviews on the basic properties of intercalators can be found in the literature<sup>70–72</sup>.

Many aspects of the DNA-daunomycin interaction, such as kinetics, self association and equilibrium binding were studied by J. B. Chaires, D. M. Crothers and coworkers in the 80's<sup>60,73,137,139</sup>. On the other hand, the DNA-EtBr interaction was characterized in various aspects by many authors, but even today one can find somewhat contradictory results about the mechanical behavior of such complexes<sup>23,56,75,77,140–142</sup>, and also for different complexes formed between DNA and other intercalators, especially when comparing results obtained from different experimental techniques<sup>56,75–77,115,140–146</sup>.

In the past years our group has studied in detail the changes in the persistence and contour lengths of DNA complexes formed with various intercalating molecules<sup>75–77,94,133</sup>, by using optical tweezers in a very low force regime ( $F < 2$  pN). We reported an abrupt structural transition in the persistence length due to drug intercalation, which is probably related to a partial denaturation of the DNA molecule due to the pulling force used to stretch the complexes<sup>76,77,133,147</sup>. In fact, recently we have explicitly shown that the persistence length of DNA-intercalators complexes is in general force-dependent<sup>148</sup>. The contour length, otherwise, does not present

this kind of behavior, increasing monotonically with drug concentration until saturation.

## 5.3 Electrostatic driven interactions

Since the DNA molecule has a high negative charge density in aqueous solution due to its phosphates (2 elementary charges per each 3.4 Å along the DNA axis), it strongly interacts with itself (*i. e.*, different DNA segments strongly repel each other thus promoting the chain swell) as well as with positively charged ligands such as ions and macro-ions, especially multivalent cations<sup>149</sup>.

DNA condensation due to multivalent cations is a classic example which shows the importance of electrostatic driven interactions in DNA solutions<sup>149–152</sup>. In this process the multivalent cations bind along the DNA double-helix, and the strong positional correlations between them start to play a role and promotes a coil-globule transition: the DNA molecule folds onto itself<sup>153–155</sup> with a high increase in the local DNA segment density at the level of both monomolecular collapse or in a multimolecular aggregation. Only cationic ligands with charge equal or superior to +3 are capable to condense the DNA molecule. As classic examples of DNA condensing agents we cite the naturally occurring amines spermine and spermidine, as well as compounds such as hexamine cobalt, multivalent metal ions and proteins. In addition, the size (contour length) of the DNA molecule is important in such process, since the bound cationic molecules act promoting DNA segment-segment attraction due to the ion-ion positional correlations<sup>153</sup>. Therefore, very small DNA fragments usually cannot be condensed at monomolecular level<sup>150</sup>.

Some models concerning electrostatic interactions between DNA and ligands were proposed along the last decades<sup>19,156,157</sup>. In fact, there are different hypotheses to explain the DNA bending mechanism by multivalent cations, including a purely electrostatic model by Rouzina and Bloomfield<sup>156</sup> and an asymmetrical phosphate neutralization model by Manning<sup>19</sup>. According to Rouzina and Bloomfield, a multivalent cation binds to the entrance of the DNA major groove, between the two phosphate strands, electrostatically repelling sodium counterions from the neighboring phosphates. The unscreened phosphates on both strands are strongly attracted to the groove-bound cation. This binding leads to groove closure, accompanied by DNA bending towards the cationic ligand<sup>156</sup>. Differently, Manning proposes that the stable double-helix structure of DNA represents an equilibrium between stretching forces (caused by interphosphates repulsion) and compressive forces (caused by attractive interaction between nucleotides). This analysis suggests that significant local interphosphate stretching forces balance compressive forces within DNA and that these stretching forces can drive DNA deformation when phosphates charge are locally neutralized.

These two approaches predict a reduction of the persistence length as the concentration of bound cations increases. In fact, the model proposed by Rouzina and Bloomfield predicts that the effective persistence length  $A_E$  of the DNA-ligand complex is given by

$$\frac{1}{A_E} = \frac{1}{A_1} + \frac{Nr}{A_2}, \quad (20)$$

where  $A_1$  is the bare DNA persistence length (when no ligands are bound  $r = 0$ ) and  $1/A_1 + 1/A_2$  is the inverse persistence length of a DNA saturated with ligands (which occurs when  $r = r_{max} = 1/N$ ). Observe that here  $N$  is the exclusion parameter of the ligand and  $r = C_b/C_{bp}$  is the ratio between the bound ligand concentration and the DNA base-pair concentration, as introduced in Section 3.

The model developed by Manning, on the other hand, predicts that the effective persistence length  $A_E$  of the DNA-ligand complex (charge neutralized DNA) is related to the original persistence length  $A_0$  (fully charged DNA) by the equation<sup>19,157</sup>

$$A_E = \frac{2}{\pi R^2} \left[ \frac{\beta A_0}{2(\xi - 1) - \ln(\kappa b)} \right]^{3/2}, \quad (21)$$

where  $R$  is the radius of the double helix,  $\beta$  is the Bjerrum length (distance between two unit charges in pure solvent - no other ions - at which the electrostatic energy is  $k_B T$ ),  $1/\kappa$  is a measure of the extent of the ion cloud around the object,  $b$  is the average axial distance between phosphates (0.17 nm) and  $\xi = \beta/b$  is a measure of the axial charge density of the DNA<sup>157</sup>.

This model predicts, for example, that for 30% of neutralized charge, the effective persistence length is  $A_E = 33.2$  nm. For 60% of neutralized charge,  $A_E = 11.1$  nm and for 100% of neutralized charge,  $A_E = 7$  nm<sup>157</sup>.

The two presented models (Eqs. 20 and 21), however, do not take into account the interaction between different DNA segments in the polymer chain and thus, rigorously speaking, are valid only for small DNA fragments. As a consequence, these models are capable to predict a reduction of the persistence length due to phosphate neutralization, but not the coil-globule transition due to cationic ligand binding.

#### 5.4 Major and minor groove ligands

Most drugs that interact electrostatically with DNA usually exhibits a preference to the major or minor groove floor of the double-helix. Many minor groove ligands are known by their antitumor and antibiotic functions. This kind of interaction is usually characterized by a combination of electrostatic, van der Waals and hydrogen bonds. Examples of minor groove ligands are the anticancer compound distamycin A, the antibiotics netropsin and berenil, and the fluorescent

stain DAPI. These drugs usually form reversible complexes with DNA, preferentially binding at AT base pairs sequences. They also induce elasticity changes on the DNA molecule, stabilizing the double-helix structure<sup>56</sup>. An extensive review on the DNA minor groove complexes can be found in ref.<sup>158</sup>.

On the other hand, major groove binding is also a kind of interaction usually characterized by electrostatic binding<sup>56</sup>.  $\alpha$ -Helical (Ac-(Leu-Ala-Arg-Leu)3-NH linker) is a peptide which interacts with DNA via major groove binding<sup>56</sup>. Other known examples are the intercalator and major groove ligand YO<sup>159</sup>, the bis-intercalator and major groove ligands YOYO and ditercalinium<sup>145,159,160</sup> and the anticancer drug neocarzinostatin<sup>160</sup>. More examples and a discussion on the main characteristics of major groove binding ligand can be found in a recent review<sup>160</sup>.

#### 5.5 Ligands with multiple binding modes

There are many ligands which can interact to DNA by different binding modes, depending on factors such as the properties of the surrounding buffer solution, the DNA base-pair sequence, external conditions such as sample illumination, etc. Some examples were already cited in the last sections. In some cases the ligand has distinct portions which interact to DNA by different modes. In other cases there is only a single binding mode for the entire ligand molecule, which can be changed upon determined conditions.

Bis-intercalators like YOYO and ditercalinium, for instance, are molecules which have two intercalating portions linked by another chemical structure which sometimes may interact with the DNA grooves<sup>145,159,160</sup>. Actinomycin D is another example of a drug with distinct portions that interact to DNA by different modes, in this case including minor groove binding and intercalation<sup>64,161-164</sup>.

Psoralen is an example of a drug which the binding mode depends on an external condition (sample illumination). As explained before, the drug initially intercalates in DNA, but forms covalent bonds with the thymines if the sample is illuminated with UVA light.

Hoechst 33258 is a fluorescent stain that can bind to DNA by intercalation or groove binding, with two different sets of physicochemical parameters. In this case the drug concentration is the factor that determines the dominant binding mode<sup>65</sup>. Some authors report a similar behavior for the intercalator doxorubicin, have founding a possibility of groove binding at AT-rich regions<sup>165</sup>.

## 6 Connecting mechanics to physical chemistry

In this section we introduce the main subject of this review, the approaches developed to establish connections between the mechanical properties and physicochemical properties of

DNA-ligand complexes. As stated before, the advantage in establishing this type of connection is the possibility to get insights on one or more properties of a certain type (physicochemical properties, for example) knowing only the behavior of a property of the other type (the persistence or contour length, for example). With such connection(s), one can considerably reduce the number of different experimental techniques necessary to perform a robust characterization of the DNA interaction(s) with a certain type of ligand. This fact thus reduces the time and cost required for getting data, since less different equipments are needed and the number of experiments that must be conducted can be considerably reduced. Furthermore, and perhaps more important, the approach opens the possibility of comparing data obtained by means of very different experimental techniques, increasing confidence in the results.

It is worth to remember that, as discussed in Section 1, both mechanical and physicochemical information about DNA complexes formed with ligands is important in the understanding of many intracellular DNA-related processes as well as in the design of new drugs and in the optimization of some current treatments of human diseases. Thus, the methodologies presented here can give important new insights both to the basic sciences as well as to the applied sciences.

In single molecule stretching experiments performed by optical or magnetic tweezers, the typical result obtained is the “force  $\times$  extension” curve of the molecule, from where the mechanical properties can be extracted by fitting an appropriate model (for DNA, the WLC model). We will show that if one knows how the contour and/or the persistence length varies as a function of the total concentration of ligand in solution ( $C_T$ ) (which is the amount of ligand added in sample preparation), it is possible to deduce physicochemical properties such as the equilibrium constants, the cooperativity degree, the exclusion number, etc.

It is worth to emphasize that all the results presented and discussed below in this manuscript were obtained by performing single molecule stretching experiments in the equilibrium situation, *i. e.*, when the number of ligand molecules bound to the DNA has achieved its equilibrium mean value (see the discussion in Section 3). This equilibrium situation can be verified experimentally by measuring the mechanical properties as a function of the time measured after initial sample incubation. In general, when one incubates bare DNA and a certain concentration of ligand, the mechanical properties of the complexes formed change in time as the ligand molecules bind to the double-helix structure, starting from the values corresponding to the bare DNA until reaching their final equilibrium values, which thus characterizes the equilibrium state. In addition, the “force  $\times$  extension” curves can be compared to the relaxation curves of the complexes in any situation. The absence of hysteresis in the stretching-relaxation curve sets is

a strong indication that the experiments are being performed in equilibrium and that the forces used are not capable to change the structure of these complexes<sup>148,166,167</sup>.

