Supporting Information

Zawadzki et al. 10.1073/pnas.1311065110

SI Text

DNA and Protein Preparation. The 1-kb DNA substrates were prepared via a PCR using two fluorescently labeled oligonucleotides as primers and a plasmid template containing directly repeated *dif* sites separated by a 1-kb KmR gene cassette (pRB10). Phusion High-Fidelity DNA polymerase (NEB) was used.

The oligonucleotides used were as follows: substrate I, forward, 5′-GTGTCGACACAXGATTTAACATAAT-3′, and reverse, 5′-CTCTAGACCATGGAXCATGTGGTGCGCATA-3′; substrate II, forward, 5′-CATGTGTCGACACATGATTTAACATAATA-TACATTATGCGCACCACATGXAGCTGAGATCTG-3′, and reverse, 5′-AGACCATGGCATGTGGTGCGCATAATGTA-TATTATGTTAAATCXTGTGGATCCAC-3′; substrate III, forward, 5′-TAGCGTCGACCTACAXGATTTAACATAATA-TACATTATGCGCACCAAATGATTCGCAGCTGAGATC-3′, and reverse, 5′-GATCGATCTCAGCTGCGAAXCATTTG-GTGCGCATAATGTATATTATGTTAAATCATGTAGGT-CGACGCTA-3′, where X indicates the position of 5-C6-amino dT. Forward and reverse oligonucleotides were labeled at the X positions with Cy5 and Cy3B, respectively. Oligonucleotides were synthesized and HPLC purified by ATDBio Ltd. Cy5 labeling was performed by ATDBio, and Cy3B labeling was performed as previously described (1). After PCRs, the products were NcoI digested and ligated to a biotin-labeled 200-bp extension (produced using PCR). Subsequently, a long (2.8-kb) DNA tail was ligated to the opposite end (following a SalI digestion) and the 4-kb substrate was gel purified. The tail acts as a loading site for FtsK; recombination was never observed in our assays in substrates without it. Substrate III was prepared with unlabeled oligonucleotides, otherwise identical to those used in substrate I preparation. Subsequently, the distal *dif* site was removed by digestion with BglII and an identical sequence, produced by the annealing of two fluorophore-labeled oligonucleotides, was ligated back.

XerC, XerD, and FtsK trimer were purified according to previously published procedures (2, 3).

Instrumentation. Single-molecule total internal reflection fluorescence (TIRF) experiments were performed on a custom-built objective type TIRF microscope. A green (532-nm Cobolt Samba) and red (635-nm Cube Coherent) laser were combined using a dichroic mirror and coupled into a fiber optic cable. The output of the fiber was focused into the back focal plane of the objective (100°—oil immersion; numerical aperture, 1.4; Olympus) and displaced perpendicular to the optical axis such that laser light was incident at the slide–solution interface at greater than the critical angle, creating an evanescent excitation field. Alternating laser excitation was implemented by directly modulating the lasers, and all data were acquired using a 10-Hz alternation rate (or 20 Hz in the case of $XerC^{KQ}D$ experiments), with excitation powers of 1 mW for each laser. Fluorescence emission was collected by the objective and separated from the excitation light by a dichroic (545 nm/650 nm; Semrock) and cleanup filters (545 mLP, Chroma; and 633/25 nm notch filter, Semrock). The emission signal was focused on a rectangular slit to crop the image and then spectrally separated, using a dichroic (630-nm DRLP; Omega), into two emission channels, which were focused side by side onto an EMCCD camera (Andor iXon 897). The EMCDD was set to an EM gain of 300, corresponding to an approximate real gain of 4.55 counts per photon.

Sample Preparation. Biotinylated DNA was immobilized to the surface of a PEG-passivated coverslip using biotin–NeutrAvidin interactions and sealed using a silicone gasket (Grace Bio-laboratories) and a second coverslip as a lid. Imaging was performed in a buffer consisting of 50 mM Tris·HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 100 $\mu g \cdot mL^{-1}$ BSA, and 1 mM UV-treated Trolox. An enzymatic oxygen scavenging system consisting of 1 $mg \cdot mL^{-1}$ glucose oxidase, 40 $\mu g \cdot mL^{-1}$ catalase, and 1.4% (wt/vol) glucose was added just before sealing the sample before image acquisition. All experiments were performed at a room temperature of 21 °C.

