

Single-Molecule DNA Biosensors for Protein and Ligand Detection**

Konstantinos Lympereopoulos, Robert Crawford, Joseph P. Torella, Mike Heilemann, Ling Chin Hwang, Seamus J. Holden, and Achilles N. Kapanidis*

Transcription factors (TFs) are sequence-specific DNA-binding proteins^[1] that control much of gene expression. TFs are natural biosensors and switches, translating chemical and physical signals (temperature shifts, light exposure, chemical concentrations, redox status) into transcriptional changes by modulating the binding of RNA polymerase to promoter DNA. Since changes in TF levels underlie fundamental biological processes such as DNA repair and cell-cycle progression, alterations in the levels of active TFs both lead to and indicate disease; for example, mutations in transcription factor p53 contribute to the rapid growth of cancer cells and, owing to their prevalence (p53 is mutated in roughly 50% of all human tumors), they have served as cancer biomarkers.^[2] Thus, methods for the sensitive detection and quantitation of TFs provide both fundamental information about gene regulation and a platform for diagnostics.

TF detection often involves gel-based assays and Western blotting; although helpful in characterizing TF–DNA interactions, these assays are tedious, expensive, and qualitative, and consume large quantities of sample. Enzyme-linked immunosorbent assays (ELISAs) are more sensitive and offer higher throughput, but they require many preparation and signal-amplification steps for the detection of low-abundance TFs. Amplification is also required in the proximity-based ligation assay,^[3] making it incompatible with TF detection in living cells and diagnostic settings that demand results within minutes.

An additional TF detection assay is based on fluorescence resonance energy transfer (FRET) between two double-stranded DNA (dsDNA) fragments containing fluorescently labeled single-stranded complementary overhangs (“molecular beacons”).^[4–6] In the presence of TF, the DNAs associate, resulting in donor fluorophore quenching as a result of FRET. This assay still requires significant amounts of sample and cannot detect low-abundance TFs; and because of the short dynamic range of FRET (1–10 nm), it also requires close proximity among the fluorophore, the quencher, and the protein–DNA interface, increasing the likelihood of steric interference with protein–DNA binding and complicating sensor design. Moreover, placing the fluorophore and the quencher on either side of the protein-binding site (usually 15–30 base pairs (bp) in length) on DNA results in very low FRET signals for most TFs.

Here, we use alternating-laser excitation (ALEX) spectroscopy^[7,8] to detect TFs and small molecules by means of the TF-dependent coincidence of fluorescently labeled DNA. Like the molecular-beacon assay, our method is based on TF-driven DNA association, is rapid, and requires no amplification. However, our assay can detect pm levels of TFs in small amounts of sample, and it is FRET-independent, bypassing the need to optimize fluorophore position or know the structural details of TF–DNA binding; this flexibility in labeling ensures unperturbed TF–DNA binding. Using ALEX, we demonstrate TF and small-molecule detection, assay multiplexing, and suitability for analysis of complex biological samples.

In our assay (Figure 1 a,b), the full DNA-binding site for a TF is split in two (as in Ref. [5]): the left half-site (H1) and the right half-site (H2). Each site contains half of the TF-binding determinants and short, complementary 3'-overhangs. H1 is labeled with a “green” fluorophore (“G”) to give half-site H1^G, whereas H2 is labeled with a spectrally distinct “red” fluorophore (“R”) to give H2^R. In the absence of TF and at DNA concentrations of roughly 10–100 pM, H1 and H2 diffuse independently and associate only transiently. In contrast, in the presence of a TF that binds to the fully assembled DNA site, H1 and H2 diffuse as a complex (H1^G–TF–H2^R; Figure 1 a, bottom).

