Intermediate states of molecular self-assembly from liquid-cell electron microscopy

Huan Wang[a], Bo Li[a], Ye-Jin Kim[a,b], Oh-Hoon Kwon[b], and Steve Granick[a,b,c,1]

[a]IBS Center for Soft and Living Matter, Institute for Basic Science, 44919 Ulsan, South Korea; [b]Department of Chemistry, Ulsan National Institute of Science and Technology, 44919 Ulsan, South Korea; and [c]Department of Physics, Ulsan National Institute of Science and Technology, 44919 Ulsan, South Korea

Traditional single-molecule methods do not report whole-molecule kinetic conformations, and their adaptive shape changes during the process of self-assembly. Here, using graphene liquid-cell electron microscopy with electrons of low energy at low dose, we show that this approach resolves the time dependence of conformational adaptations of macromolecules for times up to minutes, the resolution determined by motion blurring, with DNA as the test case. Single-stranded DNA molecules are observed in real time as they hybridize near the solid surface to form double-stranded helices; we contrast molecules the same length but differing in base-pair microstructure (random, blocky, and palindromic hairpin) whose key difference is that random sequences possess only one stable final state, but the others offer metastable intermediate structures. Hybridization is observed to couple with enhanced translational mobility and torsion-induced rotation of the molecule. Prevalent transient loops are observed in error-correction processes. Transient melting and other failed encounters are observed in the competitive binding of multiple single-stranded molecules. Among the intermediate states reported here, some were predicted but not observed previously, and the high incidence of looping and enhanced mobility come as surprises. The error-producing mechanisms, failed encounters, and transient intermediate states would not be easily resolved by traditional single-molecule methods. The methods generalize to visualize motions and interactions of other organic macromolecules.

imaging | self-assembly | liquid-cell TEM | DNA

Intermediate states and kinetic mechanisms are not possible to visualize in full conformational detail by single-molecule fluorescence imaging and traditional electron microscopy. Here, we are interested in liquid environments (1). The classical ensemble-averaged methods such as NMR, circular dichroism, ultraviolet-visible, and bulk fluorescence are not suited to dissect transitions between large families of intermediate molecular states, though they can report some intermediate states using clever labeling approaches. Förster resonance energy transfer (FRET) might be considered, but it is most powerful in two-state systems, and the assembly studied here is more complex than this. Bearing in mind that computer simulation would be the traditional alternative to approach this problem from a molecule-by-molecule standpoint, to explore proof of concept, we selected DNA longer than presently addressed by simulations (2–4), yet small enough that it is easy to obtain molecules of precisely designed chemical sequence. We consider DNA sequences 90 base pairs in length.

The revolutionary recent advances in electron microscopy, cryoelectron microscopy, three-dimensional (3D) tomography, and image reconstruction give static images. Their goal is to understand a definite structure (5). Our different goal here is to understand time dependence. Our methods build upon liquid-cell transmission electron microscopy (TEM) previously applied by others to image inorganic and metal nanoparticles (6–12), micellar and vesicle aggregates (13–15), and various methodologies (16–18). Radiation damage and electron-induced water radiolysis are normally of high concern in such experiments, but not so much in the methods implemented here, primarily because use

of deuterated ("heavy") water enabled long experiments free of radiolysis-induced bubbles (19). DNA helices involve no significant isotope effect (20). Using low electron-acceleration voltage and low electron dose, we detected no radiation damage up to 5 min (19), safely longer than processes studied in this paper. While it is true that low-energy electrons can induce DNA radiation damage in the dry state (21), this is not relevant, as our experiments were performed in water. Using a now-standard method to form liquid pockets between atomically thin graphene multilayers (7, 8, 16, 17), modified slightly in our laboratory (19), single-stranded DNA was confined between two atomically thin graphene multilayers in phosphate-buffered saline (PBS) buffer, a standard aqueous solution with salt concentration similar to that in the living cell. Experimental procedures and evidence of minimal radiation damage are elaborated on in *Materials and Methods* and *SI Appendix*. Instrumental optics did not determine the experimental resolution. Resolution was dominated by motion blurring of low-contrast organic molecules during exposure time of the camera.

A conundrum for single-molecule imaging—it has been called "molecular individuality" regarding single-molecule fluorescence imaging (22–24)—is kinetic heterogeneity. Microenvironments of individual molecules can differ significantly; they do not average out as would be the case for larger, colloidal-sized particles (25). This is why—unlike the case of colloidal systems, for which it has become routine to acquire massive statistical datasets and deduce from them statistical probability distributions (26)—for

Significance

Graphene liquid-cell electron microscopy reveals intermediate states of self-assembly—in this example, DNA when single strands form double helices. Molecules are observed for up to minutes at a time without apparent beam damage when electron energy and electron dose are low. Simultaneous in situ single-molecule imaging of conformational adaptations and motion gives more comprehensive understanding of self-assembly successes, failures, and error-producing mechanisms, confirming some earlier predictions and also presenting surprises. Loop intermediates were observed to facilitate error correction. Hybridization events accompany enhanced translational mobility and mechanistically specific persistent rotation. The information obtained goes beyond that from other single-molecule methods.

single-molecule imaging, this is not yet practical. Dozens of molecules can be imaged, but because this experiment requires judgement and an experienced operator to take measurements and analyze the resulting data, it is not yet possible to image many thousands of molecules repeatedly to specify probability distribution functions. Therefore, to organize data in this paper, we group the measurements into categories and illustrate our findings using multiple examples from representative individual time traces. Proof of concept is the main point of this paper.

Results

We compared three samples of single-stranded DNA, each of them 90 base pairs long, designed to have random, pentablock, and hairpin-inducing sequence in the double-helix state. The base-pair sequences, designed as described in *Materials and Methods*, are specified in Fig. 1. Images in liquid were acquired at 300 ms per frame, as faster acquisition lacked the sensitivity to detect single molecules.

Qualitative Patterns. We begin with what appeared to our eyes (Fig. 2*A* and Movie S1) when we first saw it to be the drastic conformational change of single-stranded DNA (no definite shape) to double-stranded DNA (rigid and rod-shaped). A tantalizing point was that it involved concerted rotation, which we quantified by considering histograms of the step sizes of rotational angles. The assembly pair rotated first in one direction during hybridization (53–58 s) and then in the opposite (58–67 s), with speed ranging from 2 to 40° per second (Fig. 2*B*). Various control experiments were performed. Individual molecules in the absence of the complementary strand were imaged as a function of time (Fig. 2*C*); one observed images consistent with the projection, onto the image plane, of a fluctuating random coil whose 2D radius was ~5 nm. This was consistent with the expected radius of gyration for this chain length (27), as the persistence length was ~2 nm (28). The projected area of these images fluctuates around a constant level, consistent with the

surface-hugging molecular conformations anticipated from old arguments in the field of polymer physics that dilute polymer chains form "pancake" surface conformations (29). Testing our interpretation that the rod-shaped object in Fig. 2*A* was a double helix, length and width of fully hybridized molecules were quantified in separate experiments (Fig. 2*D*). Projected onto the image plane, these fluctuated around the expected width (2 nm) and the expected length (30 nm) of a double helix shorter than the known persistence length of 50 nm (30). These fluctuations to lesser projected length than the average were reasonably interpreted to reflect transient tilts out of the image plane. The instances of apparent width larger than 2 nm were reasonably interpreted to reflect motion blurring. With these control experiments in mind, we interpreted the imaged hybridization process (Fig. 2*A* and Movie S1) to display two single-stranded coils of random base-pair sequence that started as separated molecules, diffused to intertwine as a fluctuating blob-shaped object, and finished as a rod-shaped double helix. Note that the transient V shape (46 s), loop with hole in the middle (51 s), and rod were fatter than a well-formed double helix (64 s).

