ChemPubSoc Europe

DOI: 10.1002/cphc.201701147

CHEMPHYSCHEM Communications

Improved Free-Energy Landscape Quantification Illustrated with a Computationally Designed Protein–Ligand Interaction

William J. Van Patten[+,a] Robert Walder[+,a] Ayush Adhikari,[a] Stephen R. Okoniewski,[a] Rashmi Ravichandran,[b] Christine E. Tinberg,[b] David Baker,[b] and Thomas T. Perkins*[a]

Quantifying the energy landscape underlying protein–ligand interactions leads to an enhanced understanding of molecular recognition. A powerful yet accessible single-molecule technique is atomic force microscopy (AFM)-based force spectroscopy, which generally yields the zero-force dissociation rate constant (k_{off}) and the distance to the transition state ($\Delta x^{\ddagger}$). Here, we introduce an enhanced AFM assay and apply it to probe the computationally designed protein DIG10.3 binding to its target ligand, digoxigenin. Enhanced data quality enabled an analysis that yielded the height of the transition state ($\Delta G^{\ddagger} = 6.3 \pm 0.2$ kcal mol^{-1}) and the shape of the energy barrier at the transition state (linear-cubic) in addition to the traditional parameters [k_{off} ($=4 \pm 0.1 \times 10^{-4}$ s^{-1}) and $\Delta x^{\ddagger}$ ($=8.3 \pm 0.1$ Å)]. We expect this automated and relatively rapid assay to provide a more complete energy landscape description of protein–ligand interactions and, more broadly, the diverse systems studied by AFM-based force spectroscopy.

Molecular recognition between proteins and ligands is fundamental to biology. Correct recognition of antigens by antibodies, substrates by enzymes, and ligands by receptors is essential to most biological processes. In addition, the ability to custom-design proteins with precise and selective molecular recognition for a target molecule would enable the development of biosensors for a wide array of biological and medical applications.

Characterizing the strength of natural and computationally designed protein–ligand interactions is usually done in bulk assays, yielding measurements of the dissociation constant (K_D). For instance, DIG10.3, which binds the steroid digoxigenin (Dig), is the first computationally designed protein to achieve a picomolar level K_D to its target ligand.[1] Indeed, DIG10.3 exhibits an affinity that rivals that of anti-Dig antibodies.[2] Molecular details of the bound state are provided by structural studies (e.g. X-ray crystallography) and have confirmed the computationally predicted binding mode.[1] However, experimental determination of the process of dissociation remains elusive. Hence, understanding of protein–ligand interactions would benefit from an expanded description of the free-energy landscape that governs dissociation, including the height ($\Delta G^{\ddagger}$) and distance ($\Delta x^{\ddagger}$) to the transition state along with the shape of the free-energy barrier at the transition state.

Single-molecule force spectroscopy (SMFS) is a powerful technique to characterize protein–ligand interactions.[3–7] In such assays, a force applied across the protein–ligand interaction promotes detachment. The resulting data, often taken with an atomic force microscope (AFM) over a range of stretching velocities and thereby loading rates ($\partial F/\partial t$),[7] yields insight into the energy landscape underlying the protein–ligand interaction projected onto the stretching axis.[8]

Standard analysis uses the Bell-Evans model, which predicts a linear relationship between the most probable rupture force and $\log(\partial F/\partial t)$ and thereby yields the dissociation rate constant at zero applied force (k_{off}) and $\Delta x^{\ddagger}$.[9] The Bell-Evans model also predicts the dissociation rate constant k increases exponentially with F, given by $k(F) = k_{off} \exp(F\Delta x^{\ddagger}/k_BT)$, where k_BT is the thermal energy. Recently, a Bell-Evans analysis was used to characterize Dig bound to anti-Dig (anti-Dig·Dig) via acoustical force spectroscopy.[10] The resulting data, similar to prior AFM studies of anti-Dig·Dig,[11] showed two linear regimes, one for low and one for high loading rates. Such data are often interpreted as two distinct energy barriers along the stretching axis (though rebinding when using short linkages at low loading rates complicates interpretation[12]).

