THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 39, pp. 29002–29012, September 28, 2007
Printed in the U.S.A.

Analysis of P Element Transposase Protein-DNA Interactions during the Early Stages of Transposition*[S]

Received for publication, May 17, 2007, and in revised form, June 29, 2007 Published, JBC Papers in Press, July 19, 2007, DOI 10.1074/jbc.M704106200

Mei Tang[‡], Ciro Cecconi[§], Carlos Bustamante[‡¶‖], and Donald C. Rio[‡1]**

*From the [‡]Department of Molecular and Cell Biology, Division of Genetics, Genomics and Development and Division of Biochemistry and Molecular Biology, Center for Integrative Genomics, University of California, Berkeley, California 94720, [§]CNR-INFM-S3 University of Modena e Reggio Emilia, 41100 Modena, Italy, the [¶]Department of Physics, University of California, Berkeley, California 94720, and the [‖]Howard Hughes Medical Institute, **Lawrence Berkeley National Laboratory, Berkeley, California 94720*

P elements are a family of transposable elements found in *Drosophila* that move by using a cut-and-paste mechanism and that encode a transposase protein that uses GTP as a cofactor for transposition. Here we used atomic force microscopy to visualize the initial interaction of transposase protein with P element DNA. The transposase first binds to one of the two P element ends, in the presence or absence of GTP, prior to synapsis. In the absence of GTP, these complexes remain stable but do not proceed to synapsis. In the presence of GTP or nonhydrolyzable GTP analogs, synapsis happens rapidly, whereas DNA cleavage is slow. Both atomic force microscopy and standard biochemical methods have been used to show that the P element transposase exists as a pre-formed tetramer that initially binds to either one of the two P element ends in the absence of GTP prior to synapsis. This initial single end binding may explain some of the aberrant P element-induced rearrangements observed *in vivo*, such as hybrid end insertion. The allosteric effect of GTP in promoting synapsis by P element transposase may be to orient a second site-specific DNA binding domain in the tetramer allowing recognition of a second high affinity transposase-binding site at the other transposon end.

Mobile genetic elements are ubiquitous among both prokaryotic and eukaryotic organisms (1). Genome sequencing projects have shown that transposable elements make up a substantial fraction of eukaryotic genomes, including 49% of the human genome (2, 3). These mobile elements can lead to mutations and genome rearrangements and appear to play a role in genome evolution (4, 5). The mechanisms by which transposons are mobilized can be grouped based upon whether there is a DNA or an RNA intermediate (1, 4, 5). P elements use a cut-and-paste mechanism with a DNA intermediate, related to those used by the Tn5, Tn10, and Tn7 prokaryotic transposons and the eukaryotic Tc1/mariner family (6–8). Studies of P element transposition *in vitro* have demonstrated cofactor requirements for both GTP and magnesium ions (Mg^{2+}) (6, 9–11). Although divalent metal ions are universally required cofactors for transposase proteins (1, 12, 13), the P element reaction is unique among this family of polynucleotidyltransferases in the use of GTP as a cofactor (9). Recent studies have shown that GTP promotes pre-cleavage synaptic complex assembly between the P element termini and the transposase protein (14). Previous *in vitro* studies had also shown that nonhydrolyzable GTP analogs support both the P element DNA cleavage and joining reactions (9). Furthermore, GTP does not effect the binding of P element transposase to the high affinity sites near the transposon termini, as assayed by DNase I footprinting (9, 15). Whereas some transposon systems have been studied with respect to the assembly state of their transposase, the oligomeric state of the active P element transposase protein and how it initially interacts with P element DNA are unknown. In cases where it has been studied, other transposases are able to function as dimers, tetramers, or hexamers to carry out the catalytic steps of transposition (1, 12, 16). For Mu transposase, DNA binding actually promotes tetramer formation (17–19). In V(D)J recombination, a dimeric complex of RAG2 and a dimer or trimer of RAG1 protein recognizes the recombination signal sequence (20, 21); for the Hermes transposase a hexamer is the active species (16); for the mariner family transposase, Mos1 is a dimer or tetramer (22–25); and for the mariner family member, Himar 1, is a tetramer (26). For Tn5 and Tn10, the protein acts as a dimer, with a single active site acting in *trans* at each transposon end (27–30). Thus, it has been of interest to understand the oligomeric state of transposase during P element transposition.

Here we use atomic force microscopy (AFM)[2] to visualize and quantitate the initial transposase protein-DNA complexes formed on P element DNA ends in the absence of GTP, prior to synapsis of the P element termini. Biochemical methods show that P element transposase exists as a tetramer in solution, in the absence of DNA. We have also used AFM to determine the volume of the transposase protein, in the absence and presence of DNA, and we used this information to determine the native

* This work was supported by National Institutes of Health Grant R01GM61987. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This paper is dedicated to our friend and colleague, Nick Cozzarelli.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1 and 2.

[1] To whom correspondence should be addressed: E-mail: don_rio@berkeley.edu.

[2] The abbreviations used are: AFM, atomic force microscopy; DTT, dithiothreitol; GTPγS, guanosine 5′-3-*O*-(thio)triphosphate; GMP-PCP, guanylyl β,γ-methylenediphosphonate; GMP-PNP, guanosine-5′-[(β,γ)-imido]triphosphate; LM-PCR, ligand-mediated PCR.

[3] Supported by NIH grant R01GM61987.

molecular weight of the protein by comparison with the relative volumes of a set of standard proteins; these measurements also indicate that the P element transposase exists as a tetramer in both conditions. This method of AFM protein volume determination has been used previously to analyze the size of protein-DNA complexes involved in transcription and DNA repair (31, 32). Our AFM data indicate that a pre-formed transposase tetramer initially interacts with either P element terminus, in the absence of GTP, prior to GTP-dependent synaptic complex formation. These studies also show that synapsis is rapid with GTP or nonhydrolyzable GTP analogs but that donor DNA cleavage occurs more slowly. These studies show initial binding of a pre-formed tetramer of P element transposase to one P element end and suggest that the GTP cofactor somehow orients a second DNA binding domain in the oligomer to interact with the high affinity transposase-binding site near the other P element end leading to synaptic complex formation. This two-stage binding/synapsis might explain how inappropriate pairing *in vivo* of two P element ends from different P elements might give rise to ectopic chromosomal rearrangements sometimes observed, such as hybrid ends insertions and deletions (33–35).

EXPERIMENTAL PROCEDURES

Protein Standards for Protein Volume Measurements—RNA polymerase holoenzyme was kindly provided by Dr. Caroline Kane (University of California, Berkeley). Alcohol dehydrogenase, bovine serum albumin, and thyroglobulin were purchased from Sigma.

Protein Purification and Excision Activity Assay—P element transposase tagged at the C terminus was purified as described previously (10, 14), and activity assays by LM-PCR to detect donor DNA cleavage were performed as described (11).

