SUPPLEMENTARY INFORMATION

METHODS

Gel Shift Assay

The electrophoretic mobility shift assay (EMSA) with Ku, LX, and XLF (Figure S2F) was performed with Cy5-labeled 60-bp dsDNA substrates made by annealing the following oligonucleotides: /5Cy5/GTA GGG CTC ATG TTA GAT CTA TCG AGC AAG TGC ATC TGC AGT ACT CAT ATG GAA TTC CCA GCT GAG and /5Phos/CTC AGC TGG GAA TTC CAT ATG AGT ACT GCA GAT GCA CTT GCT CGA TAG ATC TAA CAT GAG, with the latter oligonucleotide varied to be 5' phosphorylated or not. The EMSA performed with LX complex alone (Figure S3B) was performed using a Cy5-labeled 30-bp nicked dsDNA substrate with 5-bp mismatched tails made by annealing the following oligonucleotides: /5Cy5/TTT TTT GCG TGC GTG TGT GAT CGG CCT CCC AAA GTT TTT T, TTT TTA CTT TGG GAG GCC GA, and /5Phos/TCA CAC ACG CAC GCA TTT TT, with the latter oligonucleotide varied to be 5' phosphorylated or not. These substrates were incubated with proteins for 20 min on ice in EMSA buffer (50 mM NaCl, 75 mM KCl, 25 mM Tris pH 8, 13% Glycerol, and 0.015% X100). Samples were run on 4% (Figure S2F) or 6% (Figure S3B) polyacrylamide gels in 0.5x TBE buffer and imaged using a Typhoon (GE biosciences).

Trajectory Selection

We selected smFRET trajectories that demonstrated association/dissociation in our observation interval (<30 seconds) representing transient PECs, and smFRET trajectories that remained associated for our entire observation interval (30 seconds) representing persistent PECs. Since we employed DNA substrates capable of ligation, we predominately chose persistent PECs with at least one FRET transition, with the population that exhibited no transitions making up no more than ~7% of the total trajectories for DNA substrates with 5'P and 3'OH.

Pairing Efficiency

Pairing efficiency measurements are based on two or more independent experiments. Each independent experiment is normalized to a control reaction on each slide (substrate with both 5'P and 3'OH). These are averaged to generate the normalized pairing efficiency and SEM.

Histograms

Histograms are composed of at least 200 molecules for Persistent PECs and 50 molecules for Transient PECs.

Autocorrelation Analysis

Autocorrelation is able to extract the transition frequency in a signal by correlating the signal with itself as a function of time.

The autocorrelation of FRET efficiency was defined as:

$$AC(\tau) = \frac{\langle E_{FRET}(t) E_{FRET}(t+\tau) \rangle_t}{\langle E_{FRET}(t) \rangle_t^{\,2}}$$

Where $AC(\tau)$ is the autocorrelation function of the lagging time τ; $E_{FRET}(t)$ represents the time-trajectory of the obtained FRET efficiency. To calculate the correlation function, we applied the Fourier Transfer algorithm as:

$$AC(\tau) = \frac{1}{\langle E_{FRET}(t) \rangle_t^{\,2}} \mathrm{iFFT}\left(\left| \mathrm{FFT}\left(E_{FRET}(t) \right) \right|^2 \right)$$

Where FFT and iFFT denotes the Fast Fourier Transfer and inverse Fast Fourier Transfer (MathWorks), respectively.

The generated autocorrelation curve can then be fit with either an exponential, multi-exponential, or stretch exponential function to measure the average transition time for the PEC.

Energy calculations

Autocorrelation time constants and dwell times were calculated as previously described (1). The energies for the complexes were obtained from the parametric solution

$$E_{tot} = -RT(\ln(1/[t_{corr}) * a])) \text{ or } E_{tot} = -RT(\ln(1/[t_{dwell}) * a]))$$

Where *a* is the preexponential factor, assumed to be constant for all configurations and *T* = 25°C (room temperature), and t_{corr}/t_{dwell} are the autocorrelation and dwell times calculated. The error associated with the reported energy calculations are based on the errors from the curve fitting of the autocorrelation and dwell times.

