Single molecule detection of direct, homologous, DNA/DNA pairing

C. Danilowicza, C. H. Leea, K. Kimb, K. Hatcha, V. W. Coljeea, N. Klecknerb,1, and M. Prentissa,1

aDepartment of Physics and bDepartment of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

Contributed by Nancy Kleckner, September 30, 2009 (sent for review August 10, 2009)

Using a parallel single molecule magnetic tweezers assay we demonstrate homologous pairing of two double-stranded (ds) DNA molecules in the absence of proteins, divalent metal ions, crowding agents, or free DNA ends. Pairing is accurate and rapid under physiological conditions of temperature and monovalent salt, even at DNA molecule concentrations orders of magnitude below those found in vivo, and in the presence of a large excess of nonspecific competitor DNA. Crowding agents further increase the reaction rate. Pairing is readily detected between regions of homology of 5 kb or more. Detected pairs are stable against thermal forces and shear forces up to 10 pN. These results strongly suggest that direct recognition of homology between chemically intact B-DNA molecules should be possible in vivo. The robustness of the observed signal raises the possibility that pairing might even be the “default” option, limited to desired situations by specific features. Protein-independent homologous pairing of intact dsDNA has been predicted theoretically, but further studies are needed to determine whether existing theories fit sequence length, temperature, and salt dependencies described here.

dsDNA | sequence-dependent

Pairing of homologous DNA/chromosome regions is a central feature of many biologically important processes. Recombinational double-strand break repair and programmed homologous recombination during meiosis all involve complex series of biochemical reactions in which single-stranded DNA (ssDNA) plays a prominent role. There also exist homologous pairing reactions that seem to involve interactions between chromosomal regions whose DNAs are chemically intact double-stranded DNA (dsDNA) (1–15). In some instances, whole chromosomes pair via multiple interactions all along their lengths (1–9) or via any region present in duplicate copies (10). In other cases, pairing occurs preferentially or exclusively in particular localized regions (“pairing sites”), which tend to involve repeated sequences, specific proteins (for establishment and/or maintenance of pairing) and/or heterochromatic regions (characterized by a paucity of genes and a less “open” chromatin structure) (1, 11–15).

In contrast to recombination-related processes that are known to involve protein-mediated Watson-Crick basepairing interactions between a ssDNA and a ssDNA or dsDNA partner, the fundamental basis for “recombination-independent” pairing remains mysterious. The most obvious possibility is direct DNA/DNA interactions. Theoretical models have proposed that homology recognition arises from non-Watson-Crick hydrogen bond interactions between bases in the major or minor grooves (16). Local melting could also occur, permitting recognition via standard Watson-Crick base pairing. Other theories suggest that homology recognition can occur due to interactions between sequence-dependent charge distributions associated with neighboring DNA helices, where the charge distributions include not only the phosphates in the DNA but also other monovalent and/or divalent ions that are bound to or very near the neighboring DNA molecules (17–22). Interaction-induced correlations between the spatial distribution of charges can result in energy minimization when sequence-matched helices are in close proximity (17–21, 23–26) (SI Text). Despite its a priori attractiveness, acceptance of direct DNA/DNA pairing has been impeded by the lack of conclusive experimental evidence that such a pairing process occurs in biologically-relevant conditions. Encouragement is provided, however, by recent experiments showing evidence for preferential interactions between DNA molecules with like sequences (27, 28).

Two other general scenarios for homology recognition have been envisioned. First, information might come from local sequence information that is read out indirectly by mediating factors, e.g., by site-specific binding proteins, which then “dimerize” in trans, or by interaction of transcription complexes and/or RNAs that then carry out the inter-chromosomal interaction (29). Alternatively, homology might be recognized along the length of a chromosome via the spatial pattern of particular inter-chromosomal snaps, which, in the most extreme case, could be locally identical at every position in the array (30).

The present study investigates direct, homology-directed protein-independent dsDNA/dsDNA pairing. We provide multiple lines of evidence for homology-dependent pairwise interactions between chemically intact DNAs under biologically sensible conditions and begin to explore the dependence of pairing on sequence length, temperature, and monovalent salt concentration to provide a good basis for comparison with proposed theoretical models for pairing.

Results

Assay System. Pairing of homologous DNAs was determined using a parallel single molecule magnetic tweezers approach, similar to a sandwich assay (Fig. 1 and SI Materials and Methods). Two DNAs of interest are differentially labeled at “opposite” termini, one with biotin and the other with digoxigenin (Dig) (green circles and red diamonds, respectively). The Dig-labeled DNA can attach specifically to an anti-Dig labeled capillary, and the biotin labeled DNA can attach specifically to a streptavidin coated magnetic bead. No single DNA molecule can specifically bind a magnetic bead to the capillary surface, but if a DNA molecule labeled with Dig pairs with a DNA molecule labeled with biotin, then a paired molecule can specifically bind to both a bead and the capillary. After imposing a modest force to extend the paired molecules, the number of paired molecules is determined by counting the magnetic beads tethered to the capillary. This system has the potential to detect up to approximately 1,000 tethered beads in parallel in a single assay (31).

Our apparatus is not calibrated at forces below 2 pN; however, previous measurements suggest that 0.4 to 0.7 pN are required to extend unpaired dsDNA to between 12 and 13 µm (32), so the pairing assay measures the number of pairs that withstand at least 0.4 pN.

The distance between the tethered beads and the surface of the capillary provides additional confirmation that the pairing is homologous (Fig. 1). Pairing was examined either between a Dig-
labeled lambda and a biotinylated lambda [λDNAs (λ × λ pairing)] or a Dig-labeled λ and a smaller molecule that includes a subregion of λ and a biotin tail. In the second case, since the shorter molecules have lengths less than 1/9 of lambda, the distance between the bead and the capillary should correspond to the sequence-matched position of the biotinylated end in an unpaired λ molecule stretched by the same force. Standard conditions for pairing analysis were 150 mM NaCl phosphate buffer (PBS) at 25 °C with equal volumes of the two differently labeled molecules, always at equal concentrations, which ranged from approximately 1–3 nM (in molecules) and 30–100 μg/mL.

