JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Modular Stitching To Image Single-Molecule DNA Transport

Juan Guan,[†] Bo Wang,[†] Sung Chul Bae,[†] and Steve Granick*[,†,‡,§]

[†]Department of Materials Science and Engineering, [‡]Department of Physics, and [§]Department of Chemistry, University of Illinois, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: For study of time-dependent conformation, all previous single-molecule imaging studies of polymer transport involve fluorescence labeling uniformly along the chain, which suffers from limited resolution due to the diffraction limit. Here we demonstrate the concept of submolecular single-molecule imaging with DNA chains assembled from DNA fragments such that a chain is labeled at designated spots with covalently attached fluorescent dyes and the chain backbone with dyes of different color. High density of dyes ensures good signal-to-noise ratio to localize the designated spots in real time with nanometer precision and prevents significant photobleaching for long-time tracking purposes. To demonstrate usefulness of this approach, we image electrophoretic transport of λ-DNA through agarose gels. The unexpected pattern is observed that one end of each molecule tends to stretch out in the electric field while the other end remains quiescent for some time before it snaps forward and the stretch−recoil cycle repeats. These features are neither predicted by prevailing theories of electrophoresis mechanism nor detectable by conventional whole-chain labeling methods, which demonstrate pragmatically the usefulness of modular stitching to reveal internal chain dynamics of single molecules.

Almost a generation since Chu et al. first imaged fluorescently labeled single DNA molecules,[1] it remains challenging to resolve the submolecular dynamics of a flexible polymeric chain especially when the chain is highly coiled or looped. This limits our understanding of macromolecular transport, as motion at the level of chain segments is a fundamental measure of polymer dynamics.[2−4] To make progress, one needs to overcome three resolution challenges: the optical diffraction limit, tracking rapid motion, and knowing the relative positions of labeled segment along the chain. While the first can involve localization methods that have driven recent progress in super-resolution fluorescence imaging,[5] here we approach the latter two with modular assembly of labeled short segments into longer chains such that desired segments of a chain can be labeled with their positions resolved below the diffraction limit and monitored rapidly with respect to the main chain contour.

Our synthesis starts by constructing through restriction enzyme digestion molecular modules with defined overhanging ends. Typically a few kilobases long, they are then covalently labeled with dye and grafted to desired positions of a long parent chain with synthetic "stitching oligomers" through complementary base pairing. As hundreds of overhang sequences can be created by restriction enzymes and oligomers of any desired sequence can be synthesized, the method can potentially label segments at any desired position along a chain and even in more complex molecular architecture and networks through programmable and orthogonal modular stitching. Examples of chain architectures that we anticipate are sketched in Figure 1a, with one example implemented below.

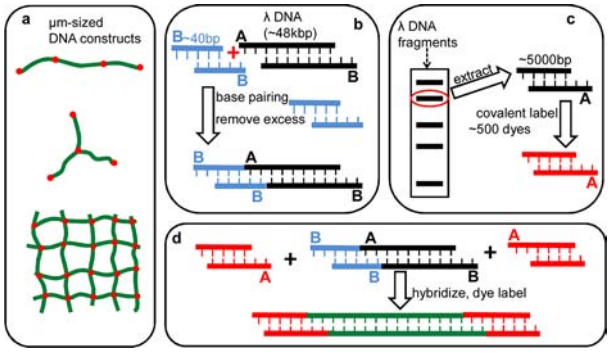

Figure 1. Scheme of submolecular modular labeling of DNA. (a) Micrometer-scale DNA constructs may be synthesized with desired chain architectures and modular labeling. These sketches show a linear chain labeled at designated spots, the ends and within the chain; a 3-arm end-labeled star; and a network labeled at the cross-links. (b−d) End-labeling procedure employed in this work. (b) First, onto the overhanging end of λ-DNA, labeled "A" (~48 kbp), is attached a shorter segment, labeled "B" (~40 bps). The new end overhangs now bear the B sequence, which is available for further reaction. (c) Next, molecules bearing overhang sequence A are purified and labeled with covalently attached fluorescent dyes. (d) Finally, the DNA fragments in (c) are attached onto the two ends of the DNA prepared in (b), and the chain backbone is labeled with dye molecules of different color.

