nature chemistry

ARTICLES

PUBLISHED ONLINE: 25 SEPTEMBER 2011 | DOI: 10.1038/NCHEM.1151

A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant

Garen G. Holman[1†], Maha Zewail-Foote[2†], Amy Rhoden Smith[1], Kenneth A. Johnson[1,3] and Brent L. Iverson[1,3*]

A long-lived and sequence-specific ligand–DNA complex would make possible the modulation of biological processes for extended periods. For this purpose, we are investigating a polyintercalation approach to DNA recognition in which flexible chains of aromatic units thread back and forth repeatedly through the double helix. Here we describe the DNA-binding behaviour of a threading tetra-intercalator. Specific binding was observed on a relatively long DNA strand that strongly favoured a predicted 14 base-pair sequence. Kinetic studies revealed a multistep association process, with sequence specificity that primarily derives from large differences in dissociation rates. The rate-limiting dissociation rate constant of the tetra-intercalator complex dissociating from its preferred binding site was extremely slow, corresponding to a half-life of 16 days. This is one of the longest non-covalent complex half-lives yet reported and, to the best of our knowledge, the longest for a DNA-binding molecule.

Molecules that bind DNA can modulate transcription, repair and replication both *in vitro* and *in vivo*, so targeting specific sequences of DNA with small molecules has been a field of continued interest for over 30 years. For example, Dervan and co-workers developed a series of polyamide molecules that bind in dimeric fashion to DNA with programmable specificity through the minor groove[1,2], with a current focus on understanding the rules for delivery and distribution *in vivo*[3,4]. Other modular approaches include oligonucleotides and peptide nucleic acids that form triple helixes[5–8].

For possible *in vivo* studies in the laboratory, as well as looking ahead to potential therapeutic applications that address long-term chronic diseases, it is of interest to develop DNA-binding molecules that exhibit very long-lived complexes, capable of modulating biological processes on timescales relevant to organism lifetimes. With this in mind, we are investigating threading polyintercalation, a binding mode that places a significant number of functional groups in both DNA grooves. For example, by linking naphthalene diimide (NDI) units, known threading intercalators[9,10], in a head-to-tail fashion with a variety of flexible peptide linkers, a growing family of unique DNA-binding molecules has been developed[11–17]. Threading bisintercalators **1** and **2** were identified previously through library-screening experiments in which the peptide linkers were randomized[12]. Structural characterization of the complexes with their respective preferred DNA-binding sites revealed that sequence-specific recognition was taking place with linkers that spanned four base pairs (bp) through the major and minor grooves of DNA for **1** and **2**, respectively (Fig. 1a,b)[12–14]. This work complements studies with a variety of other synthetic and naturally occurring mono- and bis-DNA intercalators[10,18–24].

Tetra-intercalator **3** was designed with linkers intended to bind with a minor–major–minor groove topology (Fig. 1c). The two β-alanyl-β-alanyl-β-alanyl-L-lysine (-βAla_3-Lys-) minor groove binding linkers were taken from **2** (β-alanine is the biochemical name often used for 3-aminopropanoic acid). Adipic acid with two L-lysine residues attached was chosen as the central major groove binding linker because modelling indicated that this linker would reproduce many of the structural features of the glycyl-glycyl-glycyl-L-lysine linker of **1** and also offer the advantage of introducing C_2 symmetry into **3**, which simplifies the NMR spectra and thus structural analyses. Consistent with the original design, a detailed NMR structural characterization verified threading polyintercalation and revealed the linkers to be distributed in the expected minor–major–minor groove topology (Fig. 1c)[15].

An anticipated consequence of the threading polyintercalator design of **3** is an extremely slow dissociation from its preferred DNA sequence. This expectation is based on the observation that for one of the two internal NDI units to dissociate, the adjacent terminal NDI unit must dissociate first. For the entire molecule to dissociate, the other two NDI units also have to dissociate in a coordinated fashion before rebinding occurs. The improbable nature of this highly concerted set of motions should reduce the dissociation rate significantly. One might expect the association rate to be proportionately slow as a result of the complicated nature of the bound topology. In addition to kinetics, the overall sequence specificity of the interaction has not been addressed previously using a threading tetra-intercalator such as **3**.

Herein, we report footprinting studies that verified sequence-specific binding by **3** in the context of a relatively long piece of DNA that contained a large number of other possible binding sites. Additionally, the detailed kinetic analysis of **3** binding to oligonucleotides with and without its preferred DNA sequence revealed a multistep association as well as sequence specificity and a remarkably slow dissociation process. We report here the slowest (to the best of our knowledge) dissociation half-life of any non-covalent interaction with DNA, which ultimately could lead to prolonged disruption of transcription or other biological processes.

Results

DNase I footprinting. DNase I footprinting was used to investigate the ability of **3** to bind its 14 bp preferred site within the context of a 467 bp DNA fragment. A clear footprint for **3** binding selectively to

[1]Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, USA, [2]Department of Chemistry and Biochemistry, Southwestern University, Georgetown, Texas 78626, USA, [3]Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712, USA; †These authors contributed equally to this work. *e-mail: iversonb@austin.utexas.edu