Finally, another important point deserves to be commented here. A rigorous control of the DNA integrity must be performed in single molecule experiments in order to guarantee accurate results. In the experiments performed in our group, bare DNA molecules were exhaustively tested before adding any ligand in the sample. These tests were performed by stretching the bare DNA many times in order to determine if the double-helix structure is not damaged, by measuring the mechanical parameters and comparing the results obtained to the well-known reference values. Possible DNA damage such as nicks or strand breaks are promptly identified from deviations of the persistence or/and contour lengths from the expected values. Thus, the experiments were started by adding the ligand only if the bare DNA molecule presents the expected results, and the entire experiment is performed with this same DNA molecule by changing the ligand concentration in the sample. Complete details about this procedure were previously reported<sup>64,97</sup>.

## 6.1 Historical survey

Along the past years many groups have used single molecule techniques to identify the possible binding mechanisms of DNA-ligand interactions and to extract physicochemical information of such interactions from these types of experiments<sup>56,63,65,75,90–92,97,98,137,140,159,168–175</sup>.

To the best of our knowledge, the first attempt to connect mechanics to physical chemistry for DNA-ligand systems was performed in the early 80's by the group of D. M. Crothers, who have measured the changes of the DNA contour length when interacting with various drugs (netropsin, distamycin, iremycin, daunomycin) as a function of the bound ligand fraction  $r$ , by using electric dichroism and a phase partition technique<sup>137,168</sup>. Nevertheless, they have not directly determined physicochemical properties from such data, a task which could only be performed with complementary analyzes and/or techniques.

In 1996 Coury *et al.*<sup>169</sup> has determined physicochemical properties from the contour length data of some DNA-ligand complexes, obtained using AFM<sup>169</sup>. In fact, by determining the relative increase of the contour length of DNA complexes formed with intercalators such as daunomycin and ethidium bromide, the authors were capable to estimate binding parameters such as the equilibrium constant and the exclusion number. Similar approaches were used by Mihailovic *et al.*<sup>170</sup> and Rocha *et al.*<sup>75</sup> to extract physicochemical information of DNA-intercalator complexes by measuring the relative increase of the contour length, obtained using optical tweezers.

Basically, the idea to perform such task is the follow-

ing. Call  $\Delta$  the natural distance between two DNA base pairs, which is  $\sim 0.34$  nm for B-DNA. When an intercalating molecule binds to this site, it increases such distance to a new value  $\Delta + \delta$ . Call  $L_0$  the bare DNA contour length and  $L$  the new length for a certain amount of bound ligand represented by the bound fraction  $r$ . One can promptly write the relation

$$L = L_0 + N_b \delta, \quad (22)$$

where  $N_b$  is the number of bound ligand molecules.

Observe that  $L_0 = N_{bp} \Delta$ , being  $N_{bp}$  the number of DNA base-pairs. Therefore one can write the relative change of the contour length  $\Theta$  as

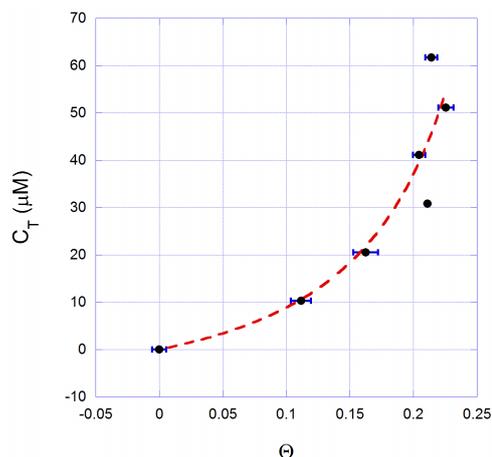
$$\Theta = \frac{L - L_0}{L_0} = \frac{N_b \delta}{N_{bp} \Delta} = \gamma r, \quad (23)$$

where  $\gamma = \delta/\Delta$ .

With Eq. 23 one can directly connect the mechanical parameter  $L$  to physicochemical parameters by expressing the bound ligand fraction  $r$  by an adequate binding isotherm. In the case of intercalators, the more convenient binding isotherm is the Neighbor Exclusion Model (NEM) (see Section 3.4), since this isotherm captures in detail the neighbor exclusion effects which always follow intercalative binding. Nevertheless, there are two problems that must be bypassed to use this approach. The first one is that in the NEM binding isotherm one cannot analytically isolate the parameter  $r$  to substitute in Eq. 23. This problem, however, can be bypassed with numerical approaches, as will be discussed soon. The second and more serious problem is that, not only NEM, but all binding isotherms are written as functions of the free ligand concentration  $C_f$ , which is not a directly accessible parameter. In fact, in general one knows only the total ligand concentration in solution  $C_T$ , the quantity used to prepare the sample, which is the sum of the bound and free ligand concentrations, *i. e.*,

$$C_T = C_f + C_b. \quad (24)$$

The partitioning of  $C_T$  into  $C_f$  and  $C_b$  is not trivial to be measured and one usually needs other experimental techniques (microcalorimetry, absorption spectroscopy, equilibrium dialysis, etc.) to evaluate such partitioning. There are, however, approaches that can be performed to bypass this problem, allowing one to use only single molecule stretching to characterize the interaction. In fact, one can estimate the bound ligand concentration from contour length changes if the length increase due to a single binding event ( $\delta$ ) is known<sup>169,170</sup>. Alternatively, as a first-order approximation one can consider  $C_f \sim C_T$  in the binding isotherm if the DNA concentration in the sample is very low (because  $C_b$  will also be very low in this case)<sup>63,170</sup>. This approximation is much used in typical tweezers experiments that tether an individual



**Fig. 2** Experimental result (*circles*) of  $C_T \times \Theta$  measured for DNA complexes formed with the intercalator diaminobenzidine, and a fitting to Eq. 25 (*dashed line*). Observe that Eq. 25 fits well to the experimental data, returning the values of the physicochemical parameters  $N = 2.5 \pm 0.6$ ,  $K_i = (1.8 \pm 0.6) \times 10^4 \text{ M}^{-1}$  and  $\gamma \sim 1$ . For this data  $C_{bp} = 2.4 \text{ } \mu\text{M}$ .

DNA molecule and then rinse away any DNA molecules in solution prior to the introduction of a ligand. Such approach is convenient because it allows one to express the binding isotherm as a function of a directly accessible parameter ( $C_T$ ), although it cannot be used always. A different approach was proposed originally by Rocha *et al.* in 2007<sup>75</sup>, which consists in manipulating Eqs. 24, 23 and 15 to write the relation

$$C_T = \frac{C_{bp}}{\gamma} \Theta + \frac{\Theta(\gamma - N\Theta + \Theta)^{N-1}}{K_i(\gamma - N\Theta)^N}. \quad (25)$$

Such approach allows one to directly fit the contour length data without any approximation: one should just plot the total ligand concentration  $C_T$  in the y-axis and the relative increase of the contour length  $\Theta$  in the x-axis, such that Eq. 25 can be used directly to fit the experimental data. In Fig. 2 we show an example of such fitting, performed originally in ref.<sup>76</sup> for DNA complexes formed with the intercalator diaminobenzidine. Other examples can be found in refs.<sup>75</sup> and<sup>133</sup> for DNA complexes formed with the intercalators ethidium bromide and psoralen, respectively.

One should note, however, that the contour length approaches discussed above can only be used for intercalators. In fact, in general only intercalators increase the DNA contour length when binding<sup>56,73,137</sup>. An exception are ligands that facilitate or inhibit base pair formation, which can also be studied by contour length approaches similar to those discussed above, using length changes as a marker of ligand binding<sup>176</sup>. The other common types of interactions between DNA and ligands, such as groove binding, electrostatic interaction

or covalent binding in general do not affect the DNA contour length. In some cases, however, these kinds of interactions can cause DNA compaction with a decrease of the “apparent contour length” measured by force spectroscopy in the low-force regime<sup>65,97,98</sup>. The concept of “apparent contour length” arises from the fact that, if the DNA molecule is partially compacted due to ligand binding, small forces in the entropic regime usually are not sufficient to fully stretch the molecule, and therefore the measured contour length will be smaller than the real one. Depending on the type of interaction, even high forces cannot be used to fully stretch the complexes and estimate the real contour length by fitting the WLC model<sup>177</sup>. The decrease of the “apparent contour length” upon increasing of the bound ligand concentration in general depends on intricate effects such as the positional correlation of bound ligands. This fact makes it difficult to directly link the contour length data to a binding isotherm, although other kinds of analyses can be performed to study such interactions.

On the other hand, the other basic mechanical property (the persistence length) is much more sensitive to different types of interactions, and in general changes for covalent binding<sup>97,98,126,133</sup>, intercalative binding<sup>56,75–77,103,140,141</sup> and groove/electrostatic binding<sup>19,56,64,65,91,92,156</sup>. This fact turns the persistence length into the ideal mechanical property to be chosen for monitoring DNA-ligand interactions and to be connected with the physical chemistry of such interactions. Nevertheless, such connection is not straightforward as the one performed for the contour length of DNA-intercalator complexes.

Only in 1998 the first attempt to connect the persistence length to physicochemical properties was performed by Rouzina and Bloomfield<sup>156</sup>, which can be synthesized in Eq. 20 presented earlier. This model however was derived in the context of electrostatic interactions and attempt to explain the changes of the persistence length due to the negative charge neutralization in the DNA phosphate backbone<sup>156</sup>. Recently, this model has been used to fit experimental data of DNA complexes formed with positively charged proteins such as HMG, HMGB1 and HMGB2<sup>91,92</sup>, with excellent agreement.

The question now is: Can the changes of the persistence length be related to physicochemical parameters for any type of interaction? In the next section we discuss an approach that can be used to perform such task.

## 6.2 A general model to connect the persistence length to physical chemistry

The DNA molecule partially covered by ligand molecules along its structure can be thought as an association of entropic springs in series. One type of spring is the bare DNA with its natural persistence length  $A_0$ , corresponding to the regions without bound ligands along the contour length of

the molecule. The other type(s) of spring(s) is(are) the local complex(es) formed between DNA and the bound ligand molecules. A simple phenomenological model to study the persistence length of DNA-ligand complexes that uses this assumption was proposed by Rocha<sup>147</sup>. Latter, it was rigorously demonstrated<sup>63</sup> that a series association of  $n$  entropic springs with persistence lengths  $A_0, A_1, A_2, \dots, A_{n-1}$  results in an effective entropic spring with the effective persistence length  $A_E$  given by

$$\frac{1}{A_E} = \frac{f_0(r)}{A_0} + \frac{f_1(r)}{A_1} + \frac{f_2(r)}{A_2} + \dots, \quad (26)$$

where  $f_0(r), f_1(r), f_2(r)$ , etc are specific functions of the bound ligand fraction  $r$ .

The function  $f_i(r)$  is in fact the probability of finding an entropic spring (a part of the DNA molecule) along the contour length with a local persistence length  $A_i$ <sup>63</sup>, which depends on the bound site fraction  $r$ <sup>63,66</sup>.