A Signal Diagnostic of Homologous λ × λ Pairing Emerges with Time of Incubation. Coincubation of differentially-labeled λDNAs under standard conditions generates an appropriate diagnostic signal that initially increases with time and then plateaus (Fig. 2A). At very short (10-min) incubation times, no pairing signal is observed above nonspecific background, confirming that no pairing is occurring during the tethering steps; thus, the signal observed after longer incubation times does not involve artifactual colocalization of DNAs on the beads or other features of the detection system. In most experiments, the number of DNA-DNA pairs formed as a function of incubation time approaches an asymptotic value (e.g., Fig. 2A), but the values at times longer than 5 h are less reproducible than the data for shorter times. Some experiments even show a significant decrease in bound pairs at very long times, possibly due to the onset of interactions between more than two λDNAs.

The λ × λ Pairing Signal Results from Pairwise Association. In a pairwise interaction, the rate of formation of beads bound at the appropriate diagnostic distance should increase with the square of the (equal) concentrations of the two differentially labeled DNAs. We measured the number of bound beads formed at various DNA concentrations after a 1 h pairing incubation because at 1 h the level of pairing is still increasing linearly with time (Fig. 2A), so the number of beads corresponds to the rate of pairing. Fig. 2B shows that this rate increases with the square of the concentration of molecules for DNA(s) concentrations of 30–100 μg/mL (1–3 nM in molecules). A possible slight deviation from this relationship occurs at the highest concentration where, even at 1 h the number of bound beads may have begun to saturate due to almost complete pairing (below). At concentrations below 30 μg/mL (1 nM in molecules) the number of tethered beads is of the order of the nonspecific binding signal. Thus, the observed bead-tethering signal represents a pairwise DNA/DNA interaction.

Homologous λ × λ Pairing Is Very Efficient. To assess the efficiency of homologous pairing we compared the number of tethered beads observed in a standard λ × λ pairing reaction with that observed for a single one with biotin attached at one end and Dig attached at the other, where tethering does not require interaction with another molecule. The two samples contained the same total amounts of DNA [80 μg/mL (2.4 nM)], and were incubated in parallel in 150 mM NaCl at 37 °C for 1 h, where the 1 μL of the sample was added to 49 μL of buffer before the fluid was inserted into the capillary, whereas in most experiments 3 μL of sample were added to 47 μL of buffer. The lower sample volume was used so that most of the bound beads in the doubly-labeled sample were not touching another bound bead, but the number of bound beads in the paired sample was also reduced. Thus, errors associated with counting noise and uncertainties in sample volume were worse in these experiments than the other experiments reported in this work. One experiment gave 820 bound beads for the doubly labeled control DNA sample and 42 bound beads for the pairing DNA sample, where the nonspecific binding is subtracted from the pairing total. Similarly, a 2-μL sample and a 3-μL sample had 79 bound beads and 125 bound beads after the nonspecific binding was subtracted; therefore, a linear scaling of the 2- and 3-μL samples would predict 39.5 and 42 bound beads, respectively. The three different pairing samples were taken on different days with significant differences in the fractional nonspecific binding, but for all three cases the pairing efficiency is approximately 20%, since perfect pairing would correspond to a bound bead number that is 25% that seen for the doubly labeled DNA*.

During the tethering steps; thus, the signal observed after longer incubation times does not involve artifactual colocalization of DNAs on the beads or other features of the detection system.

At 45 °C, the number of beads for a 3-μL sample 93 μg/mL (3 nM) of DNA exceeded 400 after the nonspecific binding was subtracted, corresponding to an efficiency of approximately 50%. We conclude that homologous pairing of two 48-kb λDNAs is an extremely efficient process, with essentially complete pairing readily achievable.

Fig. 2. Pairing of Dig-labeled λ DNA and biotin labeled λ DNA. (A) Number of tethered beads vs. time in 87 μg/mL λ DNA in PBS and incubated at 37 °C. (B) Number of tethered beads vs. square of the DNA concentration.
Five-Kilobase DNAs Pair with the Homologous Regions of Full Length λ DNA. We have also examined the interaction between λ DNA labeled with Dig and a series of biotinylated DNAs carrying 1–5 kb subregions from λ phage corresponding to selected positions along the length of the λ genome. Histograms of the fraction of the detected beads located at particular distances after 1 h incubation were determined for three different λ subregions. In each case, pairing occurs at exactly the appropriate position (Fig. 3, blue, green, and purple histograms). For comparison, the distribution of bead positions is also shown for the standard λ × λ pairing reaction and for a control sample in which biotin and Dig-labels are present at the two ends of each individual DNA (Fig. 3, yellow and gray-outlined histograms, respectively). The overlap of the yellow and gray histograms suggests that pairing does not significantly alter the extension at the applied force.

The above-5-kb DNAs included only the sequences of interest. To exclude the possibility of Watson-Crick pairing at the ends of molecules, we also examined pairing of Dig-labeled full-length λ with a biotinylated molecule in which 5 kb of λ sequence was embedded in flanking nonhomologous sequences of 400-bp and 2-kb at the bead-proximal and bead-distal ends, respectively. Pairing at the appropriate position still occurs with comparable efficiency (Fig. 3, compare magenta outline histogram with underlying blue histogram in top panel). Importantly, no pairing signal is observed between full-length λ and a nonhomologous 5 kb DNA, from pcDNA3.1, where there is no increase in the number of bound beads as a function of the incubation time (Fig. S1A). Further: (i) the number of beads bound (per unit length of capillary) is similar to that seen in control samples; and (ii) the distribution of bead positions peaks at the length of λ (≈12–13 μm), which is also the dominant bead position for nonspecific binding of a control sample containing only λ DNA end-labeled with Dig (Fig. S1B and C).