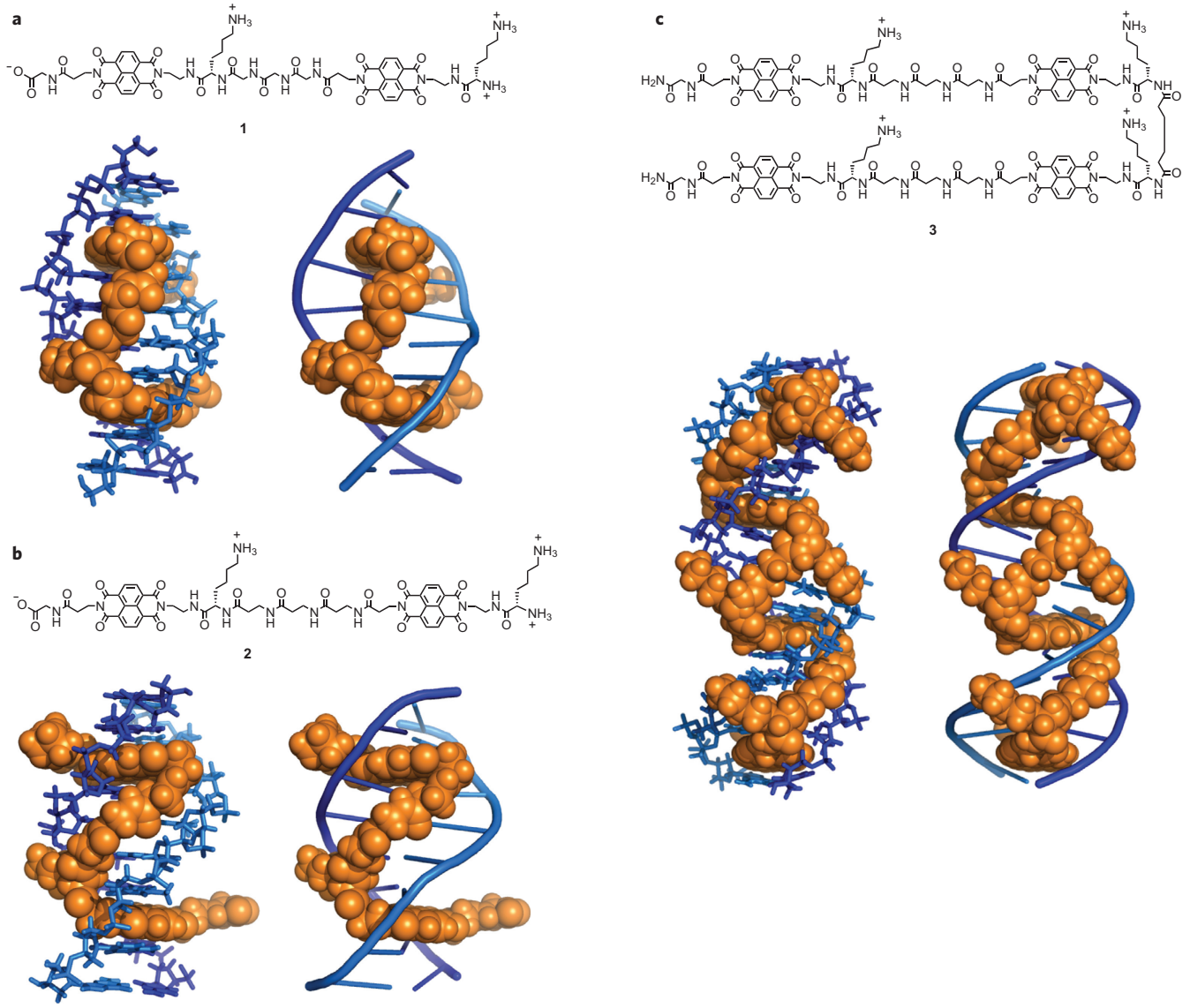

Figure 1 | Chemical structures of the intercalator ligands and models of the resulting ligand–DNA complexes. a, Bisintercalator **1** with a view from the major groove of the **1**-d(CGGTACCG)$_2$ complex[13]. **b**, Bisintercalator **2** with a view from the minor groove of the **2**-d(CGATAAGC),d(GCTTATCG) complex[14]. **c**, Tetra-intercalator **3** with a view of the **3**-d(GATAAGTACTTATC)$_2$ complex[15]. Models and illustrations were produced by the PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger.

sequence A (GATAAGTACTTATC) with high specificity was observed at 63 nM after 48 hours (Fig. 2). Nonspecific binding was not seen, even at 500 nM, which verifies the highly specific and predictable nature of DNA binding by **3**. It is assumed that **3** binds to sequence A with the same threading topology observed in the NMR structural analysis reported previously for this sequence, although for long pieces of DNA there is formally no direct experimental evidence to support this assumption[15]. Quantitative footprinting of binding thermodynamics was precluded by the extensive time needed to approach equilibrium at concentrations near our estimated dissociation constant value[25] and the limited dynamic range of the footprinting method at the low concentrations of labelled DNA required.

Dissociation kinetics. A ^{32}P-labelled oligonucleotide containing the 14 bp tetra-intercalator binding site (**oligo A**, Fig. 3a) was incubated with a stoichiometric amount of **3**, followed by separation using native polyacrylamide gel electrophoresis. The **3–oligo A** complex exhibited retarded migration on the gel. A control oligonucleotide (**oligo B**, Fig. 3a) that contained a randomly chosen sequence was also incubated with **3**, but did not produce an observable gel shift.

Gel-shift assays were used to determine the dissociation rate constant for **3** from **oligo A**, which contained the preferred binding site. In these experiments a 100-fold excess of unlabelled **oligo A** was added to a solution that contained the ^{32}P-labelled **oligo A** occupied by a stoichiometric amount of bound **3** (Fig. 3b). We assume that this technique measures the rate-limiting dissociation rate constant as **3** leaves its binding site. Gel shifts could be analysed up to 34 days before loss of radioactivity reduced the signal-to-noise ratio, and thus the dynamic range of the assay, below useful levels. The data were fitted to a standard monoexponential decay equation (Supplementary Fig. S1) and revealed a dissociation rate constant (k_d) of $5.0 \pm 0.5 \times 10^{-7}$ s^{-1} at 100 mM NaCl, which corresponds to a half-life of 16 days (Table 1). The observed k_d increased with increasing NaCl concentrations over the range 25 mM to 250 mM, which is not surprising given the four L-lysine residues within **3** that presumably make electrostatic