In general, the approach proposed in Eq. 26 can be applied by following three steps:

(a) One needs firstly to find the probability distribution of the bound ligands, *i. e.*, the set of functions  $f_i(r)$ .

(b) The second step is to choose an adequate binding isotherm that captures the physical chemistry of the system, and then plug such isotherm in Eq. 26 via the parameter  $r$ .

(c) Finally, the third step is to use the equation constructed in step (b) to fit the experimental data of the persistence length, extracting the physicochemical parameters contained in the binding isotherm and the set of local persistence lengths  $A_i$ 's.

To deduce the probability distribution mentioned in step (a), the easiest way is firstly identify how many different entropic springs one needs in the model to correct reproduce the experimental behavior of the persistence length as a function of ligand concentration. The simplest behavior of this parameter reported in the literature is a monotonic decay, found for example for the proteins HMG, HMGB1 and HMGB2<sup>91,92</sup> and for the drug cisplatin<sup>97,98</sup>. This relatively simple behavior of the persistence length can be explained with a model consisted only by two entropic springs, one representing the bare DNA (local persistence length  $A_0$ ) and the other representing the local structure formed by the ligand molecule bound to the DNA (local persistence length  $A_1$ ).

From now on let us consider a “site” the place effectively occupied by a single ligand molecule (or by a single bound cluster of molecules, in the cases in which the ligands bind to DNA forming clusters due to high positive cooperativity). One should note that even for single-ligand binding the sites are usually larger than one DNA base-pair, due to ligand size and/or neighbor-exclusion effects. A model with only two different types of entropic springs, as proposed in the last paragraph, is a *one-site quenched disorder statistical model*, since the probability distribution depends only on the occupancy of

single sites along the double-helix, *i. e.*, it does not depend on the correlation with the occupancy of nearest neighbor sites.

Consider now a particular site choose randomly along the DNA. The probability of this site to be occupied by a ligand molecule is  $x = r/r_{max}$ , with a local persistence length  $A_1$ ; and the probability of this site to be unoccupied is  $1 - x$ , with a local persistence length  $A_0$ <sup>63,66</sup>. The effective persistence length can then be written as

$$\frac{1}{A_E} = \frac{1-x}{A_0} + \frac{x}{A_1}, \quad (27)$$

and  $x = r/r_{max}$  can be directly connected to a binding isotherm.

Equation 27 was recently used by Crisafuli *et al.* to determine the physicochemical parameters of the DNA-cisplatin interaction from the persistence length data of these complexes<sup>97,98</sup>. One should observe that, since the exclusion number  $N$  is related to the saturated bound ligand fraction  $r_{max}$  by  $N = 1/r_{max}$ , the electrostatic model proposed by Rouzina and Bloomfield<sup>156</sup> (Eq. 20) is a particular case of Eq. 27 (it is just a matter of redefining the physical interpretation of the constants  $A_i$ 's).

There are other types of ligands that can induce a more intricate non-monotonic behavior for the persistence length as a function of the ligand concentration. Probably the most known example is the bacterial protein HU<sup>178</sup>, but some drugs such as cationic cyclodextrins<sup>63</sup>, actinomycin D<sup>64</sup> and hoechst 33258<sup>65</sup> also induce such behavior. To account for the persistence length changes of the DNA complexes formed with these compounds, the one-site model discussed above does not work, and one needs to introduce at least one more entropic spring with other local persistence length ( $A_2$ ), *i. e.*, one needs a *two-sites quenched disorder statistical model*, in which one must consider the probabilities associated with the occupancy of two nearest sites. In the context of a two sites model, there are therefore the following probabilities associated to the local persistence lengths: (a) two nearest sites unoccupied have local persistence length  $A_0$  and probability  $P_0 = (1-x)^2$ . (b) Two nearest sites simultaneous occupied have local persistence length  $A_2$  and probability  $P_2 = x^2$ . (c) Finally, one site unoccupied and the neighbor occupied have local persistence length  $A_1$  and probability  $P_1 = 1 - P_0 - P_2 = 2x(1-x)$ . The effective persistence length can therefore be written as

$$\frac{1}{A_E} = \frac{(1-x)^2}{A_0} + \frac{2x(1-x)}{A_1} + \frac{x^2}{A_2}, \quad (28)$$

and  $x = r/r_{max}$  can be connected to a binding isotherm as usual.

A last issue must be solved to complete the problem, both for monotonic and non-monotonic behaviors of the persistence length: one must write the binding isotherm as a function of a directly accessible parameter instead of  $C_f$ , as discussed in Section 6.1, in order to eliminate the dependence in using other experimental techniques to estimate the ligand

partitioning between the DNA ( $C_b$ ) and the solution ( $C_f$ ). Although we have discussed some approaches to perform this task for intercalators in Section 6.1, it is clear that a general approach is needed in order to contemplate the order types of ligands.

In 2012 Siman *et al.* have firstly proposed a simple iterative solution of the binding isotherm<sup>63</sup>, which was promptly generalized by Cesconetto *et al.* in 2013 with the following method. Firstly choose a particular binding isotherm, for example, the Hill binding isotherm (Eq. 14). One can plug the relations  $x = r/r_{max}$  and  $C_f = C_T - rC_{bp} = C_T - r_{max}C_{bp}x$  in this binding isotherm to write

$$x = \frac{[K_i(C_T - r_{max}C_{bp}x)]^n}{1 + [K_i(C_T - r_{max}C_{bp}x)]^n}. \quad (29)$$

Observe that this equation can be solved numerically for known values of the constants, returning  $x$  for each value of  $C_T$ . Therefore, one needs to write a simple algorithm that uses a subroutine to solve Eq. 29 for initial guessed values of the constants, and uses the results returned for  $x$  plugged into Eq. 28 or Eq. 27 to fit the experimental data of the persistence length  $A$  as a function of  $C_T$ , by using least squares fitting. With this approach the problem is completely solved. Observe that any binding isotherm can be used to get an equation similar to Eq. 29, *i. e.*, one needs only to choose a plausible binding isotherm that captures the physical chemistry of the interaction.

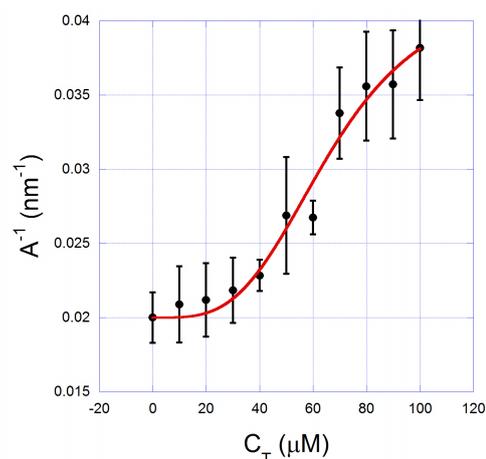
Below we revisit some results recently obtained with this approach, showing that in principle it can be used to study any type of interaction. The only requisite is that such interaction changes the DNA persistence length as the ligand binds. All the experimental data were obtained by single molecule stretching performed with optical tweezers in the entropic regime, with fittings similar to that shown in Fig. 1 (except those of Fig. 5 - see ref.<sup>178</sup>). We also discuss the main features of the physics and chemistry of the interactions revisited here. A complete detailed discussion can be found in the original articles (and the references therein). It is worth to remember that all optical tweezers measurements were performed in chemical equilibrium, waiting sufficient time for ligand equilibration before performing the stretching experiments. In addition, these measurements were performed with low forces ( $< 2$  pN) and pulling rates ( $\sim 0.1$   $\mu\text{m/s}$ ) in order to guarantee that the chemical equilibrium is not affected by the stretching forces.

In Fig. 3 we show the experimental data (*circles*) of the persistence length of DNA-cisplatin complexes as a function of drug total concentration in the sample  $C_T$ . Observe that, for convenience to fit with Eq. 27, we have plotted the inverse of the persistence length in this figure and in all subsequent ones. The fitting with the model (Eq. 27) is also shown (*red*

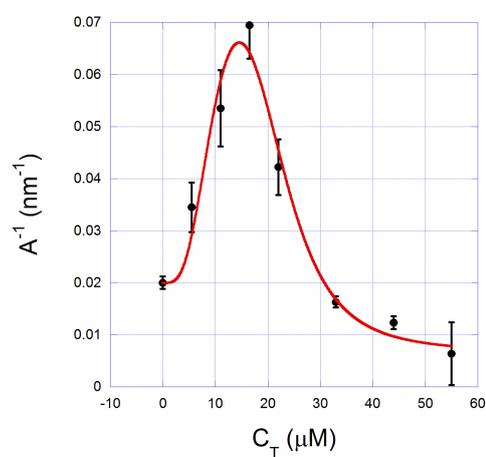
*solid line*). In this case we have used the Hill binding isotherm (Eq. 14) to perform the fitting, extracting the physicochemical parameters  $K_i = (1.6 \pm 0.2) \times 10^4 \text{ M}^{-1}$ ,  $n = 3.6 \pm 0.4$ ,  $r_{max} = 0.56 \pm 0.06$  and  $A_1 = (24 \pm 4) \text{ nm}$ . These results agree very well to those presented in ref.<sup>97</sup>, which were obtained using another fitting strategy, and as well as to results obtained from other experimental techniques<sup>179,180</sup>. In particular, the Hill exponent  $n = 3.6$  indicates that cisplatin presents positive cooperativity in its interaction with DNA.

Cisplatin and its analogues carboplatin and oxaliplatin form one of the most important class of compounds used in cancer chemotherapies, especially to treat head, neck, testicular, ovarian and non-small cell lung cancers<sup>181</sup>. In aqueous solution, two chloride ions dissociate from the compound, followed by incorporation of two water molecules. This is the active state of the drug, which can bind to DNA<sup>182</sup>. Many aspects of the DNA-cisplatin interaction are currently well established in the literature, such as the mechanism of action of the compound as an anticancer drug, which consists in damaging the DNA molecule with adducts that form interstrand and intrastrand crosslinks<sup>183</sup>. These crosslinks hinder DNA replication by introducing strong structural perturbations on the double-helix such as bendings, partial unwinding and loops<sup>126,127,183</sup>. These structural perturbations are closely related to the result found for the Hill exponent ( $n \sim 3.6$ ). In fact, a positive cooperativity could be expected in DNA-cisplatin interaction, since the crosslinks and loops induced in the DNA by the drug approximate different strand segments as the drug concentration is increased, therefore increasing the probability of forming even more crosslinks and loops as cisplatin binds<sup>97,98</sup>. A nearly similar mechanism was recently observed for the H-NS binding protein by Dame *et al.*, which have shown that a cooperative behavior in this case arises as an intrinsic property of DNA bridging due to duplex proximity<sup>184</sup>.