The synthesis scheme meets imaging-specific desiderata, although molecular cloning and DNA nanotechnology build on similar ideas but use fewer fluorescent dyes: (a) The parent chain size should exceed the diffraction limit (~300 nm or ~10 kbp), just as for many studies in prior literature.[1,6] (b) The size of labeled modular segments should be less than the diffraction limit, enabling their localization with nm precision, yet long enough to accommodate sufficient dyes to provide adequate signal-to-noise for imaging in bulk microenvironments over long time (DNA chains of 2500−5000 bps are found empirically to work well). (c) Dye pairs, on the labeled

Received: February 25, 2013
Published: April 9, 2013

   dx.doi.org/10.1021/ja4020138 | J. Am. Chem. Soc. 2013, 135, 6006−6009

segments and on the main chain contour, should possess minimal spectral overlap for simultaneous two-color imaging.

The specific example studied in this paper concerns single-molecule imaging of dynamics internal to polymer chains: the ends and main chain contour of linear DNA. We constructed a hybrid chain assembled from λ-DNA as the parent chain, ~16 μm (or 48 kbp) in length, and two segments (~5 kbp), each grafted onto an end of the parent chain. Chain ends are commonly believed to play critical roles in polymer dynamics in both equilibrium and driven systems,[2−4] and our hybrid chain allows tracking simultaneously dynamics of chain ends with respect to the main chain. The detailed procedure is now summarized. First, one end of λ-DNA (denoted as "A" in Figure 1b) is modified to bear the same sequence as the other end (denoted as "B" in Figure 1b, an overhang of 12 bases in length) by attaching a short stitching DNA oligomer (~40 bps) to end A and removing excess oligomers through a centrifuge column (Amicon column, Millipore). Meanwhile, we separate λ-DNA *Hin*dIII fragments (NEB) into bands using agarose gel electrophoresis and subsequently extract the DNA fragment containing end A from this gel (Qiagen gel extraction spin column) as *Hin*dIII generates fragments of desired length and overhang. We prepare molecular modules, kilobases long, from fragments cut by restriction enzymes because synthetic DNA oligomers are available commercially only below ~100 bps. We covalently label each extracted fragment with ~500 Cy5-derivative dye molecules (Mirus Bio) and further concentrate the solution to ~100 nM (Amicon spin column) (Figure 1c). Because ends A and B of λ-DNA are complementary in sequence, we then hybridize the labeled fragment bearing end A with modified λ-DNA bearing end B for 1 day to improve hybridization efficiency (Tris 10 mM, EDTA 1 mM, $MgCl_2$ 10 mM, pH 7.5). Last, we label the hybrid chain with diffusive DNA dyes (Sybr Gold, Invitrogen) in order to visualize the main chain (Figure 1d). For this study we ligate the modules to enhance stability of end-attachment before dispersing hybrid chains into molten agarose gel at high temperature.

To boost reaction yield, we found the following strategies useful: (a) small stitching DNA oligomers are added in >100-fold excess and later removed prior to hybridization between modules, and (b) labeled modules are concentrated to ~100 nM to improve hybridization efficiency. Overall we achieve ~30% yield.

It is worth noting alternative strategies that failed. For example, shorter DNA segments (~100 bps) attached to chain ends could not be imaged faithfully. Although this lower density of dye labeling is appropriate to DNA bar-coding and imaging static conformations on surfaces and in nanochannels,[7] imaging rapid motion far from a surface demands brighter dye-labeling.