Mean-Square Translational Displacement. After reaching double-helix form, these DNA microstructures displayed Fickian mobility, mean-square displacement proportional to elapsed time (Fig. 2*E*), suggesting not only that diffusion proceeded in a homogeneous environment, but also minimal sample damage, though experiments elapsed during continuous electron irradiation. The trajectory in Fig. 2 *E*, *Inset* concerns the same experiment as in Fig. 2*A*, motion of this molecule in the *x*–*y* plane while hybridization transpired. From inspecting histograms of step-size distribution accumulated from multiple experiments (Fig. 2*F*), the most probable step size in the time frame of 300 ms was the same in every case. Tendency (with limited statistics) toward broader distribution during hybridization is confirmed when below we evaluate the mean step during hybridization events.
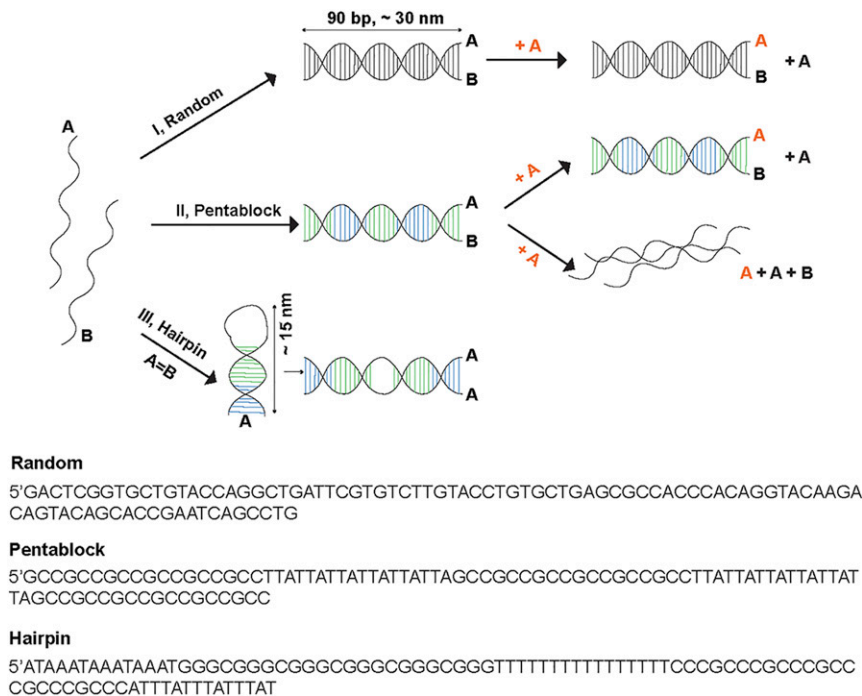

Fig. 1. DNA samples. A family of DNA 90 base pairs in length was studied with oligo sequences either random, blocky (pentablock), or palindromic hairpin. (*Upper*) Schematic depiction of their hybridization. (*Lower*) Explicit base-pair sequences for each of the three cases. Energetic considerations in their selection are described in *Materials and Methods* and *SI Appendix, SI Text*.

Random

5'GACTCGGTGCTGTACCAGGCTGATTCGTGTCTTGTACCTGTGCTGAGCGCCACCCACAGGTACAAGA
CAGTACAGCACCGAATCAGCCTG

Pentablock

5'GCCGCCGCCGCCGCCGCCTTATTATTATTATTATTAGCCGCCGCCGCCGCCGCCTTATTATTATTATTAT
TAGCCGCCGCCGCCGCCGCC

Hairpin

5'ATAAATAAATAAATGGGCGGGCGGGCGGGCGGGCGGGTTTTTTTTTTTTTTTTTTCCCGCCCGCCCGCC
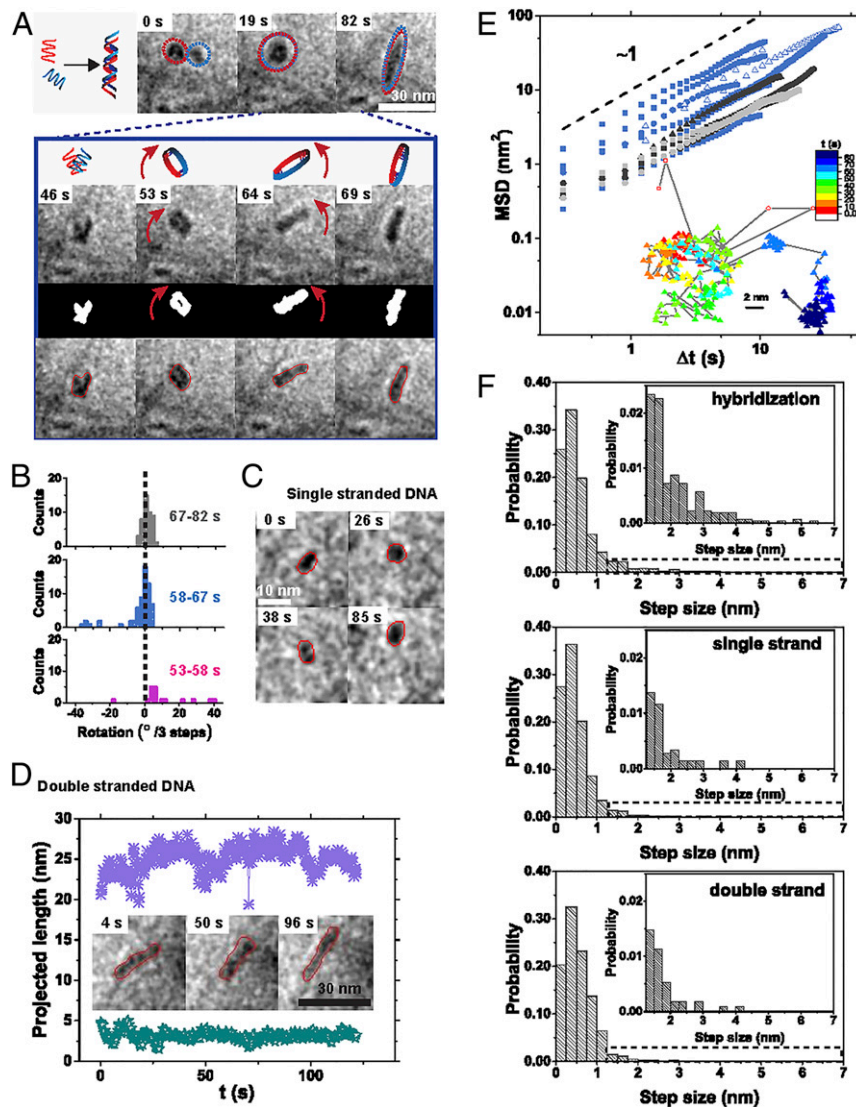CGCCCGCCCATTTATTTATTTAT

Fig. 2. DNA self-assembly and mobility in aqueous buffer. (*A*) Using detection condition 1, complementary single-stranded molecules with random base-pair sequence were imaged while they hybridized (Movie S1). The expanded view of times 46–69 s includes, for better visual contrast, accompanying binarized images obtained from threshold intensity and red circles identifying the imaging-tracking results. Dotted circles in *Upper* identify single molecules. Red arrows highlight rotations in the indicated directions at the indicated times. (*B*) Quantification of rotation steps, evaluated over three time steps (one step = 300 ms) during the indicated time intervals, for the data in *A*. Bin size is 2 angular degrees. (*C*) A single strand of random sequence DNA measured in the absence of the complementary strand; the snapshots at different times up to 85 s illustrate conformational changes and diffusion. Red circles identify the objects tracked. Detection condition 1 was used. (Scale bar, 10 nm.) (*D*) A preformed double helix of random base-pair sequence measured as a function of time, accompanied by quantification of the projected apparent length in the image plane (crosses) and width (stars). The former values, less than the known length of 30 nm, are believed to reflect tilting in the *z* direction. The latter values, larger than the known width of 2 nm, are believed to reflect motion blurring. Detection condition 1 was used. (*E*) On log–log scales, mean square displacement (MSD) of the double-helix center of mass is plotted against time with reference slope unity; each curve comes from a single experiment, and there is no systematic dependence on base-pair microstructure. Blue, hybridization events; light gray, single strand; dark gray, double strands. Circles, squares, and triangles indicate random, hairpin, and pentablock sequences, respectively. Filled symbols show detection condition 1. Open symbols show detection condition 2. *E*, *Inset* shows the projected trajectory in the *x*–*y* plane for the experiment in *A*. Open symbols refer to times when single-stranded DNA was observed; filled symbols refer to times after individual single-stranded DNA started to intertwine; the color code, from red to blue, indicates increasing time elapsed. (*F*) Probability distributions of step size during translational diffusion under detection condition 1. (*F*, *Top*) Diffusion during hybridization (2,653 images) of nine hybridization events from random, pentablock, and hairpin sequences, three of each kind. (*F*, *Middle*) Diffusion of single-stranded DNA (1,455 points from five movies). (*F*, *Bottom*) Fully formed double helix (1,149 points from four movies). Bin size is 0.25 nm (half of 1 pixel = 0.45 nm). The most probable step size is the same in every case. During hybridization, it is broadened.