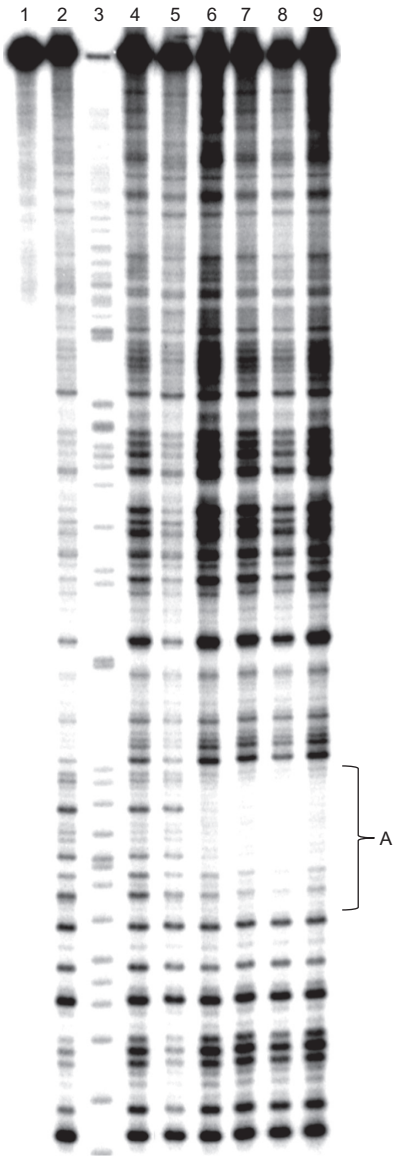

Figure 2 | DNase I footprinting of 3 bound to a 467 bp DNA segment. Lane 1 is a control lane without DNase I or **3**. Lanes 2 and 4 contain DNA digested with DNase I for seven minutes and four minutes, respectively. Lane 3 is an adenine-specific sequencing lane. In lanes 5–9, the DNA fragment was incubated with increasing concentrations of **3** (31, 63, 125, 250 and 500 nM, respectively) for 48 hours followed by digestion with DNase I for seven minutes. The preferred sequence 5′-GATAAGTACTTATC-3′ is highlighted with a bracket (A).

contacts with the phosphodiester backbone of DNA. A plot of log k_d as a function of –log [Na^+] was linear with a slope of 0.59 (Supplementary Fig. S2).

Nonspecific dissociation kinetics. To quantify the sequence-specific behaviour of **3**, the dissociation rate constant from a randomly chosen sequence, **oligo B**, was also determined. As mentioned previously, gel electrophoresis could not be used to monitor the dissociation because a gel shift could not be detected for the **3–oligo B** complex, presumably as a result of rapid dissociation. Instead, the dissociation of **3** from **oligo B** was monitored for 200 seconds using stopped-flow methods in the presence of 2% sodium dodecyl sulfate (SDS) (Fig. 3c)[18,26]. The absorbance trace obtained (at 386 nm) fitted a double exponential function, which suggests a two-step mechanism, with the observed first-order dissociation rate constants of the fast and slow steps, k_1 and k_2, found to be 0.29 s^{-1} and 0.019 s^{-1}, respectively, at 100 mM NaCl. SDS has been shown to enhance the dissociation rate for cationic intercalators that bind DNA, so values reported here should be considered as estimates of the upper limit[27]. Using the slower dissociation step for comparison, these results demonstrate that the dissociation of **3** from **oligo A**, containing the preferred binding site, is roughly 4×10^4 times slower compared to that from **oligo B**, which verifies that the differences in kinetic off-rates is largely, if not exclusively, responsible for sequence selectivity. Salt dependence of the dissociation rate was also examined for the **3–oligo B** complex (Table 1) and the log k_d versus –log [Na^+] plot was linear with a slope of 0.54 (Supplementary Fig. S3).

Association rate measurements. To study the association rates of **3**, a multipronged approach was pursued: ^{1}H NMR spectroscopy, gel mobility shift assay and stopped-flow ultraviolet–visible spectroscopy.

^{1}H NMR studies were performed with the same 14 bp binding-site oligonucleotide (**oligo C**, Fig. 3a) as used previously[15], at a final concentration of 60 μM mixed directly in an NMR tube with a stoichiometric amount of **3**. After a total of 5.6 minutes to mix and scan the sample, analysis of the spectra revealed that all the signals were identical to those of a fully equilibrated sample (Fig. 4)[15]. In other words, **3** was intercalated fully in the **oligo C** binding site with the linker in the expected minor–major–minor groove topology after 5.6 minutes. Subsequently, the **3–oligo C** NMR sample was monitored at different time points over 24 hours (Supplementary Fig. S4) with no observable change.

As a qualitative approach to monitoring association, the gel-shift assay was used with stoichiometric amounts of **3** and **oligo A** in the concentration range 250–2,500 nM. These experiments were used to gain a rough estimate of an association rate constant to be compared with, and inform, the more accurate stopped-flow spectroscopic measurements described below. Using the integrated rate equation for stoichiometric binding[28] (Supplementary equation (1)) yielded an estimated association rate constant of $1 \pm 0.5 \times 10^4$ M^{-1} s^{-1}. The advantage of the gel-shift approach is that we already know that nonspecific association does not produce a gel shift with **3**. Thus, the estimated association rate constant measured in this way is expected to represent sequence-specific binding.

The rates of association between **3** and either **oligo A** or **oligo B** were also analysed using stopped-flow methods. Hypochromism accompanies intercalation, so the time-dependent decrease at the visible absorbance maximum (386 nm) of **3** was monitored after mixing equal volumes of solutions that contained **3** or DNA. Absorbance traces were collected for fixed concentrations of **3** mixed with increasing DNA concentrations such that the final ratios after mixing ranged from 1:1 to 1:8. An initial rapid decrease in absorbance was observed, which reflects a fast association of **3** with DNA (Fig. 5a).

Single- or two-step models did not fit the data satisfactorily, and therefore kinetic traces obtained at each DNA concentration were fitted globally[29] to a minimal three-step irreversible mechanism (Fig. 5b). The model assumes an initial fast interaction followed by slower consecutive rearrangement steps in which [**3**–DNA]$_I$ and [**3**–DNA]$_{II}$ represent two different intermediate complexes[30]. Although this model describes the kinetic behaviour for both DNA sequences, the fit is best for the association of **3** with **oligo B**. In the case of **3** binding to **oligo A**, an additional step was required to account better for the fast phase, particularly at lower DNA concentrations, which is a key component of the data (Fig. 5b). The additional step fits the kinetic data, but the parameters are not sufficiently well constrained to define uniquely the rate constants for the two isomerization steps. Nonetheless, data fitted globally according to both mechanisms yielded well-defined second-order association rate constants for the two DNA sequences: $3.4 \pm 0.1 \times 10^6$ M^{-1} s^{-1} and $2.5 \pm 0.1 \times 10^6$ M^{-1} s^{-1} (k_1) for **3** binding to **oligo A** and to **oligo B**, respectively. An essentially