Oligonucleotides:

SURF_Bio_iCy5_PO4: /5Phos/CGT G/iCy5/AG AGG AGA CAG AGT GCG GGC GAA CAA CAT AAA TCG TAC CCT CGT ATG TAT CGT ATG GCT CAT GCT TAT CAG ATG CT/3Bio/

SURF_Bio_iCy5: CGT G/iCy5/AG AGG AGA CAG AGT GCG GGC GAA CAA CAT AAA TCG TAC CCT CGT ATG TAT CGT ATG GCT CAT GCT TAT CAG ATG CT/3Bio/

SURF_ddC: AGC ATC TGA TAA GCA TGA GCC ATA CGA TAC ATA CGA GGG TAC GAT TTA TGT TGT TCG CCC GCA CTC TGT CTC CTC TCA CGG CA/3ddC/

SURF_4nt comp: AGC ATC TGA TAA GCA TGA GCC ATA CGA TAC ATA CGA GGG TAC GAT TTA TGT TGT TCG CCC GCA CTC TGT CTC CTC TCA CGG CAC

SLTN_iCy3_PO4: /5Phos/TCT G/iCy3/AT AAG CAT GAG CCA TAC GAT ACA TAC GAG GGT ACG ATT TAT GTT GTT CGC CCG CAC TCT GTC TCC TCT CAC GTT TTC GTG AGA GGA GAC AGA GTG C

SLTN_iCy3: TCT G/iCy3/AT AAG CAT GAG CCA TAC GAT ACA TAC GAG GGT ACG ATT TAT GTT GTT CGC CCG CAC TCT GTC TCC TCT CAC GTT TTC GTG AGA GGA GAC AGA GTG C

SLTN_ddC: GGG CGA ACA ACA TAA ATC GTA CCC TCG TAT GTA TCG TAT GGC TCA TGC TTA TCA GAG TG/3ddC/

SLTN_4nt comp: GGG CGA ACA ACA TAA ATC GTA CCC TCG TAT GTA TCG TAT GGC TCA TGC TTA TCA GAG TGC

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Protein purity. A gel demonstrating the purity of the various proteins used for all experiments. 100 or 200 ng of wild-type LX, K273R LX, Ku70/80, and XLF was separated on a 4-12% polyacrylamide gel.

Figure S2. Mixed End Chemistry and PEC Stability. A. Representative images of donor and acceptor channels showing high (5'P/5'P), medium (SCR7), and low (5'OH/5'OH) capture efficiency from our smFRET assay. B. Diagram of dsDNA ends for blunt and 5' overhang substrates. C. Pairing efficiency of 3' overhang, blunt, and 5' overhang ends. D. Diagram of dsDNA ends with a combination of missing 5'P and 3'OH. E. Paring efficiency of DNA ends from (D) show that formation of PECs is more sensitive to ends without 5'P than 3'OH. F. EMSA for NHEJ proteins on substrates with and without 5'P. Cy5 labeled 60-bp DNA duplex (10 nM) was present in all reactions, either with 5' phosphate or not as noted above the image. Also as noted, LX complex with wild-type or K273R L4 was added at 23 nM, XLF at 40 nM, and Ku at 2 nM. The inferred protein composition of each shifted species is noted at right.

Figure S3. Ligase Inhibition and Activity. A. Pairing efficiency of NHEJ reactions when inhibited with 10 μM SCR7 or 5 μM L189 ligase inhibitors. B. Cy5 labeled nicked 30-bp DNA duplex (5 nM) was present in all reactions, either with 5' phosphate at strand break or no 5' phosphate as noted. Wild-type LX complex was added at 80 nM, 160 nM or 320 nM as noted in wedge above figure, and the LX complex with a K273R substitution in ligase IV added at 320 nM. C. Comparison of NHEJ reaction compared with XRCC4/XLF pairing on various modified end substrates.