Pairing Can Be Mediated by Regions of Homology Shorter than 5 kb. We also compared pairing between full-length λ DNA and DNAs sharing 1- to 5-kb regions of homology with or without long flanking nonhomologous tails on either side. In PBS, in the absence of tails, the pairing of a 1-kb λ DNA by full length λ does not result in bead binding that significantly exceeds a nonhomologous control (λ × pcDNA3.1) (Fig. S2). In contrast, molecules with 1 kb of homology flanked by long nonhomologous tails (Fig. S3A and B) exhibit high levels of tethered beads, comparable to those for long-tailed molecules with 5 kb of homology (compare second and fourth rows in Fig. S3C). The same results are seen for analogous substrates with 2-kb of λ homology (compare second and third rows in Fig. S3C). However, for long-tailed molecules, the resulting position distributions increase in width as the length of the homologous region decreases (from 1–5 kb) and some distributions do not even peak at the homologous position. Importantly, these pairing signals are absolutely dependent on the presence of homology: the number of bound beads observed in these reactions significantly exceeded those for a biotin-labeled molecule containing only the nonhomologous tail regions and no λ DNA (Fig. S1) or in standard control samples.

The data above suggests the following: (i) formed pairs are significantly unstable; (ii) in the presence of long nonhomologous tails, loss of homology-dependent contacts leads to “sliding” of the smaller duplex, which preserves the interaction at nonhomologous positions; (iii) in the absence of adjacent nonhomologous tails, loss of homologous contact results in unbinding. These observations imply the existence of a short-range homology-independent attractive interaction, which becomes important only after homologous interactions have brought the nonhomologous portions into sufficient proximity. Such a short-range, homology-independent interaction is consistent with experiments showing that a single dsDNA can form a tightly coiled toroid where toroid formation can be initiated by kinks that bring adjacent parts of the dsDNA molecule sufficiently close to one another (33).

Homologous Pairing Is Increased by Crowding Agents and Is Not Affected by the Presence of Competitor DNA or BSA. We studied pairing under reaction conditions closer to those present in vivo.

Molecular Crowding. Our standard pairing reactions contain approximately 0.003 pmol of each type of DNA in a 2-μL volume, for an overall total molar concentration of homologous molecules of 3 nM. This is approximately the molar concentration of two homologous DNAs in a single yeast cell nucleus (two molecules in nucleus of 1 μm diameter, i.e., a volume of 10⁻¹⁵ L). Since excluded volume effects associated with molecular crowding in vivo might increase the pairing level, we added polyethylene glycol (PEG; average molecular weight 8,000 Da), often used as a crowding agent in DNA experiments. Adding PEG increases pairing during short incubation times (1–5 h) (Fig. 4A), implying an increase in the initial pairing rate that is consistent with indications of homology-dependent dsDNA associations under crowding created by osmotic stress (27, 28).

Nonhomologous Competitor DNA. Pairing between λ and a 5 kb λ fragment (lacking tails) is not reduced by the presence of unrelated DNAs of several types including: (i) a 5 kb fragment of pcDNA3.1 at three times the concentration of the 5 kb fragment (in molecules and base pairs) (Fig. 4B); 40- to 300-kb (average 200-kb) human genomic DNA fragments at one quarter the concentration of the 5 kb fragment in molecules and a 10-fold excess in base pairs (Fig. S4); or (iii) fish sperm DNA, average length approximately 400-bp, at 250 times the concentration of the 5-kb DNA in molecules and 20 times in base pairs (Fig. 4B). Indeed, in the latter case, both the rate and final level of pairing are increased by nonhomologous DNA, a
result we attribute to excluded volume effects. We also note that the total DNA concentration (in base pairs) present in that situation, approximately 2 mg/mL, approaches the range of DNA concentrations found in a eukaryotic nucleus in vivo (10–50 mg/mL). Thus, homologous pairing is not decreased by competitive nonhomologous dsDNA at concentrations substantially exceeding those of the homologous dsDNA, for competitors of varying lengths relative to the 5-kb homologous DNA partner.

**Protein.** No evident effect on pairing is observed by inclusion of 0.1% BSA (approximately 15 μM) (Fig. S4).

Although we cannot fully reproduce the complex in vivo nuclear environment, these results show that homologous pairing is not impeded by nonhomologous sequences; homologous pairing is not suppressed in complex environments; and the rate and efficiency of pairing is increased by the crowded conditions characteristic of the cell nucleus.

**Comparison with Watson-Crick Pairing.** The above observations show that pairing can occur in regions far from the ends of either interacting molecule. Thus, pairing is not attributable to melting/fraying that gives rise to open ends and therefore not attributable to simple Watson-Crick realignment at such ends. Four additional experimental probes further argue against possible involvement of ssDNA in the observed pairing as described below.

First, we studied the effect of T4 gene32 ssDNA binding protein on λ × λ pairing. Given that T4 gene32 can bind to as few as eight bases of ssDNA occlusion by gene 32 protein should inhibit the reaction if ssDNA plays a significant role; however, the protein has no effect on the pairing.

Second, we measured the ability of base pairing interactions to mediate pairing by examining the formation of λ phage DNA dimers created from one λ labeled with biotin and a second λ labeled with Dig at its “opposite” end where the two carried complementary 12-bp ssDNA overhangs at their respective unlabeled ends. After a 2h incubation under standard pairing conditions, [60 μg/mL (~2 nM) of each DNA], the number of dimers was only 20% the number of DNA-DNA pairs measured in parallel in our standard λ × λ reaction under the same conditions.

Third, we heated a λ × λ pairing mixture to 50 °C, and then quickly quenched it, before incubation for pairing. This procedure should create ssDNA bubbles located preferentially in AT-rich regions. These ssDNA bubbles could pair with other open bubbles or interact with duplex DNA; however, no increase in the rate or level of pairing was observed.

Fourth, regions of Watson-Crick duplex as short as 12-bp are resistant to shear forces up to 25 pN (34). In contrast, molecules linked by homologous pairing over regions of 5- or 48.5-kb shear at forces of 10–20 pN. Thus, homologous associations are not maintained by single or multiple Watson-Crick base-pairing interactions of 12-bp or longer, although involvement of even shorter regions of base-pairing is not excluded.