Hybridization Pathways.

Exploring generality, expected (2–4) hybridization pathways were observed in many cases, but not all. Among these were the zipper-up (3, 4) mechanism initiated by contact of bases located near the middle of strands (Fig. 3*A* and Movie S2) and defect annealing (2, 3) (Fig. 3*B* and Movie S2). The zipper-up mechanism was observed regardless of DNA microstructure—for random (Fig. 3*A*), pentablock (*SI Appendix*, Fig. S1),

and hairpin (*SI Appendix*, Fig. S2) base-pair sequences. For example, single strands bound at the middle to form a branched structure (7 and 10 s), followed by its disappearance (11 s), as revealed by lesser width and increasing length. Quantifying these intermediate steps on a coarser scale, we plotted the time-dependent projected length, width, and angle in the laboratory frame of reference in Fig. 3*C*. Histograms of molecular rotations
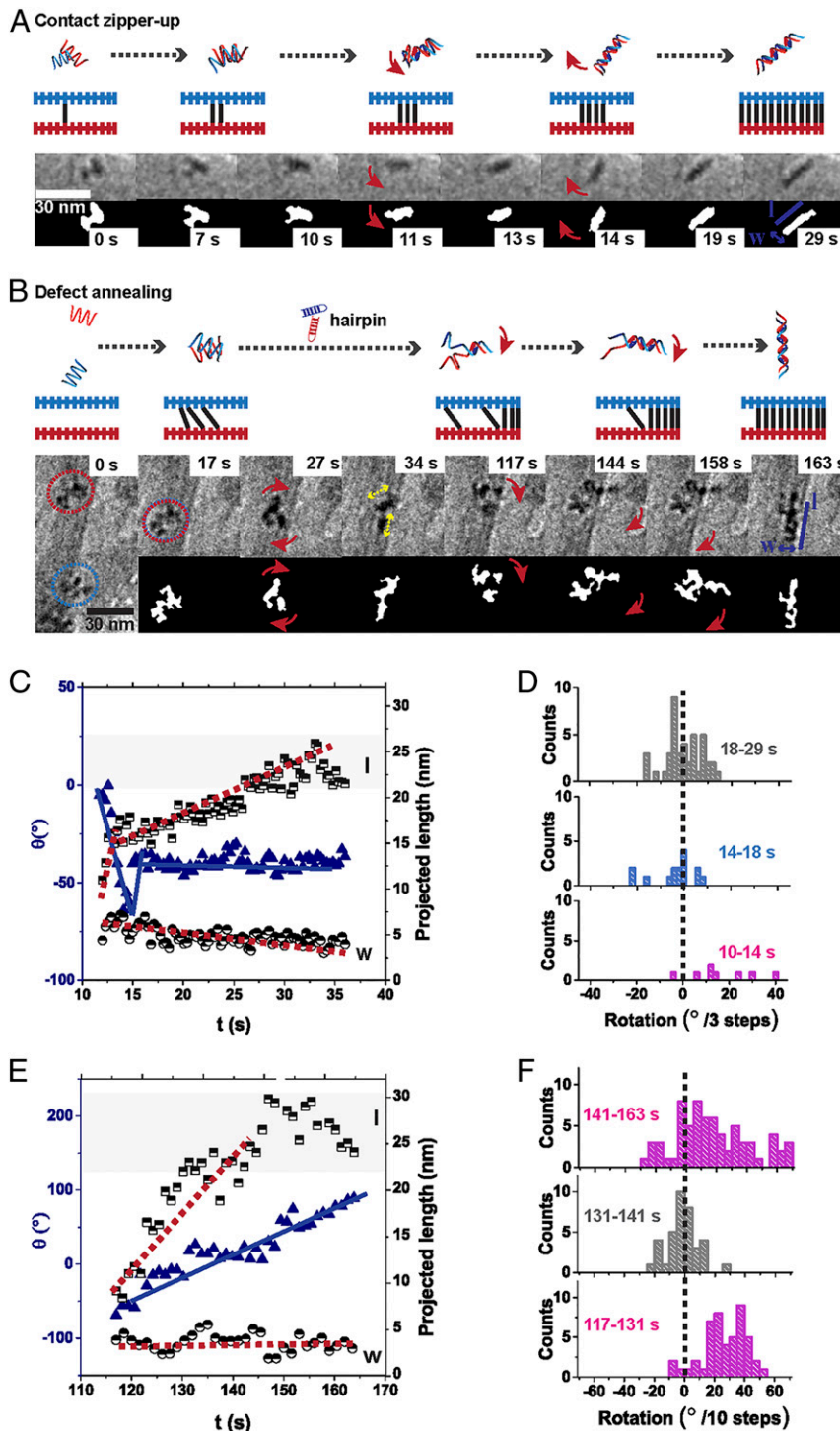
Fig. 3. Confirmation and quantification of intermediate shapes predicted from earlier literature. Detection condition 1 was used (Movie S2). (*A* and *B*) Shown are schematic diagrams of time-dependent conformations (top row), base-pair alignment (3) (second row), actual images (third row), and binarized images for better visual contrast (bottom row). (*A*) Middle-contact zipper-up for random sequence. Neighboring complementary single strands of random sequence meet and hybridize. Arrows in frames at 11 and 14 s indicate rotations in the indicated opposite directions. (Scale bar, 30 nm.) (*B*) Defect annealing for hairpin sequence. Neighboring complementary single strands of hairpin sequence meet and hybridize. The red circles at 0 and 17 s highlight individual molecules. Arrows in frames at 117, 144, and 158 s indicate persistent clockwise rotations as hybridization proceeds. Yellow double arrows at 34 s highlight formation of transient intrahairpin states. Arrows in frame 27 s indicate rotations when a hairpin forms. (Scale bar, 30 nm.) (*C*) Quantification of the zipper-up pathway. Projected length (*l*; defined in image 29 s in *A*) of double-stranded DNA (squares) grows quickly from 12 to 14 s, then more slowly until reaching *l* = 25 nm, accompanied by progressively smaller width, *w* (circles), while the relative angle θ to the horizontal (triangles) first decreases then increases in synchrony with fast growth in the initial stage, but constant after 15 s. (*D*) Quantification of rotation steps of zipper-up pathway. Histograms of net rotations, evaluated over three time steps (1 step = 300 ms), during the designated time regimes, are shown. Bin size, 2 angular degrees. (*E*) Quantification of the defect annealing pathway. Projected length (*l*; defined in image at 163 s in *B*) of double-stranded DNA (squares) grows gradually to *l* = 30 nm after one end of both strands aligns, accompanied by monotonic increase of the relative angle θ to the horizontal (triangles), while the width *w* (circles) remains constant. Lines are guides to the eye. (*F*) Quantification of rotation steps, evaluated over 10 time steps (1 step = 300 ms), of the defect annealing pathway. Bin size is 5 angular degrees.

during this process are shown in Fig. 3D. Note that rotation was persistently in one direction (10–14 s), then switched to the opposite direction (14–18 s), and finished as random fluctuation (18–29 s). But defect annealing of blocky base-pair sequences—the pentablock and hairpin sequences—was slower than for the zipper-up mechanism, and we observed transient hairpins during defect annealing of palindromic hairpin defects (Fig. 3B). For example, first, two strands met (17 s) and adapted individual "hairpin" conformations (27 and 34 s), then the correct ends aligned (117 s), and an apparent rod of constant width protruded clockwise. Quantifying these intermediate steps, the projected length, width, and angle in the laboratory frame of reference are plotted against time in Fig. 3E. Histograms of molecular rotations during this process are shown in Fig. 3F. Note that rotation was persistently in one direction (117–131 s), then stalled (131–141 s), then resumed (141–163 s).