DNA Preparation—The DNA substrate, containing a 0.6-kbp P

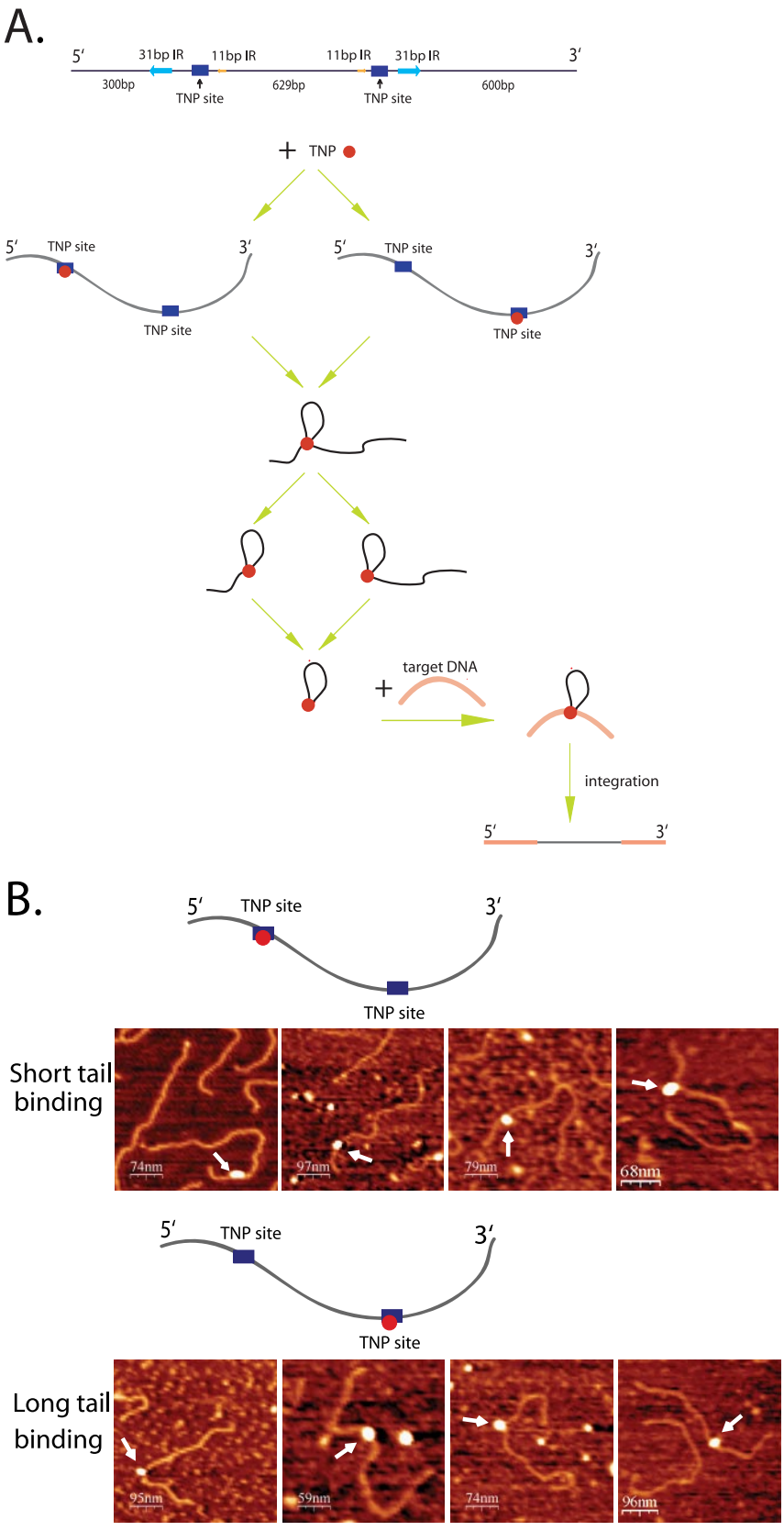

FIGURE 1. **P element transposase protein binds as a pre-formed tetramer to one of the two P element ends prior to synapsis.** *A*, diagram of P element excisim and integration; *B*, upon brief incubation of the protein with DNA (1–30 min), transposase protein binds to either one of the two transposon ends of a 629-bp P element flanked by 300 and 600-bp DNA tails (300 – 600-600). Diagrams at the *top* of each panel indicate the position at which the transposase binds along the DNA molecule. The transposase proteins bound to DNA are indicated by *white arrows*. These images were from reactions performed in the absence of GTP.

element flanked by 0.3- and 0.6-kb adjacent non-P element DNA, was prepared as described (14). The DNA substrate, containing 2.9-kb full-length P element flanked by 0.2- and 0.3-kb adjacent non-P element DNA, was prepared the same way using Pπ25.1 as a template (36), and high fidelity *Taq* polymerase was used in the PCR step.

Transposase Assembly and Excision Reactions—Assembly and excision assays were carried out as described (9, 11, 14, 37). P element donor cleavage reactions were carried out by mixing 1 μl of purified 3′ polyoma epitope-tagged transposase (~50 μg/ml) in HGKED buffer (25 mM Hepes-KOH, pH 7.6, 20% glycerol, 0.5 mM DTT, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 0.4 M KCl) with 1 ml of 50 nM P element donor DNA and 4 ml of HGED buffer (25 mM Hepes-KOH, pH 7.6, 20% glycerol, 0.5 mM DTT, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride). After incubation on ice for 10 min, the mixture was added to 14 μl of 10 mM $Mg(OAc)_2$ in HGED buffer, with or without 2 mM GTP, and incubated at 30 °C for the indicated times. The final protein concentration of protein and DNA in the 20-μl assay was 28.5 and 2.5 nM respectively. These reaction conditions are similar to those used previously for donor cleavage and strand transfer in bulk solution reactions (11, 37). Thus, in the DNA binding reaction the DNA concentration was 2.5 nM, and the protein concentration (as a monomer) was typically ~28.5 nM giving a ratio of 11 monomeric proteins per DNA molecule. In the GTP analog experiments, GTPγS, GMP-PCP, and GMP-PNP were used at a final concentration of 2 mM.

Gel Filtration Chromatography—Approximately 3.5 μg of immunoaffinity-purified transposase protein was resolved on a Superdex 200 PC3.2/30 column (GE Healthcare) on a SMART system at 4 °C in HGKED buffer (25 mM Hepes-KOH, pH 7.6, 1 mM EDTA, 300 mM KCl, 0.5 mM DTT, and 10% glycerol) at a flow rate of 40 μl/min. 40-μl fractions were collected and assayed by SDS-PAGE and immunoblotting (10, 37). Standard proteins used to calibrate the column were bovine serum albumin (66 kDa), catalase (232 kDa), and thyroglobulin (669 kDa) and were run in the same buffer as transposase.

Glycerol Gradient Velocity Sedimentation—Glycerol gradient centrifugation (38) was carried out in a 4-ml gradient with an SW60 rotor at 42,000 rpm for 16 h at 4 °C, in an L90M ultracentrifuge (Beckman). Approximately 3 μg of transposase was layered on top of a freshly poured gradient in HGEKD buffer (25 mM Hepes-KOH, pH 7.6, 1 mM EDTA, 300 mM KCl, 0.5 mM DTT, and various % glycerol as indicated in Fig. 3, *B* and *C*). 150-μl fractions were collected from the top and assayed for transposase by SDS-PAGE and immunoblotting (10, 37). Standard proteins used to calibrate the column were bovine serum albumin (66 kDa), catalase (232 kDa), and thyroglobulin (669 kDa) and were run in a parallel gradient in the same buffer as transposase. For the experiment with GTP or GTP and strand transfer oligonucleotide (P1/P2–17; 37), 0.5 mM GTP and 1 mM $Mg(OAc)_2$ were preincubated with ~3 μg of transposase at room temperature for 15 min in gradient buffer or 0.5 mM GTP and 1 mM $Mg(OAc)_2$ along with a 10-fold molar excess of P element 3′ end strand transfer oligo. Following the preincubation, the samples were loaded on top of 4-ml glycerol gradients and centrifuged as above.

Atomic Force Microscopy Imaging—The samples were imaged in air with a NanoScope III (Digital Instruments) atomic force microscope, operating in tapping mode, using tips from nanosensors (point probes, type NCH-50). The protein was equilibrated at 25 °C for 10 min in 25 mM Hepes-KOH, pH 7.6, 0.04 M KCl, 20% glycerol, 1 mM EDTA, and 0.5 mM DTT and then deposited onto freshly cleaved mica for 1 min. The samples were then washed with water for 15 s and then dried with a gentle stream of nitrogen. All images were collected at scan rates between 1.54 and 3.05 Hz with scan sizes of 1–3 μm. To investigate the effect of the water washing on the apparent transposase volume, thus on its oligomeric state, we tried two different types of water: (i) Milli-Q water (Millipore) with a pH of 4.5, and (ii) NanoPure water (Barnstead) with a pH of 6.2. In more recent experiments, once we realized the wash water pH was a problem, we used commercial water from Mallinkrodt, at pH 6.8–7.2. We also tried two different washing times as follows: 10 s (which we refer to as "quick wash") and 2 min (which we refer to as "soak wash"). For protein volume comparison experiments, such as those described in Fig. 4, extra care was taken to avoid imaging differences because of different AFM tips or microscope settings. To compare the apparent sizes of transposase oligomers between gel filtration and glycerol gradient sedimentation fractions, we used the protein peak fractions from both methods and carried out a single end binding experiment where the protein was incubated with DNA substrate DNA in the absence of GTP for 30 min, and the reaction mixture then was deposited on mica, processed, and imaged as indicated above and in Ref. 14.