Finally, we note that the long-tailed DNA segments that pair efficiently with full-length λ DNA were generated from in vivo amplified plasmid DNA, and terminally biotinylated subfragments were prepared without the use of ethanol or any chemical (e.g., phenol) known to generate deformations (Fig. S3A). Thus, pairing is unlikely to involve deformed structures along the DNAs.

**Pairing as a Function of Temperature and Monovalent or Divalent Salt. Temperature.** The pairing rate and saturated pairing level increase significantly as the temperature is raised from 5–40 °C, and then decreases strongly at higher temperatures. (Fig. 5A). Preliminary experiments further suggest that the width of the bead location distribution does not increase for temperatures between 40 and 60 °C, even though the number of paired molecules decreases dramatically. Thus, the reduced higher temperature signal still corresponds to regular homologous pairing and higher temperature does not promote nonhomologous pairing. The latter result was confirmed by measurements of the pairing between Lambda and pcDNA3.

**Monovalent and Divalent Salt.** The rate of pairing, defined by pairing levels after a 1h incubation, increases monotonically with the concentration of either NaCl or KCl over a range from 50 mM to 1 M, with slight differences for Na+ versus K+ (Fig. 5B). Below 50 mM (15 mM), pairing is of the order of the nonspecific signal. Divalent salts are well-known to promote aggregation of DNA (35). We therefore also examined homologous pairing in reaction mixtures where MgCl2 is present instead of a monovalent salt. We find significant pairing between λ and a 5-kb subregion of λ in 10 mM MgCl2 during a 10-min incubation time, whereas pairing is almost absent in 10 mM NaCl even after 1-h incubation. Further, no pairing was observed when pcDNA3.1 was paired with full λ in 10 mM MgCl2, further demonstrating that, even in MgCl2, pairing is homology-dependent.

**Implications.** These findings are compatible with occurrence of homologous DNA/DNA pairing in vivo. Pairing occurs robustly throughout the range of temperatures encountered by most living organisms and at salt concentrations corresponding to those generally thought to occur in vivo: 150 mM monovalent salt, with K+ predominant over Na+; and approximately 10 mM total Mg2+, with the concentration of free Mg2+ being considerably lower (36).

**Discussion**

The presented results reveal that in the absence of proteins, two homologous DNA segments can efficiently and rapidly identify one
another, and interact to form complexes stable against thermal motion. These findings confirm and extend the results of two other studies pointing to such a possibility (27, 28). The most important conclusion from the current work is that direct DNA/DNA interactions occur under physiologically sensible conditions; therefore such DNA/DNA interactions may underlie recombination-independent pairing in vivo.

The Mechanism of Homology-Dependent dsDNA/dsDNA Pairing. Physical theories for homologous pairing are discussed briefly below and in more detail in SI Text. Extensive discussions of theories that predict sequence dependent pairing (see 27, 28) are also excellent sources of information on possible mechanisms.

General Aspects. Any process that leads to persistent homology-dependent association of two DNA duplexes will be characterized by certain features.

First, bringing two dsDNA molecules together in solution requires one or more attractive interactions that overcome strong intermolecular repulsion due to the negative charges on the phosphodiester backbones (37). We show that the pairing rate increases with salt concentration, consistent with a reduction in repulsion, although the pairing is not simply a function of the ionic strength of the solution. Of course, changing salt concentration may also alter other interactions; however, two further findings suggest that improved screening is important: (i) an earlier work showed that, at low concentrations, Na+ screens the intramolecule backbone repulsion better than K+ (38) and this work shows that at low concentrations of salt, the pairing rate in Na+ exceeds the pairing rate in K+. (ii) Similarly, the melting temperature and unzipping force also depend on the intramolecular backbone repulsion and previous work has shown that measured values for both quantities are similar for 10 mM Mg2+ and 150 mM NaCl (39, 40), which is consistent with the observation that the pairing rates in 10 mM Mg2+ exceed those in 150 mM NaCl. We note that theories that predict homologous attraction is due to a spatial modulation of the charge distribution do not predict that the pairing will occur in monovalent salts at room temperature.

Second, the attractive interaction that brings pairs together must be dominated by homology-dependent forces because (i) we do not observe pairing between sequences without homology and (ii) pairing of homologous DNA is not suppressed by the presence of nonhomologous competitors. Strong sequence dependence of the long-range attractive interaction avoids (unwanted) nonhomologous interactions that might result from the short range nonsequence dependent attractive interaction.

Third, interactions between homologous regions must be strong enough to allow correctly bound sequences to remain together but weak enough that unmatched sequences with small regions of accidental homology unbind rapidly, thereby avoiding kinetic trapping in nonhomologous interactions.

Specific Mechanisms for Homology Recognition. Homology recognition could occur by a direct mechanism, in which attractive interactions involve the bases themselves (16), or by indirect mechanisms where recognition and attraction occur through interactions involving the helical features of the molecules that vary in correlation with base pair sequence (17–19, 21, 24). Although extensive further theoretical and experimental work is required before firm conclusions can be formed, one observation presented above helps to discriminate among models proposed thus far: we find that the rate and extent of pairing increases progressively with temperature up to 40 °C and then decreases. In general, indirect pairing models depend on the matching of the conformations of two nearby dsDNA molecules, and it has been predicted that conformation-based models [e.g., including sequence-dependent binding of ions (41, 42)] should exhibit exactly the pattern of temperature dependence that we observe. At low temperatures, homologous molecules may get frozen in nonmatching conformations that do not result in a strong attraction. Higher temperatures may allow molecules to readjust conformations in response to their neighbors, permitting the two molecules to minimize their energy by pairing. At still higher temperatures, conformational fluctuations may be too large to allow the molecules to adjust and pair. We note that the observed sequence specificity of the pairing does not decrease with temperature.

It has been suggested that pairing may result from the displacement of specifically bound water into the disordered solvent in an entropy-driven process (20). This mechanism is especially intriguing because sequence dependent hydration effects are already known to play a role in sequence dependent DNA-protein interactions (43) and RNA folding (44). It would be reasonable to assume that hydration effects also play a role in homology-dependent dsDNA/dsDNA interaction and attractive to think that such diverse situations all have a common underlying physical basis.