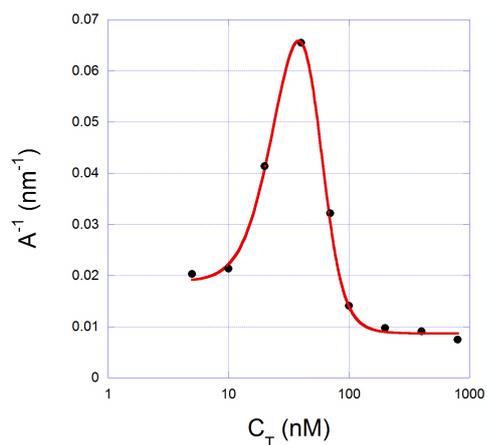
In Fig. 4 we show the experimental data (*circles*) for the inverse of the persistence length of DNA complexes formed with a monovalent cationic  $\beta$ -cyclodextrin (6-mono-deoxy-6-monoamine- $\beta$ -cyclodextrin) as a function of drug total concentration in the sample  $C_T$ , firstly presented in ref.<sup>63</sup>. Cyclodextrins (CDs) are cyclic oligosaccharides composed of D-glucose units joined by glucosidic linkages. The  $\beta$  subtype consists of seven units and has a structure that resembles a truncated cone, with hydroxyl groups localized at the outer surface of the cone. That gives CDs the property to be water-soluble and to have a relatively hydrophobic inner cavity able to partially or entirely accommodate polymers forming host-guest inclusion complexes<sup>185</sup>. Monovalent cationic  $\beta$ -cyclodextrin is usually obtained by substituting one of the hydroxyl groups by an amino group. This molecule has been used for condensing DNA and introducing it into small vesicles for gene therapy applications<sup>186</sup>.



**Fig. 3** *Circles*: inverse of the persistence length of DNA-cisplatin complexes measured by single molecule stretching experiments. *Red solid line*: a fitting to the model (Eq. 27) using the Hill binding isotherm (Eq. 14). From this fitting we have found the physicochemical parameters  $K_i = (1.6 \pm 0.2) \times 10^4 \text{ M}^{-1}$ ,  $n = 3.6 \pm 0.4$ ,  $r_{max} = 0.56 \pm 0.06$  and  $A_1 = (24 \pm 4) \text{ nm}$ . For this data  $C_{bp} = 8.9 \mu\text{M}$ .



**Fig. 4** *Circles*: inverse of the persistence length of DNA-cyclodextrin complexes measured by single molecule stretching experiments. *Red solid line*: a fitting to the model (Eq. 28) using the Hill binding isotherm (Eq. 14). From this fitting we have found the physicochemical parameters  $K_i = (9 \pm 1) \times 10^4 \text{ M}^{-1}$ ,  $n = 3.7 \pm 0.4$ ,  $A_1 = 8.4 \pm 1 \text{ nm}$  and  $A_2 = 149 \pm 21 \text{ nm}$ . For this data  $C_{bp} = 11 \mu\text{M}$ .



**Fig. 5** Circles: inverse of the persistence length of DNA-HU complexes measured by single molecule stretching experiments (experimental data by van Noort *et al.*<sup>178</sup>, error bars are not available in this case). Red solid line: a fitting to the model (Eq. 28) using the Hill binding isotherm (Eq. 14). From this fitting we have found the physicochemical parameters  $K_i = (3.4 \pm 0.4) \times 10^7 \text{ M}^{-1}$ ,  $n = 3.6 \pm 0.3$ ,  $A_1 = (8.5 \pm 1) \text{ nm}$  and  $A_2 = (115 \pm 13) \text{ nm}$ . For this data  $C_{bp}$  is unknown due to the sample preparation procedure<sup>63</sup>. It was left as an adjustable parameter, and the fitting returns  $C_{bp} \sim 110 \text{ nM}$ .

Observe in Fig. 4 that for the complexes formed between DNA and cationic  $\beta$ -cyclodextrin the persistence length exhibits a non-monotonic behavior, and therefore we have used Eq. 28 to fit the data, together with the Hill binding isotherm (red solid line). We have found the results  $K_i = (9 \pm 1) \times 10^4 \text{ M}^{-1}$ ,  $n = 3.7 \pm 0.4$ ,  $A_1 = (8.4 \pm 1) \text{ nm}$  and  $A_2 = (149 \pm 21) \text{ nm}$ . The parameter  $r_{max} = 0.67$  was known for this ligand such that we have fixed its value in the fitting<sup>63</sup>. The value obtained for the Hill exponent  $n$  again indicates that the system is positively cooperative, in this case forming bound clusters of  $\sim 4$  drug molecules at the binding sites<sup>63</sup>.

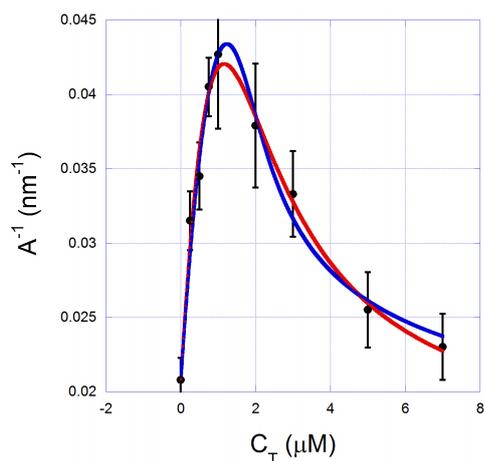
The results obtained for the bacterial protein HU in ref.<sup>63</sup> are somewhat similar, as shown in Fig. 5. We have used again Eq. 28 and the Hill binding isotherm to perform the fitting (red solid line), and the experimental data (circles) were obtained by van Noort *et al.* for this ligand<sup>178</sup>. From the fitting we have obtained the results  $K_i = (3.4 \pm 0.4) \times 10^7 \text{ M}^{-1}$ ,  $n = 3.6 \pm 0.3$ ,  $A_1 = (8.5 \pm 1) \text{ nm}$  and  $A_2 = (115 \pm 13) \text{ nm}$ . Here again,  $r_{max} = 0.11$  is a known parameter and was maintained fixed in the fitting<sup>63</sup>. For this data  $C_{bp}$  is unknown due to the sample preparation procedure<sup>63</sup>. It was left as an adjustable parameter, and the fitting returns  $C_{bp} \sim 110 \text{ nM}$ . Here the fact that the persistence length increases for high protein concentrations agrees with results obtained in AFM images, which have shown the formation of rigid filaments<sup>178</sup>.

At this point it is necessary to reflect on the use of the Hill

binding isotherm in the analysis of DNA-ligand systems. The fact that we have found a Hill exponent  $n \gg 1$  for cisplatin, cyclodextrin and HU strongly indicates that relevant positive cooperativity is present in such systems. In fact, binding isotherms with no cooperativity such as the Scatchard model (Eq. 12) or the basic Neighbor Exclusion Model (NEM) (Eq. 15) do not work in performing these fittings. The cooperative version of the neighbor exclusion model (Eq. 16) in principle could be used, but we were not successful in performing the fitting in an easy way, founding numerical problems in solving the equation analogue to Eq. 29 for this binding isotherm. In fact, the intricacy of Eq. 16 somewhat limits its applicability in the fitting approaches discussed in this review. For this reason we use the Hill binding isotherm, a much simpler equation that also takes into account cooperativity effects.

The first example of a non-cooperative system studied with our approach are the DNA complexes formed with the drug Actinomycin D (ActD), firstly presented in ref.<sup>64</sup>. This drug is a DNA ligand clinically used as an antibiotic and to treat some highly malignant cancers, such as gestational trophoblastic disease<sup>187</sup>, Wilms' tumor<sup>188</sup> and rhabdomyosarcoma<sup>189</sup>. The drug exhibits a complex interaction with double-strand DNA, presenting two distinct parts which bind to DNA by different modes: while the phenoxazine ring intercalates, preferentially at the CG base pairs, the cyclic pentapeptide chains bind to the minor groove, usually forming hydrogen bonds with the guanine bases<sup>161-164</sup>.

Here we clearly have the option of choosing different binding isotherms to perform the fitting. This fact illustrates the versatility of our approach, which returns consistent results even for different binding isotherms: it is required only to choose one that captures the basic physical chemistry of the system. In fact, if the DNA-ActD interaction is non-cooperative<sup>64</sup>, one can choose the Scatchard model or the basic (non-cooperative) neighbor exclusion model. Nevertheless, instead of the first option (Scatchard), we have chosen the Hill binding isotherm to perform the fitting. If everything is right, one should find a Hill exponent near unity ( $n \sim 1$ ), since in this case the Hill model is just equivalent to the Scatchard one. Figure 6 shows the experimental data points (circles) and the fittings to Eq. 28 with the Hill model (red solid line) and with the neighbor exclusion model (blue solid line). From the first fitting (Hill), we find  $K_i = (1.5 \pm 0.4) \times 10^6 \text{ M}^{-1}$ ,  $n = 1.1 \pm 0.2$ ,  $r_{max} = 0.11 \pm 0.01$ ,  $A_1 = (15.2 \pm 0.6) \text{ nm}$  and  $A_2 = (64 \pm 25) \text{ nm}$ . From the second fitting (NEM) we find  $K_i = (4.6 \pm 0.5) \times 10^6 \text{ M}^{-1}$ ,  $N = 4 \pm 0.5$  (the exclusion number for each bound ActD),  $A_1 = (14 \pm 2) \text{ nm}$  and  $A_2 = (140 \pm 16) \text{ nm}$ . Observe that both fittings explain well the behavior of the experimental data. The results returned for the physicochemical parameters, although somewhat dependent on the chosen binding isotherm, are realist. The relatively high variability on the values found for some of these parameters is compati-



**Fig. 6** Circles: inverse of the persistence length of DNA-ActD complexes measured by single molecule stretching experiments. Red solid line: a fitting to the model (Eq. 28) using the Hill binding isotherm (Eq. 14). From this fitting we have found the physicochemical parameters  $K_i = (1.5 \pm 0.4) \times 10^6 \text{ M}^{-1}$ ,  $n = 1.1 \pm 0.2$ ,  $r_{max} = 0.11 \pm 0.01$ ,  $A_1 = (15.2 \pm 0.6) \text{ nm}$  and  $A_2 = (64 \pm 25) \text{ nm}$ . Blue solid line: a fitting to the model (Eq. 28) using the NEM binding isotherm (Eq. 15). From this fitting we have found the physicochemical parameters  $K_i = (4.6 \pm 0.5) \times 10^6 \text{ M}^{-1}$ ,  $N = 4 \pm 0.5$  (the exclusion number for each bound ActD),  $A_1 = (14 \pm 2) \text{ nm}$  and  $A_2 = (140 \pm 16) \text{ nm}$ . For this data  $C_{bp} = 10.6 \mu\text{M}$ .

ble to the variability found when using different experimental techniques<sup>53,190–192</sup>.

A relevant question that can be raised at this point is about the accuracy of our approach to treat systems with more than one binding mode, *i. e.*, with two or more different sets of physicochemical parameters. Many compounds interact with DNA in this way, and recently we have successfully applied our fitting approach to the fluorescent dye Hoechst 33258<sup>65</sup>.