Figure S4. End Chemistry Dynamics and Stability. A. FRET histograms for ends without both 5'P and 3'OH. We see that when these ends are a mixed distribution of the histograms of ends without either 5'P or 3'OH. B. FRET histograms for reactions where L4 activity has been disrupted. Interestingly, we see that in the K273R L4 reactions there are few dynamics, suggesting that a substantial portion of the observed dynamics come from the L4 catalytic cycle. C. Comparison of wtL4 and K273R L4 reactions on various modified end substrates.

Figure S5. Energetic Stabilities of Persistent PECs. A. Representative trajectory for end lacking 5'P. B. Calculated energies from autocorrelation times show that ends without both 5'P and 3'OH are relatively stable in the PEC. C. Calculated energies from autocorrelation times show that disruption of L4 catalysis lowers the stability of DNA ends in the PEC. E. Stability of Transient PECs where calculated from the dwell times of substrates modulated for 5'P and 3'OH. F. Stability of Transient PECs with various 5'P and 3'OH are omitted.

Figure S6. Energetic Stabilities of Transient PECs. Calculated energies from dwell times for ends without both 5'P and 3'OH.

1. Rothenberg, E., Grimme, J.M., Spies, M. and Ha, T. (2008) Human Rad52-mediated homology search and annealing occurs by continuous interactions between overlapping nucleoprotein complexes. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 20274-20279.

Figure S1

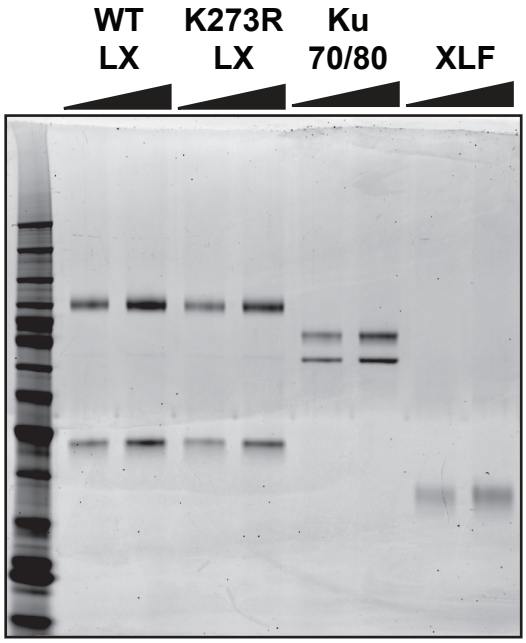
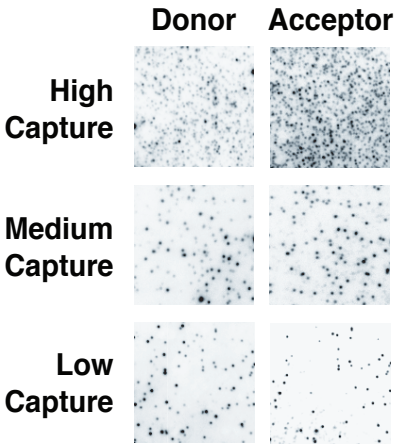
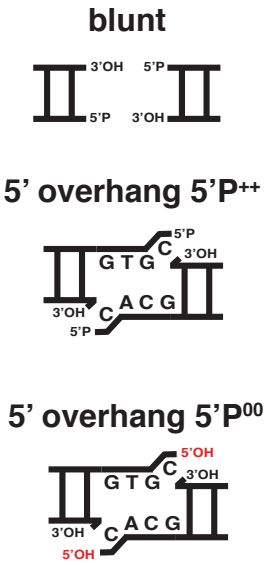