Implications for Recombination-Independent Pairing In Vivo. Pairing as observed in the present study is fully compatible with occurrence of analogous interactions in vivo. Pairing occurs efficiently between relatively short regions of homology: independent of DNA ends; under physiological conditions of monovalent salt; in the presence of complex nonhomologous competitor DNA and nonspecific protein; over reasonable time scales; at molecule concentrations comparable to that of two homologous segments in a yeast nucleus; with higher rates (or comparable rates at lower DNA concentrations) achievable by inclusion of crowding agents and by optimization of temperature. Further, the DNA concentrations involved in the observed pairing are orders of magnitude lower than those required for collapse of DNA into toroids in NaCl via nonspecific attractive interactions (33).

The robustness of the observed pairing process also raises the strong suspicion that intrinsic homology-based pairing interactions might be the “default option” in vivo. This notion is diametrically opposed to the common view that homologous pairing is a rare and unfavorable condition that must be specifically promoted by appropriate molecular features. Instead, evolution may have specifically ensured that pairing between homologous chromosomes is usually precluded genome-wide, with restrictions lifted specifically in specialized contexts where pairing is useful or, in a few organisms, left in place and accommodated as “somatic” pairing (1).

It has been argued that the most significant problem for a pairing process is not the need to find the correct partner, but rather the need to avoid getting kinetically trapped in stable nonspecific interactions, which in turn requires any general homology searching process to involve weak, transient “kissing” interactions that are then further stabilized (45). Recent theoretical work and modeling on macroscopic systems (46) also supports the usefulness of a pairing process that occurs in stages where, again, the first stage requires weak transient interactions to avoid trapping in incorrect pairs. These initial weak binding of short matching regions should be strong enough to allow neighboring regions to bind if they are matched, but weak enough that the two short matching regions will unbind if the neighboring regions do not match. Under these conditions, correctly aligned homologs will rapidly form strong bonds, but short regions of accidental homology will not trap pairs in false minima. Additional recognition stages can then promote sequence stringency over longer and longer lengths.

Such considerations could help to explain why, in several situations, in two of the most robust cases of “region-specific” pairing, for the XY chromosomes of Drosophila and for “pairing centers” in C. elegans, specificity is conferred by clusters containing multiple nonrepetitive repeats of short sequences (1, 11). For Drosophila, where 50 copies of a <250-bp repeat are normally involved in pairing, it has further been shown that, when pairing site sequences are moved to locations where flanking homology is absent, eight copies are largely sufficient to confer detectable pairing but two
It is also interesting to consider the possibility of direct DNA/DNA pairing interactions as a factor in interactions not only between different (homologous) chromosomes but between sister chromatids. Such pairing might act before, or in concert with, the known factors of protein-mediated cohesion and topological linkages. Sister chromatids will automatically tend to emerge from a replication fork into a confined joint space, and at least transiently, lack a full complement of nucleosomal features that could favor DNA/DNA interactions between sisters.

Conclusion

The current observations show that homologous dsDNAs can specifically recognize one another and pair stably enough for detection on a time scale of minutes/hours. These findings encourage future studies to assess the physical basis for such homology-dependent recognition and the relevance to, and rules for, DNA/DNA-mediated homologous pairing in vivo.

Materials and Methods

For pairing between two full-length λ DNA (NEB, 48,502 bp) two types of samples were prepared. (I) λ DNA was hybridized and ligated to an oligonucleotide complementary to the ssDNA tail at the left end of λ that contained a Dig-label and to another oligonucleotide without label at the right end. (ii) λ DNA was hybridized and ligated to a biotinylated oligonucleotide at the right end and to an oligonucleotide without label at the left end. Pairing between λ and smaller DNA molecules lacking nonhomologous tails was achieved with high fidelity PCR fragments.

In our apparatus, the magnets were held in a lateral position with respect to the microchannel on a 3-axis translation stage to exert a force perpendicular to the glass surface to which the DNA was bound. The assay system is discussed in detail in SI Text.

Pairing Reaction Protocol. Initially equal volumes of each sample are mixed and incubated for a chosen period after which an aliquot is: (i) incubated for 2 min at 37 °C with superparamagnetic (Dynabeads, 2.8-μm diameter) streptavidin-coated magnetic beads; (ii) placed for 10 min in a microchannel containing a round capillary, 0.55-mm diameter, previously coated with anti-Dig antibody. After incubating, a force is applied to the beads by bringing a permanent magnet close to the capillary to apply a constant force of less than 1 pN.

Acknowledgments. We thank A. Kornyshev and S. Leikin, participants of the Abdus Salam International Center for Theoretical Physics Conference: From DNA-Inspired Physics to Physics-Inspired Biology, 2009, Trieste, Italy, and J. Falvo. This work was supported ONR DARPA Grant N00014-01-1-0782 (to M.P.), Materials Research Science and Engineering Center, National Science Foundation Grant MR 97215 (to M.P.), and Army Research Office W911NF-04-1-0770 (to M.P.) and National Institutes of Health Grant ROI-GM-44794 (to N.K.).
To discuss possible mechanisms of DNA/DNA pairing, it is first necessary to consider (i) the attractive interactions that bring two DNA segments together in solution given that these molecules are strongly negatively charged and so would be expected to repel each other; (ii) whether the attractive interaction is present for any two molecules or only two molecules with sequence homology; (iii) whether additional interactions occur once the molecules are brought sufficiently close together; (iv) whether the additional interactions are sequence dependent. Finally, to interpret the results of our assay it is important to understand that we only measure the pairing between molecules that are bound sufficiently tightly to remain stable under 2 pN of shear stress, so the assay may miss paired molecules that were more weakly bound.

