

# Single-molecule kinetics and footprinting of DNA bis-intercalation: the paradigmatic case of Thiocoraline

Joan Camunas-Soler<sup>1,2</sup>, Maria Manosas<sup>1,2</sup>, Silvia Frutos<sup>1,2</sup>, Judit Tulla-Puche<sup>3</sup>, Fernando Albericio<sup>3</sup> and Felix Ritort<sup>1,2,\*</sup>

<sup>1</sup>Small Biosystems Lab, Departament de Física Fonamental, Facultat de Física, Universitat de Barcelona, 08028 Barcelona, Spain, <sup>2</sup>CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, 28029 Madrid, Spain and <sup>3</sup>Institute for Research in Biomedicine (IRB Barcelona), Barcelona Science Park, Baldiri Reixac 10-12, 08028 Barcelona, Spain

Received December 21, 2014; Revised January 21, 2015; Accepted January 22, 2015

## ABSTRACT

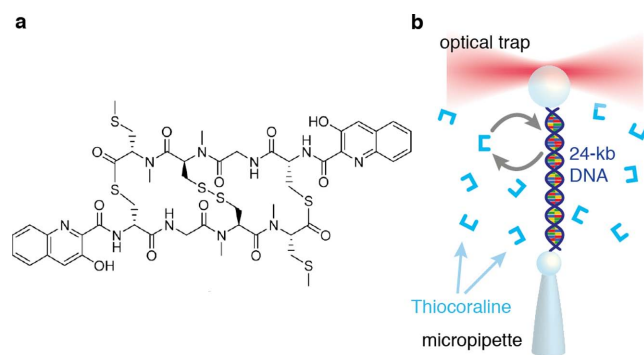
DNA bis-intercalators are widely used in molecular biology with applications ranging from DNA imaging to anticancer pharmacology. Two fundamental aspects of these ligands are the lifetime of the bis-intercalated complexes and their sequence selectivity. Here, we perform single-molecule optical tweezers experiments with the peptide Thiocoraline showing, for the first time, that bis-intercalation is driven by a very slow off-rate that steeply decreases with applied force. This feature reveals the existence of a long-lived (minutes) mono-intercalated intermediate that contributes to the extremely long lifetime of the complex (hours). We further exploit this particularly slow kinetics to determine the thermodynamics of binding and persistence length of bis-intercalated DNA for a given fraction of bound ligand, a measurement inaccessible in previous studies of faster intercalating agents. We also develop a novel single-molecule footprinting technique based on DNA unzipping and determine the preferred binding sites of Thiocoraline with one base-pair resolution. This fast and radiolabelling-free footprinting technique provides direct access to the binding sites of small ligands to nucleic acids without the need of cleavage agents. Overall, our results provide new insights into the binding pathway of bis-intercalators and the reported selectivity might be of relevance for this and other anticancer drugs interfering with DNA replication and transcription in carcinogenic cell lines.

## INTRODUCTION

Small DNA binders have become powerful tools in biomedicine due to their ability to interact with nucleic acids at specific sites and modulate key processes such as repair, transcription and replication (1–4). Among them, DNA intercalators and bis-intercalators have received continued interest during the last decades due to their particular properties that range from being excellent DNA staining probes (e.g. ethidium, YOYO) to effective anti-proliferative drugs (e.g. doxorubicin, echinomycin) (4,5). From a structural point of view, intercalators are small planar compounds that bind nucleic acids by sliding between two adjacent base pairs and bringing them apart. This process unwinds the DNA double helix and increases its contour length by  $\sim 3.4$  Å per intercalator (6). In a DNA bis-intercalator, two of such intercalative moieties are connected by a linker chain. The compound is therefore able to ‘clamp’ DNA by sandwiching two consecutive base pairs between its aromatic rings, distorting and elongating DNA a length twice that of a mono-intercalator.

Although intercalating agents are widely used in anticancer chemotherapy (7), their mechanism of action based on interfering with the transcription and repair machinery makes them also highly toxic causing undesired side effects (8). In order to improve their effectivity, a lot of effort has been put in developing new ligands that (i) exhibit improved sequence selectivity and (ii) form long-lived complexes (9–12). Increased specificity enables recognition and targeting of particular disease-related motifs that can be used for biomedical diagnostics and therapeutics (3,4). Slow off-kinetics of long-lived complexes are related to sustained response of anticancer drugs (10,13). In addition such ligands may also be useful to develop better fluorescent probes with increased imaging times (14). Hence, the development of rapid and accurate methods to elucidate the kinetics and

\*To whom correspondence should be addressed. Tel: +34 934035869; Fax: +34 934021149; Email: fritort@gmail.com



**Figure 1.** Thiocoraline bis-intercalation using optical tweezers. (a) Thiocoraline structure. (b) Scheme of the DNA stretching experiments in the optical tweezers setup.

selectivity of small ligands is paramount to advance in the rational design of less toxic and more effective compounds.

Single-molecule force-spectroscopy techniques have proven especially well-suited to investigate the structural and mechanical perturbations that intercalators induce to the DNA double helix (15,16), succeeding in characterizing the binding of very weak ligands (15). This is due to the fact that force spectroscopy uses molecular extension as the natural reaction coordinate and allows to apply force to shift the chemical equilibrium of binding. So far, most experiments have focused on deriving equilibrium information of binding (e.g. affinity) (15–18), whereas only recently attention has been paid to non-equilibrium experiments from which kinetic rates might be inferred (19–21). Moreover, despite the ubiquitous use of DNA intercalators in single-molecule experiments, there is still controversy in some fundamental aspects such as the effects of intercalation on the elasticity of dsDNA (16–17,21–23) and the existence of binding kinetic intermediates (20,23). Similarly, other key parameters such as sequence specificity have been scarcely addressed with these techniques, and still rely on classical footprinting assays that are time-consuming and complex to set up and interpret (24–26).

In this work we use optical and magnetic tweezers to study the anticancer bis-intercalator Thiocoraline (Figure 1a) (27), fully characterizing the thermodynamics, kinetics and selectivity of binding using a combination of DNA stretching and unzipping assays (28). A clear understanding of these properties was lacking due to the low solubility of this drug (29,30), making measurements difficult and leading to limited and conflicting results (31,32). We find that Thiocoraline bis-intercalates DNA with an extremely slow dissociation constant that, interestingly, decreases steeply with force. This makes possible to measure the elastic properties of intercalated DNA by performing non-equilibrium pulling experiments in which the binding fraction of ligand remains quenched. We show that bis-intercalation does not modify DNA persistence length whereas the stretch modulus does increase with the bound fraction. A series of kinetic experiments demonstrates the existence of a mono-intercalated intermediate state that contributes to the long lifetime of the complex. Finally, we

develop a single-molecule footprinting assay based on DNA unzipping, in which we show that at low concentrations the peptide binds DNA in a sequence-specific manner, with a preference to clamp CpG steps flanked by A-T base pairs. Such information is difficult to obtain in bulk assays and cannot be accessed in usual single-molecule stretching assays.

We propose that this sequence selectivity together with the force-dependent off-rate is an important feature of the mechanism of action of this subnanomolar drug ( $IC_{50} \sim 200$  pM). On a broader perspective, we provide new insights into the general mechanisms by which bis-intercalators bind DNA and their kinetic stability. The proposed single-molecule approach will be of great utility to quantify essential parameters of the specific and non-specific binding modes of ligands that target DNA (5,25).

## MATERIALS AND METHODS

### Optical tweezers setup

Optical tweezers experiments are performed with a miniaturized dual-beam setup described elsewhere (33). Briefly, a single optical trap is created by focusing two counter-propagating laser beams into a spot ( $\lambda = 845$  nm,  $P = 200$  mW) in a microfluidics chamber. Forces in an optically trapped bead are directly determined from the change in light momentum by measuring the deflection of the exiting laser beams with Position Sensing Detectors (PSDs) (33). The extension and force applied to the molecule can therefore be recorded at 1-kHz rate, and a resolution of 0.1 pN is achieved. Experiments are performed by tethering a single-DNA molecule between two polystyrene beads. One of them is confined in the optical trap, whereas the other is subjected by air suction in the tip of a glass micropipette. Tethers are made by differentially labeling each end of the DNA molecule with several biotins or digoxigenins, which bind to streptavidin- or anti-digoxigenin-coated beads, respectively. In pulling experiments, the relative molecular extension is inferred by subtracting the trap compliance ( $F/k$ ) to the relative displacement of the optical trap ( $k_{\text{trap}} = 70$  pN/ $\mu\text{m}$ ). The offset to determine the absolute molecular extension is obtained from video microscopy measurements of the bead-to-bead distance. Experimental force–extension curves (FECs) and kinetic traces were recorded at 1 kHz and filtered to 10 Hz.

### Magnetic tweezers assays

Magnetic tweezers experiments are performed with a PicoTwist magnetic tweezers instrument ([www.picotwist.com](http://www.picotwist.com)). DNA molecules are tethered between a glass surface treated with anti-digoxigenin antibody (Roche) and a 1- $\mu\text{m}$  streptavidin-coated Dynal magnetic bead (Invitrogen). DNA molecules are manipulated and stretched by capturing the bead in a magnetic trap generated by a pair of permanent magnets. The applied force is controlled by varying the distance from the magnets to the sample. Video microscopy is used to track the position of the magnetic bead in three dimensions with nanometer resolution at 30 Hz, from which the extension of the DNA molecule and the strength of the stretching force are deduced (34).

## Reagents and synthesis of DNA molecular constructs

Thiocoraline was provided by Pharmamar as lyophilized product and dissolved in DMSO as a 1-mM stock solution, which was aliquoted and stored at  $-20^{\circ}\text{C}$ . Sample concentration was checked spectrophotometrically and with HPLC. Aliquots were thawed as needed and dissolved in working buffer, adjusting the DMSO concentration to 2%. All experiments were performed in TE buffer (Tris 10 mM, ethylenediaminetetraacetic acid 1 mM) pH7.5, 100-mM NaCl, 0.01%  $\text{NaN}_3$ . Experimental controls with drug-free buffer were performed by supplementing the buffer with 2% DMSO.

The synthesis protocols for the 24-kb DNA molecule for stretching experiments, and the 6.8-kb DNA hairpin for unzipping experiments, are described in detail in (28). The 480-bp hairpin for DNA footprinting experiments is synthesized from a plasmid that contains the sequence of interest embedded between the restriction sites of Tsp45I and TspRI. This region of the plasmid is polymerase chain reaction amplified and the product is digested with the two enzymes. A set of oligonucleotides is designed to create the handles structure on the TspRI-digested end following the same approach as explained in (28). Similarly, an oligonucleotide complementary to the Tsp45I end that folds into the end-loop structure is annealed and ligated to create the final hairpin structure. A final gel purification step is added to remove competing structures.

