



Sliding on DNA: From Peptides to Small Molecules

Kan Xiong, Graham S. Erwin, Aseem Z. Ansari, and Paul C. Blainey*

Abstract: Many DNA binding proteins utilize one-dimensional (1D) diffusion along DNA to accelerate their DNA target recognition. Although 1D diffusion of proteins along DNA has been studied for decades, a quantitative understanding is only beginning to emerge and few chemical tools are available to apply 1D diffusion as a design principle. Recently, we discovered that peptides can bind and slide along DNA—even transporting cargo along DNA. Such molecules are known as molecular sleds. Here, to advance our understanding of structure–function relationships governing sequence nonspecific DNA interaction of natural molecular sleds and to explore the potential for controlling sliding activity, we test the DNA binding and sliding activities of chemically modified peptides and analogs, and show that synthetic small molecules can slide on DNA. We found new ways to control molecular sled activity, novel small-molecule synthetic sleds, and molecular sled activity in *N*-methylpyrrole/*N*-methylimidazole polyamides that helps explain how these molecules locate rare target sites.

Nature modulates biochemical reactivity and logic states by controlling dimensionality. A widely recognized example is 3D to 2D reduction-of-dimensionality (RD) by targeting factors to cellular membranes and accelerating their interaction. 3D to 1D RD occurs by targeting proteins to 1D on DNA where one-dimensional diffusion along DNA (including sliding and hopping processes) accelerates DNA target recognition.^[1] Although 1D transport of proteins along DNA has been of interest for decades, few chemical tools are available to engineer RD to 1D for applications in biotechnology, for example to enable better control of transcription, enhanced nucleic acid enzymes, and improved small-molecule activity in cells.

Recent work on adenovirus proteins led to the first identification of protein–protein interactions driven by RD to 1D^[2] and the first molecular sled, the peptide pVlc (GVQSLKRRRCF).^[3] Molecular sleds are small DNA binding molecules such as peptides that slide along sequence-nonspecific DNA and can translocate cargo along DNA. A

wide range of basic polypeptide sequences, including nuclear localization signals (NLS), exhibit these activities.^[4]

Molecules that slide on DNA are of interest as tools for accelerating and controlling biochemical processes in 1D because of their small sizes and fast diffusion.^[5] Sliding-based approaches differ from previously reported DNA-scaffolding strategies utilizing sequence-specific DNA binding proteins^[6] or nucleic acid hybridization^[7] to co-localize reacting species because interactions between molecular sleds and DNA are sequence-nonspecific, the DNA-targeting moieties are very small, and the reactants are highly mobile on the DNA template. The discovery of small peptides with robust sliding activity raises the possibility that other classes of small molecules may slide.

The mechanism by which *N*-methylpyrrole/*N*-methylimidazole polyamides, programmable sequence-specific DNA binding small molecules, localize to target sites in vivo remains unexplored—in particular, it is unclear whether polyamides undergo sliding activity on DNA, like protein transcription factors (TFs). Despite predictive models describing the affinity and specificity for target site binding in vitro,^[8] it has been difficult to predict their in vivo efficacy.^[9] Recently, a new model predicting in vivo binding based on site clustering in ~400 bp windows (including lower-affinity sites) showed superior performance.^[10] Such deviation from independent-site thermodynamic models suggests that other factors, such as mechanisms that couple binding interactions linearly along the DNA, non-equilibrium effects, epigenomic effects, and/or other active cellular processes, affect the activity of polyamides in vivo.