Grooves of the double helix were resolved when the molecule was adventitiously transiently nearly stationary on the time scale

of image acquisition (163 s in Fig. 3B). To infer pitch length would require tomography-type 3D imaging, which is too slow for the dynamical process here (18), or 3D reconstruction from 2D images, recently demonstrated as possible in liquid-cell TEM in one favorable system (31), but as DNA helical structure was not in doubt, this direction of investigation was not pursued.

Also unanticipated was that transient loops appeared to facilitate the hybridization process. This contrasts with the standard view that loops are only errors to anneal out. For example, two pentablock coils (Fig. 4A and Movie S3) met and appeared to assemble from near one end (42 s), then opened near the middle to form a loop (44 and 45 s), accompanied by clockwise rotation. The loop size then shrank, suggesting error correction of base pairing primarily near the ends. This propagated toward the chain middle as the loop disappeared and the duplex became better defined (72 s). Anticlockwise rotation accompanied final formation of the intact helix (80 s). Image analysis (Fig. 4B) showed that the projected length l quickly saturated, but the
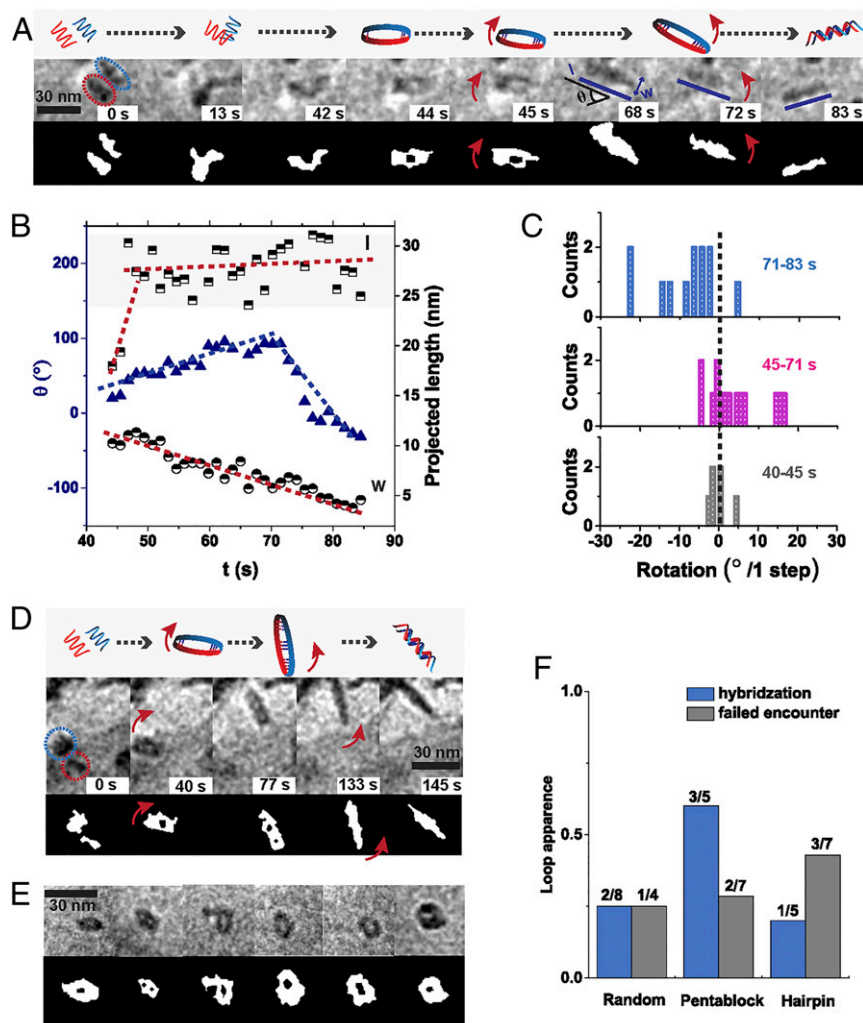
Fig. 4. Loop intermediate states. (A) Single-stranded molecules with pentablock base-pair sequence hybridize (Movie S3). The frame at 68 s identifies the angle, length, and width that are plotted in B. Note loops in frames at 44 and 45 s. Detection condition 2 was used. (Scale bar, 30 nm.) (B) In plots against elapsed time, shapes of the duplex are quantified as concern its projected length, l, and width, w (squares and circles, respectively), and the relative angle θ to the horizontal (triangles). Dashed lines are guides to the eye. (C) Quantification of rotation steps of loop annealing pathway, evaluated over 1 s, during the indicated time intervals. Bin size is 1 angular degree. (D) Images of loops observed during hybridization of a hairpin sequence; they appear at 40 and 77 s. Detection condition 1 was used. (Scale bar, 30 nm.) (E) Six additional examples of loop intermediate states: random, random, hairpin, hairpin, pentablock, and pentablock. (Scale bar, 30 nm.) (F) Histogram, normalized to unity, showing relative loop appearance for the three base-pair sequences investigated, with specification of times observed (first number) and the number of movies inspected for each case (second number).

average width w diminished monotonically throughout, echoing the progressive formation and annealing of the loop. Rotation in the laboratory frame of reference was at first random (40–45 s), then monotonic in one direction (45–71 s), but followed by a reverse (71–83 s); the histogram of rotation steps is shown in Fig. 4C. Rather than a single loop, we also observed processes that involved formation of more than a single loop, as illustrated (77 s) in Fig. 4D for a hairpin sequence. Typical images of these processes in Fig. 4E, together with the relative abundance of loops observed during hybridization and failed encounters, are summarized in Fig. 4F. Taken together, we observed loop intermediates in 12 of 36 total interaction events and 6 of 18 hybridization events.

Detailed comparisons suggest, but with a small dataset, that hybridization mediated by loop annealing was slower than hybridization by the zipper-up process. This was observed consistently. For random-sequence DNA, one sees it illustrated upon comparing Fig. 2A (82 s) to Fig. 3A (29 s). For the pentablock DNA, one sees it illustrated upon comparing Fig. 4A (80 s) to *SI Appendix*, Fig. S1 (47 s). For the hairpin DNA, one sees it upon comparing Fig. 4D (145 s) to *SI Appendix*, Fig. S2 (37 s). Interestingly, the defect-annealing process was slower when we observed it for pentablock DNA (90 s; the ninth graph in Fig. 5F) and the hairpin sequence (163 s in Fig. 3B).

It was common to observe failed encounters, by which we mean that hybridization initiated but did not proceed to completion. Most often observed for microstructure sequences that offer metastable intermediate states, we illustrate this for pentablock DNA sequences (Fig. 5A and Movie S4). Here, two strands formed rod-shaped segments (47 and 124 s), but accompanied by continual transient merger and separation, so that the assembly lacked definite overall shape. Notice that as the molecules remained within separations of 5–10 nm throughout, intermediate states of this kind would not be resolved, even by using superresolution fluorescence microscopy, even if satisfactory labeling of the molecule backbone could be achieved.