## Analysis of non-equilibrium pulling experiments: elasticity and thermodynamics

Reversible DNA pulling curves at varying Thiocoraline concentrations (0–1000 nM) with optical tweezers (magnetic tweezers) were fitted to the extensible (inextensible) Worm-like chain (WLC) model in the force range 0–40 pN (0–10 pN) using the Marko–Siggia interpolation formula

$$F(l) = \frac{k_B T}{L_p} \left( \frac{1}{4(1-l)^2} - \frac{1}{4} + l + \sum_{i=2}^{i \leq 7} \alpha_i l^i \right)$$

$$l = x/l_0 - F/S(\text{extensible})$$

$$l = x/l_0(\text{inextensible}) \quad (1)$$

where  $x$  is the molecular extension (end-to-end distance),  $l_0$  is the contour length,  $L_p$  is the persistence length,  $S$  is the stretch modulus and  $\alpha_i$  are the polynomial coefficients determined in (35). Before performing the pulling cycles to measure the metastable FECs, the fraction of bound intercalator was equilibrated to its low force value by setting the force feedback at 2 pN for 5 min. We then performed a set of pulling cycles at a pulling speed ( $v = 1.5\text{--}3 \mu\text{m/s}$ ) in which the metastable FECs were fully reversible in the investigated range of forces. For each tethered molecule we determined the  $l_0$ ,  $L_p$  and  $S$  values by fitting a set of consecutive pulling cycles (typically more than 5). The reported values at each concentration are the average of a set of different molecules ( $N \geq 7$ ).

The elongation of the DNA molecule at each Thiocoraline concentration ( $l_0^{\text{Thio}} - l_0^{\text{DNA}}$ ) is directly proportional to the number of intercalators bound to the molecule, and to

its equilibrium binding fraction ( $\nu$ ):

$$\nu = \frac{\# \text{ bound ligands}}{\# \text{ of DNA bp}} = \frac{l_0^{\text{Thio}} - l_0^{\text{DNA}}}{\Delta l_0 N_{\text{bp}}} = \frac{l_0^{\text{Thio}} - l_0^{\text{DNA}}}{2l_0^{\text{DNA}}} \quad (2)$$

where  $l_0^{\text{DNA}} = N_{\text{bp}}d$  ( $d$  being the base-pair rise, 0.34 nm) and  $\Delta l_0$  the DNA elongation induced by one bis-intercalator. We have assumed that each bis-intercalator elongates the DNA molecule a distance twice  $d$  ( $\Delta l_0 = 2d$ ) (6,15). For ligands that bind non-specifically and non-cooperatively to a one-dimensional lattice such as DNA, the McGhee Von-Hippel (MGVH) model (36) predicts that the equilibrium binding fraction satisfies the implicit equation:

$$\nu = \frac{[L]}{K_d} \frac{(1 - n\nu)^n}{(1 - n\nu + \nu)^{n-1}} \quad (3)$$

where  $K_d$  is the ligand affinity constant,  $n$  is the binding-site size (number of bp covered by each ligand) and  $[L]$  is the ligand concentration. By combining Equations (2) and (3) we could fit the titration results (Figure 2c) and obtain the equilibrium affinity constant  $K_d$  and binding-site size  $n$  of Thiocoraline. A Levenberg–Marquadt algorithm was used both for the WLC and MGVH fitting routines.

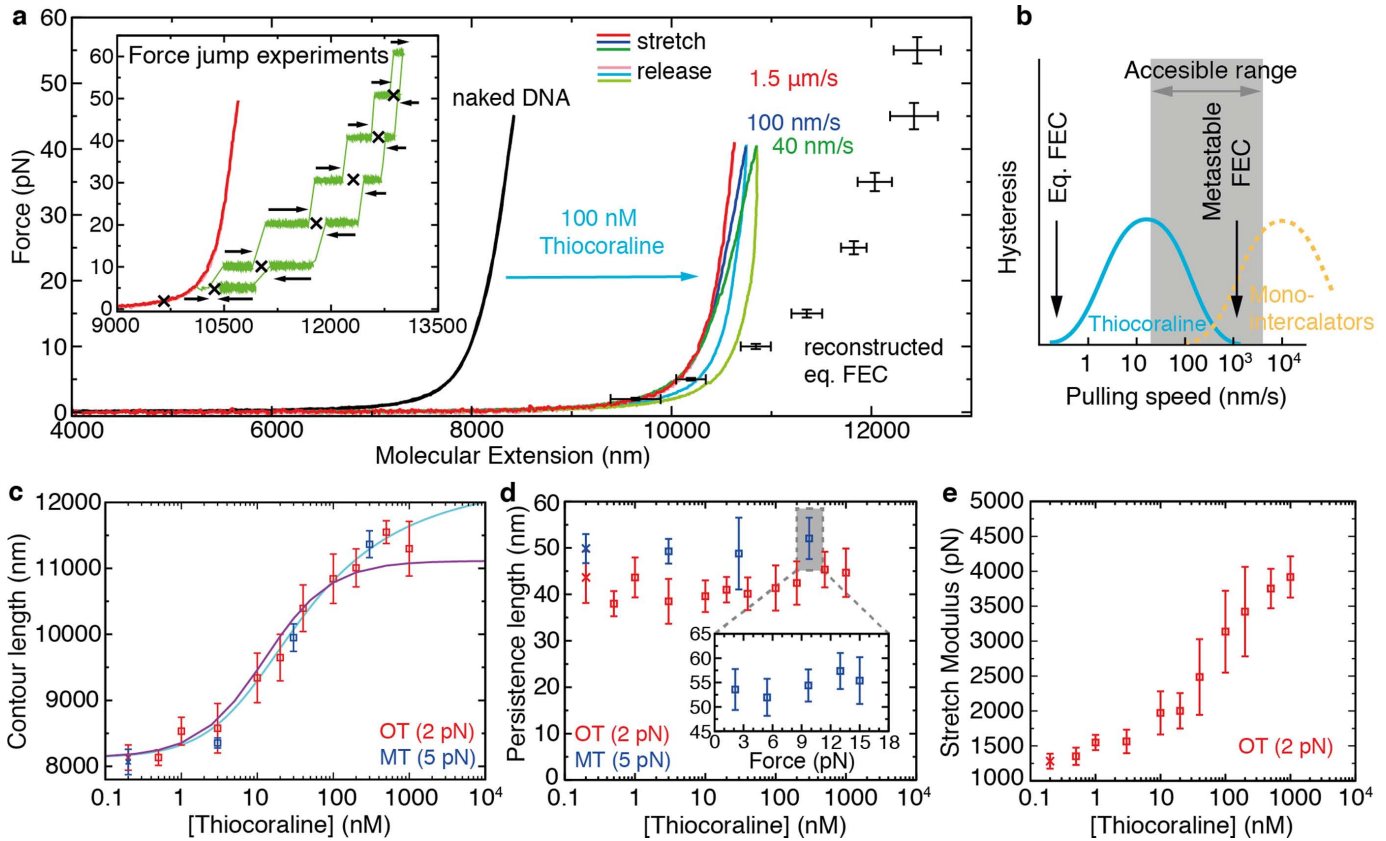
## Wash-off experiments and kinetic rates

Wash-off experiments were performed by replacing the buffer in the microfluidics chamber (containing Thiocoraline) with a scavenger DNA solution that created a fast depletion of ligand. The DNA solution was prepared by digestion of  $\lambda$ -DNA with a blunt-ended restriction enzyme (EcoRV) that leaved fragments sized 250 bp–6 kb. In this way, a randomly assorted distribution of sequences equivalent to those in the tethered DNA molecule was obtained. The digested DNA was dissolved to a 10- $\mu\text{M}$  bp concentration in the same buffer as the ligand samples (TE buffer 100-mM NaCl, pH7.5) and flowed into the experimental area using an auxiliary capillary glass tube. A low flow was maintained during the experiments ( $F_{\text{drag}} \leq 2$  pN) to ensure ligand removal. The unbinding reactions were fitted to single- and double-exponential functions.

## Statistical analysis of footprinting experiments

First, the experimental force–distance curves (FDCs) were aligned to a theoretical prediction based on the nearest-neighbor model (33). The elastic parameters ( $l_k$ ,  $S$ ) used to model the single-stranded DNA (ssDNA) were obtained from an extensible freely jointed chain (FJC) fit to the relaxation curve of the fully unfolded state ( $n = 480$  bp) and were in excellent agreement with previous results for the same ionic condition (37). A small calibration correction (below 5%) on the experimentally measured distance and force was allowed to fully overlap the theoretical and experimental FDCs between the initial rip ( $n = 27$  bp) and the fully unzipped state ( $n = 480$  bp).