Here, to establish a broader understanding of molecular sled transport along DNA, we assay a diverse set of modified peptides and DNA binding small molecules for 1D diffusion and sliding using a previously established single-molecule assay^[11] (see the Supporting Information, Figures S2 and S3, Movie S1). In this assay, biotin- λ -DNA (approximately 48 500 base pairs long) is immobilized onto a flow channel surface by one end and a laminar flow is applied to flow stretch the DNA. The stretched DNA molecules serve as spatially extended templates for the analysis of binding and transport activity of molecular sleds labeled with fluorescent dye molecules. The trajectories of single molecules are tracked by time-lapse total internal fluorescence imaging. The raw images are analyzed by using a custom single particle tracking software^[4] to identify trajectories of single molecule diffusing along DNA. The one-dimensional diffusion constant (D1) value characterizing motion along the DNA strand is then estimated by using a covariance-based method.^[12]

We previously observed that the presence of a free *N*-terminal amine increases the sliding activity and DNA affinity of peptidyl sleds (Table S1).^[4] Here we show an analogous effect upon converting a C-terminal carboxylate to an amide,

[*] K. Xiong, Prof. P. C. Blainey
Broad Institute of MIT and Harvard
Cambridge, MA 02142 (USA)
and
Department of Biological Engineering, MIT
Cambridge, MA 02142 (USA)
E-mail: pblainey@broadinstitute.org
G. S. Erwin, Prof. A. Z. Ansari
Department of Biochemistry, University of Wisconsin-Madison
Madison, WI 53706 (USA)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under <http://dx.doi.org/10.1002/anie.201606768>.

which doubles the D1 value of TMR-^LK^LR^LR^LR (Figure 1B) and increases the binding strength (see dissociation constants, K_d values in Table S2), consistent with the increased net charge on the modified peptide. We observed the same trends in D1 and K_d values for TMR-^LR^LR^LR^LR and TMR-^LR^LR^LR^LK upon C-terminal amidation (Figure 1B, Table S2), and further that this modification renders D1 more sensitive to amino acid sequence (Figure S6A) than previously observed (Supporting Information).

Knowing that either converting a C-terminal carboxylate to an amide or retaining a free N-terminal amine can enhance the sliding activity of peptidyl sleds, we were curious whether these effects could be combined. Figure 1C and Table S2 show that the D1 and K_d values of ^LK^LR^LR^LR^LC-(TMR)-NH₂, which has a free N-terminal amine and a C-terminal amide (TMR is conjugated to the Cys side chain) are not significantly different than the corresponding values for TMR-^LK^LR^LR^LR-NH₂ which has only a C-terminal amide modification, indicating that the N-terminal and C-terminal effects are non-additive.

That C-terminal amidation has similar effects on activity as removing an N-terminal R-group is a curious finding. The acid-base properties of amines and amides are distinct, but each can act as a hydrogen bond donor. These effects are also intriguing because both alterations intensify the interaction of the peptide with DNA, which might be expected to increase

“friction” for sliding and thus reduce D1 (in contrast to observations). When we previously increased peptide–DNA affinity by increasing the number of basic side chains on a peptide, we saw a decrease in D1.^[4] Thus, there seems to be a unique effect of the peptide termini on sliding kinetics, one that we speculate may be related to the hydrogen bonding potential of the terminal groups.

We assayed the activity of tetra-arginine-like peptides with side chain lengths both shorter (^LR^S: L-2-amino-3-guanidinopropionic acid) and longer (^LR^L: L-homo-arginine) than arginine. Previous structural modeling based on studies of pVlc in complex with the adenoviral proteinase (AVP; the enzyme it activates catalytically and endows with sliding activity) described the four basic residues of pVlc (KRRR) spanning the major groove with backbone in a β-sheet-like configuration.^[3] We subsequently presented an alternative model with lysine and arginine side chains inserted into the minor groove, which is much narrower than the major groove.^[4] Figure 2B shows that the K_d values of tetra-arginine-like peptides rapidly decrease as the side chain length increases. The fact that peptides with shorter side chains slide just about as fast as peptides with longer side chains that bind more tightly (Figure 2C) indicates that fast sliding does not require side chains to reach far enough into the DNA minor groove for optimal binding.