A more advanced analysis developed by Dudko and colleagues[13] yields important additional information: $\Delta G^{\ddagger}$ and the shape of the free-energy landscape at the transition state. In this model, the application of F not only alters the height of the barrier but moves the transition state towards the bound state. As a result, $k(F)$ initially increases exponentially with F like the Bell-Evans model but deviates from this dependency at higher F. While this more sophisticated model has been used to analyze macromolecular folding studies on custom-built optical-traps,[14] it has not yet been applied to the AFM-based molecular-recognition studies since its application requires higher-quality data than has been achieved using standard

[a] W. J. Van Patten,[+] Dr. R. Walder,[+] A. Adhikari, S. R. Okoniewski, Prof. Dr. T. T. Perkins
JILA, National Institute of Standards and Technology and the University of Colorado, Department of Physics and Department of Molecular, Cellular, and Developmental Biology
University of Colorado, Boulder, 440 UCB, Boulder, CO 80309-0440 (USA)
E-mail: tperkins@jila.colorado.edu

[b] R. Ravichandran, Dr. C. E. Tinberg, Prof. Dr. D. Baker
University of Washington, Seattle
Department of Biochemistry
Institute for Protein Design and Howard Hughes Medical Institute
Seattle, Washington 98195 (USA)

[+] These authors contributed equally to this work.

Supporting Information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/cphc.201701147.

commercially available AFMs.[15] Yet, in part, it's the accessibility and ease-of-use of commercial AFMs that make such AFM-based studies so popular.[7]

Here, we used SMFS to characterize the DIG10.3–Dig interaction, denoted as DIG10.3·Dig. Specifically, we developed an optical-trapping assay to initially assess the suitability of DIG10.3·Dig for SMFS studies and an AFM-based assay to quantify this interaction. The data quality in the AFM-based studies was improved by integrating low-drift cantilevers,[16] site-specific coupling,[17] and corrected pulling geometries.[18] These improvements resulted in the rapid acquisition of hundreds of constant-force traces on a commercial AFM. The resulting data directly revealed a non-exponential dependency of $k(F)$, in quantitative agreement with the predictions of the Dudko model.[13] Using this analysis, we determined $\Delta G^{\ddagger}$ and the shape of the free-energy barrier at the transition state (linear-cubic) in addition to the traditional parameters (k_{off} and $\Delta x^{\ddagger}$). Thus, this work extends AFM-based SMFS to rapidly determine $\Delta G^{\ddagger}$ in molecular-recognition studies, in particular, and AFM-based force spectroscopy studies, in general.

As a first step, we adapted a previously used surface-coupled optical-trapping assay[19] by simply replacing anti-Dig with DIG10.3 (Figure 1 A). In this assay, the DIG10.3 was passively adsorbed onto KOH-cleaned coverslips. We then stretched a 2-μm-long DNA molecule labeled at its 5′-ends with Dig and biotin via a streptavidin-coated microsphere held in an optical trap. A resulting force-extension curve acquired by moving the sample surface at constant velocity (2 μm s^{-1}) showed the canonical DNA overstretching transition[20,21] at 65 pN (Figure 1 B). With this simple assay, we also performed optical-trapping experiments at constant F. Individual traces showed DIG10.3·Dig often withstood 20 pN for more than 100 s (Figure 1 C). Hence, the DIG10.3·Dig interaction was sufficiently robust for application in many existing low-to-moderate force SMFS assays, including those used to characterize DNA-based molecular motors[22] and nucleic-acid structures.[23,24]