Data Analysis. Fluorescence intensities were extracted from images using previously described TwoTone software (4). The apparent Förster resonance energy transfer (FRET) was calculated using the area under the Gaussian fit to the donor emission under donor excitation, F_{DD}; and the area under the Gaussian fit to the acceptor emission under donor excitation, F_{DA}; using the following:

$$E^* = \frac{F_{DA}}{F_{DD} + F_{DA}}.$$

The point spread function (PSF) widths in all channels were obtained from the mean widths of the fitted elliptical Gaussians. A combination of manual and hidden Markov model (HMM)-based time trace segmentation were used to construct E^* histograms for each FRET state. Histograms were fit with a Gaussian using nonlinear least-squares fitting in MATLAB.

HMM for Data Extraction. HMM can be used as a means of resolving FRET states present in noisy data (5). In a Markov model, the probability of a transition depends only on the present state of the system. HMM refers to the hidden nature of the underlying state, which can produce a range of outputs at each time point sampled. We used Seneca, a package written by Kristofer Gryte (Oxford) (6, 7). The states extracted by HMM were categorized according to their FRET efficiency and the states that preceded them. According to these categories, data belonging to particular structural states were plotted together in histograms and fit (Table S3). Data immediately before or after a transition were discarded to minimize camera integration time effects (8).

Accurate FRET. To convert from measured emission intensities into a distance we corrected for donor leakage into the acceptor emission channel, l; direct excitation of the acceptor, d; and the combined effect of differing detection efficiencies and quantum yields between the fluorophores, γ (9). The accurate FRET is given by the following:

$$E = \frac{DA_{correct}}{\gamma F_{DD} + DA_{correct}},$$

where the corrected acceptor emission under donor excitation:

$$DA_{correct} = F_{DA} - lF_{DD} - dF_{AA},$$

and

$$\gamma = \frac{\Delta DA_{correct}}{\Delta F_{DD}},$$

where $\Delta DA_{correct}$ is the change in $DA_{correct}$ either side of a change in FRET state, and ΔF_{DD} is the change in F_{DD} either side of the same transition. Transitions, rather than photobleaching events

(10), are chosen because they are much more frequent in our data and hence allow us to confirm that variation in γ between states in our recombination pathway is small, supporting our use of a mean γ for all populations and molecules from a particular substrate.

The Förster radius, R_0, was calculated from the refractive index of the medium, $n = 1.33$; the normalized fluorescence emission spectrum of the donor, $f_D(\lambda)$; the acceptor molar extinction coefficient, $\epsilon_A(\lambda)$; the quantum yield of the donor, Q_D; and the orientation factor, κ^2, according to the following:

$$R_0 = 0.211\left(\kappa^2 n^{-4} Q_D J\right)^{1/6}$$

$$J = \int f_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda.$$

The emission and absorption spectra were measured on a fluorometer (Photon Technology International) for singly labeled central regions of substrates I and II, in the presence of WT XerCD in the same buffer as the recombination experiments, but without Trolox or the oxygen scavenging system. A literature value of 0.67 was assumed for Q_D (11). The orientation factor was assumed to be 2/3. A Perrin plot of the anisotropy was used to estimate the freedom of each fluorophore relative to its attachment point, and this allowed an estimate of the 67% confidence interval for κ^2 using FRETnps software (12). The errors were propagated (13) and the distance calculated (Table S2).

The uncertainty in each of the accurate FRET distances is ~10 Å, the majority of which can be attributed to our uncertainty in the value of the orientation factor for the fluorophores, κ^2. We can compare distances predicted using a Cre–*loxP* structure (Table S1), with our three substrates, and observe broad agreement with the predicted distances. Using multiple independent labeling schemes to investigate structure allows us to mitigate the effect of this orientational uncertainty (14).

The smallest accurate FRET that we could reliably distinguish from the background due to leakage and direct excitation was established using a Rayleigh criterion and a kinetic Monte Carlo Markov chain simulation. The simulation explicitly simulated photon emission times in each channel with the appropriate cross talks, electron multiplying gain by the CCD, read noise, and the least-squares fitting to the PSF. The photon count was chosen to match the width of the $E^* \approx 0.72$ population in substrate II, which accounts for the excess heterogeneity found in the study by Holden et al. (4). It was found that $E^* \approx 0.20$ would just be resolvable against our background, corresponding to a maximum measureable distance of around 90 Å. Hence, any states in the recombination pathway with no apparent FRET correspond to an interfluorophore distance of >90 Å.