For every data point ( $x_{\text{exp}}$ ,  $f_{\text{exp}}$ ) of an FDC, we then determined the most probable apparent number of open base pairs ( $n$ ) by finding the theoretical curve that passed closest



**Figure 2.** FECs of bis-intercalated DNA and elasticity measurements from metastable FECs. (a) DNA pulling curves before (black) and after flowing 100-nM Thiocoraline at varying pulling speeds: 1.5  $\mu\text{m/s}$  (red), 100 nm/s (blue), 40 nm/s (green). Dark (light) colors correspond to the stretching (releasing) parts of the cycle. The equilibrium FEC (black points, mean  $\pm$  SD,  $N \geq 5$ ) is recovered by performing force jump experiments. Inset: in force-jump experiments the relaxation of the molecular extension (green) is measured at constant force values starting from different initial conditions in bidirectional experiments (arrows). The equilibrium extension at each force (crosses) can then be inferred. A metastable FEC (1.5  $\mu\text{m/s}$ ) is shown in red as a reference. (b) Schematics of the hysteresis observed in FECs as a function of pulling speed. For the experimental conditions tested, Thiocoraline (cyan) shows an optimal range to measure metastable FECs, as compared to mono-intercalators (yellow) that reach equilibrium over shorter timescales. (c) Contour length ( $l_0$ ) of DNA as a function of [Thiocoraline] in optical (red squares) and magnetic tweezers (blue squares) experiments. The optical tweezers data are fitted to an MGVB binding isotherm (cyan; Equation. (3)) to obtain the equilibrium affinity and binding-site size at low force ( $K_d = 77 \pm 13$  nM,  $n = 3.84 \pm 0.12$  bp). The theoretical prediction by a three-state model (without free parameters) is shown in purple (see the ‘The kinetic off-rate ( $k_{off}$ ) reveals a binding intermediate’ section). (d) Persistence length ( $L_p$ ) of DNA as a function of [Thiocoraline].  $L_p$  remains independent of Thiocoraline concentration both in optical tweezers (red) and magnetic tweezers (blue) experiments. Inset:  $L_p$  remains also unaffected when the bound fraction is changed by equilibrating the molecule at a different force. (e) Stretch modulus ( $S$ ) of DNA as a function of [Thiocoraline].  $S$  monotonically increases with [Thiocoraline]. For panels (c)–(e) values are reported as mean  $\pm$  SD ( $N \geq 7$  independent molecules), and are determined by fitting metastable FECs ( $v = 1.5\text{--}3$   $\mu\text{m/s}$ ) to the extensible (optical tweezers) or inextensible (magnetic tweezers) WLC model. Reference values from experiments without Thiocoraline for each technique are shown at the lowest concentration (red and blue crosses).

to that point at the force  $f_{\text{exp}}$ :

$$|x_{\text{exp}} - x_{\text{th}}(n, f_{\text{exp}})| = \min_{n^*} (|x_{\text{exp}} - x_{\text{th}}(n^*, f_{\text{exp}})|). \quad (4)$$

The theoretical extension ( $x_{\text{th}}(n, f_{\text{exp}})$ ) was determined from the different contributions to the total distance:

$$x_{\text{th}}(n, f_{\text{exp}}) = x_{\text{b}}(f_{\text{exp}}) + x_{\text{h}}(f_{\text{exp}}) + x_{\text{ssDNA}}(n, f_{\text{exp}}) \quad (5)$$

where  $x_{\text{b}}$  is the displacement of the bead from the center of the optical trap,  $x_{\text{h}}$  the extension of the 29-bp dsDNA handles and  $x_{\text{ssDNA}}$  the extension of released ssDNA during unzipping. The term  $x_{\text{ssDNA}}(n, f_{\text{exp}})$  is the only term that depends on the number  $n$  of unzipped base pairs, and is modeled with an FJC of  $2n$  bases with the aforementioned elastic parameters. The other terms do not depend on  $n$  and are modeled as  $x_{\text{b}} = f_{\text{exp}}/k_{\text{trap}}$  (where  $k_{\text{trap}}$  is the stiffness of the optical trap) and  $x_{\text{h}}$  using a WLC (Equation (1)) with the

generally accepted parameters for double-stranded DNA (dsDNA) (28).

The histogram of  $n$  for an unzipping curve and the following rezipping curve was then plotted, and binding events identified as extra peaks present in the unzipping histogram. A correct matching between the other regions of the histogram was checked, and finally, the position of the binding peaks was determined from the mean value of a Gaussian fit. The binding positions  $n$  determined in this way are then correlated to the DNA sequence of the hairpin, and the normalized probability of binding for each sequence  $B_i$  is determined as

$$p_{\text{norm}}(B_i) = \frac{N_i/S_i}{\sum_j N_j/S_j} \quad (6)$$

where  $N_i$  is the number of binding events observed at the sequence  $i$  and  $S_i$  is the number of times that the sequence is found in the DNA hairpin (i.e. degeneracy).

## RESULTS

### FECs of DNA bis-intercalation

We pulled single-dsDNA molecules (half  $\lambda$ -DNA, 24 kb) in the presence of Thiocoraline in the optical tweezers setup (Figure 1b). DNA is tethered between two beads and stretched at a low force (2 pN) until it reaches its equilibrium extension. The molecule is then repeatedly pulled by moving one bead relatively to the other, and the extension measured as a function of the applied force. In this way we recorded FECs in the range 0–40 pN.

At a concentration of 100 nM, the intercalated DNA has a molecular extension significantly longer (~30%) than that of naked dsDNA (Figure 2a, red and black, respectively), as expected from an intercalative binding mode. Notably, DNA stretching experiments performed at varying pulling speeds show a hysteresis that increases with decreasing pulling speed, whereas hysteresis is not observed at the fastest speed of 1.5  $\mu\text{m/s}$  (Figure 2a, red, blue, green and Supplementary Movie SV1). The observation of hysteresis in DNA pulling curves was previously reported for the bis-intercalator dye YOYO-1 (23), and is caused by the fact that the affinity constant and kinetic rates are force-dependent. It has been shown that the affinity constant of DNA intercalators increases exponentially with applied force (15), and therefore at equilibrium more ligands are bound at higher forces. In a typical pulling cycle the force is increased from a low value (2 pN) up to a high value (40 pN) and then decreased again, enhancing intercalation during a pulling cycle. We identify two relevant timescales in a pulling experiment: the equilibration time of the binding reaction ( $\tau_{\text{eq}}$ ) and the observational time (the duration of a pulling cycle,  $\tau_{\text{obs}}$ ). If pulls were done slowly enough ( $\tau_{\text{obs}} > \tau_{\text{eq}}$ ), the equilibrium FEC would be obtained. In this case we would expect the fraction of bound ligands to increase with force. In the opposite regime ( $\tau_{\text{obs}} \leq \tau_{\text{eq}}$ ), the molecular extension cannot relax to its equilibrium value at each force, and a positive hysteresis is observed (blue and green FECs in Figure 2a). Remarkably enough, if the pulling speed is increased ( $v > 1 \mu\text{m/s}$ , red) metastable FECs are measured, where non-equilibrium but reversible curves are obtained. Along these metastable FECs the number of bound intercalators remains quenched, and the FEC does not show hysteresis (Figure 2b). To the best of our knowledge, this metastable state was not attainable with previously studied mono-intercalators and bis-intercalators due to their faster kinetic rates. In the following section, we will exploit metastability to determine the elastic properties of bis-intercalated DNA for a fixed fraction of ligand.

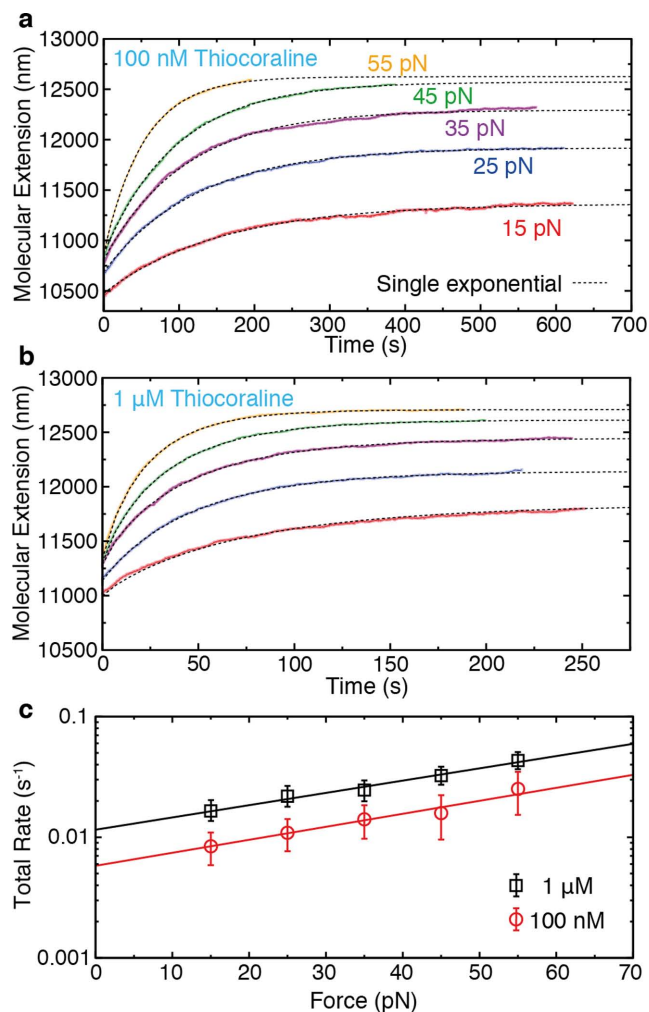
On the other hand, the slow kinetics of Thiocoraline does not allow us to obtain an equilibrium FEC using a standard pulling protocol (Figure 2b). To do so, we performed force-jump experiments in the range 5–60 pN (Figure 2a, inset). In these experiments, the molecule is maintained at a constant force using force-feedback, and the relaxation of the molecular extension is measured at each force in bidirectional experiments (black arrows in Figure 2a, inset). Relaxation

curves of force-jump experiments are well described by a single-exponential function (Supplementary Section S1) and their asymptotic value can be used to determine the equilibrium extension at each force (Figure 2a, black points). A comparison between the reconstructed equilibrium FEC and the metastable FEC (Figure 2a, red curve and black points, respectively) illustrates the pronounced effect of force in shifting the binding equilibrium of Thiocoraline. Pulling experiments performed up to a higher stretching force of 70 pN show that Thiocoraline binds DNA at forces above the overstretching transition (Supplementary Section S2).

### Binding thermodynamics and elastic properties of bis-intercalated DNA

The slow kinetics of Thiocoraline makes possible to determine the elastic properties of intercalated DNA in conditions of metastability for a fixed fraction of bound ligand. In our experiments, such fraction is initially equilibrated at a force of 2 pN and the molecule is repeatedly pulled up to 40 pN fast enough to have that fraction quenched throughout the pulling cycle. We performed pulling experiments over 3 decades of ligand concentration (Supplementary Figure S3) and fitted the metastable FECs to the extensible WLC model (see the Materials and Methods section). The contour length increased with Thiocoraline concentration (Figure 2c) reflecting the varying fraction of bound intercalator (6). A fit of the experimental data to the MGVH binding isotherm yielded a binding affinity  $K_d = 77 \pm 13 \text{ nM}$  and a binding-site size  $n = 3.84 \pm 0.12 \text{ bp}$  (see the Materials and Methods section) (15,36). The latter is in good agreement with the values previously determined for other DNA bis-intercalators and from X-ray crystallography (18,38–39).