To demonstrate the power of this approach, we imaged such chains in gel electrophoresis with emphasis on internal DNA dynamics during this process. The transport features we observe were neither detectable with a conventional whole-chain labeling approach nor predicted by prevailing theories and simulations. The λ-DNA, labeled as described above, was embedded within 1.5 wt% agarose gel (Fisher, molecular biology grade, low EEO) in the presence of 1× TBE and glucose oxidase-based anti-photobleaching buffer, and visualized using a home-built setup for two-color epifluorescence tracking (Zeiss observer.Z1, 100× objective) with time resolution of 50 ms (Andor iXon EMCCD cameras). Qualitative inspection of the raw data obtained at 9 V/cm (Figure 2a) shows that chain extension fluctuates considerably

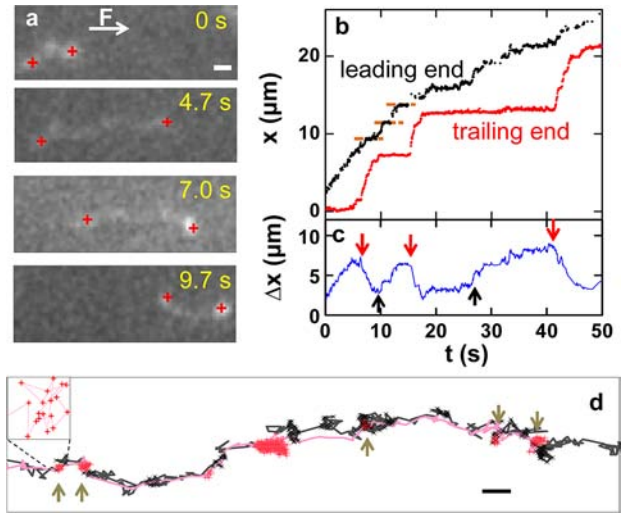

Figure 2. Illustrations of how end-labeled λ-DNA migrates through 1.5 wt% agarose gel at 9 V/cm. (a) Images of the same molecule during a time span of 10 s. The ends, marked with red crosses, are overlaid onto the chain contour (white). One end of a chain usually leads in migration. Electric force (F) points to the right. Scale bar: 1 μm. (b) An illustrative molecule's time-dependent position along electric force direction showing asymmetry between two ends. The leading end (black) advances relatively continuously, while the trailing end (red) displays longer pauses and jerky advance. The first few pausing events of leading ends are marked by dashed lines (orange). (c) The extension between two ends fluctuates, reflecting leading end stretch (↑) and trailing end recoil (↓). (d) The trajectories of the two ends, leading end (gray) and trailing end (pink), are overlaid for a representative molecule. The respective pause positions are compared, leading end (black) and trailing end (red). The two ends are seen to pause at nearly the same positions within the agarose gel, indicated by arrows. Electric force points to the right. Scale bar: 1 μm. Inset: magnified view of local motion of an end during pause. Region is 200 × 200 nm^2.

on the time scale of seconds. Rapid data acquisition is needed to see it. Such images were quantified with 10 nm resolution using home-built software based on standard methods.[8] Simultaneously tracking both chain ends (red) and the main chain contour, we find that one end of a chain usually (>90% of the time) leads the rest of the chain whereas the other end trails. This preferential end protrusion contrasts with the common expectations of so-called "hernia" or "hairpin" formations which predict random protrusion of any segment to form folded chain conformation.[9,10]

Plotting against time the end position in the electric force direction (Figure 2b), we also observe frequent pausing, even though a constant electric force is applied to the chain. Beyond intermittent motion of the center-of-mass, we also observe asymmetry between the two ends: while the leading end moves relatively continuously, the trailing end displays long pauses followed by rapid catch-up jumps. Plotting the time-dependent distance between the head and tail of a representative molecule (Figure 2c), one sees the origin of the large-scale chain length fluctuation: stretch of the leading end causes chain extension, and recoil of the trailing end causes contraction. The onsets of stretch of the leading end (↑) and recoil of the trailing end (↓) are also highlighted in this figure. A larger data set is summarized in Figure 3.