Rate Determination. The transition rates, $k = 1/\tau$, quoted in this work (Fig. 5) have been extracted using maximum-likelihood fitting in MATLAB (Figs. 2–4 and Figs. S3 and S4) using the probability density for a dwell, t, of the following:

$$p(t) = ke^{-kt}.$$

A more complicated kinetic scheme was apparent in substrate III, where reverse transitions were observed (Fig. S4*B*):

$$\text{XerC-HJ} \overset{k_-,k_+}{\longleftrightarrow} \text{XerC-Product} \xrightarrow{k_2} \text{Dissociation}$$

The rates k_+ and k_- refer to the forward and backward rates for XerC–HJ converting to XerC–P, and k_2 refers to the dissociation rate of XerC–Product. The summed dwells in XerC–P, $t_{\sum XerC\text{-}P}$ (i.e., the total time the molecule spends in XerC–P before dissociation), for each molecule should follow the probability density:

$$p\left(t_{\sum XerC-P}\right) = k_2 e^{k_2 t_{\sum XerC-P}},$$

and the dwell time spent in XerC–HJ before each transition to XerC–P, $t_{XerC-HJ}$, should follow:

$$p(t_{XerC-HJ}) = k_+ e^{-k_+ t_{XerC-HJ}}.$$

The dwell time spent in XerC–P before either a dissociation or a backward transition to XerC–HJ, t_{XerC-P} should follow:

$$p(t_{XerC-P}) = (k_2 + k_-)e^{-(k_2+k_-)t_{XerC-P}}.$$

These distributions can again be fit to the data using maximum-likelihood estimation in MATLAB, and the rates can all be deduced (Fig. S4*B*).

Reverse transitions for the other steps in the reaction pathway were rarely observed (<4 molecules in each dataset of ~200 molecules). From this, we can calculate an upper bound on the reverse rates (Fig. 5). We note that, for each state in the pathway, given a forward transition rate k_f and a backward rate k_b, the probability of observing less than m reverse transitions in N molecules is given by the following:

$$p(<m) = \sum_{0<i<m} \left(\frac{k_f}{k_b + k_f}\right)^{N-i} \left(\frac{k_b}{k_b + k_f}\right)^{i} \frac{N!}{(N-i)!i!}.$$

We choose our limiting backward rate such that were the "true" backward rate faster than our limit, our data (fewer than m reverse transitions) would have been observed less than 5% of the time. Setting $p = 5\%$; $m = 4$ and $N = 249$, 134, or 111 for substrates I, II, and III; and k_f to the forward rate for the appropriate transition, we can solve the above for k_b (Fig. 5).

1. Pinkney JNM, et al. (2012) Capturing reaction paths and intermediates in Cre-loxP recombination using single-molecule fluorescence. *Proc Natl Acad Sci USA* 109(51): 20871–20876.
2. Ferreira H, Sherratt DJ, Arciszewska LK (2001) Switching catalytic activity in the XerCD site-specific recombination machine. *J Mol Biol* 312(1):45–57.
3. Massey TH, Mercogliano CP, Yates J, Sherratt DJ, Löwe J (2006) Double-stranded DNA translocation: Structure and mechanism of hexameric FtsK. *Mol Cell* 23(4):457–469.
4. Holden SJ, et al. (2010) Defining the limits of single-molecule FRET resolution in TIRF microscopy. *Biophys J* 99(9):3102–3111.
5. Jung S, Dickson RM (2009) Hidden Markov analysis of short single molecule intensity trajectories. *J Phys Chem B* 113(42):13886–13890.
6. Uphoff S, Gryte K, Evans G, Kapanidis AN (2011) Improved temporal resolution and linked hidden Markov modeling for switchable single-molecule FRET. *ChemPhysChem* 12(3):571–579.
7. Le Reste L, Hohlbein J, Gryte K, Kapanidis AN (2012) Characterization of dark quencher chromophores as nonfluorescent acceptors for single-molecule FRET. *Biophys J* 102(11):2658–2668.
8. Bronson JE, Fei J, Hofman JM, Gonzalez RL, Jr., Wiggins CH (2009) Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. *Biophys J* 97(12):3196–3205.
9. Lee NK, et al. (2005) Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* 88(4):2939–2953.
10. Ha T, et al. (1999) Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci USA* 96(3):893–898.
11. Cooper M, et al. (2004) Cy3B: Improving the performance of cyanine dyes. *J Fluoresc* 14(2):145–150.
12. Muschielok A, et al. (2008) A nano-positioning system for macromolecular structural analysis. *Nat Methods* 5(11):965–971.
13. Uphoff S, et al. (2010) Monitoring multiple distances within a single molecule using switchable FRET. *Nat Methods* 7(10):831–836.
14. Rasnik I, Myong S, Cheng W, Lohman TM, Ha T (2004) DNA-binding orientation and domain conformation of the *E. coli* rep helicase monomer bound to a partial duplex junction: Single-molecule studies of fluorescently labeled enzymes. *J Mol Biol* 336(2): 395–408.