The data from the tetra-arginine-like series adds another example of relative independence between DNA binding affinity and D1. Previous work on NLS and cell penetrating peptide molecular sleds showed that DNA affinity and sliding activity do not always correlate.^[4] Overall, our results suggest that chemical phenomena such as hydrogen bonding at the DNA/sled/solvent interfaces may be instrumental to fast sliding and explain the sensitivity of D1 to terminal functionality. Deviations from a “friction”-based model for sliding were observed for the DNA repair protein, human oxoguanine DNA glycosylase (hOgg1), which may be affected by related chemical mechanisms.^[11]

We next turned our attention to peptides with varied backbone stereochemistry, which interact differently with the right-handed B-DNA helix. The D1 and K_d values of TMR-KRRR-NH₂ and TMR-RRRK-NH₂ sleds are hardly affected by the backbone chirality, including alternating L- and D-amino acids (Figure 2D,E, Tables S2 and S4). Inserting a flexible linker between TMR and the N-terminal amide does not impact the D1 or K_d values of D-peptides, indicating that the N-terminal TMR label interferes neither with sliding nor DNA binding properties (Figure S6B, Table S4). From these observations we conclude that backbone stereochemistry is not an important constraint on molecular sled activity.

To further examine how backbone structure affects the activities of molecular sleds, we assayed β³-peptides, which have an additional methyl group in the backbone of each monomer (Figure 2A). The D1 value of TMR-β³Rβ³Rβ³Rβ³K-NH₂ is twice that of TMR-^LR^LR^LR^LK-NH₂ and the D1 value of TMR-β³Kβ³Rβ³Rβ³Rβ³NH₂ is somewhat higher than that of TMR-^LK^LR^LR^LR-NH₂ (Figure 2G). The affinities of β³-peptides for DNA are not reduced compared with the corresponding L-peptides, suggesting that electrostatic binding interactions with DNA are analogous to standard

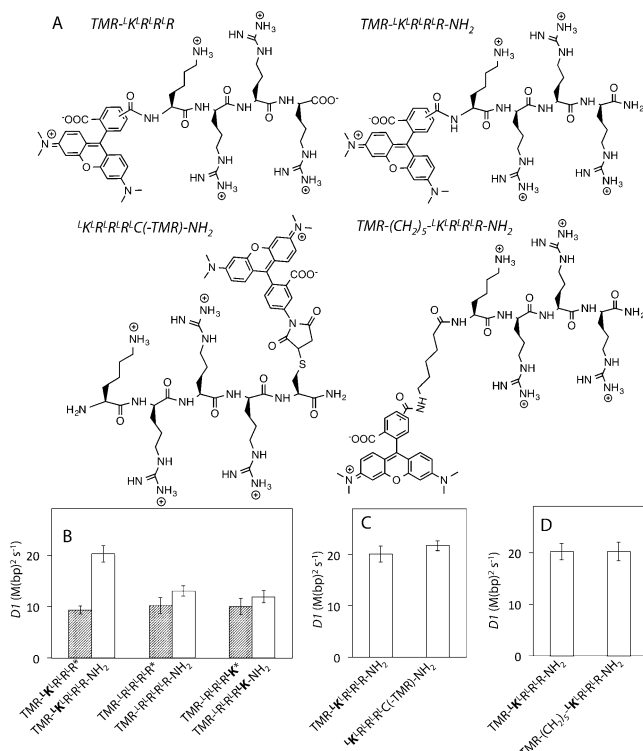


Figure 1. A) Selected chemical structures. B–D) One-dimensional diffusion constants (D1) of peptidyl sleds in 2 mM NaCl buffer (all buffers are pH 7.4). TMR was conjugated to the N-terminal amines of all peptides except for ^LK^LR^LR^LR^LC-(TMR)-NH₂. Error bars represent 95% confidence intervals in all figures. * Indicates data that were previously published^[4] and re-analyzed here. See the Supporting Information, Table S2 for D1 values.