Inspecting a representative trajectory more closely, one observes that the leading and trailing ends of the molecule

dx.doi.org/10.1021/ja4020138 | J. Am. Chem. Soc. 2013, 135, 6006−6009

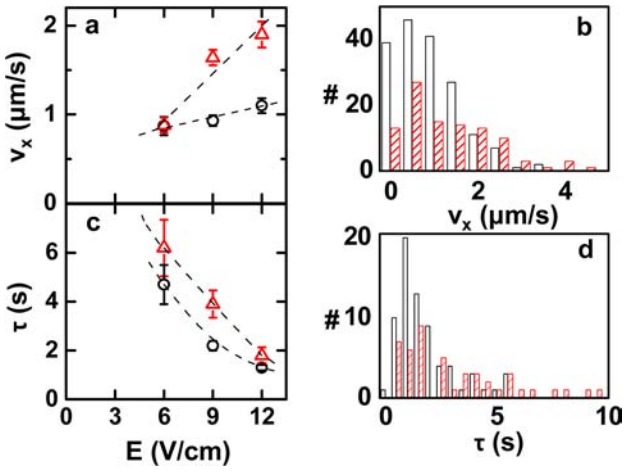

Figure 3. Electric field dependence of asymmetry between two ends quantified after averaging ∼100 events involving ∼20 molecules at each field. (a) Ensemble-averaged velocity between pauses in the electric force direction is plotted against field E for leading end during stretch events (○) and trailing end during recoil events (△). Error bars are estimated by the standard bootstrap methods. Lines are guides to the eye. (b) For E = 9 V/cm, the velocity distribution is compared for leading (open bar) and trailing (shaded bar) ends. Velocity is evaluated over 0.5 s. (c) Pause time between transport (stretch or recoil) phases is plotted against E. Symbols are the same as in panel a. Pause is identified from an apparent plateau in time-dependent displacement, such as those shown in Figure 2b. (d) For E = 9 V/cm, the pause time distribution is compared. Bars are defined as in (b).

follow almost the same migration path through the agarose gel and pause at almost the same locations within the gel, but at different times. This is illustrated in Figure 2d, where trajectories of ends are projected into the $x-y$ plane and electric force points to the right. Trajectories of the leading end (gray) and trailing end (pink) are seen to pause at nearly the same positions, as denoted by the black points (leading end) and red points (trailing end). Upon surveying multiple molecules, we find that the two ends consistently share ∼80% of the pausing locations across a range of electric field. The overlap of pausing locations between chain ends suggests the middle chain segments also pause at these locations, although we do not explicitly track the middle segments here. In other words, the resistance to transport is experienced sequentially along a migrating chain as each segment sequentially encounters the same "trap" in the network. Yet there seems to be no fixed adhesion to gel fibers: chain ends remain mobile locally over lengths consistent with estimated mesh size (Figure 2d, inset). Physically, it seems that pausing reflects tight constrictions in the static network structure that hinder transport as a chain squeezes through.

Experiments were also performed with DNA labeled at just one end. That end-labeling caused neither preferential end protrusion nor intermittency followed from our observation that the labeled ends can either lead or trail while displaying similar intermittency and stretch−recoil events. Earlier studies had shown that gel electrophoresis may involve intermittent motion and chain length fluctuation in experiments[11] and early simulations.[4,12] The novelty of this study is to reveal internal details of the constituent chain dynamics that were, to our knowledge, not identified previously. For example: intermittency was not previously so clearly attributable to passage through tight constrictions in the gel, and chain length fluctuation was not identified as stretch of leading end and recoil of trailing end.

Furthermore, submolecular labeling enables one to measure the average moving velocity of leading end during stretch and trailing end during recoil (Figure 3a), with stretch and recoil phases readily identified on the basis of their time-dependent positions between pausing events (Figure 2b). Figure 3 compares the dynamics between two ends after averaging many events involving multiple molecules under an electric field spanning from low to high. At low field (6 V/cm), the DNA chain conformation is close to that in the absence of field and never shows large extensions, whereas with increasing field strength, larger fluctuations of chain extension are observed. Although both ends move faster with increasing electric field, consistent with previous studies regarding the center-of-mass mobility of such molecules, recoil of trailing end is faster than stretch of leading end at high electric field (Figure 3a). The histogram of instantaneous velocity (Figure 3b, evaluated at 9 V/cm) further shows that the trailing end has a large portion of fast recoil events. The faster recoil of trailing ends, relative to stretch of leading ends at the highest field strengths, suggests that intrachain tension facilitates recoil by pulling trailing ends after the chain is extended. At low electric field, this difference is less prominent, as when the field is low, chains are rarely stretched to a high extension and therefore are unable to generate high tension.