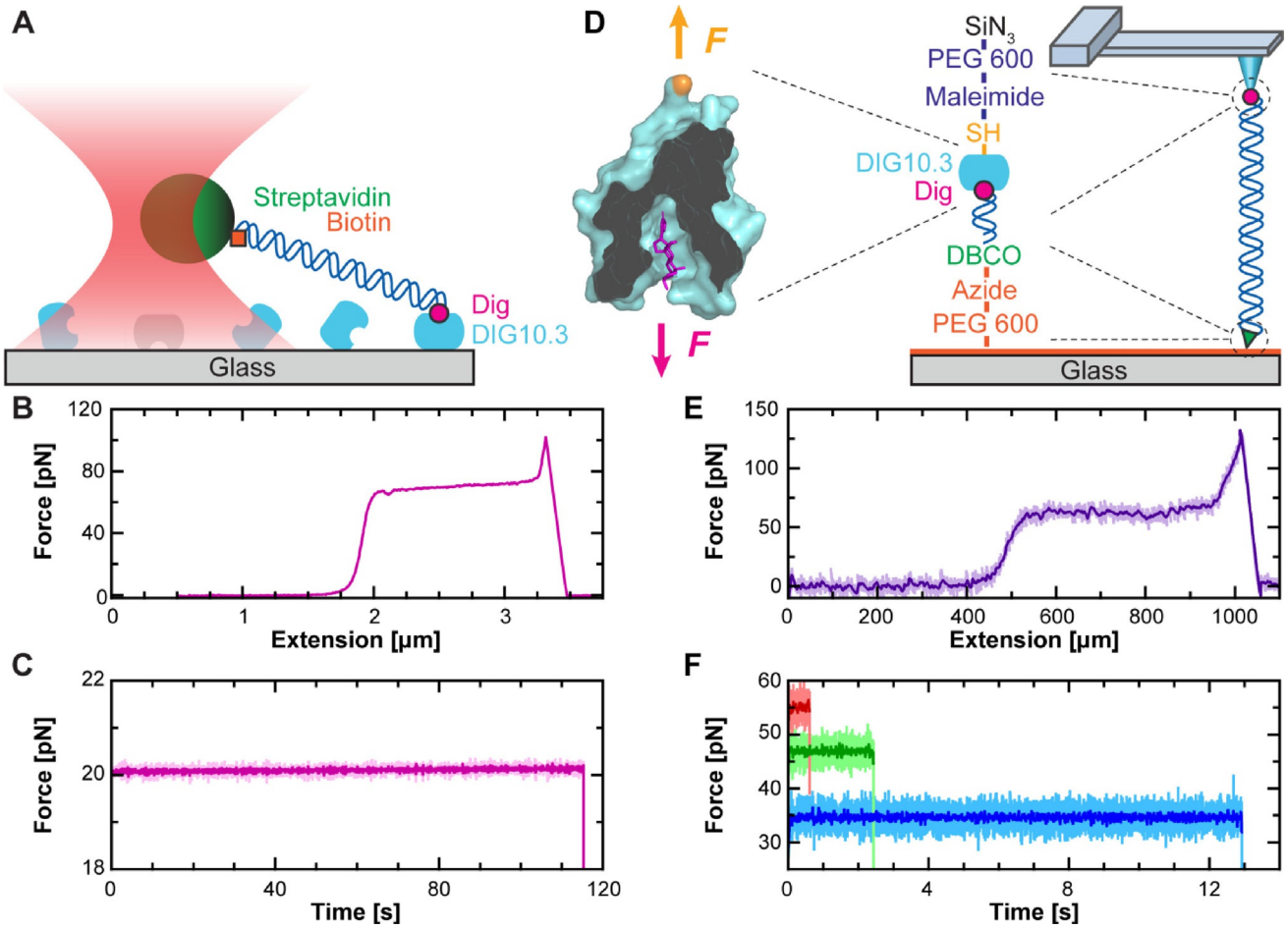

Figure 1. Single-molecule force spectroscopy (SMFS) studies using DIG10.3·Dig. A) Schematic of DIG10.3·Dig being used in an optical-trapping-based SMFS assay. B) Force-extension curve shows DNA being stretched with an optical trap undergoing the overstretching transition at 65 pN. Data smoothed to 500 Hz. C) A force-vs.-time trace shows the lifetime of DIG10.3·Dig stretched under constant force by an optical trap. Data smoothed to 40 Hz (light pink) and 10 Hz (dark pink). D) Cartoon of the coupling scheme for AFM-based assay showing force applied to the ligand and a N-terminal cysteine variant of DIG10.3. DIG10.3 was covalently coupled to a maleimide-functionalized, PEG-coated AFM tip via a maleimide-thiol bond. Dig-labeled DNA was covalently coupled to an azide-functionalized, PEG-coated coverslip using dibenzocyclooctyl (DBCO), a copper-free click chemistry reagent. Abbreviations: SiN_3, silicon nitride. E) Force-extension curve taken using an AFM also shows the DNA's overstretching transition. Data smoothed to 50 kHz (light purple) and 50 Hz (dark purple). F) Force-vs.-time traces using an AFM show the lifetime of DIG10.3·Dig stretched at constant force. Data smoothed to 1 kHz (light colors) and 50 Hz (dark colors).