Volume Measurements—Protein volumes were calculated by treating apparent proteins as solids with elliptical footprints with major and minor axes equal to protein length and protein width. The vertical cross-sections for both axes were approximated as parabolas with maximum of protein height at the center of each axis. Integration of the area of the ellipse along the height axis gave a volume equal to $p/6 \times$ (length $\times$ width $\times$ height).

RESULTS

Initial P Element Transposase-DNA Complexes Form at One Transposon End Prior to Synapsis—The P element transposition reaction occurs in stages, with initial transposase binding, synapsis, nonconcerted DNA cleavage, and the target capture and integration (Fig. 1*A*). Our previous studies have shown that P element transposase binding to its internal site-specific recognition sites on P element DNA is not affected by the presence or absence of GTP, but GTP can promote synapsis of P element ends (14). Therefore, we were interested in determining how the P element transposase interacted with the P element termini prior to synaptic complex formation. To investigate the early protein-DNA complexes that formed from 1 to 30 min, transposase protein was incubated with a linear DNA carrying a 0.6-kb P element flanked by flanking sequences of 300 bp at the 5′ end and 600 bp at the 3′ end (Fig. 1*B*, *top*). Upon brief incubation of the protein and DNA (1–30 min) at a molar ratio of ~11 transposase monomers to each P element DNA, the linear DNA was bound by transposase at either one end or the other (Fig. 1*B*) as evidenced by a protein-DNA complex formed with

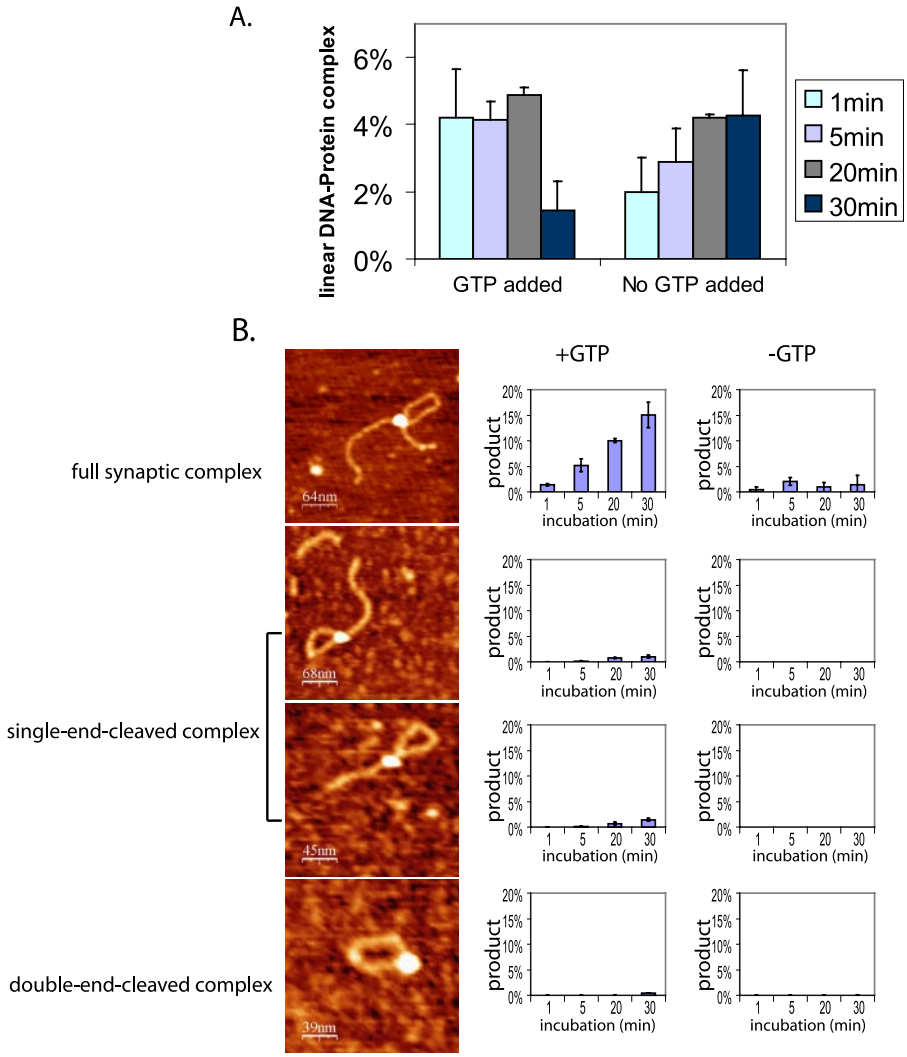

FIGURE 2. **Time course of formation of P element transposase protein-DNA complexes during the earliest stages of the reaction.** *A,* time course of the formation of the linear DNA-protein complexes shown in Fig. 1, in the presence or absence of GTP. The percentages were calculated from the ratios between linear DNA molecules bound by transposase protein to one end of the P element, and the total number of DNA molecules. After different incubation times, 1, 5, 20, or 30 min, 20 ml of the reaction was deposited on mica, rinsed with water, and imaged by AFM. These experiments were performed in the presence (*left bars*) or absence (*right bars*) of 2 mM GTP as described (14) (see "Experimental Procedures"). All single site binding events are graphed as a proportion of the total DNAs observed. *B,* time course of the formation of pre-cleavage (*top panel*) and post-cleavage (*bottom three panels*) synaptic complexes formed in the presence (+*GTP*; center panels) or absence (−*GTP*; *right panels*) of 2 mM GTP during the first 30 min of the reaction. The *error bars* represent variation between at least three independent experiments.

either a short (0.3 kb) or long (0.6 kb) tail. There appeared to be equal numbers of transposase bound to either the 5′ or 3′ transposon end, suggesting that the two ends have similar affinities for the protein, as had been suggested by earlier DNase I footprinting assays (15). In the experiments presented here, and those published previously (14), nonspecific DNA binding is seen to vary from experiment to experiment. However, in some cases transposase molecules were observed to be bound to DNA ends or at random to incorrect sites on the P element-containing DNA fragment (data not shown and see Ref. 14). These nonspecific binding events were ~10–100-fold less abundant than site-specific DNA binding and this number is consistent with previous measurements of site-specific *versus* nonspecific DNA binding by the P element transposase made by DNase I footprinting (15). More importantly, even though the ratio of transposase monomers to P element DNA is ~11:1, we did not observe two transposase complexes bound at the two transposon ends on the same DNA molecule, in over 100,000 DNA molecules examined that had transposase bound. Even when the transposase concentration was raised another 10-fold, no increase in either nonspecific binding or two transposase complexes on a single P element were observed (data not shown). Given the low occurrence of two ends being bound, it seems reasonable to propose that it is the single end-bound transposase complexes that proceed to synapsis. However, it may be the case that these stable one-end-bound complexes are stable, but more loosely bound protein-DNA complexes are lost during the sample preparation for AFM. It seems to be the case that once a single end of the transposon is bound by transposase, the second end cannot be stably bound by a second transposase complex. It may simply involve differential stability of protein-DNA interactions, since footprinting studies showed that the protein has high nonspecific DNA (15), yet nonspecific DNA binding is not frequently observed in the AFM assays (see Ref. 14 and this study). Quantitation of these time course experiments, in the presence or absence of GTP, indicated that DNA binding by the transposase protein was rapid and, as had been observed using DNase I footprinting (9, 15), occurred in both the absence or presence of GTP (Fig. 2*A*). These experiments show that initial protein-DNA complexes form first on either P element end (Fig. 2*A*), and that they form prior to the GTP-dependent synapsis observed previously (14). These findings rule out the possibility of the transposase recognizing both P element ends and oligomerizing on the DNA. In fact, measurements indicate that a pre-formed tetramer binds to one P element end prior to synapsis (see below).