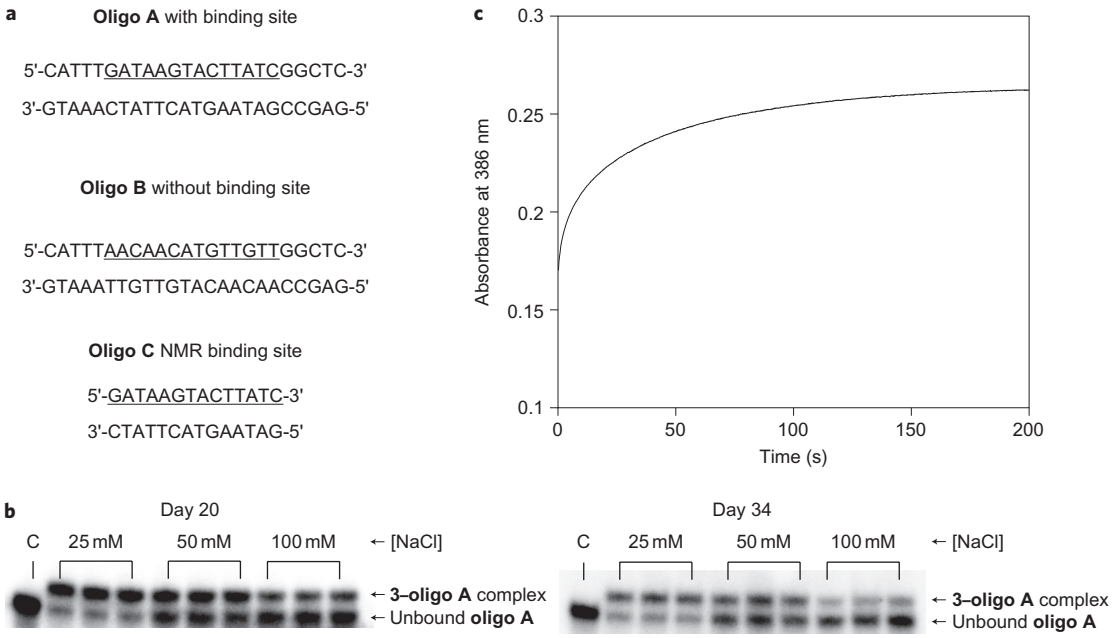

Figure 3 | Oligonucleotide duplex sequences, dissociation gel mobility shifts from oligo A and absorbance trace of the stopped-flow dissociation from oligo B. a, Duplex sequences abbreviated as **oligo A**, **oligo B** and **oligo C**, with the target sequence underlined for each duplex. **b**, **3** dissociating from **oligo A** after 20 days and 34 days of incubation, which shows the prolonged dissociation from the preferred binding site. C refers to the control lane that contains **oligo A** in the absence of **3** in 10 mM PIPES buffer (pH = 7.0), 1 mM EDTA and 100 mM NaCl. All other incubations were run as triplicate measurements in 10 mM PIPES (pH = 7.0), 1 mM EDTA with varying [Na^+] as labelled above the respective lanes. Complexed and unbound **oligo A** bands are as labelled. **c**, A stopped-flow dissociation absorbance trace of the **3**-**oligo B** complex. The equilibrated complex was mixed with equal volumes of 4% SDS in 10 mM PIPES (pH = 7.0), 1 mM EDTA and 100 mM NaCl and monitored at the visible absorbance maximum of 386 nm. The data were fit using a double-exponential equation.

Table 1 | Cumulative data for the dissociation of 3 from oligo A and oligo B at varying [Na^+].

	Oligo A		Oligo B				
[Na^+] (mM)	$k_d \times 10^{-7}$ (s^{-1})	$t_{1/2}$ (days)	k_1 (s^{-1})	A_1 (%)	k_2 (s^{-1})	A_2 (%)	k_d (s^{-1})
25	2.3	32	–	–	–	–	–
50	3.3	23	0.27	30	0.012	70	0.088
100	5.0	16	0.29	34	0.019	66	0.11
150	5.7	13	0.38	35	0.031	65	0.15
200	9.3	8	0.39	40	0.038	60	0.18
300	–	–	0.48	39	0.055	61	0.22

Dissociation rate constants of **3** from **oligo A** are calculated from the monoexponential fits to the gel shifts, with error levels estimated at ±17%. For **oligo B**, A_1 and A_2 are the relative amplitudes for the double exponential according to $k_{d,app} = A_1k_1 + A_2k_2$. The estimates of the rate constants fit a double exponential equation with a standard error less than 1%. To obtain the faster rates and corresponding amplitudes, the fits to the traces obtained at intervals of 10 seconds were coerced to the slower rates and their amplitudes obtained from the traces at 200 seconds.

irreversible rearrangement step is consistent with the remarkably slow off-rate of **3** from the target sequence. The rate constants shown in Fig. 5 are listed as irreversible steps only because there was no information to define the reverse rate constants and so these should be considered as net reaction rate-estimates only. This model can be reconciled with the apparent second-order rate constant of $1 \pm 0.5 \times 10^4\ M^{-1}\ s^{-1}$ estimated by gel-shift assays by allowing the initial binding steps to be readily reversible, which leads to the final, largely irreversible, step.

Discussion

Previous work with **3** established the ability to bind the sequence 5′-GATAAGTACTTATC-3′ in a 1:1 stoichiometry with the linkers bound in a minor–major–minor groove topology (Fig. 1c)[15]. The DNase I footprinting results presented here verify that **3** can select and bind this 14 bp sequence in the context of nearly 500 bp (Fig. 2). Although no structural verification was attempted, it is reasonable to assume a similar threading topology on a long piece of DNA in which the linkers alternate in the minor–major–minor groove topology, as with the oligonucleotide of the same sequence[15].

The observed sequence specificity could be the consequence of differences in association rates, dissociation rates or a combination of the two. Given the 4×10^4-fold slower dissociation rate constant observed for **3** binding to its preferred sequence compared to a randomly chosen sequence, it is safe to say that specificity is based largely on differences in the dissociation rates. This conclusion was reinforced by stopped-flow association data that revealed very similar association rate constants for the preferred versus control sequences.

The threading polyintercalation topology is expected intuitively to manifest an extremely slow dissociation because of the seemingly improbable molecular rearrangements required for full dissociation. In addition, threading intercalator monomers display relatively slow dissociation compared to that of non-threading monointercalators[10,11]. Indeed, using 100 mM NaCl as benchmark conditions, the observed $k_d = 5.0 \pm 0.5 \times 10^{-7}\ s^{-1}$ for **3** dissociating from its preferred binding sequence corresponds to a dissociation half-life of $t_{1/2} = 16$ days. We envision that multiple individual steps are required for full dissociation, so these values are best thought of as rate limiting through several individual rate constants.