Figure S2

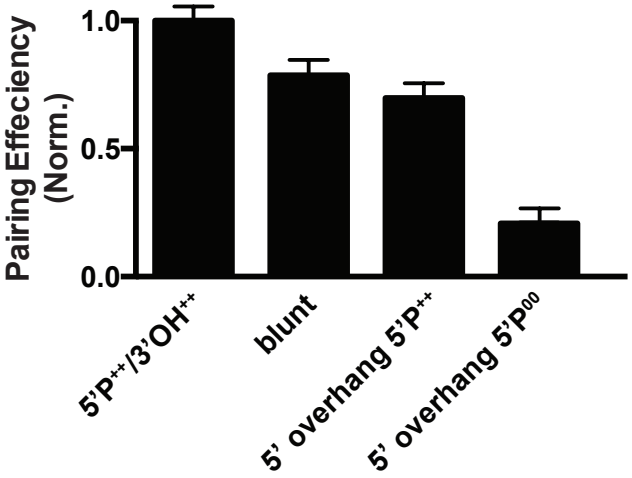
A



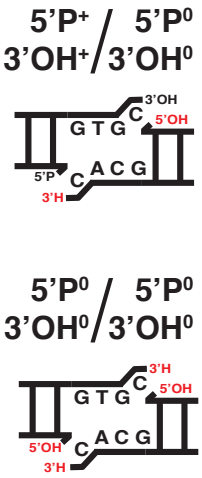
B



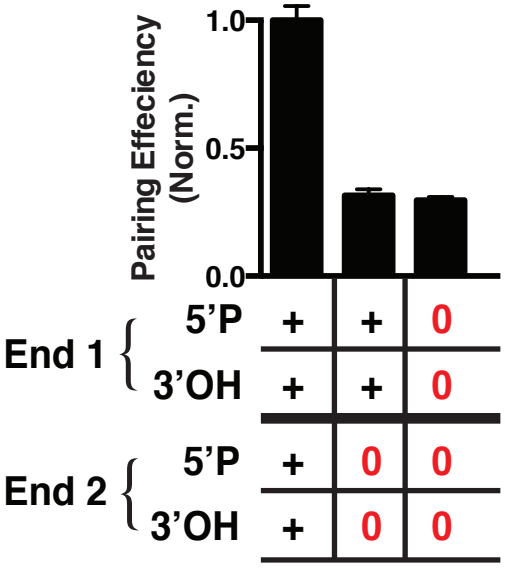
C



D



E



F

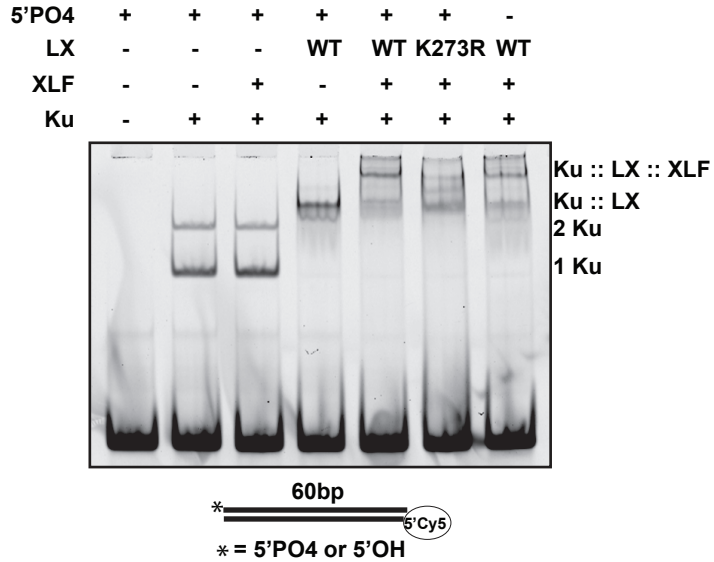
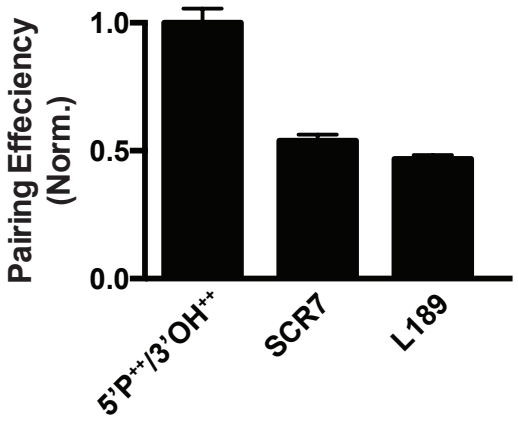
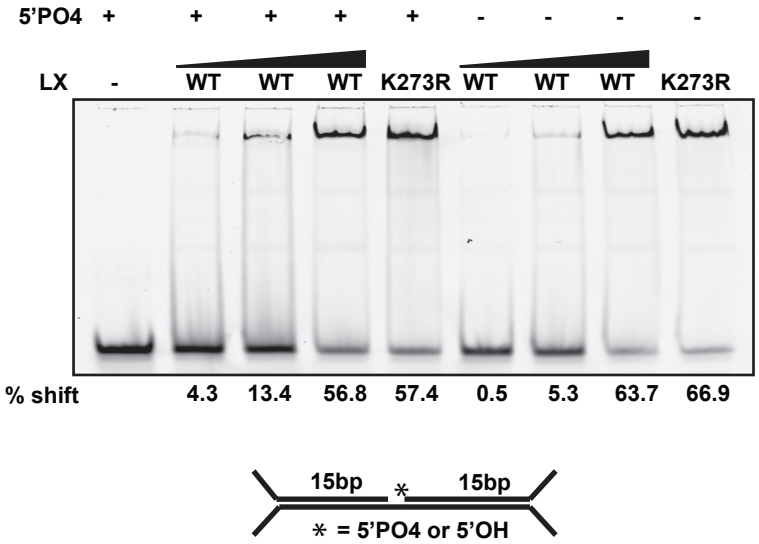