Time Course of the Formation and Reactivity of Early P Element Transposase-DNA Complexes—Our previous studies indicated that GTP allows the transposase to form synaptic complexes containing both P element ends at 30 min (14). However, we had not examined earlier time points. Given the observed single end binding (see Fig. 1 and Fig. 2*A*), we therefore quantitated formation of both synaptic and singly or dou-

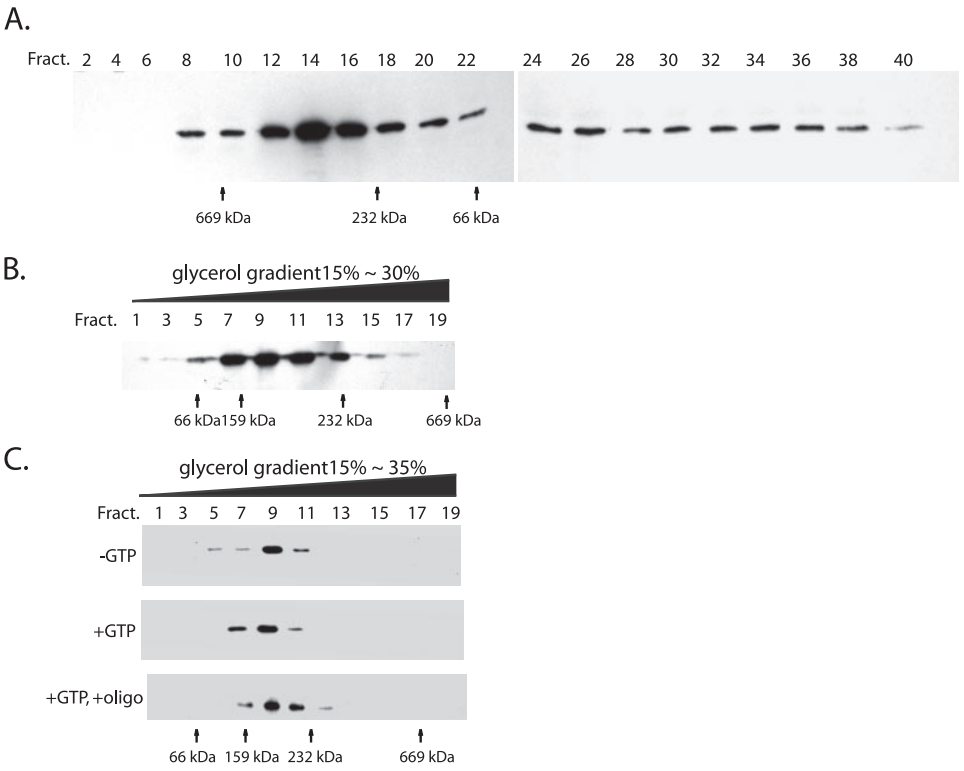

FIGURE 3. **Biochemical fractionation of P element transposase by gel filtration chromatography and glycerol gradient velocity sedimentation.** *A,* P element transposase chromatographs as a tetramer in a gel filtration experiment. About 3.5 mg of transposase was loaded onto the Superdex 200 column in 300 mM KCl buffer (see "Experimental Procedures"). Column fractions were monitored for the presence of transposase by SDS-PAGE followed by immunoblotting. Standard proteins used to calibrate the column were bovine serum albumin (66 kDa), catalase (232 kDa), and thyroglobulin (669 kDa), and their elution positions are indicated by *arrows*. Fraction (*Fract.*) numbers are indicated *above* the figure. Although the transposase protein seemed to be spreading across the column profile, perhaps because of interaction between the protein and the column matrix, the peak elution position (fractions 12–16) corresponded to an apparent molecular mass of ~350 kDa. *B,* P element transposase showed a single peak from glycerol gradient velocity sedimentation corresponding to an apparent molecular mass of ~170 kDa in several repetitions. Standard proteins (bovine serum albumin (66 kDa), alcohol dehydrogenase (159 kDa), catalase (232 kDa), and thyroglobulin (669 kDa)) were centrifuged in a parallel gradient. *C,* apparent size of P transposase determined by velocity sedimentation is not altered in the absence of GTP (*top panel*), in the presence of GTP (*middle panel*), or in the presence of GTP and an excess of a pre-cleaved strand transfer oligonucleotide (*bottom panel*). The centrifugation was performed the same as in *B*. Fractions were taken and assayed by SDS-PAGE and immunoblotting. The locations and sizes of the standard proteins are marked *below* the *bottom panel*.

bly cleaved donor DNA complexes during this early time course in the absence or presence of GTP (Fig. 2*B*). In the absence of GTP, very little synapsis was observed over the same period (Fig. 2*B*, *top panel*, *−GTP*), yet many molecules have transposase bound to one end (see Fig. 2*A*). We observed GTP-dependent, synaptic complex formation over the 1–30-min time course (Fig. 2*B*, *top panel, +GTP*). The presence of catalytically active complexes was also measured over this same early time course. In this assay, donor DNA cleavage was evident by singly tailed molecules or doubly cleaved circular excised transposon DNA fragments (Fig. 2*B*, *left panel*) (14). During this short time course, only a small fraction of the DNA went on to become cleaved at the P element termini (Fig. 2*B*, *middle* and *bottom panels, +GTP*). In the absence of GTP, no cleavage products were observed (Fig. 2*B*, *middle* and *bottom panels, −GTP*). This was expected because GTP-dependent synapsis is required prior to donor DNA cleavage (14). However, these experiments show that although transposase-dependent synapsis occurs rapidly in the presence of GTP over the 1–30-min time course, a very small fraction of the synaptic complexes undergoes DNA cleavage, suggesting that possibly a conformational change or other type of remodeling of the synaptic complex, perhaps positioning of the active sites over the appropriate positions on the P element ends, may need to occur prior to catalytic activation of the transposase-DNA synaptic complex.

Purified Native Transposase Exists as a Tetramer and Is the Form That Interacts Initially with One P Element End—One universal observation about the transposase/integrase superfamily of polynucleotidyltransferases is that they invariably act as oligomeric enzymes, with at least two active sites, one positioned at each transposon end (1, 12). Because of the unusual nature of the P element cleavage site, a staggered break with a 17-nucleotide single strand extension (11), it was of interest to determine the oligomeric state of the transposase protein in the absence of DNA and when it initially interacts with one P element terminus. Therefore, we performed gel filtration chromatography on a Superdex 200 column in 300 mM NaCl to determine the native molecular weight of the protein, monitoring the column fractions by immunoblotting (Fig. 3*A*). The peak of the transposase protein in fractions 12–16 appears to be consistent with a tetramer of ~360 kDa (four 87-kDa monomer subunits). The protein also appears to interact nonspecifically with the resin, because there is a weak protein signal across the column profile. Velocity sedimentation in 15–30% glycerol gradients was also used to ascertain the native size of the transposase protein oligomer. In this experiment, the protein sediments with an apparent molecular mass of ~150–200 kDa (Fig. 3*B*). However, this mass is not consistent with either the gel filtration (see above) or the volume measurements by AFM (see below and Fig. 4). Other oligomeric protein complexes, such as *Drosophila* origin recognition complex, behave anomalously in velocity sedimentation gradients because of aberrant hydrodynamic properties (38). Nonetheless, both the gel filtration and gradient peak fractions are active for donor DNA cleavage using an LM-PCR-based assay (data not shown; see Ref. 14). More importantly, we also directly examined both the gel filtration and gradient peak fractions to visualize DNA binding and to take volume measurements. These direct measurements of the size-fractionated transposase showed that the protein in both the gel filtration and glycerol gradient fractions was the same size and that it was a tetramer when prepared for