Fits to the WLC model also yielded the persistence length ( $L_p$ ) and stretch modulus ( $S$ ) of a bis-intercalated DNA molecule. The persistence length remained constant within experimental errors over the whole range of concentrations investigated (Figure 2d). This is in contrast with previous studies for other mono- and bis-intercalators in which a strong decrease of the persistence length with ligand concentration was reported (15–16,22–23,40). In these works, equilibrium FECs were used to obtain the elastic properties of intercalated DNA. However, in the equilibrium ensemble, the fraction of intercalated drug changes as force is increased, leading to a continuously changing contour length within a pulling cycle, thereby giving lower effective persistence lengths in the fits. A constant persistence length has also been reported for YOYO-1 in recent studies performed at low forces ( $F < 10 \text{ pN}$ ) and low salt condition (10 mM) where force-dependent intercalation is negligible (17). We confirmed our observation by performing experiments with a magnetic tweezers setup in the low-force regime (0.01–10 pN), obtaining equivalent results (Figure 2d, blue). Furthermore, we performed experiments in which the fraction of bound intercalator was increased by equilibrating the molecule at a higher initial force (Figure 2d, inset), finding again that persistence length is independent of the fraction of intercalated drug regardless of the method used to increase the binding ratio (i.e. ligand concentration or force). Finally, from our high-force ( $F > 10 \text{ pN}$ ) measurements we



**Figure 3.** Total macroscopic kinetic rate ( $k_{\text{tot}}$ ) from force-jump experiments. (a, b) DNA extension as a function of time after increasing the force from 2 pN to a preset value of: 15 pN (red), 25 pN (blue), 35 pN (purple), 45 pN (green), 55 pN (yellow). Relaxations are well described by a single-exponential function and allow us to obtain  $k_{\text{tot}}$  at the two concentrations investigated: 100 nM in panel (a) and 1  $\mu\text{M}$  in panel (b). (c) Force dependence of  $k_{\text{tot}}$  at the two investigated concentrations. The intercalation rate increases exponentially with force with similar slopes for both concentrations.

also obtained the stretch modulus of bis-intercalated DNA, finding a systematic increase in the stretching rigidity with ligand concentration (Figure 2e and Supplementary Figure S4). A fit of the reconstructed equilibrium FEC shown in Figure 2a to the WLC model gives instead  $l_0 = 12200 \pm 200$  nm,  $l_p = 8.1 \pm 0.6$  nm,  $S = 745 \pm 320$  pN, where it can be observed that the apparent persistence length and stretch modulus are markedly reduced (and the contour length increased) due to the effective effect of force-induced intercalation.

#### Total kinetic rate and direct off-rate measurements

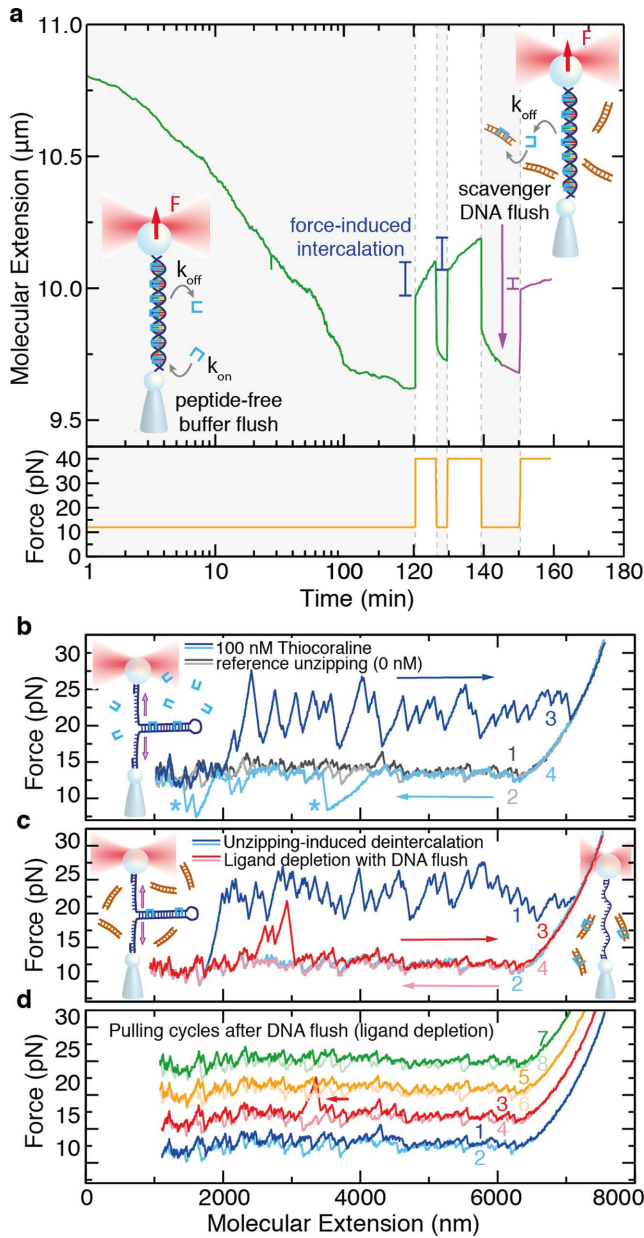
To characterize the remarkably slow kinetics of Thiocoraline we performed force-jump experiments at two different concentrations (Figure 3a and b). The relaxation of the molecular extension at each force allowed us to characterize

the total kinetic rate of intercalation as a function of force. Relaxation curves are well described by single-exponential kinetics, with a total rate that increases exponentially with force at both concentrations (Figure 3c). For a bimolecular reaction where  $k_{\text{on}}$  is much larger than  $k_{\text{off}}$ , the total rate ( $k_{\text{tot}} = k_{\text{on}} + k_{\text{off}}$ ) should be proportional to ligand concentration. However, the total kinetic rates observed at 1  $\mu\text{M}$  were found to be only  $\sim 2$ -fold faster than those obtained at 100 nM, whereas at 10 nM we observed an  $\sim 10$ -fold decrease in relation to 100 nM, suggesting that binding is not a purely bimolecular process.

To verify this we directly measured the off-rate  $k_{\text{off}}(F)$  from wash-off experiments. In these experiments, an intercalated DNA molecule is held at a constant force in the optical tweezers setup, while the chamber is flushed with peptide-free buffer (Figure 4a). However, the low solubility and hydrophobicity of Thiocoraline makes difficult to fully remove the peptide from the chamber in such experiments. In order to create an effective depletion of ligands, we flushed scavenger DNA (see the Materials and Methods section). Thus, ligands that unbind from the tethered molecule or from the microfluidics chamber are sequestered and dragged away. For instance, in the experiment shown in Figure 4a, even after 2 h of peptide-free buffer flush, force-induced intercalation occurs if the force is increased to 40 pN. However, a brief flush of 10- $\mu\text{M}$  bp DNA largely blocks this effect in a subsequent force-jump (Figure 4a, purple). This shows that ligand-free buffer flow is not sufficient to wash the chamber in constant-force experiments.

To test how reliable is this approach to generate ligand depletion, we performed unzipping experiments with a long DNA hairpin (6.8 kb). In DNA unzipping experiments the hairpin is pulled from the two ends and base pairs are sequentially disrupted (Figure 4b). In the absence of ligands, the unzipping shows a characteristic sawtooth pattern at a force  $\sim 15$  pN that is quasi-reversible at the experimental pulling speed (Figure 4b; dark gray is unzipping and light gray re-zipping). In the presence of 100-nM Thiocoraline (Figure 4b, dark blue), the unzipping pattern shows a series of peaks at  $F \geq 20$  pN due to the increased DNA stability at the positions where bis-intercalators are bound. In these experiments, the molecule can be re-zipped again by decreasing the distance between beads (Figure 4b, light blue), recovering the typical sawtooth pattern that indicates DNA hybridization. Interestingly, in the presence of Thiocoraline, a few specific locations of the hairpin (indicated with an \*) require lower forces to refold. We attribute this rare events to the formation of competing non-native structures that are locally stabilized by Thiocoraline, and delay the formation of the native hairpin. The molecule can then be repeatedly pulled in the presence of 100-nM Thiocoraline, obtaining equivalent unzipping (dark blue) and re-zipping (light blue) curves.

To test the DNA wash-off method, we flushed the scavenger solution while holding the intercalated DNA molecule partially unzipped. After a 3-min flush, the molecule is unzipped to remove all the intercalated peptide and re-zipped again (Figure 4c, blue). In the following unzipping cycle, only three binding events are detected (Figure 4c, red). Successive pulling cycles are indistinguishable from naked DNA unzipping experiments (Figure 4d),



**Figure 4.** Wash-off experiments using scavenger DNA. (a) DNA extension as a function of time (green) in a wash-off experiment performed at 12 pN (orange) using peptide-free buffer. The extension decreases with time due to the unbinding of Thiocoraline. However, upon increasing the force after 2 h, an exponential elongation is observed, showing that small amounts of ligand remained in the microfluidics chamber. This effect can be suppressed by transiently flushing a high concentration of scavenger DNA (purple). Normal timescale after 100 min. (b) Unzipping the DNA hairpin (6.8 kb) without ligand shows a characteristic quasi-reversible sawtooth pattern due to base-pair disruption (gray). In the presence of Thiocoraline, a series of force peaks is observed during pulling (dark blue) indicative of binding events. During refolding (light blue), kinetic states that temporarily block the hairpin rezipping are sometimes observed (\*). (c) To test the wash-off method, we flushed the scavenger DNA in, and removed the bound ligands by unzipping and rezipping the hairpin (blue). If the hairpin is immediately unzipped again (red), very few binding events remain. (d) Consecutive unzipping cycles performed during the following 10 min do not show binding (cycles shifted upward for clarity), and only occasionally (<20%) individual binding events are observed (arrow).

and only individual binding events (e.g. red FEC in Figure 4d) are occasionally observed (<20%). This assay is therefore particularly well suited to detect subnanomolar concentrations of ligands, as even single-binding events give a clear footprint in the DNA unzipping pattern (i.e. over 6800 potential binding sites). We conducted similar experiments starting with a lower initial concentration of 10-nM Thiocoraline with equivalent results (Supplementary Figure S5). Consequently, by using scavenger DNA in the wash-off experiments, it is possible to create a depletion of ligand in the experimental area and, as shown below, perform accurate off-rate measurements at different forces even for an initially high concentration of ligand (100 nM).