We detect TF-dependent DNA coincidence using ALEX spectroscopy,^[7–9] wherein single molecules are excited by two lasers in an alternating fashion, with each laser capable of directly exciting either a G or a R fluorophore. ALEX allows molecular sorting on two-dimensional histograms of apparent FRET efficiency E^* (a fluorescence ratio that reports on interfluorophore proximity) and probe stoichiometry S (a fluorescence ratio that reports on molecular stoichiometry). A search for all R-labeled molecules (i.e., G–R molecules

[*] Dr. K. Lympereopoulos,^[†] R. Crawford,^[†] J. P. Torella, Dr. M. Heilemann, Dr. L. C. Hwang, S. J. Holden, Dr. A. N. Kapanidis Biological Physics Research Group, Department of Physics University of Oxford, Clarendon Laboratory Parks Road, Oxford, OX1 3PU (United Kingdom) E-mail: a.kapanidis1@physics.ox.ac.uk

Dr. K. Lympereopoulos^[†]
Current address: BioQuant Institute, Cellnetworks Cluster Ruprecht-Karls Universität Heidelberg 69120 Heidelberg (Germany)

Dr. M. Heilemann
Current address: Applied Laser Physics and Laser Spectroscopy, Bielefeld University Universitätsstrasse 25, 33615 Bielefeld (Germany)

[†] These authors contributed equally to this work.

[**] We thank Dr. M. Brenowitz (Albert Einstein College of Medicine) and Dr. R. Ebricht (HHMI/Rutgers University) for plasmids, Dr. S. Weiss (UCLA) for software, and L. Sattary and J. Ghadiali for assistance. Funding was provided by the UK Bionanotechnology IRC, the EU (MIRG-CT-2005-031079), EPSRC (EP/D058775), and the Wellcome Trust (VS/06/OX/A4). M.H. was supported by a DAAD fellowship.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.200904597>.

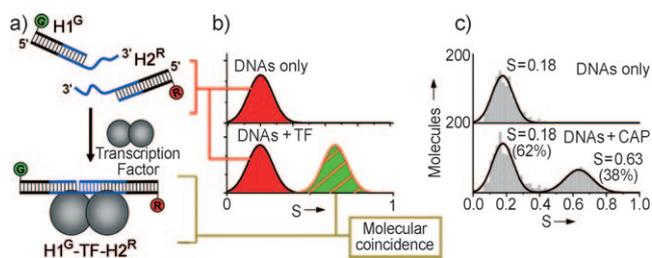


Figure 1. ALEX-based detection of transcription factors. a) Biosensor concept. In the absence of TF, two DNA half-sites, each with half of the site for TF binding and complementary 3'-overhangs, diffuse independently (top); the full site assembles transiently. When a TF binds to a full site, half-sites and TF diffuse as a complex ($H1^G$ -TF- $H2^R$, bottom). b) Schematic of an ALEX-based stoichiometry (S) histogram. The $H1^G$ -TF- $H2^R$ complex is detected as a "green-red" coincident species on ALEX histograms, distinct from free "red-only" half-sites. c) Detection of the transcriptional activator CAP. Top: without CAP, no G-R coincidence (i.e., DNA coincidence) is observed. Bottom: in the presence of CAP, CAP-specific half-sites form a complex with CAP, yielding a species with $S \approx 0.63$. Gray bars: projection of E^* - S data on the S axis; black lines: Gaussian fits. Full E^* - S histograms for CAP detection are shown in Figure S1 a,b in the Supporting Information.

plus R-only molecules) when the two half-sites diffuse independently yields a single population on the S histogram ($H2^R$, an R-only species; Figure 1 b, top). In contrast, when the two half-sites diffuse as a TF-bound complex, a species with higher S appears, reporting on TF presence (Figure 1 b, bottom). Here, we mainly use S -based sorting for biosensing, although the E^* coordinate can also be used to increase the

information content of the biosensor (see Refs. [7–9] and Figure S1 in the Supporting Information for examples of E^* - S histograms).