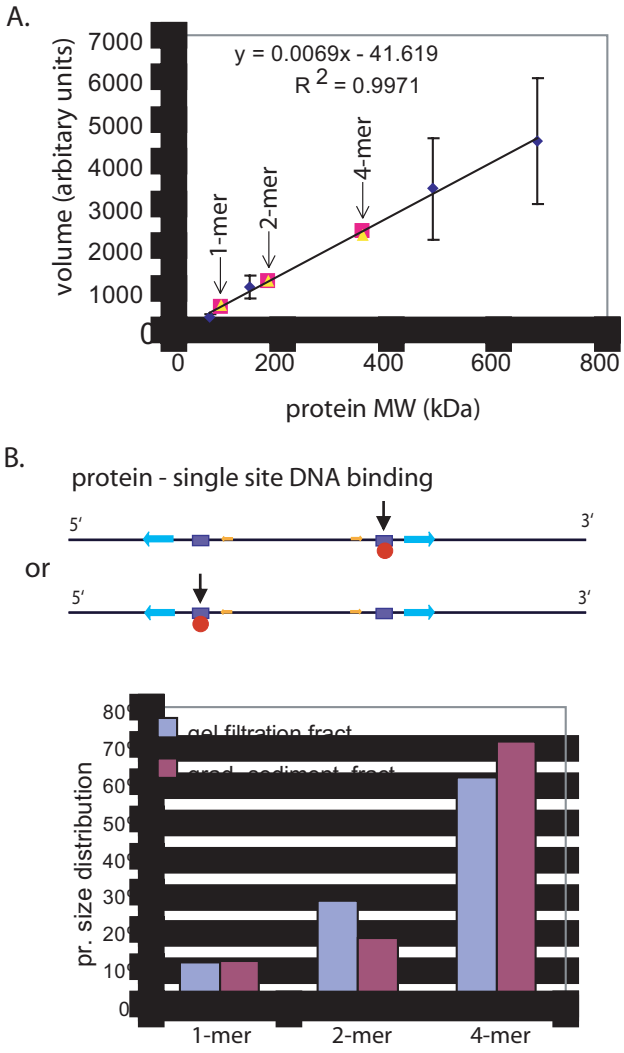

FIGURE 4. **Distribution of P element transposase oligomeric states.** *A,* molecular weight determination of transposase oligomeric states. A standard curve was obtained by measuring the volumes of four different proteins as follows: bovine serum albumin (66 kDa), alcohol dehydrogenase (140 kDa), *E. coli* RNA polymerase holoenzyme (479 kDa), and thyroglobulin (669 kDa). Both transposase and standard proteins were imaged separately and measured the same way. The *red squares* along the standard curve represent the predicted volumes for the monomeric, dimeric, and tetrameric oligomeric states of transposase; the *yellow triangles* represent the volumes experimentally measured. The molecular masses corresponding to the measured volumes are as follows: 87 kDa (monomer), 174 kDa (dimer), and 348 kDa (tetramer). Samples of P element transposase were deposited on freshly cleaved mica for 1 min, rinsed with water, and imaged by AFM. The volume measurements were carried out according to the protocol developed by V. Holmes (56). *B,* peak protein fractions from both the gel filtration chromatography and glycerol gradient velocity sedimentation experiments showed predominantly tetramers by AFM. Under our experimental conditions (a 30-min preincubation of the protein fraction with P element DNA in the absence of GTP), we observed the majority of the transposase protein binding to a single P element DNA end without synapsis. Protein size (volume) was measured as described under "Experimental Procedures." The *bar graphs* show that comparable multimeric states are observed in both fractions, with the majority of oligomers as the tetrameric form both the gel filtration and glycerol gradient methods of the protein size estimation.

AFM at pH 7 (see below and Fig. 4*B*). Additionally, the behavior of the protein in the gradients was not altered in the presence or absence of GTP or by a duplex oligonucleotide that corresponds to a pre-cleaved P element end that can be bound to the protein and is used as a substrate for strand transfer (37) (Fig. 3*C*). Taken together these data are consistent with the active form of the P element transposase being a tetramer and the oligomer being pre-formed in the absence of P element DNA.

As an independent means to determine the mass and oligomeric state of the transposase, we turned to AFM. AFM can be used to determine the volume of a protein, which by comparison to the volumes of a set of standard proteins can be used to determine the molecular weight of the protein. These sizes were determined by measuring the volumes of a set of standard proteins (bovine serum albumin, alcohol dehydrogenase, *E. coli* RNA polymerase, and thyroglobulin), plotting a standard curve and determining the relative molecular weight of the P element transposase sizes from the measured volumes (Fig. 4*A*). This type of molecular weight determination has previously been done for the DNA enhancer-binding protein NtrC and for other proteins (31, 39). Transposase protein, in the absence of DNA, was spotted on mica, rinsed with water, and imaged by AFM. The volumes of the protein species were determined and plotted as a function of numbers in the population (supplemental Fig. 1*A*). Three prevalent volume distributions were observed, consistent with monomer, dimer, and tetramer sizes, with dimeric and monomeric sizes predominating. However, we subsequently found out that low pH water washes of the mica grid prior to AFM imaging caused dissociation of the predominantly tetrameric protein into dimers and monomers (supplemental Fig. 2, *A* and *B*). This effect of low pH on the oligomeric state of transposase could also be seen with transposase-DNA complexes (supplemental Fig. 1*B*). However, and most importantly, when neutral pH (~7) wash water was used to image single end protein DNA complexes formed on P element DNA in the absence of GTP, either from the gel filtration column or glycerol gradient fractions, the predominant form of the protein bound to DNA was a tetramer (Fig. 4*B*), consistent with the observations of the protein in the absence of DNA (Fig. 4*A* and supplemental Fig. 1*A* and Fig. 2*B*) and the chromatographic behavior by gel filtration (Fig. 3*A*). Because in the presence of low pH water more monomers and dimers were observed compared with samples prepared with pH 7 water, we conclude that dissociation of tetrameric transposase complexes can occur during sample preparation with low pH water. These observations clearly show that sample preparation, and more specifically the pH of the water wash prior to imaging, could influence the level of the transposase oligomers but that the protein species that is active for transposition is a tetramer. In fact, LM-PCR assays for donor DNA cleavage (14), using both the peak gel filtration and glycerol gradient fractions that were imaged by AFM, indicate that these fractions, although dilute, exhibit donor DNA cleavage activity. Effects of pH of the water used in AFM sample preparation have been noted by others (40). A tetramer is formed even in the absence of P element DNA. Given the presence in the primary transposase amino acid sequence of a leucine zipper and a second protein interaction region (41, 42), as well as the staggered nature of the P element cleavage sites (11), it seems reasonable that the P element transposase might be active as a tetramer. Our results with the AFM volume measurements of the gel filtration and glycerol gradient fractions are indicative of higher order oligomerization of the transposase, in the absence of DNA, to a