Comparing the pause times under different electric fields, we see that it decreases rapidly for both ends with increasing electric field (Figure 3c). Since protrusion of chain ends initiates chain transport (Figure 2a) and the pause duration of leading ends decreases with increasing field (Figure 3c), electric field can promote transport by facilitating protrusion of chain ends. This is consistent with the observation of increased fraction of "hooking" events with increasing field (Figure S1), which also suggests easier end protrusion at higher field. The pause time of trailing ends exceeds that of leading ends (Figure 3c) and is more broadly distributed in the range 0.1−10 s as the chain gets caught in elongated conformations (Figure 3d). Movie S1 in the Supporting Information makes the pattern plain: one end of a chain tends to stretch out and pulls slack from the still-quiescent remainder of the chain until the other end is yanked forward.

To put this work into context: while macroscopic aspects of electrophoresis (separation efficiency and its dependence on electric field and gel type) have been understood for a long time,[4] the present new observations of microscopic chain dynamics at the single-molecule level encourage us to compare to the prevailing models on internal chain dynamics. Given the chain length and gel mesh size (∼200 nm for this agarose concentration), this system is in the long-chain regime, where at equilibrium the mass of a chain spans approximately 20 meshes. The classical theories of "reptation" biased by an electric field[4] predict neither intermittent displacement nor the large chain length fluctuations that we observe. In "hooking" (or "geometration") models, intermittency of transport and chain length fluctuation are considered to reflect the chain hooking onto a gel fiber, becoming stretched by this event, and subsequently sliding off it. But we quantified the incidence of hooking (Figure S1), and it emerges that hooking is only seen for a fraction of recoil events (10% at 6 V/cm, 30% at 9 V/cm, and <50% at 12 V/cm). Moreover, even when hooking is observed, chain transport seems to also involve the pausing mechanism discussed in Figure 2d. The "entropic trapping"

dx.doi.org/10.1021/ja4020138 | J. Am. Chem. Soc. 2013, 135, 6006−6009

model does admit intermittency, but it appears to predict random protrusion of all segments,[9,10] whereas we observe preferential protrusion at the chain ends. Predictions of the entropic barrier model are best developed for cases of single barrier crossing,[9,10] but these data suggest instead a scenario involving multiple barrier-crossing events as segments along the chain sequentially encounter a given tight constriction in the gel network (Figure 2d). It would be interesting to revisit the simulation models[4,9d,13−15] in light of the new findings reported here.

These novel observations are more pertinent to polymer dynamics than the separation technology of gel electrophoresis, which as an analytical technique already works well.[4] It is a grand challenge in polymer science to understand the motions and relaxations of individual polymer chains,[2,3] and DNA has long been appreciated as a model "polymer" for this purpose.[1] The main point of this Communication is that methods of submolecular fluorescence imaging, using the facile modular stitching approach described here, can provide new data to which prevailing models can be compared critically. Although the present data are limited so far to just a single DNA chain length (λ-DNA) and to a single agarose concentration, the characteristic transport features would not have been detectable by the conventional whole-chain labeling approach.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details and movies S1 and S2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

sgranick@illinois.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are indebted to Andrey Dobrynin and Mykyta Chubynsky for discussions. This work was supported by the U.S. Department of Energy, Division of Materials Science, under Award DEFG02-02ER46019.

■ REFERENCES

(1) Perkins, T. T.; Quake, S. R.; Smith, D. E.; Chu, S. *Science* **1994**, *264*, 822.