To rapidly quantify the 1D free-energy landscape of DIG10.3·Dig, we developed an AFM-based assay where all of the linkages were covalent, except DIG10.3·Dig (Figure 1D). Hence, any rupture could be attributed to the dissociation of DIG10.3·Dig, rather than failure of another linkage (e.g. the biotin-streptavidin coupling or the non-specific adsorption of DIG10.3 to the coverslip). Moreover, we automated data acquisition, a capability not present in our custom optical trap.[25]

Assembly of this assay started with a PCR-amplified, 635-nm-long DNA to introduce distinct 5′ labels, Dig and dibenzocyclooctyl (DBCO), a copper-free click chemistry reagent. We then covalently coupled this DNA via the DBCO to an azide-functionalized coverslip.[17] The surface density of the DNA was purposely kept low to promote attachment to a single DNA molecule. In parallel, we covalently coupled a variant of DIG10.3 with an N-terminal cysteine to maleimide-functionalized cantilevers. Non-specific adhesion was minimized by using a PEG-coating on both the cantilever and the coverslip.

To improve the overall quality of the AFM data, a number of important technical improvements were integrated. First, we improved force stability by chemically etching the metal coating off of the cantilever, achieving sub-pN force stability over 100 s.[16] Next, we improved the accuracy of the assay by positioning the DNA's anchor point to the coverslip directly below the AFM tip, assuring a vertical stretching geometry.[18] This alignment is particularly important when using stiff linkers like DNA (persistence length = 50 nm) where large lateral offsets (≈180 nm) between the tip and the DNA anchor point on the coverslip have been observed for short (650-nm) DNA.[18] Finally, we improved force precision by performing all of the measurements with a single cantilever, similar to prior work on the unfolding of proteins.[26]

We initiated this improved assay by lowering a DIG10.3-coated AFM tip towards a coverslip sparsely coated with Dig-labeled DNA and then pressing the tip gently (100 pN) into the coverslip for 2 s. This force was relatively low compared to the ≈1,000 pN generally used to promote non-specific attachment.[27] We next retracted the tip from the surface at constant velocity (1 μm s^{-1}) while using a real-time triggering scheme that stopped the retraction when $F > 20$ pN at an extension $x > 150$ nm. This scheme selected for a connection consistent with stretching DNA (rather than surface adhesion). We then performed an automated centering routine to position the DNA anchor point directly under the tip.[18] Next, we either stretched the DNA at a constant velocity (Figure 1E) or held the DNA and thereby the DIG10.3·Dig complex under constant F (Figure 1F). Upon rupture, the cantilever was lowered back to the surface and often reattached to the same individual DNA molecule, improving throughput as outlined in Figure S1.

To precisely characterize DIG10.3·Dig energetics, we needed to verify that a single DNA was stretched via DIG10.3·Dig. First, we only analyzed force-extension curves that showed a single rupture that returned to $F = 0$ pN, consistent with a single molecule attached to the AFM tip. Next, we demonstrated the vast majority of linkages to the tip were via DIG10.3·Dig by blocking digoxigenin-labeled DNA with the anti-Dig antibody, which dramatically reduced tip attachment (Figure S2). Finally, as discussed above, the only non-covalent linkage in the assay was DIG10.3·Dig, so ruptures could be attributed to the dissociation of that interaction (though dissociation arising from force-induced unfolding of DIG10.3 remains a possibility as in all SMFS protein–ligand assays).

Constant-force measurements provided a more direct comparison to theoretical predictions, although such measurements are more technically challenging to implement. In particular, both the Bell-Evans[9] and the Dudko[13] models have clear predictions for $k(F)$. Constant-velocity measurements lead to more complicated predictions of the distribution of rupture forces.[13,28] As discussed above, we met the key requirement of constant force by using cantilevers that featured sub-pN stability over 100 s.[16] In contrast, the same cantilever with its gold-coating drifts ≈50 pN over 100 s, even 2 h after immersion in liquid.[16]