Other hybridization involved mixed mechanisms: sometimes multiple independent initiation events and sometimes combinations of the mechanisms identified above. In a process that was only partially zipper-up (Fig. 5B; rotation quantified in Fig. 5C), rotation direction changed three times (21, 70, and 89 s), growth with elapsed time of rod projection was not monotonic (longer at 70 s than at 89 s), and a tiny loop intermediate was resolved (29 s). Quantifying failed encounters, we constructed histograms of successful hybridization relative to average interaction time with failed encounters included (Fig. 5D). Taken together, among the total of 18 hybridizations of 36 interaction events that we imaged, 8 were for random sequences, 5 for pentablocks, and 5 for hairpins.

Transiently Enhanced Mobility during Hybridization. For example, regarding the data shown in Fig. 2A, Fig. 5E shows step-size distribution at different times, and one observes transient shift toward larger step sizes and a broader distribution (30–45 s). The mean step size, plotted against time in Fig. 5F, rose transiently and returned to the initial value when hybridization was complete. This pattern held consistently, as one sees when the mean step size of other examples is plotted against time normalized to the hybridization time, but it was not observed for failed encounters and cases of incomplete hybridization (Fig. 5G). Taken together, we observed that for 17 events of 18 hybridization events, the mean step size was enhanced by a factor of at least 2. This is consistent with the tendency toward longer step sizes in the probability plot in Fig. 2F. It was accompanied by rotation in the laboratory frame of reference, discussed above in connection with Figs. 3 and 4. Additional examples are presented in *SI Appendix*, Figs. S1 and S2.

Three-Strand Coexistence. There is need for error correction when three strands coexist during intermediate stages of self-assembly

(A + A + B); it is interesting to notice that three-strand helices can carry biological function (32). When three such strands meet in competitive binding (Fig. 6A and Movie S5), one sees transient pairing of two strands (4 s), a triplex (6 s), and ejection of one strand (9 s), followed by repetitions of this cycle and subsequent annealing of a duplex accompanied by rotation (24 s) and final state at 39 s. It can also happen that a late-arriving third strand inserts itself into a partially hybridized duplex of two other strands. This is observed sometimes near the center of the original duplex (Fig. 6A, 13 and 16 s). In other experiments, insertion of the third strand was observed close to the geometrical end of the original duplex (Fig. 6B, 71 s; and Movie S5). This caused a loop to form near the center (80 and 84 s), while the invading strand slowly integrated into the duplex, as indicated by shortening of the visible dangling tail, to facilitate annealing with a rotation (92 s) prior to completion (93 s). Occasionally, invasion by a third strand caused the original duplex to melt. In Fig. 6C, taken from Movie S5, an invading strand approaches a rod-shaped duplex (0 s), attaches (11 s), and hops off, accompanied by rotation (17 s) probably caused by initial melting. It reattaches (37 s), then further melts (77 and 102 s). Quantitative image analysis showed that the total intensity was constant within experimental uncertainty, reflecting the same amount of electron scattering, while it originated in different events (Fig. 6D). Physically, this means that rod microstructure associated with the double helix smoothly shortened to zero, offset by growth of size and intensity of the coil-shaped object that we identify with the melted state, the kinetically stable intermediate state in which three strands coexist.

Discussion

It appears that transient adsorption of DNA to graphene walls of the sample cell slowed their motion to the point of being resolvable. Surface diffusion coefficients implied by Fig. 2E are ~1 nm$^2 \cdot$s^{-1}, with the fortunate result that conformational shapes were discerned with minimal motional blurring, despite our instrumental limitation of 300 ms per imaging frame. Noticing that the hybridization times we observed are in the range also observed in solution (33), in a vesicle (34), and for surface hybridization (35), we conclude that motion appears slowed without obvious bias, thanks to the chemically inert nature of graphene, and that rate-limiting steps do not appear unique to graphene.

Regarding the mechanism of slowing down relative to free solution, DNA is likely to interact with graphene mainly at its charged phosphate bases via π–π attraction, with magnitude of 1–2 k_BT per base pair, as can be surmised from studies of synthetic polymers (36). Weak attraction at each segmental unit adds up to substantial attraction for the entire molecule, so to observe surface-hugging configurations is reasonable. Individual base pairs thus retain conformational mobility to explore options to form hydrogen bonds with bases of complementary strands in neighboring molecules. Our 2D projections should faithfully represent surface-hugging molecules dominated by surface diffusion (37), but not exclusively, because low segmental surface sticking allows excursions into 3D solution. This may bias the kinetics toward pathways most relevant to platforms of surface-based DNA detection. When comparing the three microstructures that we designed, it is notable that failed encounters and longer hybridization times were most likely for pentablock and hairpin microstructures. An intriguing subtlety is that the interaction times of hairpins tended to be shortest when the process failed to produce successful hybridization. This may reflect the fact that hairpin strands are also capable of forming intrastrand base pairs.