We first used ALEX to detect catabolite activator protein (CAP; Figure 1c), a TF that activates lactose-utilization genes (the *lac* operon) in bacteria. In the presence of the ligand cyclic AMP, CAP can bind to a 22 bp sequence with an equilibrium dissociation constant K_D of approximately 20 pM.^[10] Upon incubation of 100 nM of CAP-specific half-sites $H1^G$ and $H2^R$ with an excess of CAP (200 nM) and dilution to approximately 10 pM DNA, the S histogram (Figure 1c, bottom) shows a high- S species with G-R stoichiometry ($S \approx 0.63$, due to $H1^G$ -CAP- $H2^R$), and a low- S species with a R-only stoichiometry ($S \approx 0.2$, due to free $H2^R$). In contrast, without CAP, no G-R molecules are detected (Figure 1c, top). We obtained also similar results (Figure S1c,d in the Supporting Information) for *lac* repressor (*lacR*; aka *lacI*), a TF that represses the *lac* operon by binding to a site distinct from that of CAP.^[11,12] Using different half-site concentrations, we have also successfully detected CAP down to pM levels (Figure S1e in the Supporting Information).

To quantify CAP, we developed a model that describes CAP biosensing using two coupled equilibria: one for half-site association, and a second for TF binding to $H1^G$ - $H2^R$ (with dissociation constants K_{D1} and K_{D2} , respectively; see the Supporting Information). We generated an expression that relates CAP concentration ($[CAP]$) to the fraction F_B of half-site bound to TF (as determined by ALEX spectroscopy), to K_{D1} , K_{D2} , and to the total concentration H_{tot} of each half-site [see Eq. (S7) in the Supporting Information]. The model suggests simple ways of optimizing assay sensitivity and