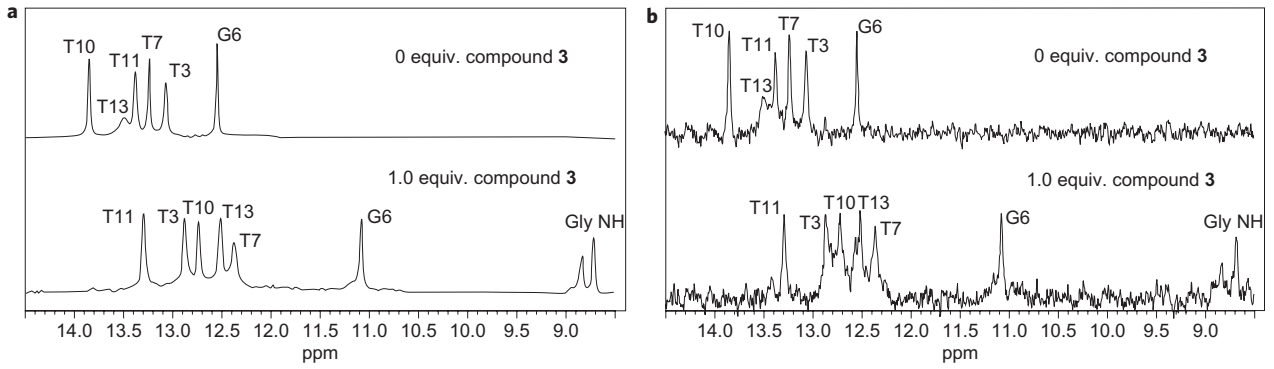

Figure 4 | ^{1}H NMR kinetic titrations of 3 binding to oligo C. a, ^{1}H NMR spectra of **3** titrated into d(GATAAGTACTTATC)$_2$ completed by Lee *et al.*[15] with the spectra for 0 equiv. (top) and 1.0 equiv. (bottom) of **3** added; the latter shows the emergence of a new set of distinct imino proton peaks. **b**, ^{1}H NMR spectra of **oligo C** with both 0 equiv. (top) and 1.0 equiv. (bottom) of **3** added. The 1:1 association kinetic run with 1.0 equiv. of **3** was produced after 5.6 minutes at a final concentration of 60 μM **3**–**oligo C** in 30 mM phosphate buffer (pH = 7.0); the distinct complexed imino proton peaks indicate structural verification of the complete association of the **3**–**oligo C** complex. All of the NMR peaks are labelled according to the single-letter designation for the DNA nucleobase as it would be numbered from 5′ to 3′ for **oligo C**.

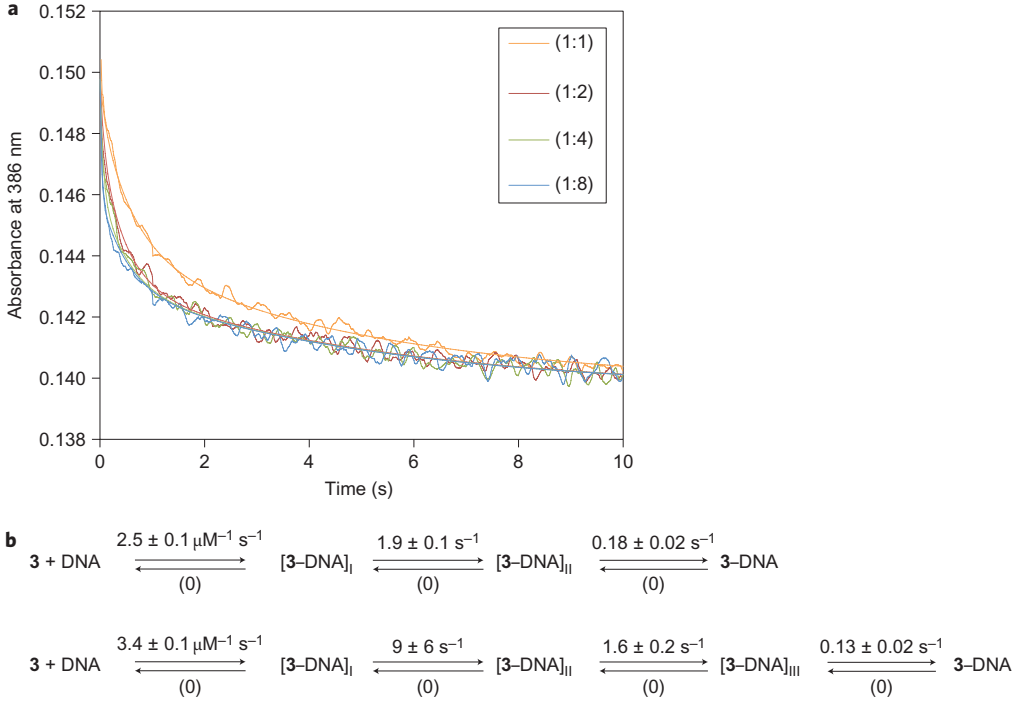

Figure 5 | Stopped-flow association absorbance traces of 3 binding to oligo B with proposed models for association of 3 to both oligo A and oligo B. **a**, Stopped-flow absorbance traces of the association of **3** (2 μM) with **oligo B** at four different DNA concentrations (2, 4, 8 and 16 μM) in 10 mM PIPES buffer (pH = 7.0), 1 mM EDTA and 100 mM NaCl. The solid lines represent the best global fit to a three-step consecutive mechanism using KinTek Explorer Software. A similar absorbance trace was obtained for the association of **3** to **oligo A** (Supplementary Fig. S5). **b**, Rate constants derived from a global fit of the concentration dependence of the data series to a three-step model. Both models assume irreversible steps towards complex formation. The data for the association of **3** with **oligo B** best fit the top model, and the data for the association of **3** with **oligo A** best fit the bottom model.