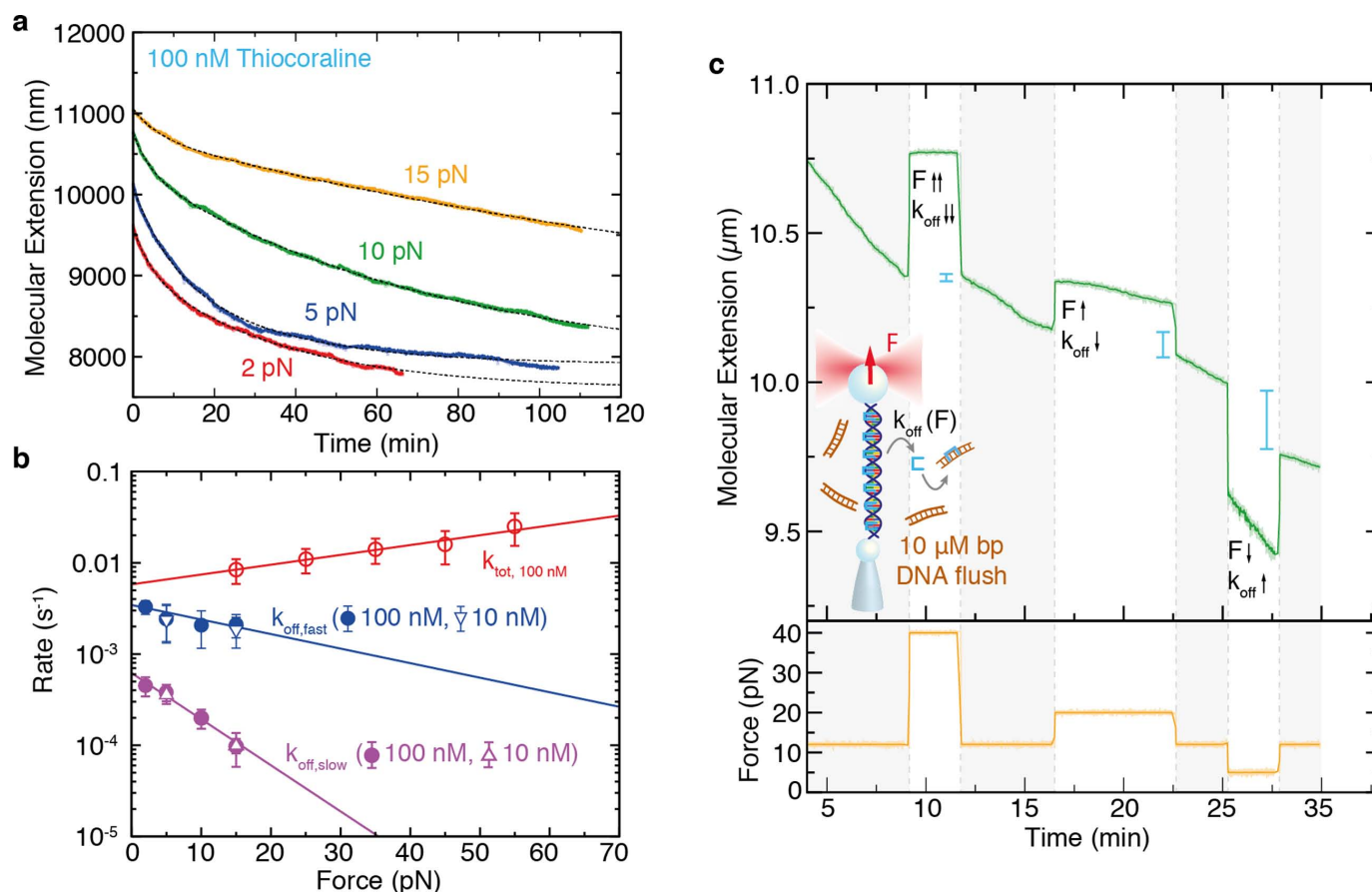
### The kinetic off-rate ( $k_{off}$ ) reveals a binding intermediate

We investigated the off-rate kinetics of Thiocoraline at different forces in the DNA stretching setup. We first followed the unbinding reaction in DNA molecules equilibrated at 100-nM Thiocoraline, where the initial binding density is large (Figure 5a). The off-rate kinetics are remarkably slow at all forces (in the order of hours), showing a strong force-dependence. A single-exponential function fails to fit the experimental data (Supplementary Figure S6a), but experiments are well described by a double-exponential relaxation (Figure 5b) whose amplitude  $A$  is dominated by the slower off-rate process ( $A_{slow} > 75\%$ , for all forces). To verify that the observed off-rates ( $k_{off, fast}$  and  $k_{off, slow}$ ) are independent of the initial binding fraction, we repeated measurements with DNA molecules equilibrated at a much lower Thiocoraline concentration (10 nM; Supplementary Figure S6b), finding rates that are compatible with the ones obtained at 100 nM (Figure 5b). The strong force dependence of the unbinding rate can be easily visualized by changing the applied force in the course of a single experiment (Figure 5c), allowing to slow down (speed up) the unbinding reaction as force is increased (decreased).

What is the origin of the two timescales observed in the wash-off experiments? The presence of two off-rates, together with the previous observation of an on-rate that is not proportional to ligand concentration, can be rationalized by considering the existence of an intermediate state in the binding reaction. A minimal model in which the binding of Thiocoraline proceeds through a mono-intercalated intermediate state (Figure 6a) and that considers three microscopic kinetic rates ( $\alpha_{on}$ ,  $\alpha'_{on}$  and  $\alpha_{off}$ ) is sufficient to reconcile all observations. The model can be analytically solved (Supplementary Section S5), and the values of  $\alpha'_{on}(F)$  and  $\alpha_{off}(F)$  directly determined from the macroscopic rates  $k_{off, fast}(F)$  and  $k_{off, slow}(F)$  measured in wash-off experiments (Figure 6b). The microscopic kinetic rates depend exponentially with the applied force  $F$  and can be fit to the expression

$$\alpha_{on'/off}(F) = \alpha_{on'/off}(0) \exp(Fx_{on'/off}^\dagger/k_B T), \quad (7)$$

where  $k_B$  is the Boltzmann constant and  $T$  is the temperature. Using the microscopic rates determined in this way, the master equations can be solved to reproduce three types of experiments with different initial conditions: (i) wash-off experiments, (ii) force-jump experiments and (iii) intercalation

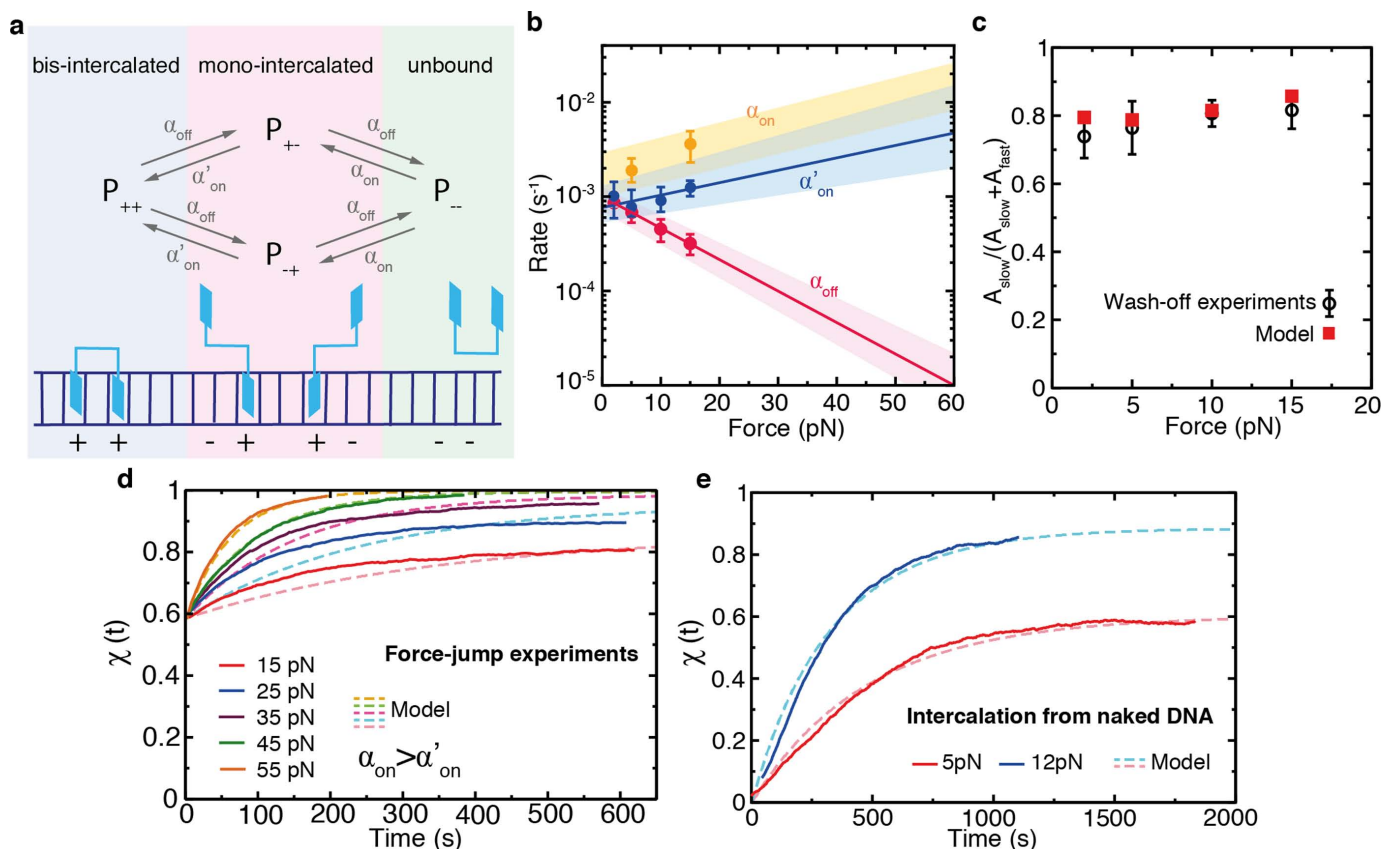


**Figure 5.** Off-rate kinetics of bis-intercalation. **(a)** DNA extension as a function of time in wash-off experiments starting at 100-nM Thiocoraline at four different forces, and fit to a double-exponential function (dotted gray). **(b)** Force dependence of the macroscopic kinetic rates: fast off-rate (blue), slow off-rate (purple). In red,  $k_{\text{tot}}$  from force-jump experiments is plotted as a reference. A fit to Equation (7) (solid lines) yields the macroscopic zero-force rates ( $k_{\text{off,slow}}(0) = 6.2(7) \cdot 10^{-4} \text{ s}^{-1}$ ,  $k_{\text{off,fast}}(0) = 3.4(4) \cdot 10^{-3} \text{ s}^{-1}$ ) and transition-state distances ( $x_{\text{off,slow}}^{\ddagger} = -0.48(5) \text{ nm}$ ,  $x_{\text{off,fast}}^{\ddagger} = -0.15(4) \text{ nm}$ ) for each process. Values are reported as mean  $\pm$  SD,  $N \geq 3$ . **(c)** Wash-off experiment in which different forces (yellow) are sequentially tested within the same kinetic trace. The molecular extension (green) is initially followed at a constant force of 12 pN. The slope of the molecular extension is proportional to the number of bis-intercalators that unbind from DNA per unit time and hence it is indicative of how the off-rate ( $k_{\text{off}}$ ) changes with force. Increasing the force from 12 pN (shaded in gray) to a higher force value (40 pN, 25 pN) causes a decrease on the slope and hence in  $k_{\text{off}}$ , whereas decreasing the force to 5 pN causes the opposite effect. We also indicated the measured change in molecular extension (blue lines) when the force is set again at 12 pN after a few minutes at 40 pN, 25 pN and 5 pN, showing that ligands unbind faster as force is decreased.