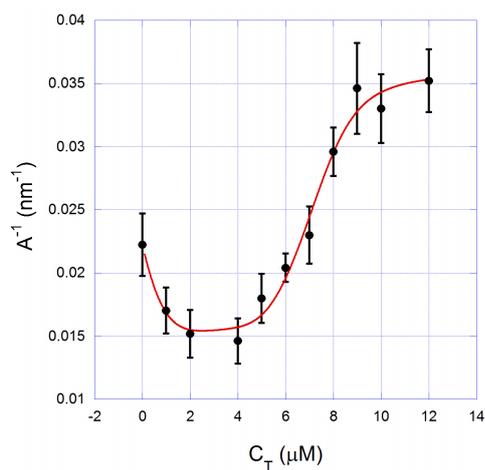
The Hoechst stains, also known as bis-benzimides, are a family of fluorescent dyes largely employed to stain the DNA molecule in molecular biology applications, allowing one to visualize DNA with fluorescence microscopy. In addition, these compounds can be potentially used as anticancer drugs<sup>193</sup>, since their strong interaction with DNA can impede the replication of the molecule. Many experimental techniques were employed over the past years to study the effects of the Hoechst 33258 subtype on the DNA molecule. In particular, it was found that the ligand binds preferentially to the DNA minor groove, especially at AT-rich regions<sup>194,195</sup>. Nevertheless, some authors have proposed that the ligand presents more than one binding mode to double-strand (ds) DNA<sup>195–197</sup>, indicating the possibility of intercalation at GC-rich regions<sup>195,198</sup>.

With our fitting approach we were able to decouple the

two main binding modes that Hoechst 33258 exhibits with DNA, by using a binding isotherm expressed as a sum of two Hill processes. We have determined the two complete sets of physicochemical parameters for each of the binding modes. In particular, we have found that the first binding mode (intercalation) is non-cooperative, with a Hill exponent  $\sim 1$ , while the second mode (groove binding) is highly positively cooperative, with a Hill exponent  $\sim 7$ . Such conclusion is in agreement with previous studies performed by other techniques (equilibrium dialysis and absorption spectroscopy)<sup>196,197</sup>. The two binding modes coexist in the entire concentration range studied here, but intercalation is dominant for  $C_T < 3 \mu\text{M}$  while groove binding is dominant for higher concentrations. Figure 7 shows the experimental data (circles) and the fitting (red solid line). We have found that, for the intercalative binding mode,  $K_1 = (1.8 \pm 0.4) \times 10^6 \text{ M}^{-1}$ ,  $n_1 = 1.1 \pm 0.3$ . On the other hand, for minor groove binding, we found  $K_2 = (2.4 \pm 0.2) \times 10^5 \text{ M}^{-1}$ ,  $n_2 = 7 \pm 3$ . Also, we found  $A_1 = (300 \pm 100) \text{ nm}$  and  $r_{max} = 0.32 \pm 0.02$ , which are global parameters independent of the binding mode. The parameter  $A_2 = 28.3 \text{ nm}$  was maintained fixed in the fitting since it is the saturation value of the persistence length, which can be directly determined from the data of Fig. 7 in this case. The error bars of the parameters obtained in this fitting are somewhat higher than those obtained for the other DNA-ligand systems presented above. This fact is due to the excess of adjustable parameters used in the fitting procedure in this case, because we have two different binding modes and consequently two sets of binding parameters.

Finally we show an example of how our fitting approach can also be used to analyze the contour length data of DNA complexes formed with ligands. GelRed is a fluorescent nucleic acid stain designed with the purpose of replacing the highly toxic ethidium bromide (EtBr) in gel electrophoresis and other experimental techniques which depends on the fluorescence of stained DNA. When bound to DNA, GelRed has the same absorption and emission spectra of EtBr and, according to its manufacturer (Biotium Inc., Hayward, CA, USA), it has the advantage of being much less toxic and mutagenic<sup>199,200</sup>.

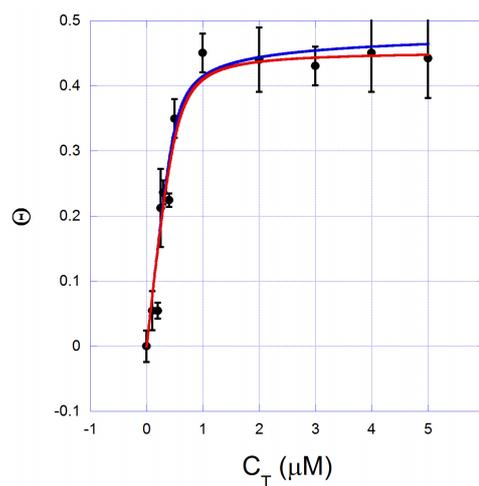
Figure 8 shows the experimental data of the relative increase of the contour length  $\Theta = (L - L_0)/L_0$  (circles), firstly presented in ref.<sup>77</sup>, and two fittings performed with Eq. 23 and two different binding isotherms: Scatchard (red solid line) and NEM (blue solid line). The two fittings are similar, returning equivalent physicochemical parameters and allowing one to conclude that GelRed interacts with DNA by bis-intercalation<sup>77</sup>. For the Scatchard fitting, we have found  $K_i = (1.8 \pm 0.4) \times 10^7 \text{ M}^{-1}$ ,  $r_{max} = 0.22 \pm 0.03$  and  $\gamma = 2.2 \pm 0.1$ . For the NEM fitting, we found  $K_i = (1.8 \pm 0.3) \times 10^7 \text{ M}^{-1}$ ,  $N = 3.7 \pm 0.4$  and  $\gamma = 1.9 \pm 0.1$ . The results obtained for this ligand with our fitting approach lead us to conclude that the GelRed dye is a bis-intercalator. In fact, the exclusion parameter  $N = 1/r_{max}$



**Fig. 7** Circles: inverse of the persistence length of DNA-hoechst complexes measured by single molecule stretching experiments. Red solid line: a fitting to the model (Eq. 28) using a sum of two Hill processes as the binding isotherm. From this fitting we decouple the two binding modes and find the physicochemical parameters  $K_1 = (1.8 \pm 0.4) \times 10^6 \text{ M}^{-1}$ ,  $n_1 = 1.1 \pm 0.3$ ,  $K_2 = (2.4 \pm 0.2) \times 10^5 \text{ M}^{-1}$ ,  $n_2 = 7 \pm 3$ ,  $A_1 = (300 \pm 100) \text{ nm}$  and  $r_{max} = 0.32 \pm 0.02$ . For this data  $C_{bp} = 20 \text{ } \mu\text{M}$ .

indicates that each bound GelRed molecule effectively occupies  $\sim 4$  DNA base-pairs, a value considerably higher than the results found for most monointercalators, and approximately twice the result for EtBr (which is  $\sim 2^{73-75}$ ). The equilibrium association constant  $K_i$  is also higher than the result obtained for typical monointercalators ( $\sim 10^5 \text{ M}^{-1}$ )<sup>73-75,147</sup>, and within the range found for most bis-intercalators ( $10^7$  to  $10^9 \text{ M}^{-1}$ )<sup>145,201-204</sup>. Finally, the result  $\gamma \sim 2$  is approximately twice the value obtained for typical monointercalators, suggesting that each bound GelRed molecule increases the DNA contour length by  $\sim 0.68 \text{ nm}$ , a result also compatible to typical bis-intercalators<sup>201,203</sup>. Observe that the bis-intercalators should increase approximately twice the DNA contour length per bound molecule, since each ligand molecule contains two intercalating portions.

In summary, we have presented many examples of DNA-ligand systems analyzed with the proposed fitting approach. All the results obtained for the physicochemical parameters are consistent with most studies found in the literature that have used many different experimental techniques, from crystallography to fluorescence resonance energy transfer<sup>53,179,190-192,196,197,205-207</sup>. Thus, our fitting approach allows a direct comparison between the results obtained from single molecule stretching experiments to those obtained from typical ensemble-averaging techniques, which are usually used to characterize the physical chemistry of DNA-ligand interactions. A weakness of the presented approach that can be



**Fig. 8** Circles: experimental data of the relative increase of the contour length  $\Theta = (L - L_0)/L_0$  of DNA-GelRed complexes. Red solid line: a fitting to Eq. 23 using the Scatchard binding isotherm (Eq. 13), from which we have found the physicochemical parameters  $K_i = (1.8 \pm 0.4) \times 10^7 \text{ M}^{-1}$ ,  $r_{max} = 0.22 \pm 0.03$  and  $\gamma = 2.2 \pm 0.1$ . Blue solid line: a fitting to Eq. 23 using the NEM binding isotherm (Eq. 15), from which we have found the physicochemical parameters  $K_i = (1.8 \pm 0.3) \times 10^7 \text{ M}^{-1}$ ,  $N = 3.7 \pm 0.4$  and  $\gamma = 1.9 \pm 0.1$ . For this data  $C_{bp} = 2.4 \text{ } \mu\text{M}$ .

pointed concerns the error bars of the physicochemical parameters obtained from the fitting procedure, which can be a bit high if the number of adjustable parameters used in the fitting is high, as in the case of multiple binding modes. Nevertheless in these cases one can maintain fixed some parameters previously measured with other techniques and perform the fitting using only the adjustable parameters of interest. In this way, the fitting approach can be used to compare and verify results obtained from very different experimental techniques, and can still be useful in the investigation of DNA-ligand interactions.

## 7 Conclusions

We have reviewed important topics of the field “DNA-ligand interactions”, from DNA mechanics to DNA-ligand physical chemistry, emphasizing how one can connect the changes of the mechanical properties of DNA induced by the binding ligand to the physicochemical information of such interaction. This type of connection is extremely relevant because it allows one to perform a robust characterization of the interaction both from the point of view of the mechanical properties and of the physical chemistry of the interaction by using only one experimental technique: single molecule stretching experiments. Moreover, the possibility of performing such connection reduces the time and cost required for getting results for

a DNA-ligand system, since less different equipments are required and the number of experiments that must be conducted can be considerably reduced. Furthermore, and more important, it opens the possibility of comparing the results obtained by means of very different experimental techniques, in special when comparing single molecule techniques to ensemble-averaging techniques.

In particular, we reviewed a fitting approach recently proposed by our group to connect the persistence length of the DNA-ligand complexes to the physical chemistry of the interaction. Such approach in principle can be used for any type of ligand, from drugs to proteins, even if there are multiple binding modes. However, a test with sequence-specific ligands<sup>208</sup> is still needed. In any case, the only requisite to try the approach is that the interaction must change the DNA persistence length as the ligand binds, which usually occurs for all types of common interactions (intercalation, covalent binding, electrostatic driven interactions and groove binding) at least for some ligand concentration range.

Finally, among the future perspectives of the field we can cite: (a) Apply the current models to other types of ligands not yet explored (sequence-specific ligands, complex proteins, etc). (b) Evaluate in detail the effects of the forces applied to perform single molecule stretching on the values obtained for the mechanical parameters, and how these forces affect the efficiency of the current model in extracting physico-chemical information of the interaction. In fact, some DNA-ligand complexes have mechanical properties that are force-dependent<sup>148</sup>. Therefore, experiments performed in the enthalpic regime, in which one usually apply forces as high as tens of picoNewtons, can in principle return different results for the mechanical properties. Thus, an open question is how the current approach will work in these situations. (c) Extend the basic ideas of this type of modeling to other mechanical parameters, for example, the stretch modulus, which can help in investigating the effects of higher applied forces, as cited above.