Figure S3

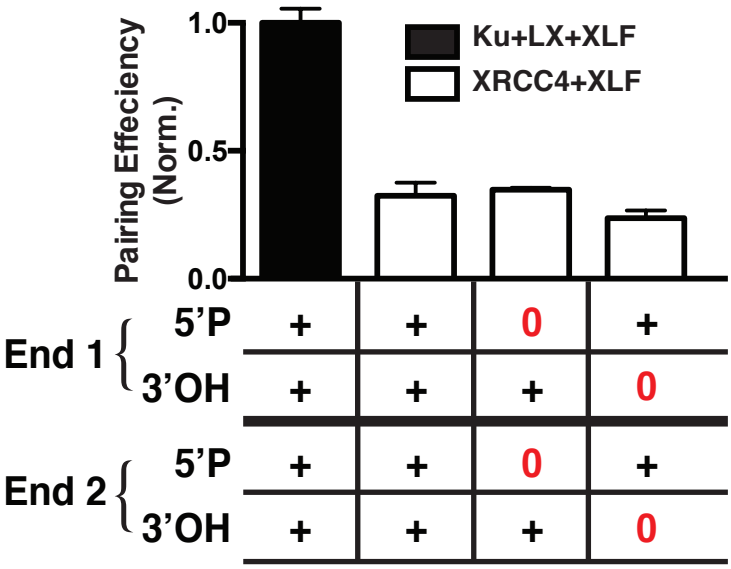
A



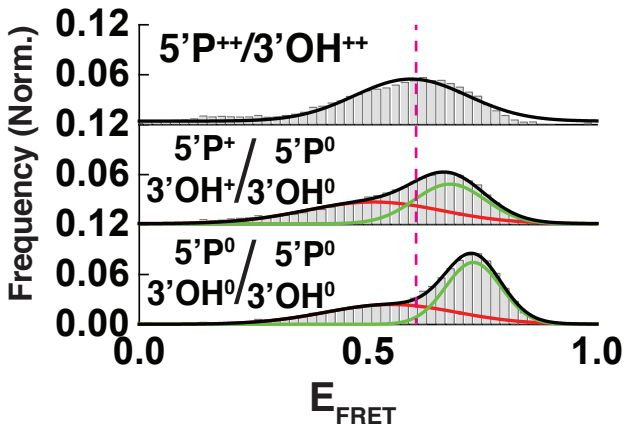
B



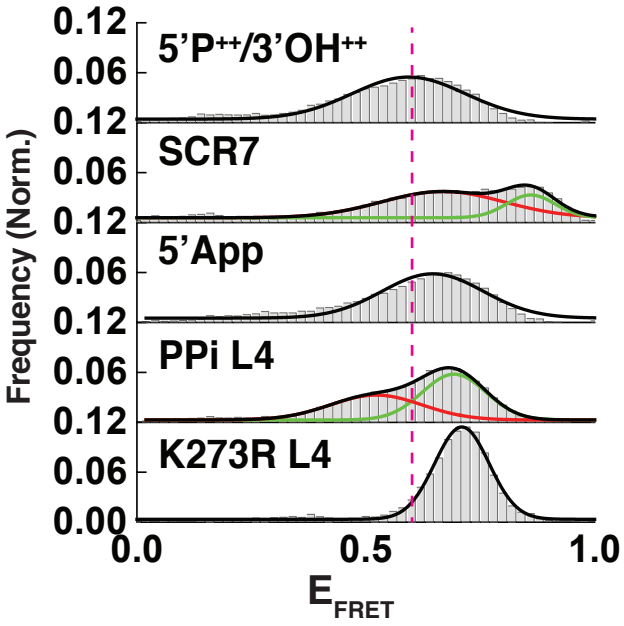
C



A


Frequency (Norm.)
0.12
0.06
0.12
0.06
0.12
0.06
0.00
5'P++/3'OH++
5'P+ / 5'P0
3'OH+/3'OH0
5'P0 / 5'P0
3'OH0/3'OH0
0.0