kinetics starting from naked DNA. Remarkably enough, the amplitudes ( $A_{\text{slow}}$  and  $A_{\text{fast}}$ ) observed in the wash-off experiments are in excellent agreement with the amplitudes predicted by the model if the intermediate state is assumed to have an extension 0.4 times that of a bis-intercalated state (Figure 6c), supporting the hypothesis that the intermediate state is mono-intercalated (with an elongation  $\sim 50\%$  that of the bis-intercalated state which is 0.68 nm). Similarly, the fractional elongation  $\chi(t)$  (0 for naked DNA, and 1 for fully bis-intercalated DNA) can be calculated from the molecular extension  $x(t)$  of kinetic experiments (see Supplementary Section S5) and compared to the prediction of the model. The model describes well the equilibrium extensions and kinetics of the force-jump experiments (Figure 6d), and of constant-force experiments that start with a naked DNA molecule (Figure 6e). Overall, this three-state model captures the essential features of the experiments providing a mechanistic explanation for the binding kinetics of Thiocoraline.

The observed increase of  $\alpha'_{\text{on}}$  with force is characterized by a distance  $x_{\text{on}}^{\ddagger} = 0.12(5) \text{ nm}$ , whereas  $\alpha_{\text{off}}$  steeply decreases with force ( $x_{\text{off}}^{\ddagger} = -0.29(2) \text{ nm}$ ). An off-rate that decreases with force is in striking contrast with previous studies of the mono-intercalator Actinomycin (19) and a threading intercalator (21), where no intermediate was observed. Whereas force facilitates both the binding and unbinding of these ligands due to base-pair destabilization (19,21), this is not the case for Thiocoraline. Our data show that DNA stretching kinetically stabilizes the bound state, suggesting that bis-intercalators behave like a molecular ‘wedge’ (i.e. applying force clamps the ligand, whereas compression pushes the intercalator out). According to our model, the characteristic elongation of each intercalating moiety can be inferred from the transition-state distances ( $x_{\text{on}}^{\ddagger}$ , and  $x_{\text{off}}^{\ddagger}$ ). For the transition between the intermediate state and the fully bound state, this yields a value ( $\Delta x_{\text{eq}} = x_{\text{on}}^{\ddagger} - x_{\text{off}}^{\ddagger} = 0.41(8) \text{ nm}$ ). This is in good agreement with the expected





**Figure 6.** A kinetic three-state model with microscopic rates  $\alpha_{on}$ ,  $\alpha'_{on}$  and  $\alpha_{off}$ . (a) Scheme of the kinetic model representing the different accessible states: unbound ligand (green), fully bis-intercalated (blue) and mono-intercalated intermediate (red). The binding of the first intercalating moiety is bimolecular (with a rate  $\alpha_{on}$  proportional to ligand concentration), whereas the transition from the mono-intercalated intermediate to a fully bis-intercalated state is unimolecular (with rate  $\alpha'_{on}$ ). On the other hand, the removal of any of the two intercalative moieties is driven by a single off-rate  $\alpha_{off}$ . (b) Microscopic rates  $\alpha'_{on}$ ,  $\alpha_{on}$  and  $\alpha_{off}$  that best describe the experiments. The values of  $\alpha'_{on}$  (blue points) and  $\alpha_{off}$  (red points) are determined from the macroscopic rates measured in wash-off experiments (Figure 5b). A fit to Equation (7) (red and blue lines) is used to extrapolate the high force values and their associated uncertainty range (shaded area). The rates  $\alpha_{on}$  are determined from intercalation experiments with naked DNA (panel (e) and Supplementary Section S5). (c) Ratio of amplitudes of the fast and slow macroscopic off-rates measured in the wash-off experiments (black) compared with the model prediction (red). (d) Fractional elongation of the force-jump experiments from Figure 3a (solid lines) and comparison to the kinetics predicted by the model (dashed lines). (e) Fractional elongation of intercalation experiments (100-nM Thiocoraline) starting from a naked DNA molecule at 5 and 15 pN (solid lines) and comparison to the kinetics predicted by the model (dashed lines).

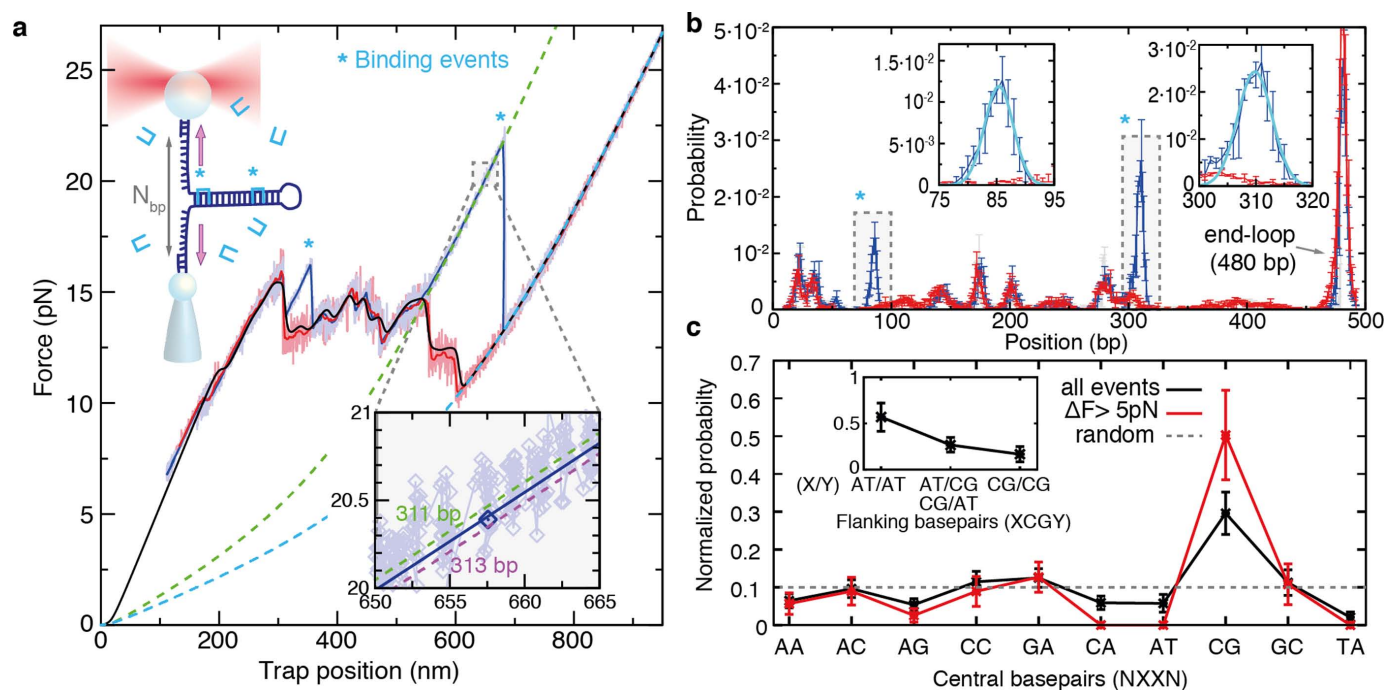
distance change for this transition  $\sim 0.4$  nm (equal to the additional 60% of elongation to reach a full bis-intercalated state). From the value of  $\alpha_{on}$  we can also infer the rate of association at zero force ( $k_a = \frac{\alpha_{on}(0)}{[Thio]} = 1.7(3) \cdot 10^4 M^{-1}s^{-1}$ ). This rate is slow if compared to mono-intercalators or charged bis-intercalators (YOYO-1) (41,42), but is in good agreement with other structurally similar non-charged bis-intercalators (43,44). Finally, the presence of a binding intermediate reproduces well the titration curve for the DNA contour length (Figure 2c).

### Single-molecule footprinting of preferred binding sites

To determine the preferred binding sites of Thiocoraline, we developed a single-molecule footprinting assay based on DNA unzipping. Unzipping-based methods have proven useful to locate binding of nucleosomes and restriction enzymes with base-pair resolution (45,46). However, to the best of our knowledge, this is the first time that they are used to directly determine the binding site of small ligands.

We repeatedly unzipped a 480-bp DNA hairpin in the presence of 5-nM Thiocoraline (Figure 7a). In the absence of ligand the unzipping and re-zipping curves fully overlap (Supplementary Figure S8a, top). However, in the presence of ligand, the unzipping curves show a series of force peaks indicative of individual binding events (Figure 7a, marked with \*). Additional pulling curves are shown in Supplementary Figure S8, where binding events to other positions along the hairpin are observed. Experiments are typically performed at a concentration where a few binding events per pulling cycle are detected. This is done to determine the position of the peaks accurately and to ensure that binding events are tested homogeneously along the DNA fragment. The overall footprinting map of the molecule and its sequence selectivity is recovered by collecting a large number of pulling cycles (Supplementary Figure S9) and correlating the position of the binding events to the DNA hairpin sequence.