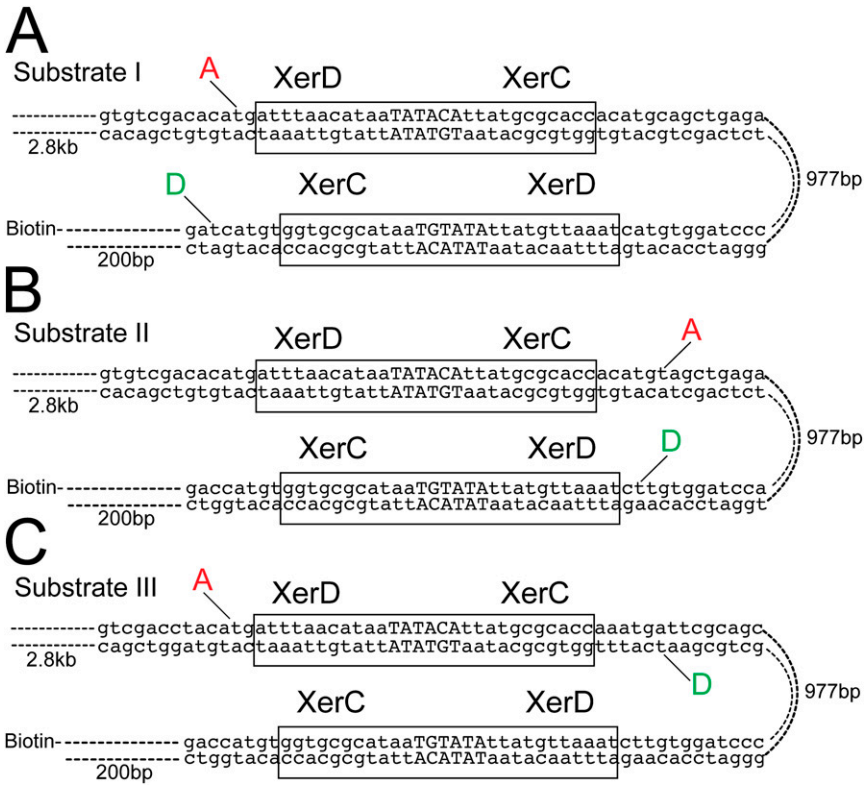

Fig. S1. Sequences and labeling positions of DNA substrates used throughout this study. *dif* sites are boxed and the 1-kbp–long DNA fragment linking two *dif* sites is represented by dashed lines. XerC and XerD binding sites are indicated with the name of the relevant protein. The red A indicates the attachment of Cy5 using a C6-linker, and the green D indicates the attachment of Cy3B using another C6-linker. (*A*) Substrate I. (*B*) Substrate II, designed to show FRET in intermediates preceding HJ isomerization and to investigate the stability of the product. (*C*) Substrate III, which was designed to show FRET in intermediates following HJ isomerization.