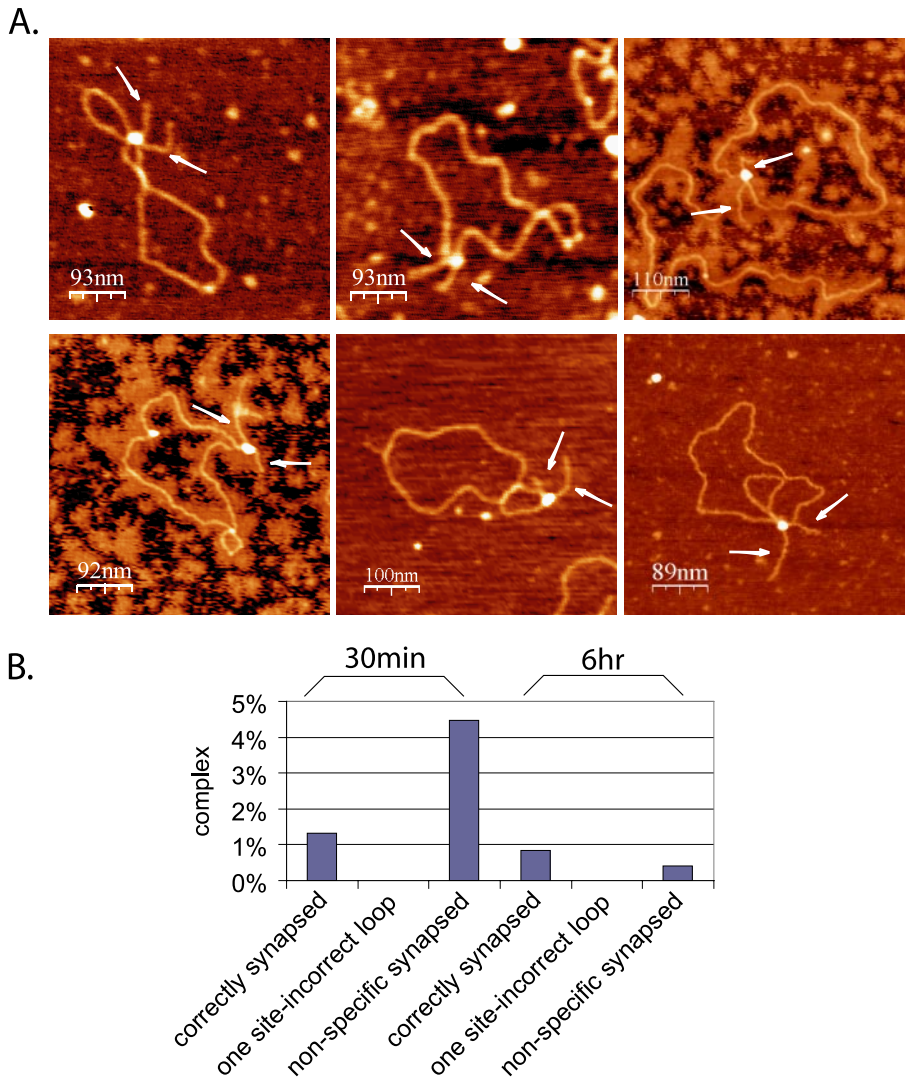

FIGURE 5. **Formation of full-length P element synaptic complexes.** *A,* 2.9-kb full-length P element sequence flanked by 200- and 300-bp DNA tails (200–2900-300) was incubated with transposase for 30 min at room temperature and then imaged by AFM. Synaptic complexes of full-length P element were observed although at much lower frequency than when a smaller 629-bp P element was used (300–600-600) (14). See Fig. 2*B* and Fig. 3*B*. The two tails of the synaptic complexes are indicated by *white arrows*. *B,* single-site binding (*i.e.* one site at one P element end bound with an incorrect loop size), nonspecifically bound synapsed or looped complexes (*i.e.* transposase not bound at either P element end), and correctly synapsed molecules for the full-length 2.9-kb P element DNA quantitated as a percentage of the total DNAs observed at 30-min and 6-h incubation times.

pre-formed tetramer and that the tetrameric form is what binds initially to one of the two P element ends prior to synapsis.

Correct Synapsis of Full-length 2.9-kb P Elements—Our previous studies by AFM of P element transposase used small, but mobile, internally deleted P elements (14) (see Fig. 1, above). We therefore wanted to see if we could visualize the synapsis of full-length 2.9-kb P elements, and at the same time, if we could visualize smaller looped DNA molecules that might represent molecules in the process of "sliding," a possible mechanism by which DNA-binding proteins find binding sites on DNA (43, 44). This mode of DNA site recognition contrasts to intersegmental transfer in which repeated collisions between proteins bound at one site and other parts of the DNA allow a bound protein to locate a second binding site (43, 44). Therefore, a 3.4-kb DNA was amplified carrying a full-length 2.9-kb P element along with 0.2- and 0.3-kb flanking DNA "tails" (200–2900-300), and this DNA molecule was used to form synaptic complexes with P element transposase. Although these images are more difficult to analyze because of the "crossing" of the long internal portion of the P element DNA, we did observe normal synaptic complexes with the appropriate tail length and the correctly sized P element segment (Fig. 5*A*). However, quantitation of these synaptic complexes at 30 min showed only 1.3% correctly synpased molecules (Fig. 5*B*), in contrast to 16.7% for the 0.6-kb P element (300–600-600) (see Fig. 2*B*, above). Additionally for this long DNA, although essentially no incorrectly sized loops (*i.e.* one site bound with an incorrect loop size) were observed, initial time points indicated nonspecific complexes (loops with neither tail of the correct length) were formed (4.47%), which dissociated over time (0.42%) (Fig. 5*B*). Because in these complexes there is no tail length of the correct size for engagement at one of the two P element ends, we believe these represent nonspecific protein binding, which has also been observed for this protein in DNase I footprinting assays (15). This type of incorrect complex was not observed (<0.01%) for the small 0.6-kb P element (data not shown). Thus, even with full-length P elements we can observe correctly synapsed molecules but at a lower frequency than was observed with a smaller 0.6-kb P element. Taken together, these data indicate that with both small and full-length P elements correctly synapsed molecules could be observed, and essentially no incorrectly sized loops with one P element end bound (<0.01%) were observed. For the longer P element, transposase complexes bound nonspecifically to DNA, but these complexes were not engaged at either P element end. These protein-DNA complexes were observed at early times but were lost with time, suggesting that incorrectly synapsed looped molecules are not stable and that these unstable, incorrect loops form more often on longer P element DNA segments, perhaps because of a higher local concentration of nonspecific DNA-binding sites for transposase (15).

Nonhydrolyzable GTP Analogs Support Normal Synapsis and Donor DNA Cleavage—Our initial characterization of the P element transposition *in vitro* showed that both GTP and nonhydrolyzable GTP analogs supported both catalytic steps of P ele-

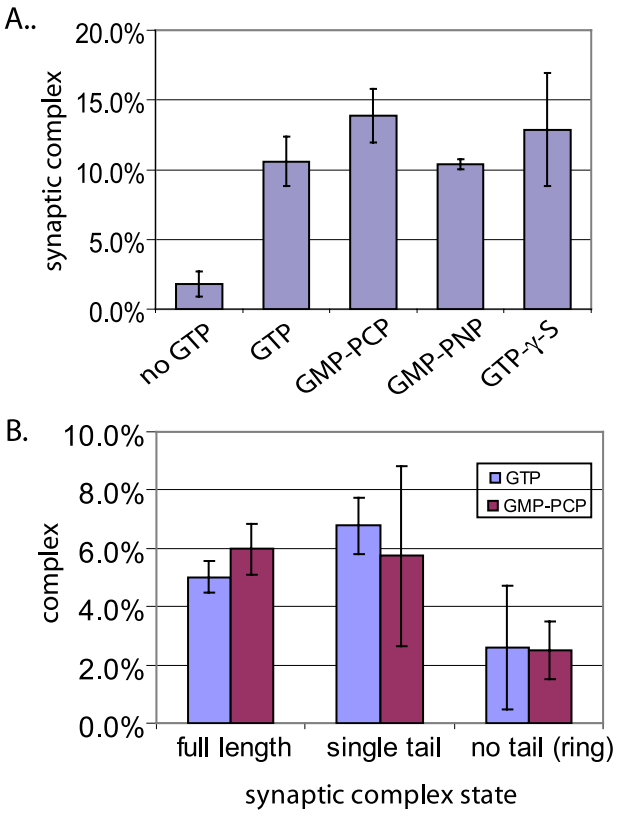

FIGURE 6. **Nonhydrolyzable GTP analogs support normal levels of transposase-mediated P element synapsis and excision activity.** *A,* effect of 2 mM nonhydrolyzable GTP analogs GMP-PCP, GMP-PNP, and GTPγS on synaptic complex formation, and excision activity was tested in a standard assay and examined by AFM. In the graphs, the total synaptic complexes (uncleaved, singly cleaved, and doubly cleaved) that were generated during a 1-h incubation in the presence or absence of 2 mM GTP or GTP analogs are shown. *B,* comparison between the percentages of pre- and post-cleavage synaptic complexes corresponding to full-length uncleaved (two tails), singly cleaved (*single tail*) or doubly cleaved (*no tail* (*ring*)) formed during a overnight incubation in the presence of 2 mM GTP or 2 mM GMP-PNP.

ment transposition, donor DNA cleavage, and strand transfer (9). Thus, it was of interest to test the ability of these analogs to support synapsis and donor cleavage. By quantitating the total fraction of synpased and singly/doubly cleaved P element DNA molecules, it was evident that all three nonhydrolyzable GTP analogs tested, GTPγS, GMP-PCP, and GMP-PNP, can support both synapsis and donor DNA cleavage at levels equivalent to that observed with regular GTP (Fig. 6*A*). The fact that little activity was observed without added GTP, as noted previously, argues that all three analogs actually bind to the transposase protein to activate it for synapsis and donor DNA cleavage. Additionally, we quantitatively compared each class of synaptic complex formed with either GTP or GMP-PCP (Fig. 6*B*). In this case, the levels of uncleaved (full-length, two tails), singly cleaved (single tail), or doubly cleaved (no tail ring) were similar, again indicating that the GTP analog can full substitute for GTP in the assembly and donor DNA cleavage stages of the reaction. Thus, as observed previously in bulk biochemical assays, these AFM studies show GTP binding, but not hydrolysis, is required for P element transposase activity and that these nonhydrolyzable analogs support the conformational effects on the transposase protein that GTP normally provides to cause synapsis of the P element ends, ultimately resulting in the catalytic activation of the protein once assembled on P element DNA.