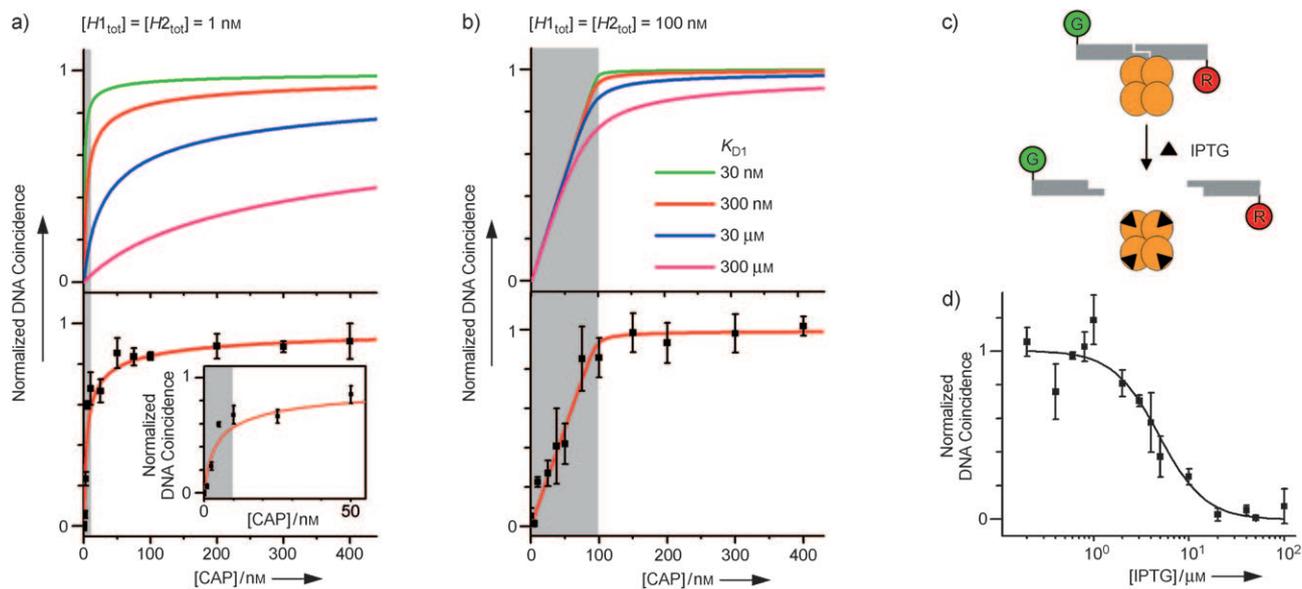


Figure 2. Quantifying transcription factors and their ligands. a,b) Normalized DNA coincidence as a function of $[CAP]$ using a) 1 nM and b) 100 nM DNA half-site solutions. Top: Model predictions as a function of K_{D1} . Bottom: fit of model (red curve) to ALEX data. The inset in (a) shows a close-up of the assay's dynamic range. Gray area: dynamic range of the measurement; errors bars: standard error of the mean. c) Small-molecule detection: sensing IPTG (triangles). At low IPTG, the presence of *lacR* results in DNA coincidence (top). At high IPTG, *lacR* dissociates from DNA, reducing DNA coincidence (bottom). d) Titration of the $H1^G$ -*lacR*- $H2^R$ complex with IPTG shows loss of DNA coincidence at increased IPTG levels. The IPTG concentration at 50% binding inhibition is roughly 5 μ M.

dynamic range (Figure S2 in the Supporting Information). When the H_{tot} is on the order of K_{D1} , half-site hybridization is efficient and K_{D2} controls CAP detection: for TFs with a low K_{D2} , the bound fraction is a linear function of [CAP], allowing CAP detection between 0 and H_{tot} (Figure 2b, top). Conversely, by setting $H_{\text{tot}} \ll K_{\text{D1}}$, the bound fraction increases hyperbolically, shifting the dynamic range and increasing sensitivity (Figure 2a). Assay sensitivity and dynamic range can thus be adjusted by altering H_{tot} , K_{D1} , and K_{D2} (Figure S2 in the Supporting Information); altering H_{tot} is trivial, and rational changes in K_{D1} (e.g., by modifying the length of the 3'-overhangs^[13,14]) or K_{D2} (by using mutant TF-binding sites) can easily be made.

To test our ability to quantify TFs, we titrated either 1 nM or 100 nM half-site DNA with 0–400 nM CAP, diluted to approximately 10 μM , and measured the normalized DNA coincidence, F_{B} (Figure 2a,b, bottom). To compare the data determined with the 1 nM solution with predictions (Figure 2a, bottom), we supplied the model [Eq. (S7) in the Supporting Information] with $H_{\text{tot}} = 1$ nM and $K_{\text{D1}} = 300$ nM (from ensemble measurements; see the Supporting Information), and fitted for K_{D2} ; the best-fit value for K_{D2} was roughly 10 μM , in good agreement with the published value (approximately 20 μM ^[10]). For the 100 nM DNA half-site solutions (Figure 2b bottom), we observe the expected linear increase of F_{B} with CAP concentration until saturation at $[\text{CAP}] \approx H_{\text{tot}}$, and an excellent agreement between our experiments and the model predictions for $K_{\text{D1}} = 300$ nM and $K_{\text{D2}} = 10$ μM (red line, Figure 2b, top).

Small-molecule detection, like TF detection, can also form the basis of a useful bioassay. Since TFs act as natural sensors for small molecules (e.g., sugars, nucleotides, metals, amino acids), we tested whether our assay can detect small molecules. As a proof of principle, we detected the lactose analogue isopropyl β -D-1-thiogalactopyranoside (IPTG), which binds to lacR and induces a conformational change that reduces its DNA-binding affinity by 1000-fold^[15,16] (Figure 2c). We incubated fixed concentrations of the half-sites and lacR with 0–100 μM IPTG. As expected, at low IPTG concentration, lacR leads to high DNA coincidence (Figure 2d). Increasing the IPTG concentration reduced DNA coincidence to the levels observed without lacR (Figure 2d). Analysis of the binding-inhibition curve yields a half-maximum binding-inhibition value of approximately 5 μM , in excellent agreement with published values of roughly 3 μM in similar buffers.^[17]

The ability to detect multiple analytes simultaneously (multiplexing, often achieved using microarrays^[18]) leverages diagnostic assays. To determine whether we could multiplex our assay in solution, we generated CAP-specific half-sites yielding a high- S species when bound to CAP, and lacR-specific half-sites yielding a lower- S species when bound to lacR (see the Supporting Information); this strategy led to simultaneous detection of both TFs (Figure S3a,b in the Supporting Information). We also explored E^* -based multiplexing based on three-color ALEX,^[19] wherein CAP-bound half-sites yielded low- E^* species, and lacR-bound half-sites yielded high- E^* species (Figure S3c,d in the Supporting Information). Our results provide support for developing

assays with increased resolution between TF-bound species (e.g., S-based coding using arrays of G or R fluorophores).