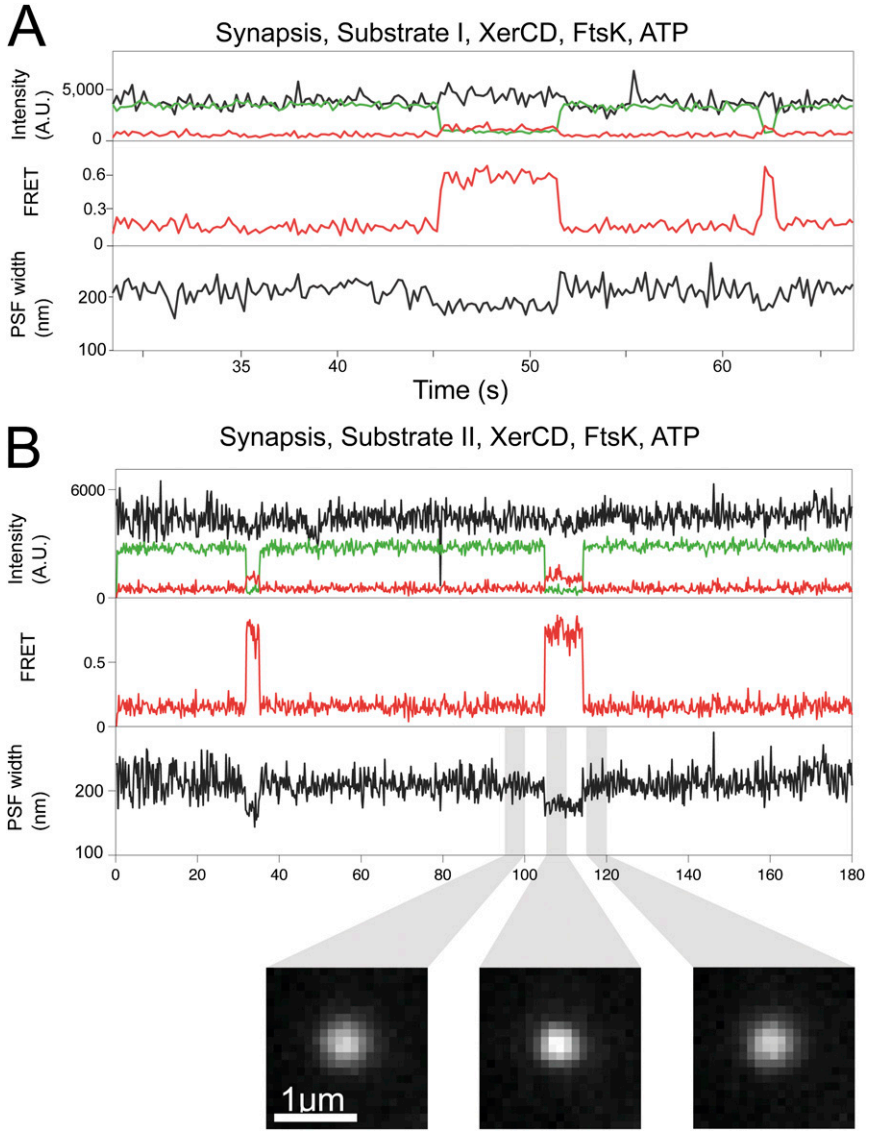

Fig. S2. Synapsis formation within substrates I and II in the presence of FtsK and ATP. (*A*) Time trace showing reversible synapsis formation in substrate I. Loading of FtsK was a limiting factor in recombination on the surface and substrate I would undergo repetitive synapsis formation until activated by FtsK. (*B*) Time trace showing synapsis events in the presence of FtsK in substrate II. PSF examples before, during, and after synapsis were made by averaging 25 frames (highlighted in gray). During synapsis, the PSF is narrower, which acts as our tethered fluorophore motion observable.