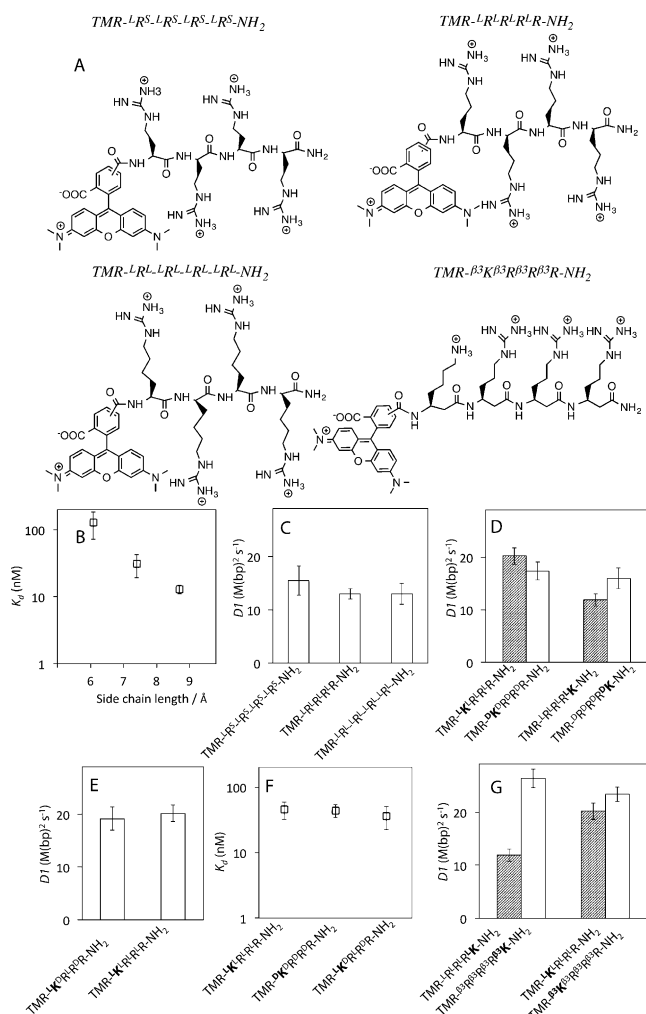


Figure 2. A) Selected chemical structures. B) Dependence of K_d values on guanidinium side chain length. The fully extended side chain lengths of ^LR^S, ^LR and ^LR^L are 6.1, 7.4 and 8.7 Å, respectively. C, D and F, G) D1 of peptide analogs in 2 mM NaCl buffer. E) The effect of the peptide backbone chirality on K_d . See the Supporting Information, Table S4 for D1 and K_d values.

peptides. The faster sliding activity of β^3 -peptides might result from the more flexible backbone which allows these molecules to adopt conformations that better accommodate the geometric constraints of sliding transition state complexes.

Based on the findings of our structure–function work with modified peptides, we set out to design a new minimal molecular sled unrelated to peptides and smaller than known molecular sleds. Previous studies showed that polyamine–DNA binding is predominantly electrostatic and DNA sequence nonspecific,^[13] characteristics that mirror those of known peptidyl molecular sleds. We selected pentaethylene-hexamine (PEHA, ~230 Da) as a test case based on its small size, the number and spacing of charges, and molecular symmetry that would yield a homogeneous population with a free terminal amine after mono-labeling with Cy3B-NHS ester. Cy3B-PEHA indeed binds DNA in line with the charge-affinity trends we have observed across peptides (Table S5).^[4] Cy3B-PEHA was found to be capable of fast

1D translocation on DNA with an estimated D1 value of $8.2 \pm 0.5 \text{ M(bp)}^2 \text{ s}^{-1}$, \pm standard error of the mean (SEM) (Figure 3B). Increasing the salt concentration does not increase the D1 value of PEHA further (Figure 3B), indicating that PEHA predominantly translocates on DNA by sliding and almost certainly undergoes rotation-coupled helical sliding.^[11,14] These results suggest that a wide variety of other polyamines and basic small molecules may slide on DNA. Polyamines are known to play important roles in DNA function such as DNA condensation, regulation of genome structure, and protection of DNA from external agents and radiation damage,^[13a] but it remains to be discovered what role sliding plays in these biological processes.