Using an automated data-acquisition process, we recorded Dig10.3·Dig lifetimes over a range of F. Histograms of the resulting lifetimes ($N \approx 50$–250 per F) were analyzed to deduce a dissociation rate constant $k(F)$ at each F from a fit to a single exponential (Figure 2A). Larger F led to shorter lifetimes and thus a higher dissociation rate constant (Figure 2B). Below 40 pN, $k(F)$ scaled exponentially with F and therefore increased linearly on a log-linear plot (Figure 2B, dashed line). Yet, at higher F, $k(F)$ deviated from this linearity. To verify that this curvature was reproducible, we took two independent sets of

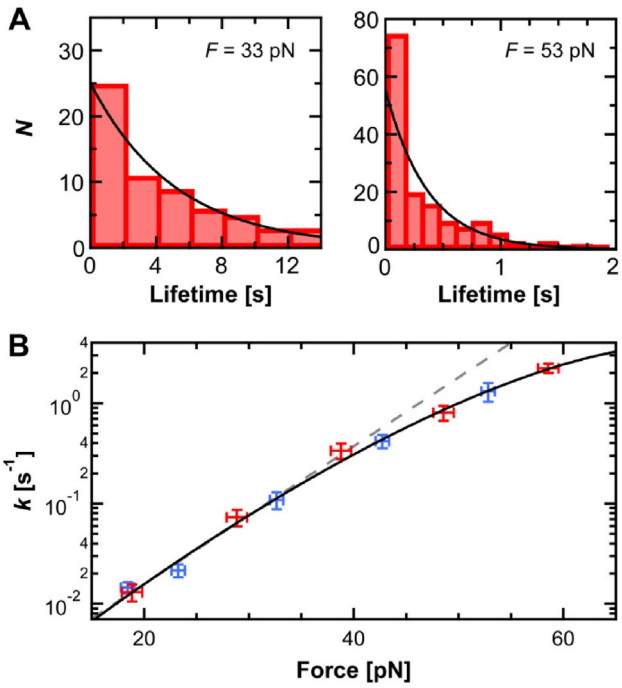

Figure 2. Quantifying DIG10.3·Dig under force. A) Histograms of DIG10.3·Dig lifetime at two different forces (33 and 53 pN, $N = 65$ and 159 respectively) were analyzed by fitting to an exponential (black line). B) The dissociation rate constant k as a function of F shows curvature on a log-linear plot. Two data sets taken on consecutive days are shown (red and blue). The resulting data was fit to the Dudko model ([Eq. (1)], black line). Error bars represent the standard error of k ($N \approx 50$–250 per point; $N_{total} = 1{,}081$) and standard deviation of F. Gray dashed line represents a Bell-Evans fit to the data below 40 pN.

data over two sequential days using the same individual cantilever for improved precision (Figure 2B, red and blue). The resulting data showed quantitative agreement, providing assurance that this curvature was reproducible.

To see if the observed non-exponential scaling of $k(F)$ was consistent with the Dudko analysis,[13] we fit the data using [Eq. (1)]

$$k(F) = k_{off}\left(1-\frac{vF\Delta x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{\frac{1}{v}-1}\exp\left(\frac{\Delta G^{\ddagger}}{k_B T}\left[1-\left(1-\frac{vF\Delta x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{\frac{1}{v}}\right]\right) \quad (1)$$

The fit showed excellent agreement between our data and this model (Figure 2B, black line). The resulting parameters were $k_{off}=4.1\pm0.4\times10^{-4}\ s^{-1}$, $\Delta x^{\ddagger}=8.3\pm0.1$ Å, $\Delta G^{\ddagger}=6.3\pm0.2$ kcal mol^{-1} and $v=0.67\pm0.02$. Uncertainties correspond to the standard errors for the fitting parameters. This value of v quantitatively agrees with the value of 2/3 predicted by Dudko and colleagues[13] for a linear cubic potential at the transition state (whereas a cusp potential would correspond to $v=0.5$ and the simpler Bell-Evans model corresponds to $v=1$).