0.5
1.0
EFRET

B


Frequency (Norm.)
0.12
0.06
0.12
0.06
0.12
0.06
0.12
0.06
0.00
5'P++/3'OH++
SCR7
5'App
PPi L4
K273R L4
0.0
0.5
1.0
EFRET

C

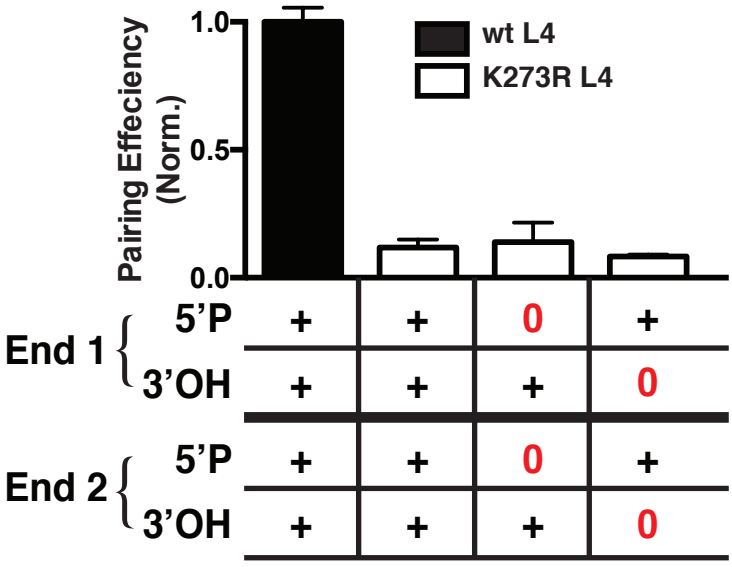
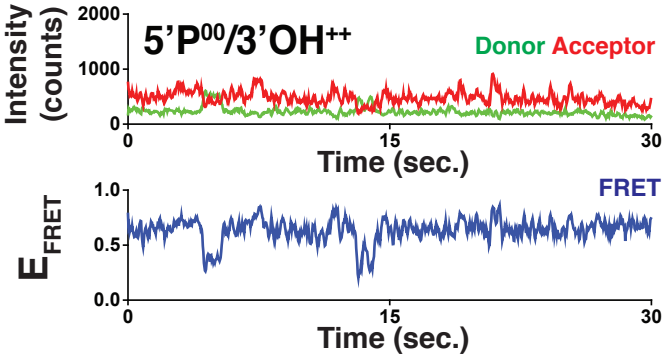
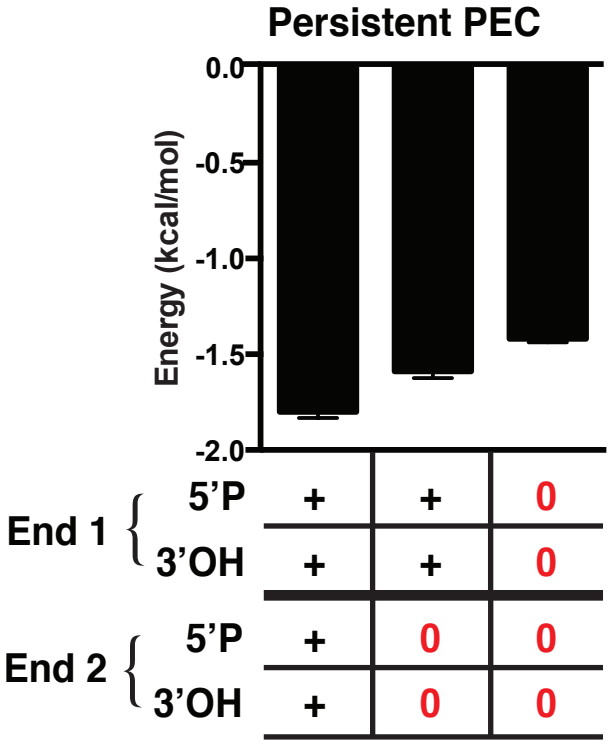