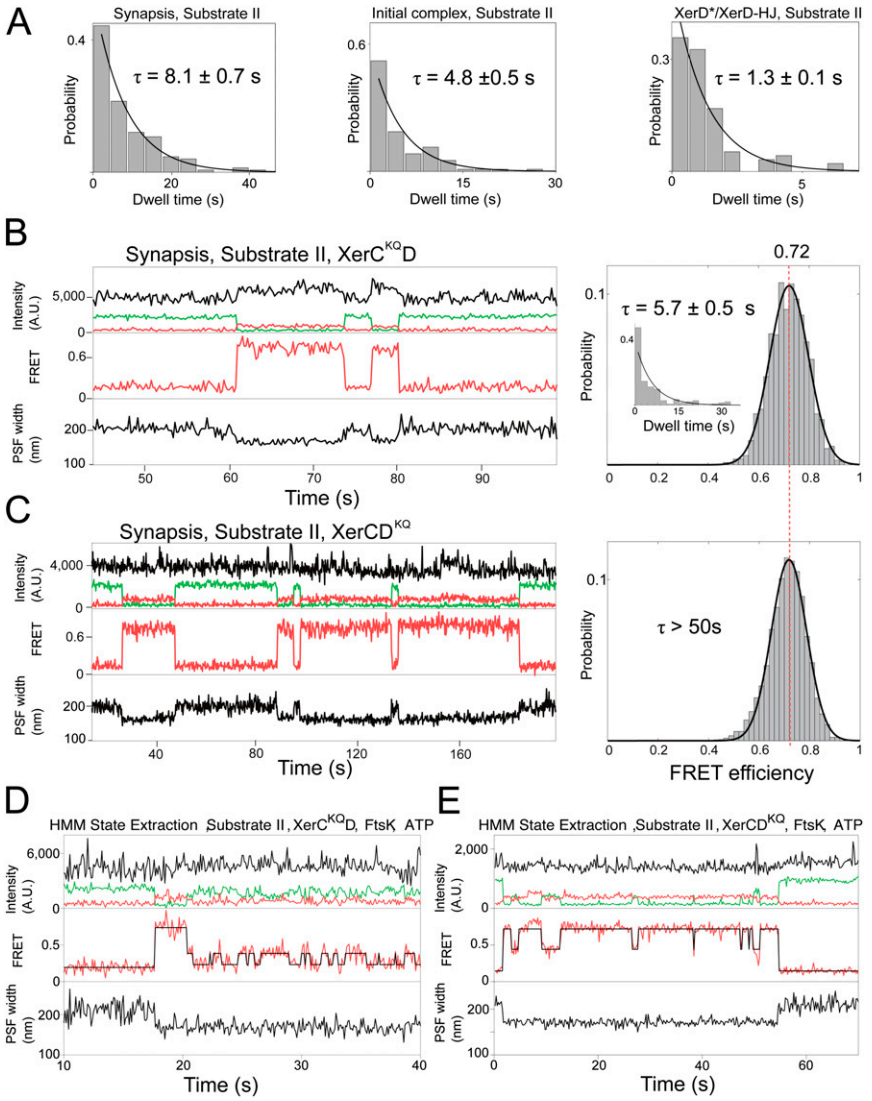

Fig. S3. Substrate II: dwell time analysis, synapsis with XerC^{KQ} and XerD^{KQ} mutants, and HMM for state extraction. (*A*) Dwell time histograms corresponding to lifetimes in Fig. 3 *B* and *C*. The mean lifetime, τ, was obtained from a one parameter exponential fit to the data using MATLAB. (*B* and *C*) In the absence of FtsK–ATP the XerC^{KQ}D and XerCD^{KQ} protein combinations formed complexes similar to the WT XerCD. (*D*) Time trace showing oscillations between XerD–HJ ($E^* \approx 0.37$) and XerC–HJ ($E^* \approx 0.18$), in an experiment using XerD and a catalytic mutant, XerC^{KQ}. The black line in the *Middle* shows the HMM idealized time trace, used to identify transition times for data extraction. (*E*) Time trace showing repeated activation of the synaptic complex formed by XerC and XerD^{KQ}. The black line in the *Middle* shows the HMM idealized time trace, used to identify transition times for data extraction.