How do SMFS measurements of DIG10.3·Dig compare to those of anti-Dig·Dig? As the DIG10.3·Dig data were taken at constant force, we compared our results to the lower-loading rate barrier in the recent anti-Dig·Dig experiments.[10] In particular, k_{off} for DIG10.3·Dig is 6-fold higher than anti-Dig·Dig, despite a lower K_D, implying a faster k_{on} since $K_D=k_{off}/k_{on}$. Interestingly, $\Delta x^{\ddagger}$ for DIG10.3·Dig was about half of that for anti-Dig·Dig ($\Delta x^{\ddagger}=15.4\pm0.7$ Å). As a result, whereas DIG10.3·Dig has a higher intrinsic k_{off}, the lifetime of DIG10.3·Dig at $F>10$ pN is predicted to be longer than for anti-Dig·Dig (Figure S3). This result arises from the difference in force sensitivity due to a smaller $\Delta x^{\ddagger}$. On a practical note, the overall similarity in interaction lifetime over the 5–20 pN range suggests that DIG10.3, which was expressed in *E. Coli*, can easily substitute for anti-Dig in a wide range of current SMFS assays.

The determination of $\Delta G^{\ddagger}$ demonstrated here in a commercial AFM opens the door to a more complete and accessible determination of free-energy landscape parameters of diverse protein–ligand systems. However, we note that resolving the curvature visualized in Figure 2B—essential to constraining $\Delta G^{\ddagger}$ and v within the Dudko model—required advances in both the quality and quantity of data. Quality was improved by using cantilevers that featured sub-pN stability,[16] which was essential since $k(F)$ scales exponentially with F at low F. Moreover, a sufficient quantity of high-quality data was equally critical. Initial studies with ≈20 lifetimes at each F did not show the curvature seen in Figure 2B. Only when we acquired ≈50–250 lifetimes at each F was the curvature well defined and reproducible. Finally, while using DNA may seem inconvenient, the distinct advantage of a long spacer over the much shorter PEG linkers commonly used in many SMFS studies of molecular recognition is that the extended DNA rapidly recoils upon rupture, preventing rebinding artifacts that can alter interpretation.[12]

In summary, we developed an enhanced constant-force AFM-based assay to provide an expanded description of the free-energy landscape underlying protein–ligand interactions. We demonstrated it by characterizing a computationally designed protein–ligand interaction and applying a more sophisticated analysis to yield $\Delta G^{\ddagger}$ and the shape of the energy barrier at the transition state in addition to the traditional parameters (k_{off} and $\Delta x^{\ddagger}$). By doing so in an automated and relatively rapid manner (≈2 days), we anticipate this assay will provide a more complete energy landscape description of diverse natural protein–ligand interactions as well as experimental feedback to further optimize computationally designed interactions. In turn, enhanced computationally designed protein–ligand interactions offer promising capabilities in diagnostics and therapeutics. Finally, our improved AFM-based assay is not limited to protein–ligand interactions but can be immediately adopted to studies of protein–protein interactions[29] and proteins unfolding under constant force.[30]

Acknowledgements

We thank Lyle Uyetake for DNA preparation. This work was supported by an NIH Molecular Biophysics Training Grant awarded to S.R.O. (T32 GM-065103), a Graduate Research Fellowship from the NSF [DGE 1144083 to S.R.O.], HHMI to D.B., the NSF (DBI-1353987 to T.T.P.; MCB-1716033 to T.T.P., Phy-1734006), and NIST. T.T.P is a staff member of NIST's quantum physics division.

Conflict of interest

The authors declare no conflict of interest.

Keywords: atomic force microscopy · energy landscape · protein design · protein–ligand interactions · single-molecule force spectroscopy