## 8 Acknowledgements

This work was supported by the Brazilian agencies: Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). The author also thanks the collaborators and students that have contributed to many of the presented results: E. B. Ramos, O. N. Mesquita, F. A. P. Crisafuli, E. F. Silva, E. C. Cesconetto and L. Siman.

## References

- 1 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 964–967.
- 2 B. Alberts, A. Johnson, J. Lewis, M. Raff and K. Roberts, *Molecular Biology of the Cell*, Garland Science, New York, 5th edn, 2007.
- 3 C. Bustamante, J. F. Marko, E. D. Siggia and S. Smith, *Science*, 1994, **265**, 1599–1600.
- 4 J. F. Marko and E. D. Siggia, *Macromolecules*, 1995, **28**, 8759–8770.
- 5 T. Odijk, *Macromolecules*, 1995, **28**, 7016–7018.
- 6 M. D. Wang, H. Yin, R. Landick, J. Gelles and S. M. Block, *Biophys. J.*, 1997, **72**, 1335–1346.
- 7 P. G. de Gennes, *Scaling concepts in Polymer Physics*, Cornell University Press, Ithaca, 1st edn, 1979.
- 8 A. Bates and A. Maxwell, *DNA Topology*, Oxford University Press, Oxford, 2nd edn, 2005.
- 9 A. Vologodskii, *Statistical-mechanical analysis of enzymatic topological transformations in DNA molecules. Mathematics of DNA Structure, Function and Interactions. The IMA volumes in Mathematics and its Applications*, Springer, New York, 2009.
- 10 A. Vologodskii, *Nucl. Acids Res.*, 2009, **37**, 3125–3133.
- 11 L. H. Hurley, *Nature Rev. Cancer*, 2002, **2**, 188–200.
- 12 C. D. Scripture and W. D. Figg, *Nature Rev. Cancer*, 2006, **6**, 546–558.
- 13 S.-D. Li and L. Huang, *Gene Therapy*, 2006, **13**, 1313–1319.
- 14 C. Sheridan, *Nature Biotech.*, 2011, **29**, 121–128.
- 15 H. G. Hansma, R. Golan, W. Hsieh, C. P. Lollo, P. Mullen-Ley and D. Kwok, *Nucl. Acids Res.*, 1998, **26**, 2481–2487.
- 16 M. M. O. Sullivan, J. J. Green and T. M. Przybycien, *Gene Therapy*, 2003, **10**, 1882–1890.
- 17 This information can be promptly verified in any citation database of peer-reviewed literature, such as Scopus<sup>TM</sup> ([www.scopus.com](http://www.scopus.com)) or Web of Science<sup>TM</sup> (<http://webofknowledge.com/>).
- 18 T. Odijk, *J. Polym. Sci.*, 1977, **15**, 477–483.
- 19 G. S. Manning, *Biophys. J.*, 2006, **91**, 3607–3616.
- 20 C. G. Baumann, S. B. Smith, V. A. Bloomfield and C. Bustamante, *Proc. Natl. Acad. Sci. USA.*, 1997, **94**, 6185–6190.
- 21 J. R. Wenner, M. C. Williams, I. Rouzina and V. A. Bloomfield, *Biophys. J.*, 2002, **82**, 3160–3169.
- 22 T. Strick, J.-F. Allemand, D. Bensimon, R. Lavery and V. Croquette, *Physica A*, 1999, **263**, 392–404.
- 23 S. B. Smith, L. Finzi and C. Bustamante, *Science*, 1992, **258**, 1122–1126.
- 24 D. Boal, *Mechanics of the cell*, Cambridge University Press, 2002.
- 25 A. Vologodskii and M. D. Frank-Kamenetskii, *Nucl. Acids Res.*, 2013, **41**, 6785–6792.
- 26 A. Vologodskii, Q. Du and M. Frank-Kamenetskii, *Artif DNA PNA XNA*, 2013, **4**, 1–3.
- 27 S. Geggier and A. Vologodskii, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 15421–15426.
- 28 J. A. Schellman, *Biopolymers*, 1974, **13**, 217–226.
- 29 P. A. Wiggins and P. C. Nelson, *Phys. Rev. E*, 2006, **73**, Art. No. 031906.
- 30 P. A. Wiggins, T. van der Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker and P. C. Nelson, *Nat. Nanotech.*, 2006, **1**, 137–141.
- 31 C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block and V. Croquette, *Biophys. J.*, 1999, **76**, 409–413.
- 32 M. D. Frank-Kamenetskii and S. Prakash, *Phys. Life Rev.*, 2014, **11**, 153–170.
- 33 S. Geggier, A. Kotlyar and A. Vologodskii, *Nucl. Acids Res.*, 2011, **39**, 1419–1426.
- 34 M. D. Frank-Kamenetskii, *Advances in Sensing with Security Applications. NATO Science Series II: Mathematics, Physics and Chemistry.*, 2006, **218**, 295–326.

- 35 M. D. Frank-Kamenetskii, *Mol. Biol.*, 2002, **36**, 232–235.
- 36 H. Clausen-Schaumann, M. Rief, C. Tolksdorf and H. E. Gaub, *Biophys. J.*, 2000, **78**, 1997–2007.
- 37 M. Rief, H. Clausen-Schaumann and H. E. Gaub, *Nat. Struct. Mol. Biol.*, 1999, **6**, 346–349.
- 38 M. C. Williams, I. Rouzina and M. J. McCauley, *Proc. Natl. Acad. Sci. USA.*, 2009, **106**, 18047–8.
- 39 T. Strick, J.-F. Allemand, V. Croquette and D. Bensimon, *Prog. Biophys. Mol. Bio.*, 2000, **74**, 115–140.
- 40 G. Charvin, J.-F. Allemand, T. R. Strick, D. Bensimon and V. Croquette, *Contemp. Phys.*, 2004, **45**, 383–403.
- 41 A. Mossa, M. Manosas, N. Forns, J. M. Hugué and F. Ritort, *J. Stat. Mech (Theor. and Exp.)*, 2009, **2009**, P02060.
- 42 M. Manosas, A. Mossa, N. Forns, J. M. Hugué and F. Ritort, *J. Stat. Mech (Theor. and Exp.)*, 2009, **2009**, P02061.
- 43 N. F. J. M. Hugué and F. Ritort, *Phys. Rev. Lett.*, 2009, **103**, 248106.
- 44 A. Alemany and F. Ritort, *Biopolymers*, 2014, **101**, 1193–1199.
- 45 S. B. Smith, Y. J. Cui and C. Bustamante, *Science*, 1996, **271**, 795–799.
- 46 C. Bustamante, S. B. Smith, J. Liphardt and D. Smith, *Curr. Opin. Struct. Biol.*, 2000, **10**, 279–285.
- 47 C. G. Baumann, V. A. Bloomfield, S. B. Smith, C. Bustamante, M. D. Wang and S. M. Block, *Biophys. J.*, 2000, **78**, 1965–1978.
- 48 C. Bustamante, Z. Bryant and S. B. Smith, *Nature*, 2003, **421**, 423–427.
- 49 J. Gore, Z. Bryant, M. Nollmann, M. U. Le, N. R. Cozzarelli and C. Bustamante, *Nature*, 2006, **442**, 836–839.
- 50 Z. Bryant, M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli and C. Bustamante, *Nature*, 2003, **424**, 338–341.
- 51 Z. Bryant, F. C. Oberstrass and A. Basu, *Curr. Opin. Struct. Biol.*, 2012, **22**, 304–312.
- 52 K. Pant, R. L. Karpel, I. Rouzina and M. C. Williams, *J. Mol. Biol.*, 2004, **336**, 851–870.
- 53 T. Paramanathan, I. Vladescu, M. J. McCauley, I. Rouzina and M. C. Williams, *Nucleic Acids Res.*, 2012, **40**, 4925–4932.
- 54 T. Paramanathan, F. Westerlund, M. J. McCauley, I. Rouzina, P. Lincoln and M. C. Williams, *J. Am. Chem. Soc.*, 2008, **130**, 3752–3.
- 55 M. Cruceanu, R. J. Gorelick, K. Musier-Forsyth, I. Rouzina and M. C. Williams, *J. Mol. Biol.*, 2006, **363**, 867–77.
- 56 A. Sischka, K. Tönsing, R. Eckel, S. D. Wilking, N. Sewald, R. Rios and D. Anselmetti, *Biophys. J.*, 2005, **88**, 404–411.
- 57 C. Kleimann, A. Sischka, A. Spiering, K. Tönsing, N. Sewald, U. Diederichsen and D. Anselmetti, *Biophys. J.*, 2009, **97**, 2780–2784.
- 58 M. Slutsky and L. A. Mirny, *Biophys. J.*, 2004, **87**, 4021–4035.
- 59 D. R. Phillips and D. M. Crothers, *Biochemistry*, 1986, **25**, 7355–7362.
- 60 J. B. Chaires, N. Dattagupta and D. M. Crothers, *Biochemistry*, 1985, **24**, 260–267.
- 61 G. Scatchard, *Ann. N. Y. Acad. Sci.*, 1949, **51**, 660–672.
- 62 A. V. Hill, *Proc. Physiol. Soc.*, 1910, **40**, iv–vii.
- 63 L. Siman, I. S. S. Carrasco, J. K. L. da Silva, M. C. Oliveira, M. S. Rocha and O. N. Mesquita, *Phys. Rev. Lett.*, 2012, **109**, 248103.
- 64 E. C. Cesconetto, F. S. A. Junior, F. A. P. Crisafuli, O. N. Mesquita, E. B. Ramos and M. S. Rocha, *Phys. Chem. Chem. Phys.*, 2013, **15**, 11070–11077.
- 65 E. F. Silva, E. B. Ramos and M. S. Rocha, *J. Phys. Chem. B*, 2013, **117**, 7292–6.
- 66 J. D. McGhee and P. H. von Hippel, *J. Mol. Biol.*, 1974, **86**, 469–489.
- 67 M. S. Rocha, *Biopolymers*, 2010, **93**, 1–7.
- 68 M. J. McCauley, E. M. Rueter, I. Rouzina, L. J. M. III and M. C. Williams, *Nucleic Acids Res.*, 2013, **41**, 167–181.
- 69 D. Murugesapillai, M. J. McCauley, R. Huo, M. H. N. Holte, A. Stepanyants, L. J. M. III, N. E. Israeloff and M. C. Williams, *Nucleic Acids Res.*, 2013, **42**, 8996–9004.
- 70 L. D. Williams, M. Egli, Q. Gao and A. Richa, *Structure and Function - Volume One. Proc. 7th Conversation in Biomolecular Stereodynamics.*, 1992.
- 71 M. F. B. na, M. Cacho, A. Gradillas, B. de Pascual-Teresa and A. Ramos, *Curr. Pharmac. Des.*, 2001, **7**, 1745–1780.
- 72 M. K. Gofar, N. M. Kor and Z. M. Kor, *Int. J. Adv. Biol. Biom. Res.*, 2014, **2**, 811–822.
- 73 J. B. Chaires, N. Dattagupta and D. M. Crothers, *Biochemistry*, 1982, **21**, 3933–3940.
- 74 B. Gaugain, J. Barbet, N. Capelle, B. P. Roques and J. L. Pecq, *Biochemistry*, 1978, **17**, 5078–5088.
- 75 M. S. Rocha, M. C. Ferreira and O. N. Mesquita, *J. Chem. Phys.*, 2007, **127**, Art. No. 105108.
- 76 L. A. Reis, E. B. Ramos and M. S. Rocha, *J. Phys. Chem. B*, 2013, **117**, 14345–14350.
- 77 F. A. P. Crisafuli, E. B. Ramos and M. S. Rocha, *Eur. Biophys. J.*, 2015, **44**, 1–7.
- 78 F. Ritort, *J. Phys. - Condens. Mat.*, 2006, **18**, R531–R583.
- 79 K. C. Neuman and A. Nagy, *Nat. Methods*, 2008, **5**, 491–505.
- 80 K. C. Neuman, T. Lionnet and J.-F. Allemand, *Ann. Rev. Mat. Res.*, 2007, **37**, 33–67.
- 81 A. Ashkin, *Phys. Rev. Lett.*, 1970, **24**, 156&.
- 82 A. Ashkin and J. M. Dziedzic, *Science*, 1987, **235**, 1517–1520.
- 83 A. Ashkin, *P. Natl. Acad. Sci. USA*, 1997, **94**, 4853–4860.
- 84 M. S. Rocha, *Am. J. Phys.*, 2009, **77**, 704–712.
- 85 J. R. Moffitt, Y. R. Chemla, S. B. Smith and C. Bustamante, *Annu. Rev. Biochem.*, 2008, **77**, 205–228.
- 86 D. G. Grier, *Nature*, 2003, **424**, 810–816.
- 87 K. Svoboda and S. M. Block, *Annu. Rev. Bioph. Biom.*, 1994, **23**, 247–285.
- 88 K. C. Neuman and S. M. Block, *Rev. Sci. Instrum.*, 2004, **75**, 2787–2809.
- 89 I. Heller, T. P. Hoekstra, G. A. King, E. J. G. Peterman and G. J. L. Wuite, *Chem. Rev.*, 2014, **114**, 3087–3119.
- 90 K. R. Chaurasiya, T. Paramanathan, M. J. McCauley and M. C. Williams, *Phys. Life Rev.*, 2010, **7**, 299–341.
- 91 M. J. McCauley and M. C. Williams, *Biopolymers*, 2007, **85**, 154–168.
- 92 M. J. McCauley and M. C. Williams, *Biopolymers*, 2009, **91**, 265–282.
- 93 G. V. Shivashankar, G. Stolovitzky and A. J. Libchaber, *Appl. Phys. Lett.*, 1998, **73**, 291–293.
- 94 M. S. Rocha, N. B. Viana and O. N. Mesquita, *J. Chem. Phys.*, 2004, **121**, 9679–9683.
- 95 M. W. Allersma, F. Gittes, M. J. deCastro, R. J. Stewart and C. F. Schmidt, *Biophys. J.*, 1998, **74**, 1074–1085.
- 96 J.-C. Meiners and S. R. Quake, *Phys. Rev. Lett.*, 2000, **84**, 1–7.
- 97 F. A. P. Crisafuli, E. C. Cesconetto, E. B. Ramos and M. S. Rocha, *Integr. Biol.*, 2012, **2012**, 568–574.
- 98 F. A. P. Crisafuli, E. C. Cesconetto, E. B. Ramos and M. S. Rocha, *Appl. Phys. Lett.*, 2012, **100**, 083701.
- 99 I. de Vlaminck and C. Dekker, *Annu. Rev. Biophys.*, 2012, **41**, 453–72.
- 100 D. Kilinca and G. U. Lee, *Integr. Biol.*, 2014, **6**, 27–34.
- 101 J. Lipfert, X. Hao and N. H. Dekker, *Biophys. J.*, 2009, **96**, 5040–5049.
- 102 C. Gosse and V. Croquette, *Biophys. J.*, 2002, **82**, 3314–3329.
- 103 D. Salerno, D. Brogioli, V. Cassina, D. Turchi, G. L. Beretta, D. Seruggia, R. Ziano, F. Zunino and F. Mantegazza, *Nuc. Acids Res.*, 2010, **38**, 7089–7099.
- 104 J. D. Moroz and P. Nelson, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 14418–14422.
- 105 A. Celedon, D. Wirtz and S. Sun, *J. Phys. Chem. B*, 2010, **114**, 16929–16935.
- 106 R. Nambiar, A. Gajraj and J.-C. Meiners, *Biophys. J.*, 2004, **87**, 1972–1980.
- 107 W. J. Greenleaf, M. T. Woodside, E. A. Abbondanzieri and S. M. Block,