While solution-based multiplexing may allow simultaneous detection of multiple TFs, surface-based techniques can enable truly high-throughput assays (i.e. by using microarrays^[18]) capable of detecting hundreds of TFs and small molecules. To test whether our assay is compatible with surface immobilization, we attached biotinylated half-site H1^{G} on a glass surface, incubated it with lacR and H2^{R} , and imaged the surface with a CCD camera (Figure 3a, left panel). In the presence of lacR, the aligned images show colocalization of G and R fluorophores (Figure 3a, right panel), whereas without lacR, no colocalized spots appear (Figure 3b). Using this method, we have detected lacR down to the 100 pM level (Figure 3c). Similar results were obtained for CAP (Figure S4 in the Supporting Information).

For either gene-expression analysis or point-of-care diagnostics, biosensors must perform in biological fluids that contain nucleases, proteases, and other molecules that may interfere with detection, producing false-positive or false-negative results. To test the robustness of our assay, we first detected TFs in human cell extracts, we added CAP and CAP-

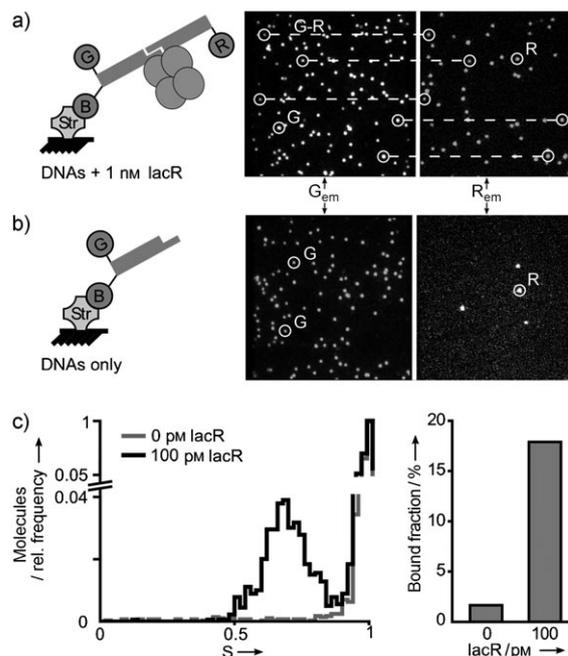


Figure 3. Enabling multiplexing. a,b) TF detection on surfaces. a) A biotinylated H1^{G} was attached on streptavidin-coated glass and incubated with 1 nM lacR and 1 nM H2^{R} ; after washing, the surface was imaged. In the presence of lacR, the images show diffraction-limited spots where G fluorophores (left frames) and R fluorophores (right frames) colocalize. G-only spots are due to H1^{G} -only, plus H1^{G} -lacR- H2^{R} complexes with bleached R fluorophore; R-only spots are due to H1^{G} -lacR- H2^{R} complexes with bleached G fluorophore, plus non-specifically bound H2^{R} molecules. b) Without lacR, few R-only spots appear (due to nonspecific surface binding of H2^{R}). c) TF detection on the surface with high sensitivity. Left: Relative molecular frequency vs. stoichiometry in the presence and absence of 100 pM lacR. Right: Bound fraction (the fraction of G fluorophores that colocalize with R fluorophores) calculated by summing the molecular frequency in the S histogram, for $S < 0.87$ (colocalized G and R) and $S > 0.87$ (G-only).

specific half-sites to HeLa cell nuclear extracts, incubated for 10 min, diluted to approximately 50 pM DNA, and performed ALEX experiments. Without CAP (Figure 4a, top), limited DNA coincidence is observed. In the presence of CAP, however, strong DNA coincidence is detected (Figure 4a, bottom).