In a vacuum, two identical objects with fixed homogeneous charge distributions will always repel each other due to electrostatics; however, if the objects have a spatially varying charge distribution that includes regions of positive charge separated by regions of negative charge, the objects can attract each other when the separation between them is of the order of the characteristic size of the spatial variation in the charge distribution, although at larger separations the objects will still repel each other. The spatial variation can be static or dynamic. The attractive interaction between identical salt crystals is a case where the charge distributions are static. For such crystals, the attractive force is only significant when the spacing between the crystals is of the order of the lattice spacing, and the attractive force depends strongly on the exact alignment of the two crystal lattices to form one single larger crystal. Misaligned or mismatched lattices will not result in an attractive force. The van der Waals attractive force associated with London dispersion is a case where the spatial variation in the charge distribution is dynamic and due to correlated fluctuations in the charge distributions of neighboring objects.

If two identical objects are immersed in a liquid containing counterions the interactions are much richer. Hydration forces (1) and the pressure exerted by crowding agents (2, 3) can provide significant effective attractive forces that do not occur for objects in a vacuum. In addition, even if the objects have a fixed homogeneous charge distribution, the charge distribution of the mobile counterions in solution can create a spatially dependent charge distribution on or around the object (4–6), so that the net force can be either attractive or repulsive (7). The spontaneous organization of point-like, multivalent counterions into an orderly Wigner-crystal-like lattice can provide an attraction (7, 8). A recent review has suggested that such models are not relevant for attractions between DNA molecules in monovalent or divalent ions in water at room temperature (4). Double stranded DNA in solution is a very highly charged helical molecule with a linear charge density of approximately −0.6 e/Å or equivalently one negative charge every 0.17 nm, where the charge is distributed along the two backbones of a helix containing a major and a minor groove with widths of approximately 2 and 1.4 nm, respectively; therefore, dsDNA will have a spatially varying charge distribution even without Wigner crystallization. It is also known that the exact helical structure of dsDNA (9, 10) and the location of counterions in the helix can depend on sequence (11–13). Thus, homologous sequences would have matching charge distributions, but mismatched sequences would require significant DNA deformation to match the charge distributions (4–6); consequently, the sequence dependent structure of dsDNA may provide an attractive electrostatic interaction that will be much stronger when sequences are matched. Fluctuations in the spatial distribution of bound counterions and DNA backbone may result in both sequence-dependent and sequence independent attractive interactions that initially increase with temperature and then subsequently decrease with temperature, but the attraction between homologous ones should be stronger since it does not require deformation to match the charge distribution. The low temperature increase is due to an increase in the variation in the spatial distributions of charges that allows the two structures to readjust in response to each other, and the high temperature decrease is due to excessive deviations between the spatial structures of the two molecules (4, 14, 15).

In what follows, we will review the results of some previous experiments that have observed attractive interactions between dsDNA molecules. Many experiments have shown that despite the large average negative charge on dsDNA, the spatial variations in the charge density that result from interactions with multivalent and polyvalent counterions, can condense a dsDNA molecule so it collapses on itself, often forming an orderly toroid; therefore, there must be a net attractive interaction between different parts of the same highly charged dsDNA molecule (16, 17). In divalent salts, there is an attractive force where some divalent ions such as Mn²⁺ and Cd²⁺ can condense individual dsDNA molecules, whereas others such as Ca²⁺ and Mg²⁺ do not result in condensation; however, divalent ions that do not result in condensation can still result in aggregation (16). Some studies that appear to consider ions with a higher valency may be misleading; for example, spermine has four charged groups with one positive charge each, but it is not a 4⁺ point-like ion. At low concentrations multivalent counterions totally neutralize the surface charge whereas at larger concentrations they invert the net macroion charge (18).

Previous experiments have not observed the aggregation of dsDNA molecules in monovalent salts even at 1 M salt concentrations unless crowding agents, such as PEG, were present to provide an attractive interaction (19, 20); however, although some theory has predicted that an attractive interaction could occur in monovalent salts in the absence of crowding agents if temperature is sufficiently low, the temperatures involved so far were below the freezing point of water so the results are not relevant to in vivo conditions or in vitro experiments done in liquid water (21).

Of course, sequence-dependent pairing requires more than an attractive interaction. It requires that there be a sequence-specific recognition step, which may or may not be incorporated in the initial attractive interaction that is required to overcome the Coulomb repulsion that is present at long distances. The possibility that the sequence recognition is included in the initial long range attractive interaction is supported by calculations (15, 22) that show that the sequence dependence of the structure of the DNA backbone results in the charge distribution forming a sequence-dependent bar code, where Coulombic forces and hydration forces may play a role in a sequence-dependent attractive interaction (4, 6, 15, 22). Counterion binding to DNA is also sequence-specific, potentially enhancing this barcode (13). In contrast, it is possible that an initial weak long range attraction is not sequence-dependent, but that there is a subsequent step in which there is an interaction between the bases that is stable if and only if the bases match. Possible mechanisms include both Watson-Crick pairing and non-Watson-Crick pairing (23–25). Thus, for the model in which the initial attraction is sequence independent but the second interaction is sequence dependent it is assumed that the initial weak interaction will not hold molecules together in the absence of the second sequence dependent interaction.

Simple theories that do not consider water and assume that
pairing depends on the static matching of Wigner crystalline structures, predict that the pairing should either increase or decrease monotonically with temperature, as the ratio of the Coulombic potential to the thermal energy either increases or decreases with temperature (7); however, the interactions with water make predictions of the temperature dependence of the pairing more challenging, so more theoretical work would be required to provide a prediction of the temperature dependence of a Wigner-crystal based interaction in the presence of water. In contrast, theories that suggest that the sequence dependent pairing is a function of correlations between the dynamic charge distributions of the two helices do make clear predictions about the temperature dependence of pairing. These theories predict that the pairing will show an initial increase with temperature followed by a decrease with temperature at higher temperatures (4). Our experimental results suggest that correlations between dynamic charge distributions do enhance pairing.

SI Materials and Methods

Apparatus. In our apparatus, the magnetic field gradient is produced by one stack of five permanent magnets each of 6.4 × 6.4 × 2.5 mm³ dimensions (26). The total magnetic field is approximately that of a solenoid with its long axis in the z direction thus the resulting force on a given bead in the sample is almost exclusively in the z direction, and varies by less than 1% over the region of the liquid sample monitored in the experiment. The magnets were held in a lateral position with respect to the microchannel on a 3-axis translation stage to exert a force perpendicular to the glass surface to which the DNA was bound. The magnitude of the force applied on the beads was determined by the distance between the magnet and the glass surface.