First, the position of the binding events was determined from a statistical analysis of the FDCs (see the Materi-



**Figure 7.** Single-molecule footprinting reveals preference for CpG steps. (a) DNA unzipping (blue) and re-zipping (red) of a 480-bp DNA hairpin in the presence of 5-nM Thiocoraline. Pulling speed is 70 nm/s and data are recorded at 1 kHz (light colors) and filtered to 10 Hz (dark colors). The binding of a ligand brings an extra stabilization energy, causing a force peak in the unzipping curve (\*). A theoretical prediction of the equilibrium FDC is shown in black and used for alignment. An FJC curve corresponding to a fully unzipped hairpin ( $n = 480$  bp, cyan dashed) and one that passes close to a binding event ( $n = 311$  bp, green dashed) are plotted as a reference. Inset: for each datapoint of the unzipping/re-zipping curves we found the most probable number of open base pairs  $n$ , by finding the theoretical curve that passes closest to it. For the highlighted datapoint (blue), the distance is minimized for  $n = 312$  bp (blue line). (b) Histogram of  $n$  values of the unzipping curve (blue) and re-zipping curve (red) from panel (a). The histograms match each other except for the positions in which there is a binding event (highlighted in gray, \*). Insets: zoom of a binding peak and fit to a Gaussian function (cyan). (c) Normalized probability of clamping sites XX: including all binding events (black), or by selecting events with rupture forces 5 pN higher than the average unzipping force (red). The dashed line indicates the distribution expected for a ligand that uniformly binds all sites. Inset: normalized probability of observing A-T bp or C-G bp flanking the preferred binding site XCGY. ( $N = 142$  binding events, 104 unzipping curves and  $\sim 1$  binding event per cycle).

als and Methods section) (28,47). Briefly, the experimental FDCs were aligned to the theoretical prediction based on the nearest-neighbor model (Figure 7a, black) (33). For each experimental datapoint we then determined the most probable number of unzipped base pairs  $n$ , by finding the elastic curve (see the Materials and Methods section) that passes closest to the datapoint (Figure 7a, inset) (28,37,47). In this way, we assigned a number of open base pairs  $n$  to each experimental datapoint (Figure 7b). FDCs are reversible at this pulling speed, and consequently the unzipping and re-zipping histograms (Figure 7b, red and blue, respectively) overlap except for those positions in which there is a binding event (Figure 7b and Supplementary Figure S8b, blue peaks). By fitting these extra peaks to Gaussian distributions the position of each binding event could be identified (Figure 7b, insets). In this way, we determined the probability of binding for each different dinucleotide pair (Figure 7c, black). Remarkably, we find that Thiocoraline bis-intercalates all 10 possible dinucleotides, but shows a clear preference to clamp CpG steps. Moreover, by analyzing the neighboring bases we also find that flanking A-T base pairs are favored (i.e. ACGT, TCGA, ACGA/TCGT) (Figure 7c, inset). The sensitivity of the technique can be increased by only selecting events that show rupture forces greater than a given threshold (e.g. 5 pN) above the local unzipping force. As expected, this analysis shows a higher peak

at the specific binding sites (Figure 7c, red), confirming the selectivity of Thiocoraline for clamping CpG steps. An estimation of the measurement error can be found by considering alternative distances between the position of the binding peaks and the clamping positions (Supplementary Figure S10).

So far, the existence of specific binding sites for Thiocoraline was rather controversial. Although the peptide showed a general preference for G-C-rich regions, protected sites were not observed in classical DNA footprinting experiments, suggesting that it had little sequence selectivity (31). On the other hand, fluorescent melting experiments with short oligonucleotides showed a clear preference for G-C-rich oligos and particularly those containing CpG steps (32). Our direct measurement of binding positions from unzipping experiments demonstrates that Thiocoraline has a specific affinity for CpG steps (particularly those flanked by A-T base pairs), although it can also clamp a large number of alternative dinucleotide steps, in agreement with our stretching results that show binding saturation at high enough concentrations (Figure 2c). Moreover, this binding pattern seems remarkably similar to that exhibited by the quinoxaline bis-intercalator Echinomycin that shares several structural similarities with Thiocoraline that might confer specificity for these motifs (5,48–49). In addition, information on cooperative binding should also be accessible

with this technique. The presence of cooperative effects between contiguous binding sites is expected to show both: (i) a higher frequency of binding events at these positions and (ii) increased rupture forces than in non-cooperative sites. In fact, the DNA fragment used in these experiments contains a motif (TCGTACGA) that corresponds to two juxtaposed specific binding sites. Remarkably, binding events of Thiocoraline at this position are particularly abundant and show high rupture forces, suggesting cooperative binding. Although more experiments are needed to confirm this result, general synergistic effects should be detectable with our technique.

Overall, our footprinting assay gives direct access to individual binding events of ligands due to the higher thermodynamic stability that they induce upon binding DNA. Consequently, specific and non-specific binding sites are observed in a single experiment, without relying on indirect measurements such as differential access of cleavage agents, shifts in gel-mobility (difficult for small ligands) and/or radiolabeling (24–26). Similarly, DNA templates can be much longer than typical gel footprinting assays, and a statistically significant amount of data can be collected in a few hours. We therefore anticipate that this method may be widely applicable to determine the sequence specificity of a large number of small ligands that are difficult to characterize with traditional assays.

## DISCUSSION

Thiocoraline, an anticancer peptide that bis-intercalates DNA, exhibits an unusual slow kinetics of binding. We have exploited this feature to clarify the mechanical properties of bis-intercalated DNA in the low- and high-force stretching regimes (16–17,23,40). Contrary to previous studies in which equilibrium FECs were used, we have worked with reversible metastable FECs in which the fraction of ligand is kept constant. We found that the bending rigidity of DNA remains unchanged, which we interpret in terms of a compensation between the electrostatic effect of unwinding the DNA double helix (larger interphosphate distance) and the straightening induced by base-stacking interactions with the intercalated chromophores.

We also showed that Thiocoraline has a high binding affinity to DNA finding compatible values from equilibrium and kinetic measurements ( $K_d = 77 \pm 13$  nM) and reaching the nearest-neighbor exclusion limit of  $n = 4$  at submicromolar concentration (50). Remarkably enough, the affinity to DNA is dominated by a very slow macroscopic off-rate ( $k_{\text{off}}(0) = 6.2(7) \cdot 10^{-4} \text{ s}^{-1}$ ) that steeply decreases with force. This allows us to kinetically control ligand unbinding by changing the applied force, obtaining extremely long-lived complexes at even moderate stretching forces ( $F < 20$  pN). From a structural point of view, this indicates that a pulling force further accommodates the bound ligand, stabilizing the bis-intercalated tetranucleotide motif. This might be common to other cytotoxic DNA bis-intercalators (e.g. Echinomycin, Triostin), and we hypothesize it is very likely related to the high toxicity of this family of anticancer drugs that show activity at concentrations much below their  $K_d$  values (e.g. Thiocoraline  $\text{IC}_{50} \sim 200$  pM (31)). To accurately determine these unbinding rates, we have used a combina-

tion of stretching and unzipping experiments that allowed us to ensure that subnanomolar traces of bis-intercalator did not remain in the wash-off experiments. This methodology might be particularly relevant in future studies of other highly hydrophobic and low solubility ligands (28). Our experiments also reveal that the binding pathway proceeds through a long-lived mono-intercalated intermediate (lifetime  $\sim 10$  min) which elongates DNA by 0.28 nm (which is 0.4 times the elongation of a fully bis-intercalated ligand, 0.68 nm) contributing to the long lifetime of the peptide (hours). The stretching-unzipping methodology allows us to extract the force dependence of the microscopic rates ( $\alpha_{\text{on}}(F)$ ,  $\alpha'_{\text{on}}(F)$  and  $\alpha_{\text{off}}(F)$ ) from the characteristic timescales of kinetic experiments performed with varying conditions.

To determine the sequence selectivity of the ligand, we developed a novel DNA unzipping assay to perform single molecule DNA footprinting experiments. Using this assay we could directly locate single-binding events finding that Thiocoraline shows a higher preference to clamp CpG steps, particularly those flanked by A-T base pairs. We foresee that this assay might be of extreme utility to determine the binding site of small ligands as it gives direct access to thermodynamically stabilized sites without requirement of cleavage enzymes and radiolabeling techniques, allowing to rapidly recover the binding sites in DNA fragments 10-fold longer than typical footprinting templates (25,26).

So far, Thiocoraline has been observed to interfere with primer extension by DNA polymerase  $\alpha$ , showing a very prolonged inhibition of DNA replication even after 48-h removal of the drug (51), in contrast with the typical topoisomerase inhibition mechanism of most DNA intercalators (e.g. doxorubicin). However conclusive studies are lacking, being unclear how its selectivity for cancer cells arises (49,51–52). In view of these results, we suggest that the increased binding stability of the ligand at moderate stretching forces might create long-lived kinetic roadblocks that stall the progression of replication and transcription enzymes at picomolar drug concentrations. Moreover, we have shown that these roadblocks might also happen through the stabilization of non-native kinetic structures that are particularly present in DNA-melted regions during transcription and replication. More importantly, the determined sequence specificity suggests that Thiocoraline might accumulate at specific genomic locations rich in CpG steps. Notably, CpG sites are sparsely distributed in human genomes, accumulating in many regulatory regions and having much lower frequencies elsewhere (53). We hypothesize that this inhomogeneous distribution of CpG islands might be related to the selectivity of Thiocoraline for carcinogenic cell lines (51).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank A. Martinez-Monge and M. Ribezzi-Crivellari for support on the theoretical FDC prediction, and V. Croquette for a critical reading of the manuscript.

## FUNDING

The European Research Council [Infernos 308850, MagReps 267 862]; Icrea Academia 2013. Funding for open access charge: European Research Council [FP7 Infernos 308850].

*Conflict of interest statement.* None declared.