Pairing Effeciency (Norm.)
1.0
0.5
0.0
wt L4
K273R L4
End 1
5'P + + 0 +
3'OH + + + 0
End 2
5'P + + 0 +
3'OH + + + 0

Figure S5

A



B



C

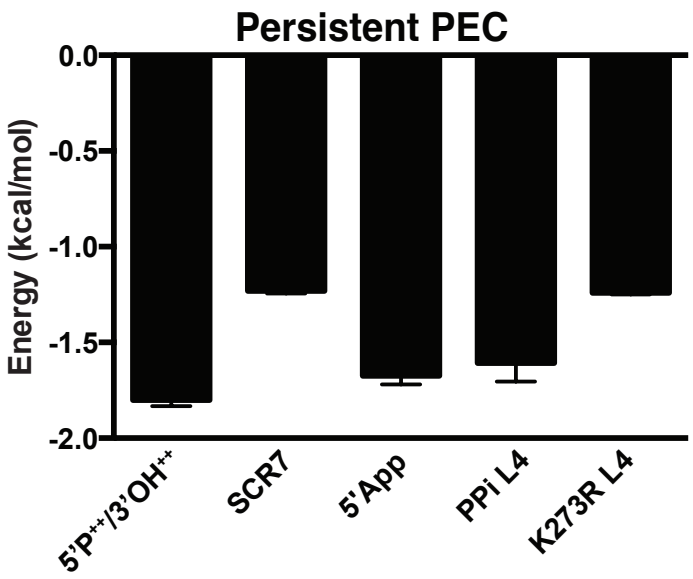
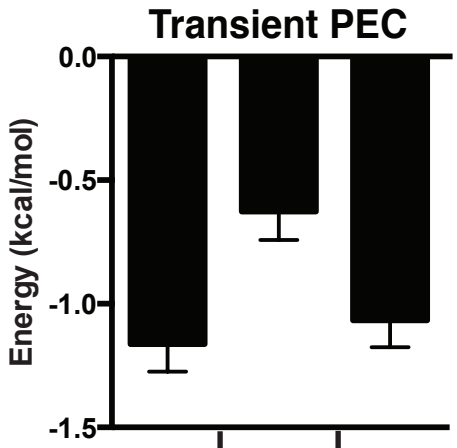


Figure S6



End 1	5'P	+	+	0
	3'OH	+	+	0
End 2	5'P	+	0	0
	3'OH	+	0	0