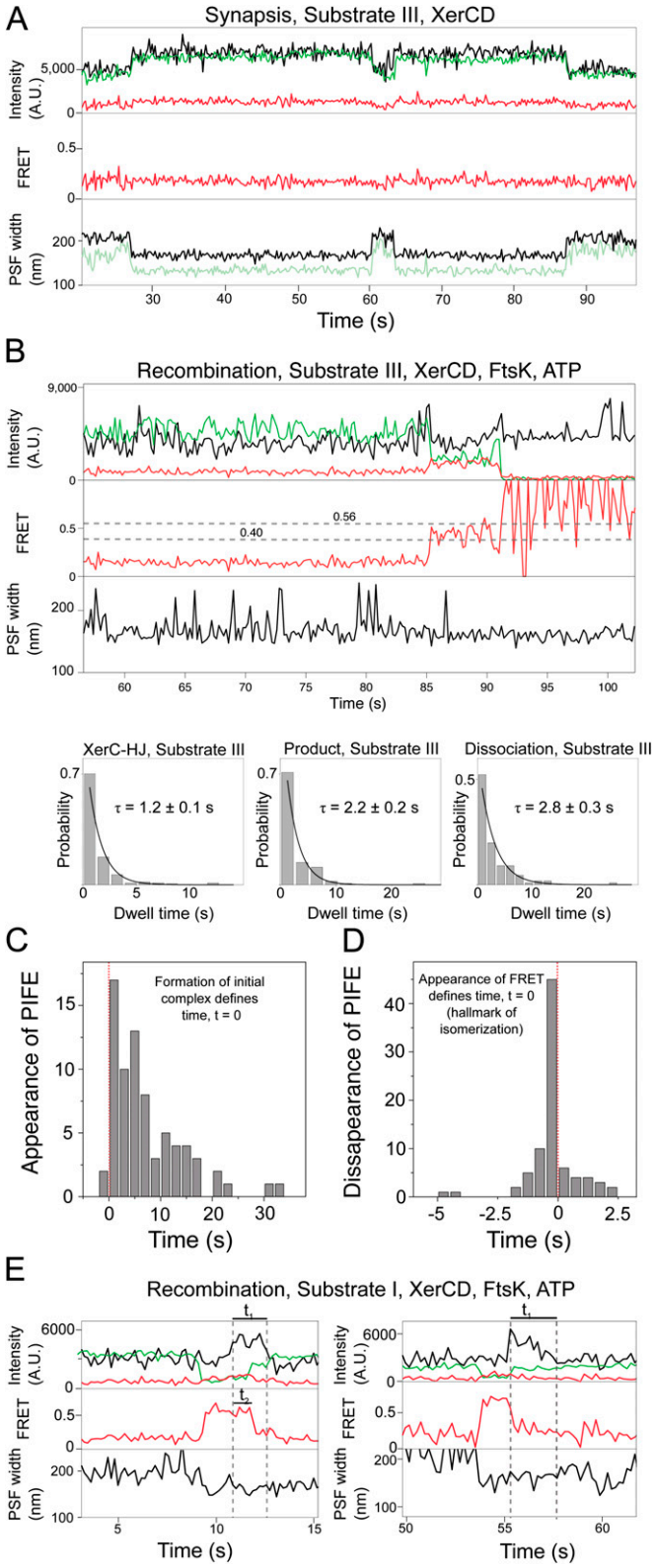

Fig. S4. Substrate III synapsis and dwell time analysis; and protein-induced fluorescence enhancement (PIFE). (*A*) Simultaneous narrowing of donor and acceptor PSFs indicates the formation of synaptic complex in substrate III. Additionally, increases in the intensity of both fluorophores indicate that they are closer to the surface within the evanescent excitation field. Without FtsK and ATP, these complexes reverted back to substrate. (*B*) Example time trace and dwell time histograms for substrate III. The time trace shows transitions between XerC–HJ ($E^* \approx 0.56$) and XerC–P ($E^* \approx 0.40$) before dissociation. All data were fit with single-parameter exponential decays in MATLAB. Occasional interconversions back to XerC–HJ from product DNA were observed. To quantify this rate, the dwells for product transitioning to either dissociation or back to XerC–HJ were fit to recover the total transition rate from product (*Middle*). The dwell time for product to dissociate was also fit (*Right*), and the rate of backward stepping from product to XerC–HJ was determined as the difference between these rates.

Legend continued on following page

(*C*) A histogram of the timing of the appearance of PIFE relative to the timing of formation of the initial synapsis. The distribution indicates that FtsK arrives after the formation of the initial complex, showing that the initial synapsis is the substrate for FtsK activation. (*D*) Disappearances of PIFE relative to HJ isomerization (appearance of FRET). In the majority of events, PIFE ceases just before HJ isomerization. (*E*) PIFE in substrate I. Around 50% of events were accompanied by PIFE in substrate I. Time t_1 corresponds to presence of FtsK near the synaptic complex. Time t_2 is the time needed for the induction of the conformational switch from initial complex to XerD*.