DISCUSSION

The results presented here show that the initial protein-DNA complexes formed between P element transposase and the transposon ends occur by interaction of a pre-formed tetramer with either one of the two termini. Many DNA-binding proteins that act on two sites on DNA are composed of multiple subunits. These proteins often can come either preassembled as oligomers, assemble on the DNA, or assemble on protein (43, 44). In the case of P elements, it appears that the initial contact of the protein with one end of the transposon DNA occurs via interactions of one site-specific DNA binding domain within the tetramer with one high affinity transposase-binding site located internally near the transposon terminus (15, 42). It is this pre-assembled oligomer that then synapses the P element ends, presumably by a second site-specific protein-DNA interaction at the opposite transposon end that carries a transposase-binding site (15, 42). Because of the unusual nature of the P element DNA cleavage, leaving a staggered double strand break, with a 17-nucleotide 3′ extension (11), it would seem to make sense that two catalytic domains would be required to cleave each transposon end, so that a tetramer would donate two subunits to each terminal donor DNA cleavage. Additionally, at high protein concentrations the truncated KP repressor protein, which carried the N-terminal THAP DNA binding domain, can also bind to the sub-terminal 11-bp inverted repeats (42). However, in footprinting assays with single transposon ends, the full-length transposase protein does not appear to stably interact with these internal repeats (15). Nonetheless, it seems possible that two DNA binding domains in the tetrameric transposase synaptic complex might interact with two DNA-binding sites at each transposon end to stabilize the complex.

In other transposon systems, different strategies are used to assemble synaptic complexes (12, 18, 45, 46). In the case of Tn5 and Tn10, a dimeric complex is active at the transposon ends (27–30), yet the state of the protein in solution and whether monomers bind each transposon end with subsequent synapsis or whether dimers bind to one end first is not known (29, 47). In the case of bacteriophage Mu, whereas each transposon end has three binding sites for the monomeric transposase, binding of monomeric subunits to these sites results in the DNA-induced assembly of an active and stable tetramer, in which two catalytically active subunits perform DNA cleavage and joining in *trans* (17, 19, 45, 48–51). It has been shown recently that the eukaryotic Hermes transposase exists in solution and in crystals as a hexamer (16). Another eukaryotic transposase, MosI, seems to function as a dimer or tetramer (22–25), whereas a related mariner transposase, Himar 1, functions as a tetramer (26). Unlike the case of bacteriophage Mu, where a monomeric protein undergoes DNA-induced tetramer formation (18, 19, 49), we have no evidence for oligomerization of the P element transposase protein on P element DNA, because we consistently see what appears to be a tetramer on one of the P element ends prior to synapsis and in fact can visualize tetramers of P element transposase in the absence of DNA. Although the stoi-

chiometry of the V(D)J recombinase RAG1-RAG2 in the synaptic complex is not clear (for a review see Ref. 52), it is clear, as we have observed here with P element transposase, that *in vitro* with short oligonucleotide substrate DNAs the RAG1/RAG2 oligomer binds first to one recombination signal sequence and then assembles a paired synaptic complex, with the second recombination signal sequence joining as naked DNA (21, 52). Previous studies have noted similarities between V(D)J recombination and P element transposition, notably the natural pairing of two signal/transposon ends with different spacer lengths (6, 11). Thus, our studies shed light on the mechanism of how P element DNA is recognized by and assembles with the transposase protein and also highlight the fact that a number of distinct strategies are used in different transposition systems to assemble catalytically active protein-DNA complexes.

Studies of a truncated, naturally occurring repressor form of the P element transposase, termed the KP protein, carries the N-terminal site-specific DNA binding domain and can inhibit transposase activity *in vitro* by binding to the transposase-binding sites (41, 42). The KP protein also contains two distinct protein-protein interaction domains, a leucine zipper, and a second unclassified dimerization motif. Chemical cross-linking experiments showed that the KP protein dimerized in solution through these motifs (41, 42). The KP protein lacks both the GTP binding and catalytic domains found in the full-length transposase (6). Hence, our observation that the full-length transposase can exist as a tetramer suggests that perhaps the GTP binding or catalytic domains might mediate a higher order oligomerization. It is known that nucleotide binding domains, such as those found in the AAA^+ superfamily of ATP-binding/hydrolyzing proteins, can form higher order oligomers (53, 54). Thus, although the N-terminal region of the transposase could mediate protein dimerization, it seems likely that the central and C-terminal regions of the protein could promote higher order oligomerization.

The formation of this initial P element transposase protein-DNA complex that we have observed occurs independently of the GTP cofactor, yet our previous studies showed that GTP promotes synapsis of the P element ends (14). The data presented here indicate that a pre-formed tetramer, in the absence of GTP, binds to one P element end but does not allow synaptic complex formation. These observations suggest that GTP may function to orient a second DNA binding domain within the transposase oligomer to allow synapsis to occur via the binding of one of the other three available DNA binding domains in the tetramer to the high affinity transposase-binding site on the opposite P element end (6, 15). Although we do observe correct, but relatively inefficient, synapsis of the two P element ends in a full-length 2.9-kb P element, we also observe incorrectly looped molecules of smaller sizes. These are only observed with the full-length 2.9-kb but not the 0.6-kb P element. Such smaller loops might support the idea of sliding or "tracking" of the protein along DNA to find its second binding site on the P element ends. However, these incorrectly looped molecules decay with time and do not give rise to a concomitant increase in correctly synapsed full-length 2.9-kb P elements, suggesting that they are not stable intermediates on the pathway to correctly synapsed molecules. Tracking is the normal mode of finding second sites performed by "motor proteins," often accompanied by hydrolysis of nucleoside triphosphates, as occurs with the type I restriction endonucleases (43, 44). We have no evidence for this type of mechanism, and moreover, we have shown here that synapsis of the P element ends occurs normally with nonhydrolyzable GTP analogs (see above and Ref. 9). By contrast, "looping" or "intersegmental transfer" is the mechanism used by FokI, a type IIS restriction enzyme, whose activity for double strand cleavage by a dimeric complex at one recognition site is stimulated by the presence of a second recognition site elsewhere on the same DNA molecule (43, 44, 55). The data presented here are consistent with a mechanism involving intersegmental transfer, at least with smaller P elements, and this search mechanism is less efficient on longer P element DNAs because of a higher effective concentration of nonspecific binding sites between the two P element ends. These observations are consistent with and provide an explanation for why longer recombinant P elements transpose less efficiently *in vivo* (6, 7).

Our data indicate that, prior to synapsis of the P element ends, the transposase tetramer binds to one transposon end. In the presence of GTP *in vitro*, the two P element ends undergo synapsis (14). P elements are known to undergo some unusual and unexplained rearrangements *in vivo* (33–35). In particular, when the termini from two different P elements located on different DNA molecules, either sister chromatids or homologs, undergo synapsis, cleavage, and transposition, a process termed hybrid end-insertion (33) occurs that results in both chromosomal deletions and duplications (34). These ectopic P element rearrangements could be explained by synapsis of two P element ends from different DNA molecules (35). The observations reported here that one P element end binds transposase tetramers initially, prior to synapsis, may explain the molecular basis for hybrid end-insertion. Again, unlike bacteriophage Mu, P element transposase arrives at the P element ends as a preformed tetramer prior to synapsis. The ability of the transposase, bound to one P element end, to synapse with a second P element end on a nearby sister chromatid or homolog and then undergo normal transposition would explain the underlying basis for these abnormal chromosomal rearrangements. Although all transposase proteins need to bring distant transposon ends together and generally are not active until synapsis occurs, it appears that there are different ways by which different transposase proteins initially interact with the transposon DNA and assemble to generate a synaptic complex prior to DNA cleavage and joining. It is also more generally interesting to note that both differential stability and conformational changes of these protein-DNA assemblies are known to modulate the activities and outcomes of these DNA recombination reactions by these coordinated molecular machines (1, 12, 19, 49).