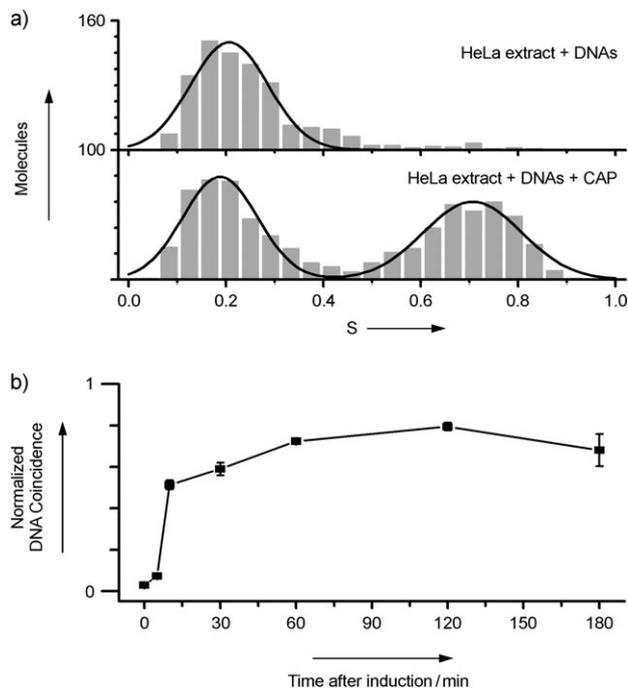


Figure 4. Detecting transcription factors in biological samples. a) Detection of 10 nM CAP in HeLa nuclear extracts. Without CAP (top), little DNA coincidence is seen; in the presence of CAP (bottom), substantial DNA coincidence is seen (species with $S \approx 0.7$). b) Detection of gene expression in bacterial lysates. CAP expression was induced at 0 h, and DNA coincidence was measured over time using 100 nM half-sites; here, DNA coincidence of 0.8 corresponds to roughly 80 nM active CAP. Errors bars: standard error of the mean.

We then tested the ability of our sensors to detect changes in gene expression in *E. coli* whole-cell lysates for increased CAP synthesis, which was induced in cells containing a plasmid carrying the CAP gene; cells were then analyzed at various points before and after induction. Using ALEX spectroscopy, we observed an increase in active CAP concentration from a basal level before induction, to a saturated level within roughly 1 hour (Figure 4b). Accounting for dilutions during lysate preparation, the active CAP concentration at saturation corresponds to about 300 pM (roughly 3×10^5 molecules of active CAP per cell, or 20–30% of the total cellular protein^[19]), consistent with direct quantification of total CAP by SDS-PAGE (roughly 40% of total protein).

In conclusion, we have developed a versatile assay for detecting TFs and small molecules. The compatibility with surface-based detection opens exciting prospects for multiplexing, for example using arrays of spatially addressable half-sites. The assay is compatible with important eukaryotic TFs

that possess high-to-moderate DNA affinity, such as p53 ($K_d \approx 1.6$ nM^[20]), NF- κ B ($K_d \approx 8$ pM^[21]), and estrogen receptor ($K_d \approx 1.8$ nM^[22]). Since some of these TFs dissociate from the DNA on the 1–5 minute timescale, it will be best to detect them using a TIRF format (TIRF = total internal reflection) that allows rapid, parallel sampling of more than 1000 molecules within seconds and accumulates enough statistics to establish the concentration and dissociation rate of complexes. Our ability to detect 10 nM of a TF in nuclear extracts (corresponding to roughly 10000 TF molecules in a “typical” mammalian cell) compares favorably to the copy number of many disease-associated TFs (e.g., p53 ranges from 30000–200000 copies per cell in many cancer cell lines^[23]).