All known molecular sleds have a total net charge of $\geq +3$ and bind DNA mainly through electrostatic interactions.^[3,4] Likewise, essentially all proteins that slide on DNA have significant electrostatic interactions with nonspecific DNA. We were curious whether lower-charge molecules that achieve a preponderance of their interaction energies by a different means can also undergo 1D diffusion. Polyamides

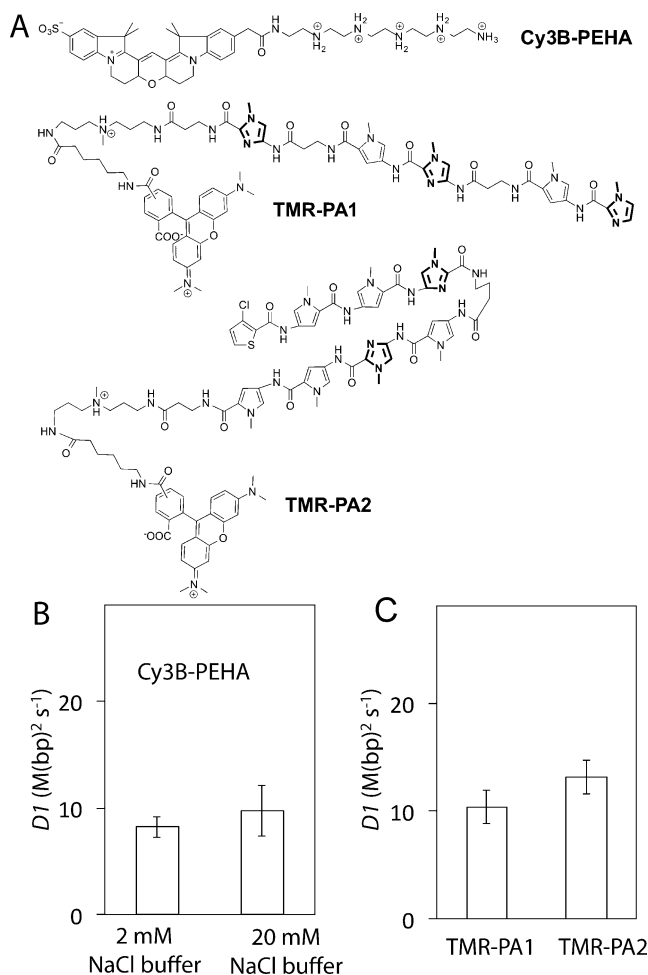


Figure 3. Molecular sled activity of synthetic small molecules. A) Chemical structures of Cy3B-PEHA, TMR-PA1 and TMR-PA2. DNA is colored by electrostatic potential. B) The D1 values of Cy3B-PEHA in 2 mM NaCl and 20 mM NaCl buffers. C) The D1 values of TMR-PA1 and TMR-PA2 in 2 mM NaCl buffer. See the Supporting Information, Table S5 for D1 values.

can be rationally designed to bind the DNA minor groove sequence-specifically through hydrogen bonding interactions and shape-selective recognition (Figure S5) and are pursued as candidates for chemical regulation of gene expression.^[8c, 10, 15, 16]

We observed that a linear polyamide,^[17] PA1, and a hairpin polyamide,^[18] PA2 (Figure 3 A, S5), each have a net charge of +1 and both exhibit fast 1D diffusion on DNA, with estimated D1 values of 10.4 ± 0.8 and 13.1 ± 0.8 m(bp)²s⁻¹, respectively (Figure 3 C, S3H,I). Despite their structural dissimilarity to peptides, both polyamides slide with rates similar to peptidyl sleds, raising the possibility of commonalities between their sequence non-specific binding modes and 1D diffusion mechanisms.