Tentatively, we attribute in-plane rotations and enhanced translational mobility during hybridization to release of torsion energy (24). Model systems show that when the DNA helix is perturbed, overwound, or underwound, this generates torque (38), and the same is reasonable to expect as strands progressively twist
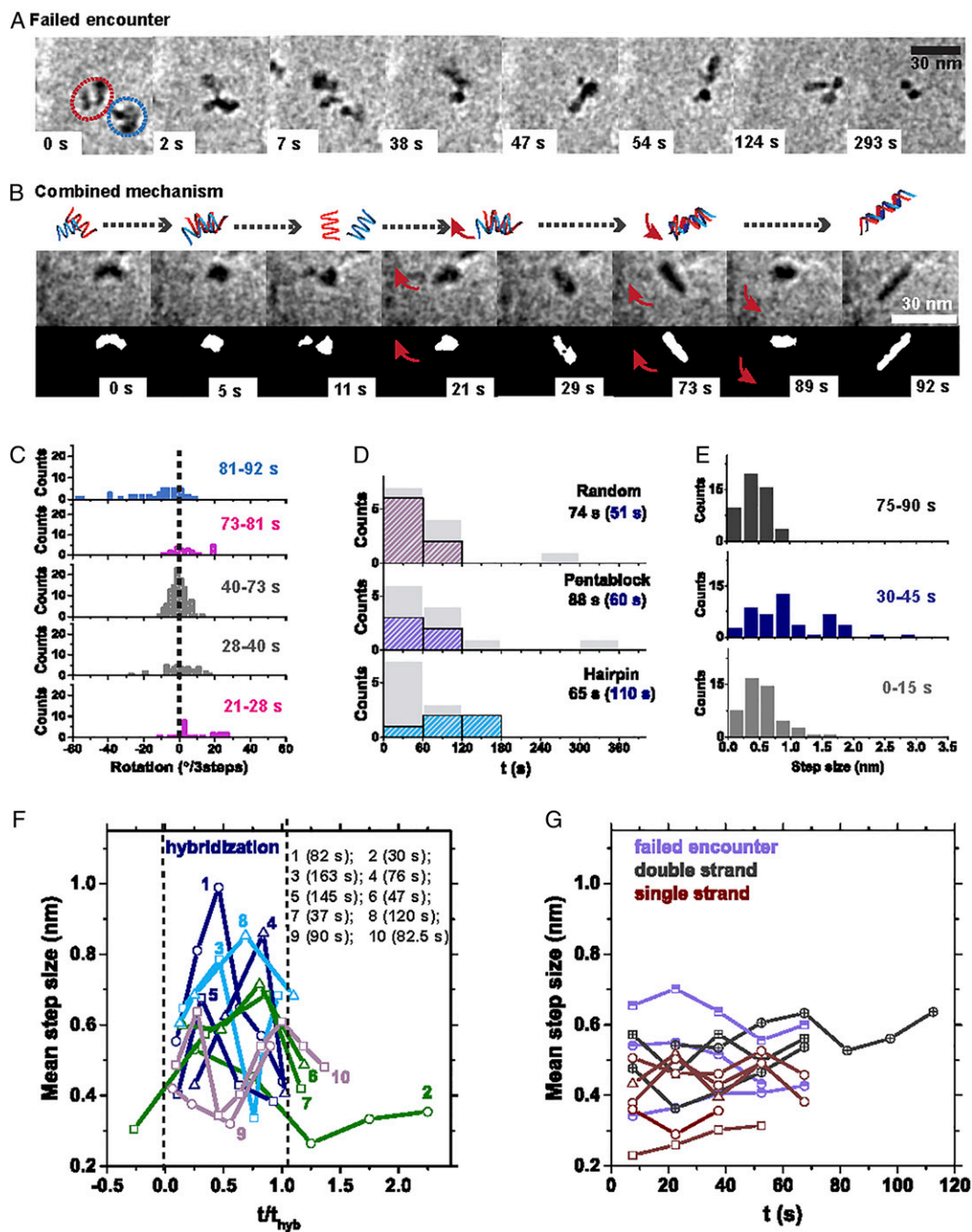
Fig. 5. Hybrid and failed mechanisms illustrated and quantified. (*A*) An example of a failed encounter, the partial assembly followed by disassembly of complementary pentablock base pair sequences (Movie S4). Detection condition 2 was used. (Scale bar, 30 nm.) (*B*) An example in which different hybridization mechanisms combine. Molecules interact via a middle-contact path (0 and 5 s) but fail to zip up (11 s), recombine (29 s) and, ultimately, a well-formed duplex results (92 s). (*B, Top*) Schematic diagrams of time-dependent conformations. (*B, Middle*) Actual images. (*B, Bottom*) Binarized images for better visual contrast. Arrows in frames at 21, 73, and 89 s indicate rotations. Detection condition 1 was used. (Scale bar, 30 nm.) (*C*) Quantification of rotation steps of data in *A*. Histograms of net rotations, evaluated over three time steps (1 step = 300 ms), toward positive, negative, and zero during the indicated time regimes. Bin size is 2 angular degrees. (*D*) Histogram comparing successful and failed encounter times for the three base-pair sequences studied. Successful events are defined from first superposition of two molecules to time at which a rod shape of constant length is observed. Failed events are defined from when two molecules first present themselves as a single feature to the time of their permanent separation. Hatched and gray bars show the mean time of hybridization and failed encounter, also specified numerically in parentheses and without parentheses, respectively. (*E*) Histogram of step distribution, over the time intervals indicated, for the hybridization event analyzed in Fig. 2*A*. Bin size is 0.25 nm. (*F*) Mean step sizes of translational diffusion, accompanied by the projected area of the molecule, are plotted against time normalized by the total hybridization time, t_{hyb}, for the hybridization events. The time interval of each step is 300 ms; the mean step size was averaged over 50 frames. Circles, triangles, and squares refer to random, pentablock, and hairpin sequences, respectively. Color codes are navy for loop annealing, olive for middle-contact zipper-up, blue for defect annealing, and pink for combined. These data summarize all data presented in the text from detection condition 1. (*G*) Mean step sizes of translational diffusion, accompanied by the projected area of the molecule, are plotted against time for no hybridization cases: single strand (brown), double strand (dark gray), and failed encounters (purple). The time interval of each step is 300 ms; the mean step size was averaged over 50 frames. Circles, triangles, and squares refer to random, pentablock, and hairpin sequences, respectively. These data summarize all data presented in the text from detection condition 1.
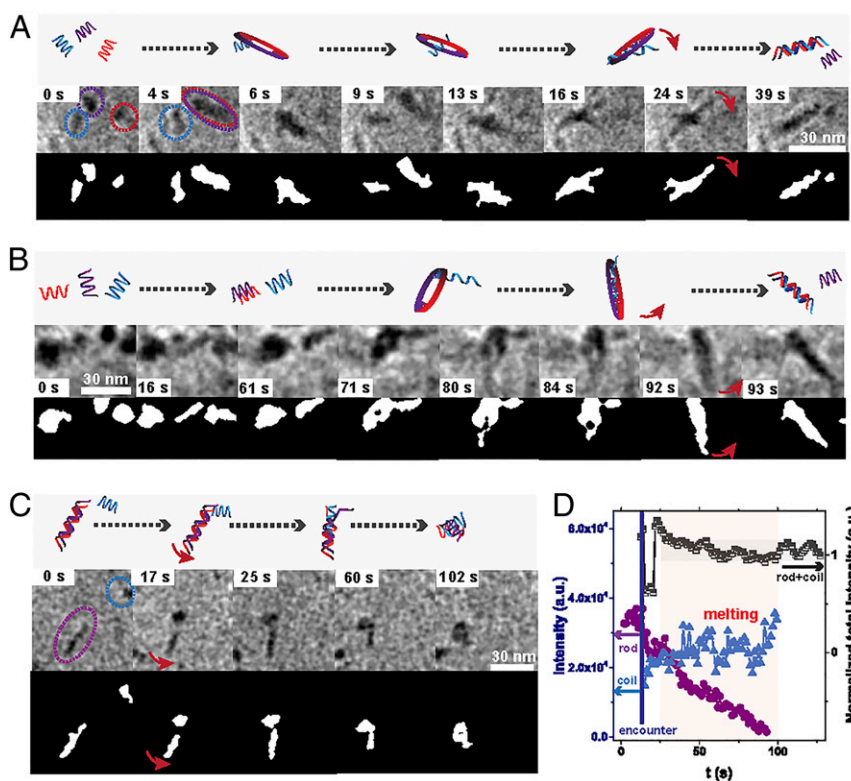
Fig. 6. Transient three-molecule helices (A + A + B). (*A–C*) The images (*Middle*) are accompanied by schematic diagrams of time-dependent conformations (*Top*) and binarized images for better visual contrast (*Bottom*). Dotted circles highlight individual molecules. Red arrows highlight sudden rotation (Movie S5). (*A*) Random base-pair sequence, detection condition 1. An invading blue strand attaches to the loop near the center of the original duplex (red + purple). (Scale bar, 30 nm.) (*B*) Pentablock strands, detection condition 2. An invading blue strand attaches to the duplex at one end, facilitating loop formation (71, 80, and 84 s). (Scale bar, 30 nm.) (*C*) Pentablock base-pair sequence, detection condition 2. An invading blue strand attaches to a rod-shaped pentablock duplex and melts it. (Scale bar, 30 nm.) (*D*) Intensity on the electron detector plotted against time. The original rod-shaped double helix (circles) and the invading molecule (triangles) are quantified as mean pixel intensity times their projected area. Note that the total intensity (rod + coil) is close to constant during melting. The shaded times highlight times when three strands interact. a.u., arbitrary units.

to form complementary helices. Regarding our observation of frequent loop intermediates, which is not predicted by earlier simulation work (2–4), DNA "breathing" (39) and bubble dynamics in double-strand duplex (40) are known, and our observations are in this same spirit. Whether interactions with the graphene surface might enhance loop appearance is not known. No quantitative explanation is offered at this time.

From the observation of competition between three strands to form a single double helix, one concludes that, while a competing third strand can successfully anneal an original pair of strands from errors by attachment to either its ends or near center with a loop, it can also induce melting and, in this way, efficiently create errors. The process contributes extra complexity to the heterogeneity in hybridization from the known varying mechanisms and kinetics. We did not further explore the issue at this time, but on physical grounds, we expect competition of this sort when DNA strands self-assemble in practically relevant technological situations of DNA profiling, origami, and DNA-controlled colloidal/nanoparticle self-assembly (41–43). Three-strand experiments may shed some light on strand-displacement mechanisms, which are important in DNA nanotechnology. There has been work on modeling the kinetics of such reactions (44).