The robustness of our assay in cell lysates provides a starting point for the detection of TFs in living cells. Since single fluorophores have been detected in bacteria,^[24] and as 1 nM corresponds to a single molecule in a single *E. coli* cell, our ability to detect pM levels of TFs may allow detection down to a few protein copies in single cells without amplification; short, labeled DNAs can be introduced in bacteria using electroporation. Use of existing technologies for introducing short DNAs in mammalian cells (e.g., electroporation, use of lipofectamine, liposomal transfection) may permit detection of low-abundance biomarkers in populations of cells where only few cells signify diseased states. Finally, combining our assay with compact or portable instruments will aid rapid, on-site diagnostics.

Received: August 18, 2009

Revised: November 3, 2009

Published online: January 13, 2010

Keywords: biosensors · laser spectroscopy · protein–DNA interactions · single-molecule studies · transcription factors

- [1] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York, 1999.
- [2] B. Vogelstein, D. Lane, A. J. Levine, *Nature* **2000**, *408*, 307.
- [3] S. Fredriksson, M. Gullberg, J. Jarvius, C. Olsson, K. Pietras, S. M. Gustafsdottir, A. Ostman, U. Landegren, *Nat. Biotechnol.* **2002**, *20*, 473.
- [4] E. Heyduk, Y. Fei, T. Heyduk, *Comb. Chem. High Throughput Screening* **2003**, *6*, 347.
- [5] E. Heyduk, E. Knoll, T. Heyduk, *Anal. Biochem.* **2003**, *316*, 1.
- [6] T. Heyduk, E. Heyduk, *Nat. Biotechnol.* **2002**, *20*, 171.
- [7] A. N. Kapanidis, N. K. Lee, T. A. Laurence, S. Doose, E. Margeat, S. Weiss, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8936.
- [8] Y. Santoso, L. C. Hwang, L. Le Reste, A. N. Kapanidis, *Biochem. Soc. Trans.* **2008**, *36*, 738.
- [9] N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, S. Weiss, *Biophys. J.* **2005**, *88*, 2939.
- [10] R. H. Ebright, Y. W. Ebright, A. Gunasekera, *Nucleic Acids Res.* **1989**, *17*, 10295.
- [11] M. Fried, D. M. Crothers, *Nucleic Acids Res.* **1981**, *9*, 6505.
- [12] C. M. Falcon, K. S. Matthews, *Biochemistry* **2000**, *39*, 11074.
- [13] K. J. Breslauer, R. Frank, H. Blocker, L. A. Marky, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 3746.
- [14] M. J. Lane, T. Paner, I. Kashin, B. D. Faldasz, B. Li, F. J. Gallo, A. S. Benight, *Nucleic Acids Res.* **1997**, *25*, 611.
- [15] M. D. Barkley, A. D. Riggs, A. Jobe, S. Burgeois, *Biochemistry* **1975**, *14*, 1700.
- [16] M. Lewis, *C. R. Biol.* **2005**, *328*, 521.

- [17] A. D. Riggs, R. F. Newby, S. Bourgeois, *J. Mol. Biol.* **1970**, *51*, 303.
- [18] E. Southern, K. Mir, M. Shchepinov, *Nat. Genet.* **1999**, *21*, 5.
- [19] S. B. Zimmerman, S. O. Trach, *J. Mol. Biol.* **1991**, *222*, 599.
- [20] N. M. Nichols, K. S. Matthews, *Biochemistry* **2001**, *40*, 3847.
- [21] M. de Lumley, D. J. Hart, M. A. Cooper, S. Symeonides, J. M. Blackburn, *J. Mol. Biol.* **2004**, *339*, 1059.
- [22] M. S. Ozers, J. J. Hill, K. Ervin, J. R. Wood, A. M. Nardulli, C. A. Royer, J. Gorski, *J. Biol. Chem.* **1997**, *272*, 30405.
- [23] L. Ma, J. Wagner, J. J. Rice, W. Hu, A. J. Levine, G. A. Stolovitzky, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14266.
- [24] J. Elf, G. W. Li, X. S. Xie, *Science* **2007**, *316*, 1191.
-