We were not able to differentiate a sliding vs. hopping mechanism for PA1 and PA2 by measuring the salt dependence of their D1 values as they bind DNA mainly through non-electrostatic interactions. An alternative approach to detect hopping is to test for drift of the sliding molecules in the direction of the buffer flow used to stretch the DNA templates. Molecules undergoing 1D diffusion may drift in the flow direction because of large size that presents significant flow drag^[19] or because of hopping that enables the molecule to freely drift when not bound to DNA (see estimation of the drift of molecules hopping on DNA in the Supporting Information). Despite the high flow rate used in our measurements of PA1 and PA2, we detected no significant drift in the flow direction, consistent with a model where the diffusing molecule remains bound to DNA more than 99.99 % of the time and very close to the DNA surface where the molecule is partially protected from the flow (Figure S7). We estimated K_d values for PA1 and PA2 to be 363 ± 76 and 432 ± 87 nM and salt-independent (Table S6), in-line with previously reported values of PA1 and PA2 binding to non-cognate DNA sequences.^[20] The average sliding lengths for PA1 and PA2 were calculated (from the apparent off-rate and D1; see Figure S8) to be 2.2 kbp and 2.7 kbp, respectively. These findings lead us to predict that polyamides are capable of sliding hundreds to thousands of base pairs at physiological salt concentrations.

Polyamide sliding provides a biophysical mechanism for target site localization within the genome by RD to 1D and facilitated diffusion like DNA binding proteins.^[1] Sliding also helps explain the preference of polyamides for clustered sites in vivo^[10] by providing a mechanism to couple the affinity of linked sites on the ~400 bp length scale where clustering effects are observed.^[10] We imagine that polyamides become locally trapped in clusters of binding sites, leading to increased occupancy at these loci.

One limitation of polyamides is their poor uptake to cells and nuclei. Previous work showed improved cellular permeability of polyamides tagged with NLS sequences.^[21] We recently reported that cell penetrating peptides and NLS sequence peptides exhibit molecular sled function, suggesting that conjugating protease-resistant peptide analogs with cell penetrating and nuclear uptake activity to polyamides might improve their efficacy.^[4]

Our observations that small molecules with a wide range of chemical functionalities slide on DNA, that specific

modifications can be used to tune crucial properties like DNA affinity and sliding speed, and that polyamides with sequence-recognition capability exhibit sliding activity open the way for engineering facilitated diffusion. In addition to sequence-specific polyamides with improved properties, small molecule inhibitors with molecular sled activity could prove efficacious as therapeutics. For example, basic compounds with chromatin-proximal targets could be optimized for cell penetrating, nuclear localization, DNA binding, and sliding activities to improve potency and reduce off-target activity outside the nucleus. Other compounds could be conjugated to minimal molecular sleds like PEHA to similarly target chromatin and suppress off-target interactions.

Acknowledgements

This work was supported by startup funding and a Career Award at the Scientific Interface (to P.C.B.) from the Burroughs Wellcome Foundation and GM133508 (to A.Z.A.) from the National Institutes of Health. We also thank Nihal Korkmaz at University of Wisconsin-Madison for help with structural modeling.

Keywords: drug delivery · molecular sleds · polyamide · single molecule imaging · small molecule

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 15110–15114
Angew. Chem. **2016**, *128*, 15334–15338