Conclusion

The capability to observe and quantify whole-molecule interactions and assembly in real time can be generalized to other self-assembling organic macromolecules. Earlier, this laboratory demonstrated that single molecules of synthetic polymers can be imaged by using essentially these same techniques (45), but

interpretation of those single-molecule imaging experiments was problematic, as the molar mass of synthetic macromolecules is polydisperse, unlike the DNA studied here. Regarding DNA, we have not yet introduced DNA-associating enzymes. Origami-type assembly is also of interest. By using methods introduced here, the pathway to both systems is clear.

It is appropriate to contrast the information generated by using this technique with that obtainable by using fluorescence-based measurements, such as FRET, which also generates single-molecule traces as a function of time. The FRET technique is normally applied when conformational options are binary and are exceedingly informative in this regard (34), but intermediate states of the kind studied in this paper are multidimensional rather than binary. FRET measures average separations, but not conformations that link those separations, so the present TEM-based method is complementary and can be informative when the system presents more than binary options.

The usefulness of the information is not to identify final equilibrium states, which are known already from other perspectives, but to understand error production and correction along the paths to the end. The methods may have relevance to nucleic acid analytic biotechnologies (DNA-based sensing, sequencing, and mapping); DNA technology for drug delivery (encapsulation, release, targeting use sequence specificity, and origami) and precision medicine, and DNA-controlled colloidal assembly, which likewise suffer from error production.

Current impediments to acquiring larger datasets may be surmounted when methods are developed to automate these measurements; the current limitation to a small number of molecules

is set by operator skill to assemble the graphene liquid cell and operator experience, judgement, and patience to analyze the resulting data. Another limitation is that structural identification is based on projections of 3D structures onto the imaging plane. The 3D reconstruction, recently accomplished regarding a rigid biomolecule of definite shape (31), will be difficult to generalize to the present problems, in which shape changes in time. The current limitation of time resolution by direct single-molecule contrast can be improved by the use of phase plate and more sophisticated image processing concerning dynamic blurring. At the present stage of advancement of the technique, the 3D shape of the molecules imaged is inferred by the person who analyzes movies and retains memories of how images change with time.

Materials and Methods

DNA Oligos. Single-stranded DNA oligos 90 base pairs long were synthesized (Integrated DNA Technologies) with polyacrylamide gel electrophoresis purification and used as received. Sequences are shown in Fig. 1. Several design criteria were applied, as discussed in *SI Appendix, SI Text.*

The oligos were dissolved in $1\times$ PBS/D_2O buffer to make 100 μM stock solution. For random and pentablock DNA oligos, stock solutions of complementary sequences were mixed at 1:1, and this mixture was immediately encapsulated by graphene layers for TEM imaging. As the hairpin sequence can self-dimerize, stock solution was annealed and then diluted with D_2O/PBS buffer at ratio of 1:1 to produce pockets.

Preparation of Graphene Liquid Pockets. Our liquid pockets formed by creases when one graphene sheet was laid as a sandwich on top of another that trapped the liquid, as reported (19). The method generally produces 50- to 100-nm-thick water pockets that are 100–300 nm in width and 200–1,000 nm in length. Specifically, we first transferred commercial chemical vapor deposition-grown two-layer graphene/Cu (ACS Materials) onto a holey carbon gold TEM grid (QUANTIFOIL, Electron Microscopy Sciences and SPI Supplies), followed by depositing by micropipette ~0.5-μL liquid droplets to maximize the chances to form multiple liquid pockets. Quickly and gently, using tweezers, we placed the droplet-bearing, graphene-coated grid onto a graphene sheet that floats on a copper etchant solution, 0.1 M $(NH_4)_2S_2O_8$, and left it for ~10 min for graphene sheets to encapsulate and tightly seal the sample liquid between creases. TEM grids were then picked up and placed on filter papers with sample side facing up for 10 min to form pockets. Reagents (Sigma-Aldrich) were purchased with the highest purity available and used as received.

Electron-Microscopy Conditions. Two TEM instruments were employed, a high-end JEOL-2100 TEM equipped with a direct electron-detection camera with single-electron sensitivity (46) (K2 Summit, Gatan) and a common service instrument, the JEOL-1400 TEM equipped with a US1000 camera in the University Central Research Facility at Ulsan National Institute of Science and Technology. In both instruments, the electron beam was emitted from a lanthanum hexaboride LaB_6 cathode (JEOL). Limited sensitivity to single molecules caused us to use 300 ms per frame for both detectors.

Detection condition 1. Using the JEOL-2100, we imaged at electron acceleration voltage 80 keV and electron dose rate 2–5 e·Å^{-2}·s^{-1} (read from direct software display, counted mode), with temporal resolution 0.3 s and no lag between frames. A 4×4 binning was applied to the camera, resulting in

images of 958×926 pixels with pixel size 0.45×0.45 nm. This detection condition was used for Fig. 2, Fig. 3, Fig. 4 *D* and *E*, Fig. 5*B*, and Fig. 6*A* and *SI Appendix,* Fig. S2.

Detection condition 2. Using the JEOL-1400, we imaged at electron acceleration voltage 120 keV and electron dose rate 10 e·Å^{-2}·s^{-1}, calculated from the current density display (4–5 pA·cm^{-2} at 60,000× magnification). A 4×4 binning was applied to the camera, resulting in images of 512×512 pixels, with pixel size 0.74 nm × 0.74 nm. Images were captured with 300-ms exposure time with 1-s lag between each of them, limited by camera–computer transfer speed. This detection condition was used for Figs. 4*A*, 5*A*, and 6 *B* and *C* and *SI Appendix,* Fig. S1

Considerations of possible electron-beam influences are discussed in *SI Appendix, SI Text.*

Image Analysis. Electron micrographs obtained with detection condition 1 were converted to .tiff from .dm4 by using Digital Micrograph software provided by Gatan Inc., then imported to ImageJ and applied drift correction using "template matching." Electron micrographs obtained with detection condition 2 were directly imported to ImageJ with .dm3 format. To identify features of individual molecules and trajectories, home-written Interactive Data Language codes were used as follows.

To increase signal noise. We began by cropping a small square region, typically 150×150 pixels, that contained solely molecules of interest. Onto this, a real-space band-pass filter was applied to suppress small-scale pixel noise and large-scale brightness fluctuations.

To binarize images. Binarization was accomplished by setting to 255 the values of pixels with brightness larger than a threshold value, usually 90. Other pixels were set to 0.

To obtain center of mass. The boundary of the DNA molecule in images obtained from binarization was featured by applying the chain-code algorithm (47), with an example given in Fig. 2*A*. The center of mass of the molecule was obtained from the mean values of boundary points.

False detection was minimized by setting the empirical criteria that the perimeter of a detected loop must exceed a threshold value, usually 20–40 pixels, and that the distance between mass center of different loops should be physically meaningful, usually larger than 5 pixels and less than 20 pixels. Finally, we connected the discrete boundary points into a closed loop by polygon fitting. The area enclosed by the loop was identified as the projected size of the DNA molecule.

Width and Length of DNA. These quantities (Figs. 3 *C* and *E* and 4*B*) were determined from full-width half-maximum of section analysis by using ImageJ using boundary-featured images. Angle was obtained from the pixel coordinates of the ends of rod-shaped duplex molecules, $\tan\theta = (y_1 - y_2)/(x_1 - x_2)$.

Mean square displacement (e.g., Fig. 2*E*) was calculated for each hybridization event by using oversampling, as is traditional to do for colloidal particles (26). The number of frames was typically 300–1,000 frames for detection condition 1 and 100–300 frames for detection condition 2.

Data Availability. All data discussed in the paper are presented in the figures and supporting information will be made available to readers upon request.