(2) Dealy, J.; Larson, R. *Structure and Rheology of Molten Polymers*; Hanser Publications; University Heights, IL, 2006.

(3) Rubeinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press Inc.: New York, 2007.

(4) (a) Dorfman, K. D. *Rev. Mod. Phys.* **2010**, *82*, 2903. (b) Viovy, J.-L. *Rev. Mod. Phys.* **2000**, *72*, 813.

(5) For example: (a) Huang, B.; Bates, M.; Zhuang, X. *Annu. Rev. Biochem.* **2009**, *78*, 993. (b) Biteen, J. S.; Thompson, M. A.; Tselentis, M. K.; Bowman, G. R.; Shapiro, L.; Moerner, W. E. *Nat. Methods* **2008**, *5*, 947. (c) Hell, S. W. *Science* **2007**, *316*, 1153.

(6) For example: (a) Bustamante, C. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 415. (b) Volkmuth, W. D.; Austin, R. H. *Nature* **1992**, *358*, 600. (c) Balducci, A.; Hsieh, C.-C.; Doyle, P. S. *Phys. Rev. Lett.* **2007**, *99*, 238102. (d) Maier, B.; Radler, J. O. *Phys. Rev. Lett.* **1999**, *82*, 1911. (e) Bonthuis, D. J.; Meyer, C.; Stein, D.; Dekker, C. *Phys. Rev. Lett.* **2008**, *101*, 108303. (f) Robertson, R. M.; Laib, S.; Smith, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7310.

(7) For example: (a) Levy-Sakin, M.; Ebenstein, Y. *Curr. Opin. Biotechnol.* **2013**, DOI: 10.1016/j.copbio.2013.01.009. (b) Jo, K.; Dhingra, D. M.; Odijk, T.; de Pablo, J. J.; Graham, M. D.; Runnheim, R.; Forrest, D.; Schwartz, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2673.

(8) Wang, B.; Kuo, J.; Bae, S. C.; Granick, S. *Nat. Mater.* **2012**, *11*, 481.

(9) (a) Wong, C. T. A.; Muthukumar, M. *Biophys. J.* **2008**, *95*, 3619. (b) Muthukumar, M. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 435. (c) Long, D.; Viovy, J.-L. *Phys. Rev. E* **1996**, *53*, 803. (d) Smith, S. B.; Heller, C.; Bustamante, C. *Biochemistry* **1991**, *30*, 5264. (e) Olvera de la Cruz, M.; Gersappe, D.; Shaffer, E. O. *Phys. Rev. Lett.* **1990**, *64*, 2324.

(10) (a) Levy, S. L.; Mannion, J. T.; Cheng, J.; Reccius, C. H.; Craighead, H. G. *Nano Lett.* **2008**, *8*, 3839. (b) Han, J.; Turner, S. W.; Craighead, H. G. *Phys. Rev. Lett.* **1999**, *83*, 1688. (c) Akerman, B. *Phys. Rev. E* **1996**, *54*, 6685. (d) Rousseau, J.; Drouin, G.; Slater, G. W. *Phys. Rev. Lett.* **1997**, *79*, 1945. (e) Nykypanchuk, D.; Strey, H. H.; Hoagland, D. A. *Science* **2002**, *297*, 987.

(11) (a) Schwartz, D. C.; Koval, M. *Nature* **1989**, *338*, 520. (b) Akerman, B. *Electrophoresis* **1996**, *17*, 1027. (c) Gurrieri, S.; Rizzarelli, E.; Beach, D.; Bustamante, C. *Biochemistry* **1990**, *29*, 3396.

(12) Shaffer, E. O.; Olvera de la Cruz, M. *Macromolecules* **1989**, *22*, 1351.

(13) Duke, T. A. J. *J. Chem. Phys.* **1990**, *93*, 9049.

(14) Zimm, B. H. *J. Chem. Phys.* **1991**, *94*, 2187.

(15) Laachi, N.; Dorfman, K. D. *J. Chem. Phys.* **2010**, *133*, 234104.