- [1] O. G. Berg, R. B. Winter, P. H. von Hippel, *Biochemistry* **1981**, *20*, 6929–6948.
- [2] a) V. Graziano, G. B. Luo, P. C. Blainey, A. J. Perez-Berna, W. J. McGrath, S. J. Flint, C. S. Martin, X. S. Xie, W. F. Mangel, *J. Biol. Chem.* **2013**, *288*, 2068–2080; b) P. C. Blainey, V. Graziano, A. J. Perez-Berna, W. J. McGrath, S. J. Flint, C. S. Martin, X. S. Xie, W. F. Mangel, *J. Biol. Chem.* **2013**, *288*, 2092–2102; c) S. A. Stratmann, S. R. Morrone, A. M. van Oijen, J. Sohn, *Elife* **2015**, *4*, e11721.
- [3] W. F. Mangel, W. J. McGrath, K. Xiong, V. Graziano, P. C. Blainey, *Nat. Commun.* **2016**, DOI: 10.1038/ncomms1020.
- [4] K. Xiong, P. C. Blainey, *Nucleic Acids Res.* **2016**, *44*, 2266–2273.
- [5] A. Turkin, L. Zhang, A. Marozzi, W. F. Mangel, A. Herrmann, A. M. van Oijen, *Chem. Sci.* **2016**, *7*, 916–920.
- [6] a) T. A. Ngo, E. Nakata, M. Saimura, T. Morii, *J. Am. Chem. Soc.* **2016**, *138*, 3012–3021; b) S. Slomovic, J. J. Collins, *Nat. Methods* **2015**, *12*, 1085–1090.
- [7] X. Li, D. R. Liu, *Angew. Chem. Int. Ed.* **2004**, *43*, 4848–4870; *Angew. Chem.* **2004**, *116*, 4956–4979.
- [8] a) S. Nishijima, K. Shinohara, T. Bando, M. Minoshima, G. Kashiwazaki, H. Sugiyama, *Bioorg. Med. Chem.* **2010**, *18*, 978–983; b) T. G. Uil, H. J. Haisma, M. G. Rots, *Nucleic Acids Res.* **2003**, *31*, 6064–6078; c) C. D. Carlson, C. L. Warren, K. E. Hauschild, M. S. Ozers, N. Qadir, D. Bhimsaria, Y. Lee, F. Cerrina, A. Z. Ansari, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4544–4549.
- [9] J. A. Raskatov, N. G. Nickols, A. E. Hargrove, G. K. Marinov, B. Wold, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16041–16045.
- [10] G. S. Erwin, D. Bhimsaria, A. Eguchi, A. Z. Ansari, *Angew. Chem. Int. Ed.* **2014**, *53*, 10124–10128; *Angew. Chem.* **2014**, *126*, 10288–10292.

- [11] P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, X. S. Xie, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5752–5757.
- [12] C. L. Vestergaard, P. C. Blainey, H. Flyvbjerg, *Phys. Rev. E* **2014**, *89*, 022726.
- [13] a) A. A. Ouameur, H. A. Tajmir-Riahi, *J. Biol. Chem.* **2004**, *279*, 42041–42054; b) H. Deng, V. A. Bloomfield, J. M. Benevides, G. J. Thomas, Jr., *Nucleic Acids Res.* **2000**, *28*, 3379–3385.
- [14] P. C. Blainey, G. B. Luo, S. C. Kou, W. F. Mangel, G. L. Verdine, B. Bagchi, X. S. Xie, *Nat. Struct. Mol. Biol.* **2009**, *16*, 1224–U34.
- [15] F. Yang, N. G. Nickols, B. C. Li, G. K. Marinov, J. W. Said, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1863–1868.
- [16] D. S. Pilch, N. Poklar, E. E. Baird, P. B. Dervan, K. J. Breslauer, *Biochemistry* **1999**, *38*, 2143–2151.
- [17] a) R. Burnett, C. Melander, J. W. Puckett, L. S. Son, R. D. Wells, P. B. Dervan, J. M. Gottesfeld, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11497–11502; b) A. R. Urbach, J. J. Love, S. A. Ross, P. B. Dervan, *J. Mol. Biol.* **2002**, *320*, 55–71.
- [18] a) A. Viger, P. B. Dervan, *Bioorg. Med. Chem.* **2006**, *14*, 8539–8549; b) D. M. Chenoweth, P. B. Dervan, *J. Am. Chem. Soc.* **2010**, *132*, 14521–14529.
- [19] P. C. Blainey, Ph.D. thesis, Harvard University, **2007**.
- [20] J. W. Puckett, K. A. Muzikar, J. Tietjen, C. L. Warren, A. Z. Ansari, P. B. Dervan, *J. Am. Chem. Soc.* **2007**, *129*, 12310–12319.
- [21] T. P. Best, Ph.D. thesis, California Institute of Technology, **2005**.

Received: July 12, 2016

Revised: October 1, 2016

Published online: November 4, 2016