Sample Preparation. For pairing between two full-length λ DNAs (NEB; 48,502 bp) two types of samples were prepared. (i) λ DNA was hybridized and labeled to an oligonucleotide complementary to the ssDNA tail at the left end of λ that contained a Dig-label (3’ Dig TT TCC AGC GGC GGG 5’). (ii) λ DNA was hybridized and labeled to a biotinylated oligonucleotide (5’ AGG TCG CCG CCC TTT Biotin 3’). Pairing between λ and smaller DNA molecules used sample (i) above plus a DNA carrying a terminal biotin label at its left end (3’ Dig TT TCC AGC GGC GGG 5’). Pairing of longer DNA molecules, the background arises from the nonspecific binding of the DNA to the unlabeled end of the λ Dig molecule, with the resulting signals found at a distance corresponding to fully-extended λ DNA (e.g., Fig. S1). Although the shorter DNA may also nonspecifically bind at one end, their fully extended lengths are so short that they were not easily distinguished from beads that adhere directly to the surface of the capillary, which are not counted in the assay. In all of the presented experiments, the number of tethered beads was two times the number in either control sample, typically more than five.

Typical binding numbers for control samples at 80 μg/mL were 10–40 bound beads for the sample with the λ-biotin molecule, whereas the number of bound beads for the λ-Dig molecule was less than five. Thus, experiments using short sequences that were paired with λ-Dig molecules had very low nonspecific binding. If the scaling of the nonspecific binding were linear as a function of DNA concentration, then the nonspecific binding in the pairing experiments would be half the binding in the control experiments since the control experiments use twice the DNA concentration of the actual pairing experiments. Since the number of bound beads in the control experiments with λ-Dig molecules was less than five, the counting noise was of the order of the number of bound beads. At DNA concentrations below 50 μg/mL the counting noise dominates most measurements of the nonspecific binding even for the λ-biotin molecules. At concentrations from 50–100 μg/mL the nonspecific binding for the controls increases with concentration, but even there the counting noise and sample to sample variation is sufficient that the scaling is not unambiguously linear.

It is likely that the nonspecific binding is dominated by the number of defects in the protein coverage on the capillaries since the nonspecific binding for samples taken on the same day with the same capillary preparation is very reproducible but the variation between samples at the same concentration measured using different batches of capillaries is as large or larger than the measured concentration dependence of the nonspecific binding.

Preparation of 5-kb Fragments. Briefly, the fragments were amplified using Pfu Ultra II fusion (Stratagene) in a thermocycler. Typical conditions were as follows: 1 ng λ DNA, 1 X buffer, 1 mM MgCl₂, 0.5 μM dNTPs, and 1 U Pfu Ultra II fusion. The cycling protocol was 5 min at 95 °C, 30 cycles of 30 s at 95 °C, 30 s at 55 °C, 3 min at 72 °C, and 15 min at 72 °C. The oligonucleotides had the following sequences: fragment 5-kb comprising bp 21,613–26,598 and another molecule with a biotin-labeled end are detected (Fig. 1 A and B), and the number of tethered beads separated from the edge of the inner capillary by a DNA-DNA pair is counted.

To maximize reproducibility in our measurements, the mixture of labeled dsDNA samples was separated in several tubes before incubation and every sample assayed for pairing level is taken as an aliquot from a different tube. This procedure avoids variations due to shear forces created by pipetting from the same tube several times. Further, to minimize variation in the efficiencies of association of terminal tags with the bead and/or the antibody-coated slide, comparisons of pairing efficiencies between/among different situations were always carried out in parallel, on the same day, with the same preparations of all non-DNA reagents.

Control Experiments. Control experiments are performed for every pairing experiment. In the controls, each of the two types of labeled DNA is incubated separately with the beads and inside the capillary. In such reactions, no molecule contains both Dig- and biotin labels and thus no bead can be specifically bound to both a bead and the surface; however, a low level nonspecific binding does result in some tethered beads. In experiments comparing two complete λ molecules, the background arises from the nonspecific binding, with the nonspecific binding of the biotin labeled DNA giving a higher background than the nonspecific binding of the Dig-labeled DNA. In experiments that considered the pairing of shorter sequences with λ, the background arises primarily from nonspecific binding of the bead to the unlabeled end of the λ-Dig molecule, with the resulting signals found at a distance corresponding to fully-extended λ DNA (e.g., Fig. S1). Although the shorter DNA may also nonspecifically bind at one end, their fully extended lengths are so short that they were not easily distinguished from beads that adhere directly to the surface of the capillary, which are not counted in the assay. In all of the presented experiments, the number of tethered beads in the pairing signal was two times the number in either control sample, typically more than five.

Typical binding numbers for control samples at 80 μg/mL were 10–40 bound beads for the sample with the λ-biotin molecule, whereas the number of bound beads for the λ-Dig molecule was less than five. Thus, experiments using short sequences that were paired with λ-Dig molecules had very low nonspecific binding. If the scaling of the nonspecific binding were linear as a function of DNA concentration, then the nonspecific binding in the pairing experiments would be half the binding in the control experiments since the control experiments use twice the DNA concentration of the actual pairing experiments. Since the number of bound beads in the control experiments with λ-Dig molecules was less than five, the counting noise was of the order of the number of bound beads. At DNA concentrations below 50 μg/mL the counting noise dominates most measurements of the nonspecific binding even for the λ-biotin molecules. At concentrations from 50–100 μg/mL the nonspecific binding for the controls increases with concentration, but even there the counting noise and sample to sample variation is sufficient that the scaling is not unambiguously linear.

It is likely that the nonspecific binding is dominated by the number of defects in the protein coverage on the capillaries since the nonspecific binding for samples taken on the same day with the same capillary preparation is very reproducible but the variation between samples at the same concentration measured using different batches of capillaries is as large or larger than the measured concentration dependence of the nonspecific binding.