- Phys. Rev. Lett.*, 2005, **95**, 208102.
- 108 C. Bustamante and C. Rivetti, *Annu. Rev. Biophys. Biomol. Struct.*, 1996, **25**, 395–429.
- 109 D. Anselmetti, J. Fritz, B. Smith and X. Fernandez-Busquets, *Single Mol.*, 2000, **1**, 53–58.
- 110 C. Rivetti, M. Guthold and C. Bustamante, *J. Mol. Biol.*, 1996, **264**, 919–932.
- 111 Y. Yang, L. E. Sass, C. Du, P. Hsieh and D. A. Erie, *Nucleic Acids Res.*, 2005, **33**, 4322–4334.
- 112 T. Strunz, K. Oroszlan, R. Schäfer and H.-J. Güntherodt, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 11277–11282.
- 113 J. Zlatanovaa, S. M. Lindsayy and S. H. Leuba, *Prog. Biophys. Mol. Biol.*, 2000, **74**, 37–61.
- 114 A. Engel, Y. Lyubchenko and D. Müller, *Trends Cell Biol.*, 1999, **9**, 77–80.
- 115 K. Yoshikawa, Y. Matsuzawa, K. Minagawa, M. Doi and M. Matsumoto, *Bioch. Biophys. Res. Commun.*, 1992, **188**, 1274–1279.
- 116 Y. Yoshikawa, K. Yoshikawa and T. Kanbe, *Biophys. Chem.*, 1996, **61**, 93–100.
- 117 N. Yoshinaga, T. Akitaya and K. Yoshikawa, *Biochem. Biophys. Res. Commun.*, 2001, **286**, 264–267.
- 118 A. Kurtz, E. T. Kool and W. E. Moerner, *Biophys. J.*, 2005, **88**, 350A–350A.
- 119 K. Yoshikawa, S. Hirota, N. Makita and Y. Yoshikawa, *Phys. Chem. Lett.*, 2010, **1**, 1763–1766.
- 120 I. Amitani, B. Liu, C. C. Dombrowski, R. J. Baskin and S. C. Kowalczykowski, *Methods Enzymol.*, 2010, **472**, 261–291.
- 121 P. G. J. van Mameren and, G. Farge, P. Hooijman, M. Modesti, M. Falkenberg, G. J. L. Wuite and E. J. G. Peterman, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 18231–18236.
- 122 P. G. G. A. King and, U. Bockelmann, M. Modesti, G. J. L. Wuite and E. J. G. Peterman, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 3859–3864.
- 123 C. Zhang, A. Hernandez-Garcia, K. Jiang, Z. Gong, D. Guttula, S. Y. Ng, P. P. Malar, J. A. van Kan, L. Dai, P. S. Doyle, R. de Vries and J. R. C. van der Maarel, *Nucleic Acids Res.*, 2013, **41**, e189.
- 124 K. G. Strothkamp and S. J. Lippard, *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 2536–2540.
- 125 S. M. Cohen and S. J. Lippard, *Prog. Nucleic Acid Res.*, 2001, **67**, 93–130.
- 126 N.-K. Lee, J.-S. Park, A. Johner, S. Obukhov, J.-Y. Hyon, K. J. Lee and S.-C. Hong, *Phys. Rev. E*, 2009, **79**, 041921.
- 127 X.-M. Hou, X.-H. Zhang, K.-J. Wei, C. Ji, S.-X. Dou, W.-C. Wang, M. Li and P.-Y. Wang, *Nucleic Acids Res.*, 2009, **37**, 1400–1410.
- 128 D. W. Ussery, R. W. Hoepfner and R. R. Sinden, *Method. Enzymol.*, 1992, **212**, 242–262.
- 129 W. McNeely and K. L. Goa, *Drugs*, 1998, **56**, 667–690.
- 130 T. R. Coven, F. P. Murphy, P. Gilleaudeau, I. Cardinale and J. G. Krueger, *Arch. Dermatol.*, 1998, **134**, 1263–1268.
- 131 D. Tran, Y. K. Kwok and C. L. Goh, *Photodermatol. Photo.*, 2001, **17**, 164–167.
- 132 H. P. Spielmann, T. J. Dwyer, S. S. Sastry, J. E. Hearst and D. E. Wemmer, *P. Natl. Acad. Sci. USA*, 1995, **92**, 2345–2349.
- 133 M. S. Rocha, A. D. Lúcio, S. S. Alexandre, R. W. Nunes and O. N. Mesquita, *Appl. Phys. Lett.*, 2009, **95**, Art. No. 253703.
- 134 L. S. Lerman, *J. Mol. Biol.*, 1961, **3**, 18–30.
- 135 V. Luzzati, F. Masson and L. S. Lerman, *J. Mol. Biol.*, 1961, **3**, 634–639.
- 136 I. C. Lin, O. A. von Lilienfeld, M. D. Coutinho-Neto, I. Tavernelli and U. Rothlisberger, *J. Phys. Chem. B*, 2007, **111**, 14346–14354.
- 137 H. Fritzsche, H. Triebel, J. B. Chaires, N. Dattagupta and D. M. Crothers, *Biochemistry*, 1982, **21**, 3940–3946.
- 138 S. Nafisi, A. A. Saboury, N. Keramat and H.-A. T.-R. J.-F. Neault, *J. Mol. Struct.*, 2007, **827**, 35–43.
- 139 J. B. Chaires, N. Dattagupta and D. M. Crothers, *Biochemistry*, 1982, **21**, 3927–3932.
- 140 I. Tessmer, C. G. Baumann, G. M. Skinner, J. E. Molloy, J. G. Hoggett, S. J. B. Tendler and S. Allen, *J. Mod. Optic.*, 2003, **50**, 1627–1636.
- 141 V. Cassina, D. Seruggia, G. L. Beretta, D. Salerno, D. Brogioli, S. Manzini, F. Zunino and F. Mantegazza, *Eur. Biophys. J.*, 2010, **40**, 59–68.
- 142 J. Lipfert, S. Klijnhout and N. H. Dekker, *Nucl. Acids Res.*, 2010, **38**, 7122–7132.
- 143 Y. Matsuzawa and K. Yoshikawa, *Nucleos. Nucleot.*, 1994, **13**, 1415–1423.
- 144 S. R. Quake, H. Babcock and S. Chu, *Nature*, 1997, **388**, 151–154.
- 145 T. Berge, N. S. Jenkins, R. B. Hopkirk, M. J. Waring, J. M. Edwardson and R. M. Henderson, *Nucl. Acids Res.*, 2002, **30**, 2980–2986.
- 146 N. Kaji, M. U. M and Y. Baba, *Electrophoresis*, 2001, **22**, 3357–3364.
- 147 M. S. Rocha, *Phys. Biol.*, 2009, **6**, Art. No. 036013.
- 148 R. F. Bazoni, C. H. M. Lima, E. B. Ramos and M. S. Rocha, *Soft Matter*, 2015, **11**, 4306–4314.
- 149 V. A. Bloomfield, *Biopolymers*, 1991, **31**, 1471–1481.
- 150 V. A. Bloomfield, *Biopolymers*, 1997, **44**, 269–282.
- 151 P. G. Arscott, A. Z. Li and V. A. Bloomfield, *Biopolymers*, 1990, **30**, 619–630.
- 152 G. E. Plum, P. G. Arscott and V. A. Bloomfield, *Biopolymers*, 1990, **30**, 631–643.
- 153 N. Grønbech-Jensen, R. J. Mashl, R. F. Bruinsma and W. M. Gelbart, *Phys. Rev. Lett.*, 1997, **78**, 2477–2480.
- 154 R. G. Winkler, M. Gold and P. Reineker, *Phys. Rev. Lett.*, 1998, **80**, 3731–3734.
- 155 M. Khan and B. Jönsson, *Biopolymers*, 1999, **49**, 121–125.
- 156 I. Rouzina and V. A. Bloomfield, *Biophys. J.*, 1998, **74**, 3152–3164.
- 157 A. Podestá, M. Indrieri, D. Brogioli, G. S. Manning, P. Milani, R. Guerra, L. Finzi and D. Dunlap, *Biophys. J.*, 2005, **89**, 2558–2563.
- 158 B. H. Geierstanger and D. E. Wemmer, *Annu. Rev. Biophys. Biomol. Struct.*, 1995, **24**, 463–493.
- 159 R. Eckel, R. Ros, A. Ros, S. D. Wilking, N. Sewald and D. Anselmetti, *Biophys. J.*, 2003, **85**, 1968–1973.
- 160 P. L. Hamilton and D. P. Arya, *Nat. Prod. Rep.*, 2012, **29**, 134–143.
- 161 D. J. Patel, S. A. Kozlowski, J. A. Rice, C. Broka and K. Itakura, *Proc. Natl. Acad. Sci. USA*, 1981, **78**, 7281–7284.
- 162 W. Müller and D. M. Crothers, *J. Mol. Biol.*, 1968, **35**, 251–290.
- 163 H. M. Sobell, S. C. Jain, T. D. Sakore and C. E. Nordman, *Nature. New Biol.*, 1971, **231**, 200–205.
- 164 F. Takusagawa, M. Dabrow, S. Neidle and H. M. Berman, *Nature*, 1982, **296**, 466–469.
- 165 C. Pérez-Arnaiz, N. Busto, J. M. Leal and B. García, *J. Phys. Chem. B*, 2014, **118**, 1288–1295.
- 166 M. C. Williams, I. Rouzina and V. A. Bloomfield, *Acc. Chem. Res.*, 2002, **35**, 159–166.
- 167 R. Krautbauer, H. C. Shaumann and H. E. Gaub, *Angew. Chem. Int. Ed.*, 2000, **39**, 3912–3915.
- 168 N. Dattagupta, M. Hogan and D. M. Crothers, *Biochemistry*, 1980, **19**, 5998–6005.
- 169 J. E. Coury, L. McFail-Isom, L. D. Williams and L. A. Bottomley, *P. Natl. Acad. Sci. USA*, 1996, **93**, 12283–12286.
- 170 A. Mihailovic, I. Vladescu, M. J. McCauley, E. Ly, M. C. Williams, E. M. Spain and M. E. Nunez, *Langmuir*, 2006, **22**, 4699–4709.
- 171 R. Krautbauer, L. H. Pope, T. E. Schrader, S. Allen and H. E. Gaub, *FEBS Lett.*, 2002, **510**, 154–8.
- 172 R. Krautbauer, S. Fischerländer, S. Allen and H. E. Gaub, *Single Mol.*, 2002, **3**, 97–103.
- 173 S. Husale, W. Grange and M. Hegner, *Single Mol.*, 2002, **3**, 91–96.
- 174 J. Camunas-Soler, S. Frutos, C. V. Bizarro, S. de Lorenzo, M. E.