The complex topology of a bound threading polyintercalator also raises questions about the detailed mechanism of association. We can predict a rapid and nonspecific association between positively charged **3** and negatively charged DNA driven largely by electrostatic attraction. After this, a large number of distinct states can be envisioned as the molecule samples various topologies and sequences before ending up in the final threading polyintercalation binding mode. The stopped-flow measurements confirm at least a three-step process, consistent with expectation. Perhaps surprisingly, there were significant similarities between the association rate profiles and rate constants for **3** binding to its preferred versus a random sequence, although detailed curve fitting indicated there may be some subtle mechanistic differences between the two.

Unfortunately, no detailed mechanistic interpretation is possible based on the stopped-flow data alone because spectrophotometric monitoring cannot be used to gain an unambiguous structural understanding of the process. For example, it is possible that the observed decreases in NDI absorbance result from a step in which the NDI units are simply buried in a DNA groove and not yet intercalated. In addition, some mechanistic features could be very hard to detect as distinct steps because it is possible that several different

bound topologies with just two NDI units intercalated could have very similar overall absorbance values.

Lack of structural information in the stopped-flow experiment was addressed by using ^{1}H NMR studies and qualitative gel-shift measurements. A detailed analysis of the entire spectrum and especially the imino proton signals verified that, within the limits of NMR detection (greater than 95%), after the 5.6 minute mixing time **3** was bound fully to its preferred site in the expected threading polyintercalator topology previously assigned (Fig. 4)[15]. Unfortunately, because the complex was formed fully before the first scan was acquired, no detailed rate information could be obtained. However, the gel-shift association experiments were used to provide an estimate of the association rate constant for **3** binding to its preferred sequence in a manner that survives gel electrophoresis. Recall that the gel shift was observed only for **3** binding its preferred sequence, not a randomly chosen sequence, so this measurement is assumed to verify binding to the oligonucleotide in a sequence-specific fashion. The approximate association rate constant obtained in this way was $1 \pm 0.5 \times 10^4$ M^{-1} s^{-1}, which is consistent with full binding under the conditions of the NMR experiment, but is still considerably slower than the slower of the two rate constants determined by stopped flow. It is probable that the stopped-flow assay reports the initial, and readily reversible, association steps, which are followed by a largely irreversible step to form the tight complex observed by NMR spectroscopy and the gel-shift assay.

It is interesting to compare the dissociation rate constant measured here to those for other reported long-lived complexes. To the best of our knowledge, a $k_d = 5.0 \pm 0.5 \times 10^{-7}$ s^{-1}, which corresponds to a half-life of 16 days, compares favourably to all reported dissociation rate constants for molecules bound to DNA. In fact, we know of only a handful of reported slower dissociation rate constants, including that for the synthetic multivalent tris-vancomycin-tris-D-Ala complex ($t_{1/2} \approx 200$ days)[31], a few highly engineered antibodies ($t_{1/2} \approx 5$–26 days)[32–34] and the avidin–biotin system ($t_{1/2} \approx 89$–107 days)[35,36].

Given the lower estimate for the apparent second-order rate constant of $1 \pm 0.5 \times 10^4$ M^{-1} s^{-1} that governs the formation of the specific complex and the slow dissociation rate, we estimate a net dissociation constant of 50 pM for **3** binding to its preferred site on **oligo A**. It is almost certain that binding to a preferred site on a long piece of DNA will be complicated by additional mechanistic steps not seen with the relatively short oligonucleotides used here. Currently, we are embarking on a comprehensive mechanistic study intended to uncover the rates and individual steps involved when a threading polyintercalator such as **3** 'scans' long segments of DNA, locates a preferred site and then fully intercalates into it. The results of these analyses will be reported in due course.

In summary, a roughly 4×10^4-fold preference for a specific 14 bp sequence was determined for **3** relative to a randomly chosen sequence. This specificity was almost entirely because of differences in dissociation rate constants. Importantly, a half-life of 16 days at 100 mM NaCl is the slowest we know of DNA-binding molecules and is presumably sufficiently slow to disrupt biological processes, such as transcription, repair and replication, for prolonged periods. As adding between 4 and 5 kcal mol^{-1} of binding energy to the dissociation rate constant reaches a bound half-life of 90 years, we will attempt to exploit the modular nature of our threading polyintercalators to make longer molecules with expected bound half-lives relevant to mammalian lifetimes.

Methods

Tetra-intercalator (3) synthesis. The 9-fluorenylmethoxycarbonyl-Lys-NDI-(β-Ala)$_3$-Lys-NDI-Gly resin was prepared as described previously[14,37] and thereafter the tetra-intercalator synthesis procedure of Lee *et al.* was followed[15]. ^{1}H NMR spectra were in agreement with previously published data[15] and high-resolution mass spectroscopy–electron spray ionization gave: $C_{128}H_{150}N_{30}O_{34}$ $[M + 2H]^{+2}$ predicted 1325.5466, found 1325.8; $C_{128}H_{151}N_{30}O_{34}$ $[M + 3H]^{+3}$ predicted 884.0336, found 884.5; $C_{128}H_{152}N_{30}O_{34}$ $[M + 4H]^{+4}$ predicted 663.681, found 663.6. Tetra-intercalator concentration was determined by the absorbance at 386 nm using an extinction coefficient of 51,300 M^{-1} cm^{-1}.

DNase I footprinting. A double-stranded oligonucleotide that contained the tetra-intercalator binding site was inserted between the SfiI restriction enzyme sites of the pMoPac16 vector[38]. The sequence of the construct was confirmed prior to use. The 467 bp DNA fragment was amplified by polymerase chain reaction (PCR) using a primer that was 5′-^{32}P end-labelled. The PCR product was purified on a 5% native gel and the bands were excised and eluted in TE buffer (10 mM Tris-Cl, 1 mM ethylenediaminetetraacetic acid (EDTA), pH = 7.0). The labelled fragment was incubated for 48 hours with **3** at the designated concentration followed by DNase I digestion in 20 mM phosphate buffer (pH = 7.0) and 2 mM $MgCl_2$ for seven minutes. A detailed protocol is described elsewhere[39]. To determine the footprinting sequence, the digestion was compared with an adenine-specific sequencing lane[40].