## REFERENCES

- Helleday, T., Petermann, E., Lundin, C., Hodgson, B. and Sharma, R.A. (2008) DNA repair pathways as targets for cancer therapy. *Nat. Rev. Cancer*, **8**, 193–204.
- Koster, D.A., Palle, K., Bot, E.S., Bjornsti, M.-A. and Dekker, N.H. (2007) Antitumour drugs impede DNA uncoiling by topoisomerase I. *Nature*, **448**, 213–217.
- Chenoweth, D.M. and Dervan, P.B. (2009) Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 13175–13179.
- Pang, B., Qiao, X., Janssen, L., Velds, A., Groothuis, T., Kerkhoven, R., Nieuwland, M., Ovaa, H., Rottenberg, S., vanTellingen, O. *et al.* (2013) Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. *Nat. Commun.*, **4**, 1908.
- Leng, F., Chaires, J.B. and Waring, M.J. (2003) Energetics of echinomycin binding to DNA. *Nucleic Acids Res.*, **31**, 6191–6197.
- Lerman, L. (1961) Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.*, **3**, 18–30.
- Tse, W.C. and Boger, D.L. (2004) Sequence-selective DNA recognition: natural products and nature's lessons. *Chem. Biol.*, **11**, 1607–1617.
- Hurley, L.H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 188–200.
- Adhireksan, Z., Davey, G.E., Campomanes, P., Groessler, M., Clavel, C.M., Yu, H., Nazarov, A.A., Yeo, C. H.F., Ang, W.H., Dröge, P. *et al.* (2014) Ligand substitutions between ruthenium–cymene compounds can control protein versus DNA targeting and anticancer activity. *Nat. Commun.*, **5**, 3462.
- Holman, G.G., Zewail-Foote, M., Smith, A.R., Johnson, K.A. and Iverson, B.L. (2011) A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant. *Nat. Chem.*, **3**, 875–881.
- Warren, C.L., Kratochvil, N.C., Hauschild, K.E., Foister, S., Brezinski, M.L., Dervan, P.B., Phillips, G.N. and Ansari, A.Z. (2006) Defining the sequence-recognition profile of DNA-binding molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 867–872.
- Fechter, E.J., Olenyuk, B. and Dervan, P.B. (2004) Design of a sequence-specific DNA bisintercalator. *Angew. Chem. Int. Ed. Engl.*, **43**, 3591–3594.
- Phillips, D.R. and Crothers, D.M. (1986) Kinetics and sequence specificity of drug-DNA interactions: an in vitro transcription assay. *Biochemistry*, **25**, 7355–7362.
- vanMameren, J., Gross, P., Farge, G., Hooijman, P., Modesti, M., Falkenberg, M., Wuite, G.J. and Peterman, E.J. (2009) Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 18231–18236.
- Vladescu, I.D., McCauley, M.J., Nuñez, M.E., Rouzina, I. and Williams, M.C. (2007) Quantifying force-dependent and zero-force DNA intercalation by single-molecule stretching. *Nat. Methods*, **4**, 517–522.
- Lipfert, J., Klijnhout, S. and Dekker, N.H. (2010) Torsional sensing of small-molecule binding using magnetic tweezers. *Nucleic Acids Res.*, **38**, 7122–7132.
- Günther, K., Mertig, M. and Seidel, R. (2010) Mechanical and structural properties of YOYO-1 complexed DNA. *Nucleic Acids Res.*, **38**, 6526–6532.
- Kleimann, C., Sischka, A., Spiering, A., Tönsing, K., Sewald, N., Diederichsen, U. and Anselmetti, D. (2009) Binding kinetics of bisintercalator Triostin A with optical tweezers force mechanics. *Biophys. J.*, **97**, 2780–2784.
- Paramanathan, T., Vladescu, I., McCauley, M.J., Rouzina, I. and Williams, M.C. (2012) Force spectroscopy reveals the DNA structural dynamics that govern the slow binding of Actinomycin D. *Nucleic Acids Res.*, **40**, 4925–4932.
- Paik, D.H. and Perkins, T.T. (2012) Dynamics and multiple stable binding modes of DNA intercalators revealed by single-molecule force spectroscopy. *Angew. Chem. Int. Ed. Engl.*, **51**, 1811–1815.
- Almaqashi, A., Paramanathan, T., Lincoln, P., Rouzina, I., Westerlund, F. and Williams, M. (2014) Strong DNA deformation required for extremely slow DNA threading intercalation by a binuclear ruthenium complex. *Nucleic Acids Res.*, **42**, 11634–11641.
- Sischka, A., Toensing, K., Eckel, R., Wilking, S.D., Sewald, N., Ros, R. and Anselmetti, D. (2005) Molecular mechanisms and kinetics between DNA and DNA binding ligands. *Biophys. J.*, **88**, 404–411.
- Murade, C., Subramaniam, V., Otto, C. and Bennink, M.L. (2009) Interaction of oxazole yellow dyes with DNA studied with hybrid optical tweezers and fluorescence microscopy. *Biophys. J.*, **97**, 835–843.
- Galas, D.J. and Schmitz, A. (1978) DNAase footprinting a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.*, **5**, 3157–3170.
- Hampshire, A.J., Rusling, D.A., Broughton-Head, V.J. and Fox, K.R. (2007) Footprinting: a method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods*, **42**, 128–140.
- Connaghan-Jones, K.D., Moody, A.D. and Bain, D.L. (2008) Quantitative DNase footprint titration: a tool for analyzing the energetics of protein–DNA interactions. *Nat. Protoc.*, **3**, 900–914.
- Tulla-Puche, J., Góngora-Benítez, M., Bayó-Puxan, N., Francesch, A.M., Cuevas, C. and Albericio, F. (2013) Enzyme-labile protecting groups for the synthesis of natural products: solid-phase synthesis of Thiocoraline. *Angew. Chem. Int. Ed. Engl.*, **125**, 5838–5842.
- Camunas-Soler, J., Frutos, S., Bizarro, C.V., deLorenzo, S., Fuentes-Perez, M.E., Ramsch, R., Vilchez, S., Solans, C., Moreno-Herrero, F., Albericio, F. *et al.* (2013) Electrostatic binding and hydrophobic collapse of peptide–nucleic acid aggregates quantified using force spectroscopy. *ACS Nano*, **7**, 5102–5113.
- Albericio, F., Tulla-Puche, J., Auriemma, S. and Falciani, C. (2013) Orthogonal chemistry for the synthesis of thiocoraline-triostin hybrids. Exploring their structure-activity relationship. *J. Med. Chem.*, **56**, 5587–5600.
- Zolova, O.E., Mady, A.S. and Garneau-Tsodikova, S. (2010) Recent developments in bisintercalator natural products. *Biopolymers*, **93**, 777–790.
- Boger, D.L., Ichikawa, S., Tse, W.C., Hedrick, M.P. and Jin, Q. (2001) Total syntheses of thiocoraline and BE-22179 and assessment of their DNA binding and biological properties. *J. Am. Chem. Soc.*, **123**, 561–568.
- Negri, A., Marco, E., García-Hernández, V., Domingo, A., Llamas-Saiz, A.L., Porto-Sandá, S., Riguera, R., Laine, W., David-Cordonnier, M.-H., Bailly, C. *et al.* (2007) Antitumor activity, X-ray crystal structure, and DNA binding properties of thiocoraline A, a natural bisintercalating thiodipeptide. *J. Med. Chem.*, **50**, 3322–3333.
- Huguet, J.M., Bizarro, C.V., Forn, N., Smith, S.B., Bustamante, C. and Ritort, F. (2010) Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 15431–15436.
- Gosse, C. and Croquette, V. (2002) Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys. J.*, **82**, 3314–3329.
- Bouchiat, C., Wang, M., Allemand, J., Strick, T., Block, S. and Croquette, V. (1999) Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.*, **76**, 409–413.
- McGhee, J.D. and vonHippel, P.H. (1974) Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.*, **86**, 469–489.
- Bosco, A., Camunas-Soler, J. and Ritort, F. (2014) Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res.*, **42**, 2064–2074.
- Ughetto, G., Wang, A.H., Quigley, G.J., van derMarel, G.A., vanBoom, J.H. and Rich, A. (1985) A comparison of the structure of echinomycin and triostin A complexed to a DNA fragment. *Nucleic Acids Res.*, **13**, 2305–2323.

39. Spielmann, H.P., Wemmer, D.E. and Jacobsen, J.P. (1995) Solution structure of a DNA complex with the fluorescent bis-intercalator TOTO determined by NMR spectroscopy. *Biochemistry*, **34**, 8542–8553.
40. Rocha, M., Ferreira, M. and Mesquita, O. (2007) Transition on the entropic elasticity of DNA induced by intercalating molecules. *J. Chem. Phys.*, **127**, 105108.
41. Wakelin, L. and Waring, M. (1980) Kinetics of drug-DNA interaction: dependence of the binding mechanism on structure of the ligand. *J. Mol. Biol.*, **144**, 183–214.
42. Reuter, M. and Dryden, D.T. (2010) The kinetics of YOYO-1 intercalation into single molecules of double-stranded DNA. *Biochem. Biophys. Res. Commun.*, **403**, 225–229.
43. Fox, K.R. and Waring, M.J. (1984) Stopped-flow kinetic studies on the interaction between echinomycin and DNA. *Biochemistry*, **23**, 2627–2633.
44. Boger, D.L. and Saionz, K.W. (1999) DNA binding properties of key sandramycin analogues: systematic examination of the intercalation chromophore. *Bioorg. Med. Chem.*, **7**, 315–321.
45. Koch, S.J., Shundrovsky, A., Jantzen, B.C. and Wang, M.D. (2002) Probing protein-DNA interactions by unzipping a single DNA double helix. *Biophys. J.*, **83**, 1098–1105.
46. Hall, M.A., Shundrovsky, A., Bai, L., Fulbright, R.M., Lis, J.T. and Wang, M.D. (2009) High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.*, **16**, 124–129.
47. Huguet, J., Forns, N. and Ritort, F. (2009) Statistical properties of metastable intermediates in DNA unzipping. *Phys. Rev. Lett.*, **103**, 248106.
48. Van Dyke, M.M. and Dervan, P.B. (1984) Echinomycin binding sites on DNA. *Science*, **225**, 1122–1127.
49. Dawson, S., Malkinson, J.P., Paumier, D. and Searcey, M. (2007) Bisintercalator natural products with potential therapeutic applications: isolation, structure determination, synthetic and biological studies. *Nat. Prod. Rep.*, **24**, 109–126.
50. Persil, Ö. and Hud, N.V. (2007) Harnessing DNA intercalation. *Trends Biotechnol.*, **25**, 433–436.
51. Erba, E., Bergamaschi, D., Ronzoni, S., Faretta, M., Taverna, S., Bonfanti, M., Catapano, C., Faircloth, G., Jimeno, J. and D'Incalci, M. (1999) Mode of action of thiocoraline, a natural marine compound with anti-tumour activity. *Br. J. Cancer*, **80**, 971–980.
52. Wyche, T., Dammalapati, A., Cho, H., Harrison, A., Kwon, G., Chen, H., Bugni, T. and Jaskula-Sztul, R. (2014) Thiocoraline activates the Notch pathway in carcinoids and reduces tumor progression in vivo. *Cancer Gene Ther.*, **21**, 518–525.
53. Esteller, M. (2008) Epigenetics in cancer. *N. Engl. J. Med.*, **358**, 1148–1159.