1. J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, London, 2011).
2. M. J. Hoefert, E. J. Sambriski, J. José de Pablo, Molecular pathways in DNA-DNA hybridization of surface-bound oligonucleotides. *Soft Matter* **7**, 560–566 (2011).
3. T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye, A. A. Louis, DNA hybridization kinetics: Zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* **41**, 8886–8895 (2013).
4. T. J. Schmitt, J. B. Rogers, T. A. Knotts, 4th, Exploring the mechanisms of DNA hybridization on a surface. *J. Chem. Phys.* **138**, 035102 (2013).
5. D. Elmlund, H. Elmlund, Cryogenic electron microscopy and single-particle analysis. *Annu. Rev. Biochem.* **84**, 499–517 (2015).
6. T. J. Woehl, J. E. Evans, I. Arslan, W. D. Ristenpart, N. D. Browning, Direct in situ determination of the mechanisms controlling nanoparticle nucleation and growth. *ACS Nano* **6**, 8599–8610 (2012).
7. J. M. Yuk *et al.*, High-resolution EM of colloidal nanocrystal growth using graphene liquid cells. *Science* **336**, 61–64 (2012).
8. Q. Chen *et al.*, 3D motion of DNA-Au nanoconjugates in graphene liquid cell electron microscopy. *Nano Lett.* **13**, 4556–4561 (2013).
9. H.-G. Liao *et al.*, Nanoparticle growth. Facet development during platinum nanocube growth. *Science* **345**, 916–919 (2014).
10. S. Keskin *et al.*, Visualization of multimerization and self-assembly of DNA-functionalized gold nanoparticles using in-liquid transmission electron microscopy. *J. Phys. Chem. Lett.* **6**, 4487–4492 (2015).
11. F. M. Ross, Opportunities and challenges in liquid cell electron microscopy. *Science* **350**, aaa9886 (2015).
12. N. D. Loh *et al.*, Multistep nucleation of nanocrystals in aqueous solution. *Nat. Chem.* **9**, 77–82 (2017).
13. L. R. Parent *et al.*, Directly observing micelle fusion and growth in solution by liquid-cell transmission electron microscopy. *J. Am. Chem. Soc.* **139**, 17140–17151 (2017).
14. A. Ianiro *et al.*, Liquid-liquid phase separation during amphiphilic self-assembly. *Nat. Chem.* **11**, 320–328 (2019).
15. M. T. Proetto *et al.*, Dynamics of soft nanomaterials captured by transmission electron microscopy in liquid water. *J. Am. Chem. Soc.* **136**, 1162–1165 (2014).
16. C. Wang, Q. Qiao, T. Shokuhfar, R. F. Klie, High-resolution electron microscopy and spectroscopy of ferritin in biocompatible graphene liquid cells and graphene sandwiches. *Adv. Mater.* **26**, 3410–3414 (2014).
17. S. Keskin, N. de Jonge, Reduced radiation damage in transmission electron microscopy of proteins in graphene liquid cells. *Nano Lett.* **18**, 7435–7440 (2018).

18. W. J. Dearnaley *et al*., Liquid-cell electron tomography of biological systems. *Nano Lett.* **19**, 6734–6741 (2019).
19. H. Wang, K. H. Nagamanasa, Y.-J. Kim, O.-H. Kwon, S. Granick, Longer-lasting electron-based microscopy of single molecules in aqueous medium. *ACS Nano* **12**, 8572–8578 (2018).
20. H. Lederer, R. P. May, J. K. Kjems, G. Baer, H. Heumann, Solution structure of a short DNA fragment studied by neutron scattering. *Eur. J. Biochem.* **161**, 191–196 (1986).
21. G. Khorsandgolchin, L. Sanche, P. Cloutier, J. R. Wagner, Strand breaks induced by very low energy electrons: Product analysis and mechanistic insight into the reaction with TpT. *J. Am. Chem. Soc.* **141**, 10315–10323 (2019).
22. H. P. Lu, L. Xun, X. S. Xie, Single-molecule enzymatic dynamics. *Science* **282**, 1877–1882 (1998).
23. M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, T. Ha, Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.* **86**, 2530–2537 (2004).
24. P. Kosuri, B. D. Altheimer, M. Dai, P. Yin, X. Zhuang, Rotation tracking of genome-processing enzymes using DNA origami rotors. *Nature* **572**, 136–140 (2019).
25. W. B. Rogers, W. M. Shih, V. N. Manoharan, Using DNA to program the self-assembly of colloidal nanoparticles and microparticles. *Nat. Rev. Mater.* **1**, 16008 (2016).
26. J. Yan *et al*., Reconfiguring active particles by electrostatic imbalance. *Nat. Mater.* **15**, 1095–1099 (2016).
27. A. Y. L. Sim, J. Lipfert, D. Herschlag, S. Doniach, Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **86**, 021901 (2012).
28. B. Tinland, A. Pluen, J. Sturm, G. Weill, Persistence length of single-stranded DNA. *Macromolecules* **30**, 5763–5765 (1997).
29. P. G. d. Gennes, *Scaling Concepts in Polymer Physics/Pierre-Gilles de Gennes* (Cornell University Press, Ithaca, NY, 1979).
30. P. J. Hagerman, Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 265–286 (1988).
31. G. Marchello, C. De Pace, N. Wilkinson, L. Ruiz-Perez, G. Battaglia, 4D liquid-phase electron microscopy of ferritin by Brownian single particle analysis. arXiv:1907.03348 (7 July 2019).
32. F. A. Buske, J. S. Mattick, T. L. Bailey, Potential in vivo roles of nucleic acid triple-helices. *RNA Biol.* **8**, 427–439 (2011).
33. J. X. Zhang *et al*., Predicting DNA hybridization kinetics from sequence. *Nat. Chem.* **10**, 91–98 (2018).
34. I. I. Cisse, H. Kim, T. Ha, A rule of seven in Watson-Crick base-pairing of mismatched sequences. *Nat. Struct. Mol. Biol.* **19**, 623–627 (2012).
35. E. M. Peterson, M. W. Manhart, J. M. Harris, Single-molecule fluorescence imaging of interfacial DNA hybridization kinetics at selective capture surfaces. *Anal. Chem.* **88**, 1345–1354 (2016).
36. P. Frantz, S. Granick, Kinetics of polymer adsorption and desorption. *Phys. Rev. Lett.* **66**, 899–902 (1991).
37. S. A. Sukhishvili *et al*., Materials science. Diffusion of a polymer 'pancake'. *Nature* **406**, 146 (2000).
38. S. Forth, M. Y. Sheinin, J. Inman, M. D. Wang, Torque measurement at the single-molecule level. *Annu. Rev. Biophys.* **42**, 583–604 (2013).
39. P. H. von Hippel, N. P. Johnson, A. H. Marcus, Fifty years of DNA "breathing": Reflections on old and new approaches. *Biopolymers* **99**, 923–954 (2013).
40. G. Altan-Bonnet, A. Libchaber, O. Krichevsky, Bubble dynamics in double-stranded DNA. *Phys. Rev. Lett.* **90**, 138101 (2003).
41. N. C. Seeman, H. F. Sleiman, DNA nanotechnology. *Nat. Rev. Mater.* **3**, 17068 (2017).
42. J. A. Weinstein, A. Regev, F. Zhang, DNA microscopy: Optics-free spatio-genetic imaging by a stand-alone chemical reaction. *Cell* **178**, 229–241.e16 (2019).
43. G. Yamankurt *et al*., Exploration of the nanomedicine-design space with high-throughput screening and machine learning. *Nat. Biomed. Eng.* **3**, 318–327 (2019).
44. D. Y. Zhang, G. Seelig, Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* **3**, 103–113 (2011).
45. K. H. Nagamanasa, H. Wang, S. Granick, Liquid-cell electron microscopy of adsorbed polymers. *Adv. Mater.* **29**, 1703555 (2017).
46. Y. M. Lee, Y. J. Kim, Y.-J. Kim, O.-H. Kwon, Ultrafast electron microscopy integrated with a direct electron detection camera. *Struct. Dyn.* **4**, 044023 (2017).
47. W. Shahab, H. Al-Otum, F. Al-Ghoul, A modified 2D chain code algorithm for object segmentation and contour tracing. *Int. Arab J. Inf. Technol.* **6**, 250–257 (2009).