Preparation of 5-kb Fragments. Briefly, the fragments were amplified using Pfu Ultra II fusion (Stratagene) in a thermocycler. Typical conditions were as follows: 1 ng λ DNA, 1 X buffer, 1 mM MgCl₂, 0.5 μM dNTPs, and 1 U Pfu Ultra II fusion. The cycling protocol was 5 min at 95 °C, 30 cycles of 30 s at 95 °C, 30 s at 55 °C, 3 min at 72 °C, and 15 min at 72 °C. The oligonucleotides had the following sequences: fragment 5-kb comprising bp 21,613–26,598
on λ DNA, 5′-XUUUTGCTCATGCCACACAAAGTG-3′ where X is biotin and the U’s are 2′Ome RNA bases and reverse 5′-GAAAGCGTGCTTAACACCTC-3′. Fragment 5-kb comprising bp 1,302–16,322 on λ DNA, 5′-GGGGCGGACCCUCUGTGTTGCTCATGCCACACAAAGTG-3′, where the first 12 bp are identical to a λ end and complementary to a biotinylated oligo (AGG TCG CCC TTC TTT Biotin 5′) and the U is a 2′Ome RNA base, and reverse 5′-GCCATGTGTTGCTGTGATGC-3′. Fragment 5-kb comprising bp 116–5,525 on λ DNA, 5′-XUUUAAACGACAGGTGCTGAAGC-3′ where X is biotin and the Us are 2′Ome RNA bases and reverse 5′-CCCTCTACAGTTGAGGGTAC-3′. Pfu and many other proof reading enzymes do not copy 2′Ome RNA base pairs and thus an overhang is created during PCR where the sequence, as well as the length of the sequence are predetermined by the designed oligonucleotide sequence. Following PCR, the fragments were separated via gel electrophoresis on a 1× TBE 1% agarose gel. The fragments were gel purified using a Nucleospin kit (Machery and Nagel).

Calibration at Low Forces. We note that any change in elasticity due to the pairing of the short sequence would not significantly affect the elasticity of the rest of the dsDNA molecule that is not paired. Our apparatus is not calibrated at forces below 2 pN, but previous work suggests that a 12-μm extension of λ corresponds to approximately 0.4 pN of applied force (28). Thus, at an applied force of about 0.4 pN, we observe that paired dsDNA has an extension that is similar to unpaired dsDNA since the position of beads corresponding to the pairing of two complete λ molecules is approximately the same as the approximately 12-μm length of unpaired λ.

Fig. S1. Pairing of 6 μg/mL biotinylated 5 kb fragments and 60 μg/mL Dig-labeled λ phage molecules. (A) Biotinylated fragments were copied from λ phage subregion at position 21,613 (blue) and pcDNA3.1 vector as template (light blue). (B) Histograms for the measurements after 1 h incubation period from the curves shown in (A). (C) The pcDNA3.1 histogram shown in light blue in (B) is shown on a larger scale so that the position distribution is clear. The distribution is dominated by the length characteristic of a λ molecule and they probably correspond to λ molecules nonspecifically bound to the beads.
Fig. S2. Pairing of Dig-labeled DNA samples 60 μg/mL in PBS buffer with biotinylated DNA molecules of different length: 60 μg/mL full λ DNA (red), 6 μg/mL 5 kb subregion of λ DNA position 21,613 (blue), 1.2 μg/mL 1 kb subregion of λ DNA position 21,613 (light blue), 6 μg/mL approx. 5 kb copied from pcDNA3.1 vector (brown).
Fig. S3. Pairing of Dig-labeled λ DNA incubated with biotin-labeled fragments in which regions of homology are flanked by long nonhomologous tails. (A) Gentle preparation of biotinylated DNA fragments comprising λ DNA plus flanking non-lambda sequences at each end (for details see SI Materials and Methods). In brief: the desired λ DNA segment was inserted into pRS305. The series of λ segments used is shown in (B). For each segment, plasmid DNA was prepared in the absence of phenol extraction, digested with the indicated restriction enzymes and a biotinylated linker ligated onto one end. Between the digestion and ligation steps, and after ligation, DNAs were purified and concentrated through spin columns. DNA species were then separated on an agarose gel and extracted from the gel by electroelution (see SI Materials and Methods). Note that the nonbiotinylated end of the substrate fragment is blunt end (SfoI) to prevent Watson-Crick base pairing with another such molecule during the pairing reaction. Also, the bead is separated from its adjacent dsDNA terminus by a 3' ssDNA link that permits free bead rotation. Lambda DNA segments ranged in size from 1–5 kb; flanking non-λ regions were the same 400- and 4,500-bp segments in all cases. (B) DNA substrates containing λ DNA segments of various lengths from different positions of the λ sequence. Schematic diagrams are shown for constructs of 5-, 2-, and 1-kb λ DNA containing flanking regions from pRS305 plasmid as described in (A). (C) Distribution of extensions for pairing between Dig-labeled λ DNA incubated with biotinylated fragments without flanking tails (first row) and with long flanking nonhomologous tails attached to homolog subregions of decreasing length: 5-kb (second row), 2-kb (third row), and 1-kb (fourth row). Positions of each subregions are detailed on each plot and correspond to nos. 21,613, 11,302, and 116; blue, green, and red/orange, respectively. The lavender outlines show the histogram for a sequence where the nonhomologous tails are directly joined without any homolog subregion. The number of bound beads in the controls was more than an order of magnitude lower than the number bound for the homologs, but the control distribution does not peak at the length characteristic of λ. Experiments were performed under standard pairing reaction conditions analogous to those in Fig. 4 (text). Distances were measured for about 100 tethered beads in each experiment.

Danilowicz et al. www.pnas.org/cgi/content/short/0911214106
Fig. S4. Pairing of 6 μg/mL biotinylated 5-kb fragments and 60 μg/mL Dig-labeled λ phage molecules in the absence (blue) and presence of nonspecific competitors: 50 μg/mL human genomic dsDNA (red), and 0.1% m/v BSA (BSA) (green).