Gel mobility shift assays to monitor association and dissociation. The top strand of **oligo A** was labelled with ^{32}P at the 5′-end, annealed to its compliment strand and gel purified. After excising the band from the gel, the duplex oligonucleotide was precipitated in ethanol and resuspended in double-distilled H_2O. The DNA concentration was quantified with a microvolume spectrophotometer (Nanodrop ND100). For association experiments, a stoichiometric amount of **3** was added to a solution that contained both radiolabelled (50 nM) and unlabelled **oligo A** such that the final DNA concentration ranged from 250 to 2,500 nM in piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES, 10 mM) buffer (pH = 7.0), 1 mM EDTA and 100 mM NaCl in a final volume of 12 μl. Aliquots taken at representative times from each DNA concentration were mixed with calf thymus DNA to a final concentration of 250 μM bp to prevent further association. Samples were then mixed with loading buffer (0.25% Bromophenol Blue, 0.25% Xylene Cyanol, 40% sucrose) and loaded on an 8% non-denaturing gel. The gel was dried, visualized digitally and the amount of bound and unbound radiolabelled **oligo A** was quantified using ImageQuant 4.1 imaging software. For dissociation experiments, radiolabelled **oligo A** (1 μM) was incubated with **3** (1 μM) in 10 mM PIPES buffer (pH = 7.0), 1 mM EDTA and NaCl concentrations ranging from 25 to 200 mM. Complete association was verified first by gel electrophoresis before conducting the time-point dissociation experiments. Once confirmed, a 100-fold excess of non-labelled **oligo A** was added to the **3**–**oligo A** complex. Aliquots taken at various time points were separated by gel electrophoresis and the bands were quantified as above.

Stopped-flow kinetics. Kinetic data were acquired using a KinTek stopped-flow instrument equipped with an optical cell of path length 2.5 cm. The absorbance change at 386 nm was monitored for the association and dissociation of **3** with **oligo A** and **oligo B**. Kinetic experiments of DNA association were conducted with a constant concentration of **3** and varying DNA oligonucleotide concentration. DNA oligos (ranging from 4 to 32 μM) in buffer that contained 10 mM PIPES (pH = 7.0), 1 mM EDTA and 100 mM NaCl were mixed with equal volumes of **3** (4 μM) in the same buffer. No absorbance change occurred when **3** was mixed with buffer alone. For the dissociation experiments, pre-incubated **3**–**oligo B** complexes in 10 mM PIPES (pH = 7.0), 1 mM EDTA and the designated NaCl concentration (50–300 mM) were mixed rapidly with a 4% SDS solution in the same buffer. Data collected at each timescale consisted of 1,000 data points. Global data fitting of the averaged kinetic traces was done using the KinTek Explorer software (KinTek Corp, Austin, Texas, USA).

^{1}H NMR sample preparation and spectroscopy. Solutions were prepared in 90% H_2O/10% D_2O containing 30 mM phosphate buffer (pH = 7.0) and 100 mM NaCl. One solution contained **oligo C** (120 μM), concentration verified by nanodrop spectroscopy (ε = 143,400 M^{-1} cm^{-1}, molecular mass = 8,523.6 g mol^{-1}) and the other solution contained **3** (120 μM), as verified by ultraviolet–visible spectroscopy (ε = 51,300 M^{-1} cm^{-1}) (ref. 11). Control spectra were obtained for **oligo C** at 60 μM final concentration. The kinetic association was monitored by mixing a stoichiometric amount of **3** (prepared as above) directly into an NMR tube that contained **oligo C** (prepared as above) to a final 1:1 concentration of 60 μM. Experiments were performed using a Varian DirectDrive 600 MHz spectrometer. ^{1}H NMR spectra in 90% H_2O/10% D_2O were taken at 27 °C using a jump–return solvent suppression method. All spectra were processed using VNMRJ (Varian, Inc.).

Received 19 May 2011; accepted 17 August 2011; published online 25 September 2011

References

1. Dervan, P. B. Design of sequence-specific DNA-binding molecules. *Science* **232,** 464–471 (1986).
2. Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **9,** 2215–2236 (2001).
3. Nickols, N. G. & Dervan, P. B. Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl Acad. Sci. USA* **104,** 10418–10423 (2007).