[1] C. E. Tinberg, S. D. Khare, J. Dou, L. Doyle, J. W. Nelson, A. Schena, W. Jankowski, C. G. Kalodimos, K. Johnsson, B. L. Stoddard, D. Baker, *Nature* **2013**, *501*, 212–216.
[2] S. Y. Tetin, K. M. Swift, E. D. Matayoshi, *Anal. Biochem.* **2002**, *307*, 84–91.
[3] V. T. Moy, E. L. Florin, H. E. Gaub, *Science* **1994**, *266*, 257–259.
[4] G. U. Lee, D. A. Kidwell, R. J. Colton, *Langmuir* **1994**, *10*, 354–357.
[5] Y. Gilbert, M. Deghorain, L. Wang, B. Xu, P. D. Pollheimer, H. J. Gruber, J. Errington, B. Hallet, X. Haulot, C. Verbelen, P. Hols, Y. F. Dufrene, *Nano Lett.* **2007**, *7*, 796–801.
[6] M. Rangl, A. Ebner, J. Yamada, C. Rankl, R. Tampe, H. J. Gruber, M. Rexach, P. Hinterdorfer, *Angew. Chem. Int. Ed.* **2013**, *52*, 10356–10359; *Angew. Chem.* **2013**, *125*, 10546–10549.
[7] P. Hinterdorfer, Y. F. Dufrene, *Nat. Methods* **2006**, *3*, 347–355.
[8] R. Merkel, P. Nassoy, A. Leung, K. Ritchie, E. Evans, *Nature* **1999**, *397*, 50–53.
[9] E. Evans, K. Ritchie, *Biophys. J.* **1997**, *72*, 1541–1555.
[10] G. Sitters, D. Kamsma, G. Thalhammer, M. Ritsch-Marte, E. J. G. Peterman, G. J. L. Wuite, *Nat. Methods* **2015**, *12*, 47–50.
[11] G. Neuert, C. Albrecht, E. Pamir, H. E. Gaub, *FEBS Lett.* **2006**, *580*, 505–509.
[12] R. W. Friddle, A. Noy, J. J. De Yoreo, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13573–13578.
[13] O. K. Dudko, G. Hummer, A. Szabo, *Phys. Rev. Lett.* **2006**, *96*, 108101.
[14] W. J. Greenleaf, K. L. Frieda, D. A. Foster, M. T. Woodside, S. M. Block, *Science* **2008**, *319*, 630–633.
[15] K. C. Neuman, A. Nagy, *Nat. Methods* **2008**, *5*, 491–505.

[16] A. B. Churnside, R. M. Sullan, D. M. Nguyen, S. O. Case, M. S. Bull, G. M. King, T. T. Perkins, *Nano Lett.* **2012**, *12*, 3557 – 3561.
[17] R. Walder, M.-A. LeBlanc, W. J. Van Patten, D. T. Edwards, J. A. Greenbert, A. Adhikari, S. R. Okoniewski, R. M. A. Sullan, D. Rabuka, M. C. Sousa, T. T. Perkins, *J. Am. Chem. Soc.* **2017**, *139*, 9867 – 9875.
[18] R. Walder, W. J. Van Patten, A. Adhikari, T. T. Perkins, *ACS Nano* **2017**, https://doi.org/10.1021/acsnano.7b05721.
[19] Y. Seol, J. Li, P. C. Nelson, T. T. Perkins, M. D. Betterton, *Biophys. J.* **2007**, *93*, 4360 – 4373.
[20] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J. L. Viovy, D. Chatenay, F. Caron, *Science* **1996**, *271*, 792 – 794.
[21] S. B. Smith, Y. Cui, C. Bustamante, *Science* **1996**, *271*, 795 – 799.
[22] Y. R. Chemla, *Phys. Chem. Chem. Phys.* **2010**, *12*, 3080 – 3095.
[23] J. Liphardt, B. Onoa, S. B. Smith, I. J. Tinoco, C. Bustamante, *Science* **2001**, *292*, 733 – 737.
[24] E. Pfitzner, C. Wachauf, F. Kilchherr, B. Pelz, W. M. Shih, M. Rief, H. Dietz, *Angew. Chem. Int. Ed.* **2013**, *52*, 7766 – 7771; *Angew. Chem.* **2013**, *125*, 7920 – 7925.
[25] A. R. Carter, Y. Seol, T. T. Perkins, *Biophys. J.* **2009**, *96*, 2926 – 2934.
[26] M. Otten, W. Ott, M. A. Jobst, L. F. Milles, T. Verdorfer, D. A. Pippig, M. A. Nash, H. E. Gaub, *Nat. Methods* **2014**, *11*, 1127 – 1130.
[27] T. Bornschlogl, M. Rief, *Methods Mol. Biol.* **2011**, *783*, 233 – 250.
[28] E. Evans, K. Ritchie, *Biophys. J.* **1999**, *76*, 2439 – 2447.
[29] S. W. Stahl, M. A. Nash, D. B. Fried, M. Slutzki, Y. Barak, E. A. Bayer, H. E. Gaub, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 20431 – 20436.
[30] J. M. Fernandez, H. Li, *Science* **2004**, *303*, 1674 – 1678.

Manuscript received: October 23, 2017
Accepted manuscript online: October 25, 2017
Version of record online: December 4, 2017