- Fuentes-Perez, R. Ramsch, S. Vilchez, C. Solans, F. Moreno-Herrero, F. Albericio, R. Eritja, E. Giralt, S. B. Dev and F. Ritort, *ACS Nano*, 2013, **7**, 5102–5113.
- 175 J.-F. Allemand, D. Bensimon and V. Croquette, *Curr. Opin. Struc. Biol.*, 2003, **13**, 266–274.
- 176 K. R. Chaurasiya, M. J. McCauley, W. Wang, D. F. Qualley, T. Wu, S. Kitamura, H. Geertsema, D. S. B. Chan, A. Hertz, Y. Iwatani, J. G. Levin, K. Musier-Forsyth, I. Rouzina and M. C. Williams, *Nat. Chem.*, 2014, **6**, 28–33.
- 177 M. S. Rocha, A. G. Cavalcante, R. Silva and E. B. Ramos, *J. Phys. Chem. B*, 2014, **118**, 4832–4839.
- 178 J. van Noort, S. Verbrugge, N. Goosen, C. Dekker and R. T. Dame, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 6969–6974.
- 179 P. M. Takahara, C. A. Frederick and S. J. Lippard, *J. Am. Chem. Soc.*, 1996, **118**, 12309–12321.
- 180 D. B. Zamble and S. J. Lippard, *Trends Biochem. Sci.*, 1995, **20**, 435–439.
- 181 S. G. Chaney, S. L. Campbell, B. Temple, E. Bassett, Y. Wu and M. Faldu, *J. Inorg. Biochem.*, 2004, **98**, 1551–1559.
- 182 A. L. Pinto and S. J. Lippard, *Bioschim. Biophys. Acta.*, 1985, **780**, 167–180.
- 183 K. Stehlikova and H. Kostrhunova, *Nucleic Acids Res.*, 2002, **30**, 2894–2898.
- 184 R. T. Dame, M. C. Noom and G. J. L. Wuite, *Nature*, 2006, **444**, 387–390.
- 185 M. A. Spies and R. L. Schowen, *J. Am. Chem. Soc.*, 2002, **124**, 14049–14053.
- 186 G. D. Tavares, C. M. Viana, J. G. V. C. Araujo, G. A. Ramaldes, W. S. Carvalho, J. L. Pesquero, J. Vilela, M. S. Andrade and M. C. de Oliveira, *Chem. Phys. Lett.*, 2006, **429**, 507–512.
- 187 R. Osathanondh, D. P. Goldstein and G. B. Pastorfide, *Cancer*, 1975, **36**, 863–866.
- 188 G. J. D'Angio, A. E. Evans, N. Breslow, B. Beckwith, H. Bishop, P. Feigl, W. Goodwin, L. L. Leape, L. F. Sinks, W. Sutow, M. Tefft and J. Wolff, *Cancer*, 1976, **38**, 633–646.
- 189 D. Pinkel and J. Pickren, *JAMA - J. Am. Med. Assoc.*, 1961, **175**, 293–298.
- 190 J. R. Neto and M. F. Colombo, *Biopolymers*, 2000, **53**, 46–59.
- 191 F. Sha and F. M. Chen, *Biophys. J.*, 2000, **79**, 2095–2104.
- 192 J. Goodisman, R. Reh fuss, B. Ward and J. C. Dabrowiak, *Biochemistry*, 1992, **31**, 1046–1058.
- 193 B. S. Reddy, S. K. Sharma and J. W. Lown, *Curr. Med. Chem.*, 2001, **8**, 475–508.
- 194 M. Saito, M. Kobayashi, S. Iwabuchi, Y. Morita, Y. Takamura and E. Tamiya, *J. Biochem.*, 2004, **136**, 813–823.
- 195 C. Bailly, P. Colson, J. Henichart and C. Houssier, *Nucl. Acids Res.*, 1993, **21**, 3705–3709.
- 196 J. Bontemps, C. Housster and E. Fredericq, *Nucl. Acids Res.*, 1975, **2**, 971–984.
- 197 T. Stokke and H. B. Steen, *J. Histochem. Cytochem.*, 1985, **33**, 333.
- 198 P. Colson, C. Houssier and C. Bailly, *J. Biomol. Struct. Dyn.*, 1995, **13**, 351–366.
- 199 *Nucleic Acid Detection Technologies* - <http://www.biotium.com>.
- 200 Q. Huang, L. Baum and W. L. Fu, *Clin. Lab.*, 2010, **56**, 149–152.
- 201 K. Günther, M. Mertig and R. Seidel, *Nucl. Acids Res.*, 2010, **38**, 6526–6532.
- 202 C. U. Murade, V. Subramaniam, C. Otto and M. L. Bennink, *Biophys. J.*, 2009, **97**, 835–843.
- 203 M. Maaloum, P. Mullera and S. Harlepp, *Soft Matter*, 2013, **9**, 11233.
- 204 C. Garbay-Jaureguiberry, P. Laugãa, M. Delepierre, S. Laalami, G. Muzard, J. B. L. Pecq and B. P. Roques, *Anticancer Drug Des.*, 1987, **1**, 323–335.
- 205 D. Sagi, N. Friedman, C. Vorgias, A. B. Oppenheim and J. Stavans, *J. Mol. Biol.*, 2004, **341**, 419–428.
- 206 S. M. Rappaport and Y. Rabin, *Phys. Rev. Lett.*, 2008, **101**, 038101.
- 207 A. L. Galo, J. R. Neto, D. P. Brognaro, R. C. Caetano, F. P. Souza and M. F. Colombo, *J. Phys. Chem. B*, 2011, **115**, 8883–8890.
- 208 S. Neidle and M. J. Waring, *Sequence-specific DNA Binding Agents.*, RSC Publishing, 2006.

**Insight Box** – Review article “Extracting physical chemistry from mechanics: a new approach to investigate DNA interactions with drugs and proteins in single molecule experiments” by M. S. Rocha.

---

In this review we focus on the idea of establishing connections between the mechanical properties of DNA-ligand complexes and the physical chemistry of DNA-ligand interactions. In particular, we show that if one knows how a mechanical property changes as a function of the ligand concentration in the sample, many insights on the physical chemistry of the interaction can be promptly obtained. Such method opens the possibility of characterizing the interactions between DNA and ligands by using only one experimental technique: single molecule stretching. Furthermore, it also opens new possibilities in comparing results obtained by very different approaches, in special when comparing single molecule techniques to ensemble-averaging techniques.

---