Acknowledgments—We thank members of the Rio laboratory for suggestions and advice throughout the course of this work and Dr. Caroline Kane for the E. coli RNA polymerase holoenzyme.

REFERENCES

1. Craig, N. L., Cragie, R., Gellert, M., and Lambowitz, A. M. (eds) (2002) *Mobile DNA II*, American Society for Microbiology, Washington, D. C.
2. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., Szustakowki, J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y. J. (2001) *Nature* **409,** 860–921
3. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigo, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001) *Science* **291,** 1304–1351
4. Curcio, M. J., and Derbyshire, K. M. (2003) *Nat Rev Mol. Cell Biol.* **4,** 865–877
5. Kazazian, H. H., Jr. (2004) *Science* **303,** 1626–1632
6. Rio, D. C. (2002) in *Mobile DNA II* (Craig, N. L., Cragie, R., Gellert, M., and Lambowitz, A. M., eds) pp. 484–518, American Society for Microbiology, Washington, D. C.
7. Engels, W. R. (1989) in *Mobile DNA* (Berg, D. E., and Howe, M. M., eds) pp. 437–484, American Society Microbiology, Washington, D. C.
8. Craig, N. L. (1995) *Science* **270,** 253–254
9. Kaufman, P. D., and Rio, D. C. (1992) *Cell* **69,** 27–39
10. Mul, Y. M., and Rio, D. C. (1997) *EMBO J.* **16,** 4441–4447
11. Beall, E. L., and Rio, D. C. (1997) *Genes Dev.* **11,** 2137–2151
12. Rice, P. A., and Baker, T. A. (2001) *Nat. Struct. Biol.* **8,** 302–307
13. Gellert, M. (2002) *Annu. Rev. Biochem.* **71,** 101–132
14. Tang, M., Cecconi, C., Kim, H., Bustamante, C., and Rio, D. C. (2005) *Genes Dev.* **19,** 1422–1425
15. Kaufman, P. D., Doll, R. F., and Rio, D. C. (1989) *Cell* **59,** 359–371
16. Hickman, A. B., Perez, Z. N., Zhou, L., Musingarimi, P., Ghirlando, R., Hinshaw, J. E., Craig, N. L., and Dyda, F. (2005) *Nat. Struct. Mol. Biol.* **12,** 715–721
17. Lavoie, B. D., Chan, B. S., Allison, R. G., and Chaconas, G. (1991) *EMBO J.* **10,** 3051–3059
18. Baker, T. A. (1993) *Curr. Opin. Genet. Dev.* **3,** 708–712
19. Baker, T. A., and Mizuuchi, K. (1992) *Genes Dev.* **6,** 2221–2232
20. Rodgers, K. K., Villey, I. J., Ptaszek, L., Corbett, E., Schatz, D. G., and Coleman, J. E. (1999) *Nucleic Acids Res.* **27,** 2938–2946
21. Mundy, C. L., Patenge, N., Matthews, A. G., and Oettinger, M. A. (2002) *Mol. Cell. Biol.* **22,** 69–77
22. Auge-Gouillou, C., Brillet, B., Germon, S., Hamelin, M. H., and Bigot, Y. (2005) *J. Mol. Biol.* **351,** 117–130
23. Auge-Gouillou, C., Brillet, B., Hamelin, M. H., and Bigot, Y. (2005) *Mol. Cell. Biol.* **25,** 2861–2870
24. Richardson, J. M., Dawson, A., O'Hagan, N., Taylor, P., Finnegan, D. J., and Walkinshaw, M. D. (2006) *EMBO J.* **25,** 1324–1334
25. Zhang, L., Dawson, A., and Finnegan, D. J. (2001) *Nucleic Acids Res.* **29,** 3566–3575
26. Lipkow, K., Buisine, N., Lampe, D. J., and Chalmers, R. (2004) *Mol. Cell. Biol.* **24,** 8301–8311

27. Bhasin, A., Goryshin, I. Y., Steiniger-White, M., York, D., and Reznikoff, W. S. (2000) *J. Mol. Biol.* **302,** 49–63
28. Bolland, S., and Kleckner, N. (1996) *Cell* **84,** 223–233
29. Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Rayment, I. (2000) *Science* **289,** 77–85
30. Steiniger-White, M., Rayment, I., and Reznikoff, W. S. (2004) *Curr. Opin. Struct. Biol.* **14,** 50–57
31. Wyman, C., Rombel, I., North, A. K., Bustamante, C., and Kustu, S. (1997) *Science* **275,** 1658–1661
32. de Jager, M., van Noort, J., van Gent, D. C., Dekker, C., Kanaar, R., and Wyman, C. (2001) *Mol. Cell* **8,** 1129–1135
33. Gray, Y. H., Tanaka, M. M., and Sved, J. A. (1996) *Genetics* **144,** 1601–1610
34. Preston, C. R., Sved, J. A., and Engels, W. R. (1996) *Genetics* **144,** 1623–1638
35. Venken, K. J., and Bellen, H. J. (2005) *Nat. Rev. Genet.* **6,** 167–178
36. O'Hare, K., and Rubin, G. M. (1983) *Cell* **34,** 25–35
37. Beall, E. L., and Rio, D. C. (1998) *EMBO J.* **17,** 2122–2136
38. Gossen, M., Pak, D. T., Hansen, S. K., Acharya, J. K., and Botchan, M. R. (1995) *Science* **270,** 1674–1677
39. Ratcliff, G. C., and Erie, D. A. (2001) *J. Am Chem. Soc.* **123,** 5632–5635
40. Moreno-Herrero, F., de Jager, M., Dekker, N. H., Kanaar, R., Wyman, C., and Dekker, C. (2005) *Nature* **437,** 440–443
41. Lee, C. C., Beall, E. L., and Rio, D. C. (1998) *EMBO J.* **17,** 4166–4174
42. Lee, C. C., Mul, Y. M., and Rio, D. C. (1996) *Mol. Cell. Biol.* **16,** 5616–5622
43. Halford, S. E., and Marko, J. F. (2004a) *Nucleic Acids Res.* **32,** 3040–3052
44. Halford, S. E., Welsh, A. J., and Szczelkun, M. D. (2004b) *Annu. Rev. Biophys. Biomol. Struct.* **33,** 1–24
45. Craigie, R. (1996) *Cell* **85,** 137–140
46. Gueguen, E., Rousseau, P., Duval-Valentin, G., and Chandler, M. (2005) *Trends Microbiol.* **13,** 543–549
47. Reznikoff, W. S. (2003) *Mol. Microbiol.* **47,** 1199–1206
48. Aldaz, H., Schuster, E., and Baker, T. A. (1996) *Cell* **85,** 257–269
49. Mizuuchi, M., Baker, T. A., and Mizuuchi, K. (1995) *Cell* **83,** 375–385
50. Williams, T. L., and Baker, T. A. (2004) *J. Biol. Chem.* **279,** 5135–5145
51. Yang, J. Y., Jayaram, M., and Harshey, R. M. (1996) *Cell* **85,** 447–455
52. Jones, J. M., and Gellert, M. (2004) *Immunol. Rev.* **200,** 233–248
53. Hanson, P. I., and Whiteheart, S. W. (2005) *Nat. Rev. Mol. Cell Biol.* **6,** 519–529
54. Lupas, A. N., and Martin, J. (2002) *Curr. Opin. Struct. Biol.* **12,** 746–753
55. Catto, L. E., Ganguly, S., Milsom, S. E., Welsh, A. J., and Halford, S. E. (2006) *Nucleic Acids Res.* **34,** 1711–1720
56. Holmes, V. (2001) *Biochemical and Structural Studies of Recombination and Chromosome Condensation Proteins*. Ph.D. dissertation, University of California, Berkeley