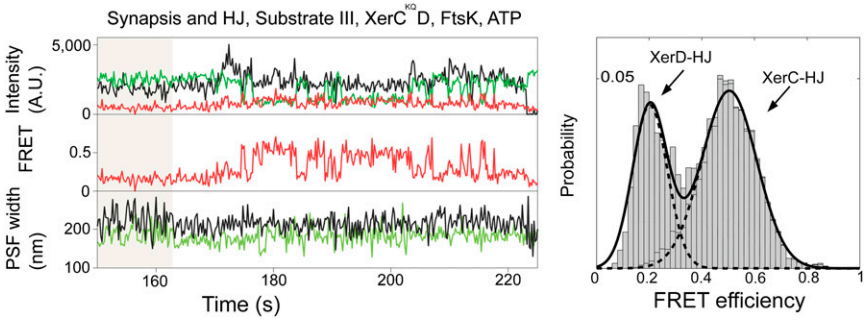

Fig. S5. HJ isomerization in substrate III with XerC^{KQ}, WT XerD, FtsK, and ATP. We see isomerization between a state with $E^* \approx 0.20$ and $E^* \approx 0.50$, and we conclude that $E^* \approx 0.20$ corresponds to XerD–HJ in this substrate. $E^* \approx 0.50$ is lower than the $E^* \approx 0.56$ in the WT complexes, because rapid transitions to XerD–HJ and back within a single frame will skew the distribution toward a lower apparent FRET.

Table S1. Substrate design and predicted interfluorophore distances for possible states in the recombination pathway

	Distance, Å			
Substrate	XerC*–Substrate	XerD*–Substrate	XerC*–Product	XerD*–Product
Substrate I	65	104	107	82
Substrate II	95	85	93	106
Substrate III	107	82	65	104

Distances between fluorophores were inferred from the Cre crystal structure in different nucleoprotein conformations. We performed three radii accessible volume calculations using software described in Sindbert et al. (1) using 2HOI (2). We used the mean fluorophore positions to estimate their separation. The DNA arms were extrapolated and their position relative to the active/inactive cleavage site was calculated by identifying which strand would be next cleaved in any structure and counting the appropriate number of base pairs on the correct strand in the right direction.

1. Sindbert S, et al. (2011) Accurate distance determination of nucleic acids via Förster resonance energy transfer: Implications of dye linker length and rigidity. *J Am Chem Soc* 133(8):2463–2480.
2. Ghosh K, Guo F, Van Duyne GD (2007) Synapsis of loxP sites by Cre recombinase. *J Biol Chem* 282(33):24004–24016.

Table S2. Accurate FRET efficiencies and distances calculated for all identifiable states in all substrates

Population	Accurate FRET	SD	Distance, Å	Uncertainty in distance, Å
Substrate I				
Synapsis	0.68	0.02	58.6	8.8
Recombination	0.70	0.02	57.6	8.7
Substrate II (XerCD)				
Synapsis	0.84	0.01	51.6	7.7
Initial complex	0.82	0.02	53.2	8.0
XerD–HJ	0.39	0.06	73.5	11.4
XerC–HJ/XerC–P	<0.2		>90	
Substrate II ($XerC^{KQ}D$)				
Initial start	0.83	0.01	52.4	7.8
Initial end	0.83	0.01	52.6	7.9
XerD–HJ	0.42	0.05	72.0	10.9
XerC–HJ	<0.2		>90	
Substrate II ($XerCD^{KQ}$)				
Initial complex	0.83	0.01	52.4	7.8
XerD*	0.53	0.03	66.8	10.1
Substrate III				
XerC–HJ	0.72	0.01	56.7	8.6
XerC–P	0.54	0.02	64.7	9.8

Table S3. Parameters for the Gaussian fits to population histograms of E* FRET efficiencies (Figs. 2–4)

Population	Mean	SEM	SD	SEM of SD
Substrate I				
Synapsis	0.549	0.001	0.082	0.001
Recombination	0.571	0.002	0.111	0.002
Substrate II (XerCD)				
Synapsis	0.720	0.002	0.086	0.002
Initial complex	0.717	0.002	0.101	0.002
XerD*	0.372	0.003	0.106	0.003
XerC–HJ/XerC–P	0.174	0.001	0.050	0.001
Substrate II ($XerC^{KQ}D$)				
Initial start	0.727	0.002	0.095	0.002
Initial end	0.718	0.002	0.097	0.002
XerD–HJ	0.366	0.002	0.079	0.002
XerC–HJ	0.217	0.001	0.061	0.001
Substrate II ($XerCD^{KQ}$)				
Initial complex	0.718	0.001	0.081	0.001
XerD*	0.438	0.004	0.094	0.004
Substrate III				
XerC–HJ	0.559	0.003	0.096	0.003
XerC–P	0.400	0.002	0.078	0.002

The data were extracted manually and using HMM and plotted in a histogram, and then fit with a Gaussian using a nonlinear least-squares method in MATLAB. Then mean and SD are from the fitted curve. The SE in the mean (SEM) is defined as the 1σ-confidence interval on the parameter fit.