4. Muzikar, K. A., Nickols, N. G. & Dervan, P. B. Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl Acad. Sci. USA* **106,** 16598–16603 (2009).
5. Fox, K. R. Targeting DNA with triplexes. *Curr. Med. Chem.* **7,** 17–37 (2000).
6. Arya, D. P. New approaches toward recognition of nucleic acid triple helices. *Acc. Chem. Res.* **44,** 134–146 (2010).
7. Egholm, M., Buchardt, O., Nielsen, P. E. & Berg, R. H. Peptide nucleic acids (PNA). Oligonucleotide analogues with an achiral peptide backbone. *J. Am. Chem. Soc.* **114,** 1895–1897 (1992).
8. Janowski, B. A., Hu, J. & Corey, D. R. Silencing gene expression by targeting chromosomal DNA with antigene peptide nucleic acids and duplex RNAs. *Nature Protocols* **1,** 436–443 (2006).
9. Yen, S. F., Gabbay, E. J. & Wilson, W. D. Interaction of aromatic imides with deoxyribonucleic acid, spectrophotometric and viscometric studies. *Biochemistry* **21,** 2070–2076 (1982).
10. Tanious, F. A., Yen, S. F. & Wilson, W. D. Kinetic and equilibrium analysis of a threading intercalation mode: DNA sequence and ion effects. *Biochemistry* **30,** 1813–1819 (1991).
11. Lokey, R. S. *et al.* A new class of polyintercalating molecules. *J. Am. Chem. Soc.* **119,** 7202–7210 (1997).
12. Guelev, V. M., Harting, M. T., Lokey, R. S. & Iverson, B. L. Altered sequence specificity identified from a library of DNA-binding small molecules. *Chem. Biol.* **7,** 1–8 (2000).
13. Guelev, V. *et al.* Peptide bis-intercalator binds DNA via threading mode with sequence specific contacts in the major groove. *Chem. Biol.* **8,** 415–425 (2001).
14. Guelev, V., Sorey, S., Hoffman, D. W. & Iverson, B. L. Changing DNA grooves – a 1,4,5,8-naphthalene tetracarboxylic diimide bis-intercalator with the linker $(\beta\text{-Ala})_3$-Lys in the minor groove. *J. Am. Chem. Soc.* **124,** 2864–2865 (2002).
15. Lee, J., Guelev, V., Sorey, S., Hoffman, D. W. & Iverson, B. L. NMR structural analysis of a modular threading tetraintercalator bound to DNA. *J. Am. Chem. Soc.* **126,** 14036–14042 (2004).
16. Chu, Y., Sorey, S., Hoffman, D. W. & Iverson, B. L. Structural characterization of a rigidified threading bisintercalator. *J. Am. Chem. Soc.* **129,** 1304–1311 (2007).
17. Chu, Y., Hoffman, D. W. & Iverson, B. L. A pseudocatenane structure formed between DNA and a cyclic bisintercalator. *J. Am. Chem. Soc.* **131,** 3499–3508 (2009).
18. Chaires, J. B., Dattagupta, N. & Crothers, D. M. Kinetics of the daunomycin–DNA interaction. *Biochemistry* **24,** 260–267 (1985).
19. Wilson, W. D. *et al.* DNA sequence dependent binding modes of 4′,6-diamidino-2-phenylindole (DAPI). *Biochemistry* **29,** 8452–8461 (1990).
20. Tanious, F. A., Veal, J. M., Buczak, H., Ratmeyer, L. S. & Wilson, W. D. DAPI (4′,6-diamidino-2-phenylindole) binds differently to DNA and RNA: minor-groove binding at AT sites and intercalation at AU sites. *Biochemistry* **31,** 3103–3112 (1992).
21. Wilson, W. D., Krishnamoorthy, C. R., Wang, Y. H. & Smith, J. C. Mechanism of intercalation: ion effects on the equilibrium and kinetic constants for the interaction of propidium and ethidium with DNA. *Biopolymers* **24,** 1941–1961 (1985).
22. Westerlund, F., Wilhelmsson, L. M., Nordén, B. & Lincoln, P. Monitoring the DNA binding kinetics of a binuclear ruthenium complex by energy transfer: evidence for slow shuffling. *J. Phys. Chem. B* **109,** 21140–21144 (2005).
23. Westerlund, F., Nordell, P., Nordén, B. & Lincoln, P. Kinetic characterization of an extremely slow DNA binding equilibrium. *J. Phys. Chem. B* **111,** 9132–9137 (2007).
24. Leng, F., Priebe, W. & Chaires, J. B. Ultratight DNA binding of a new bisintercalating anthracycline antibiotic. *Biochemistry* **37,** 1743–1753 (1998).
25. Hampshire, A. J., Rusling, D. A., Broughton-Head, V. J. & Fox, K. R. Footprinting: a method for determining sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods* **42,** 128–140 (2007).
26. Müller, W. & Crothers, D. M. Studies of the binding of actinomycin and related compounds to DNA. *J. Mol. Biol.* **35,** 251–290 (1968).
27. Westerlund, F., Wilhelmsson, L. M., Nordén, B. & Lincoln, P. Micelle-sequestered dissociation of cationic DNA-intercalated drugs: unexpected surfactant-induced rate enhancement. *J. Am. Chem. Soc.* **125,** 3773–3779 (2003).
28. Anslyn, E. V. & Dougherty, D. A. *Modern Physical Organic Chemistry* (University Science Books, 2006).
29. Johnson, K. A., Simpson, Z. B. & Blom, T. Global Kinetic Explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* **387,** 20–29 (2009).
30. Bevilacqua, P. C., Kierzek, R., Johnson, K. A. & Turner, D. H. Dynamics of ribozyme binding of substrate revealed by fluorescence-detected stopped-flow methods. *Science* **258,** 1355–1358 (1992).
31. Rao, J., Lahiri, J., Isaacs, L., Weis, R. M. & Whitesides, G. M. A trivalent system from vancomycin-D-Ala-D-Ala with higher affinity than avidin–biotin. *Science* **280,** 708–711 (1998).
32. Boder, E. T., Midelfort, K. S. & Wittrup, K. D. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc. Natl Acad. Sci. USA* **97,** 10701–10705 (2000).
33. Graff, C. P., Chester, K., Begent, R. & Wittrup, K. D. Directed evolution of an anti-carcinoembryonic antigen scFv with a 4-day monovalent dissociation half-time at 37 °C. *Protein Eng. Des. Sel.* **17,** 293–304 (2004).
34. Rajpal, A. *et al.* A general method for greatly improving the affinity of antibodies by using combinatorial libraries. *Proc. Natl Acad. Sci. USA* **102,** 8466–8471 (2005).
35. Green, N. M. Avidin: 1-The use of [^{14}C]biotin for kinetic studies and for assay. *Biochem. J.* **89,** 585–591 (1963).
36. Piran, U. & Riordan, W. J. Dissociation rate constant of the biotin–streptavidin complex. *J. Immunol. Meth.* **133,** 141–143 (1990).
37. Guelev, V. M., Cubberley, M. S., Murr, M. M., Lokey, R. S. & Iverson, B. L. Design, synthesis, and characterization of polyintercalating ligands. *Methods Enzymol.* **340,** 556–570 (2001).
38. Hayhurst, A. *et al.* Isolation and expression of recombinant antibody fragments to the biological warfare pathogen *Brucella melitensis*. *J. Immunol. Methods* **276,** 185–196 (2003).
39. Trauger, J. W. & Dervan, P. B. Footprinting methods for analysis of pyrrole–imidazole polyamide/DNA complexes. *Methods Enzymol.* **340,** 450–466 (2001).
40. Iverson, B. L. & Dervan, P. B. Adenine specific DNA chemical sequencing reaction. *Nucleic Acids Res.* **15,** 7823–7830 (1987).

Acknowledgements

This work was supported by the Robert A. Welch Foundation (grant F1188 to B.L.I., F1604 to K.A.J. and departmental grant AF-0005 to M.Z.F.) and the National Institutes of Health (grant GM-069647 to B.L.I.). Acknowledgement is made to the Donors of the American Chemical Society Petroleum Research Fund for partial support of this research (M.Z.F.). M.Z.F. was also supported by the Southwestern University faculty sabbatical program. We thank Steven Sorey for his help with the ^{1}H NMR spectra.

Author contributions

G.G.H., M.Z.F. and A.R.S. performed the experiments. G.G.H., M.Z.F., A.R.S., K.A.J. and B.L.I. designed the experiments and analysed the data. G.G.H., M.Z.F. and B.L.I. co-wrote the paper.

Additional information

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/naturechemistry. Supplementary information and chemical compound information accompany this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to B.L.I.