Simulative and experimental characterization of a pH-dependent clamp-like DNA triple-helix nanoswitch

Federico Iacovelli,†# Andrea Idili,‡# Alessandro Benincasa,† Davide Mariottini,‡ Alessio Ottaviani,† Mattia Falconi,† Francesco Ricci‡ and Alessandro Desideri†**

† Department of Biology, University of Rome, Tor Vergata, Rome, Italy

‡ Department of Chemistry, University of Rome, Tor Vergata, Rome, Italy

*Corresponding authors, email: *E-mail: francesco.ricci@uniroma2.it; desideri@uniroma2.it

This Supplementary Information section includes:

1) Urea fitting equation

Measuring binding free energies between triplex clamp-switch and its DNA target

In this work we use urea titration curves to estimate the effect of the pH on the stability of the DNA triple helix formed by the binding between the clamp-switch and its 6-base DNA target by analyzing and fitting the urea titration curves (Figure 7) and by using a two-state denaturation model.[1-3]

Below we give a brief description of the rationale behind this model.

The equilibrium between the clamp-switch (P) and its 6-base DNA target (T) can be described as:

$$PT \underset{}{\overset{\text{UREA}}{\rightleftharpoons}} P + T$$

This equilibrium gradually shifts towards the unfolded state with increasing urea concentration. At any given urea concentration, we can express the fluorescence of the sample (F) as the sum of the fluorescence of the signaling clamp-switch in the folded (bound state) and unfolded (not-bound state) states:

$$F = F_F \cdot \frac{[PT]}{P_0} + F_{UN} \cdot \frac{[P]}{P_0} \qquad (1)$$

Where F_F and F_U represent the fluorescence intensities of the signaling clamp-switch in the folded and unfolded states, respectively. Because P_0 represents the clamp-switch concentration used during the experiment and [PT] and [P] the concentrations of the clamp-switch in the folded and unfolded state, respectively we have:

$$P_0 = [PT] + [P] \qquad (2)$$

At any specific urea concentration we can thus describe the binding equilibrium constant (K_B^{urea}) and the folding free energy ($\Delta G°_B{}^{urea}$) of the signaling clamp-switch:

$$K_B^{urea} = \frac{[P]\cdot[T]}{[PT]} \qquad (3)$$

$$\Delta G_B^{\circ\, urea} = -RT \cdot \ln K_B^{urea} \qquad (4)$$

When the total concentration of 6-base DNA target (T_{tot}) is much greater than the total concentration of the clamp-switch (P_0), $[T_{tot}]$ can be used instead of [T] in Eq. (3). The ratios $[PT]/P_0$ and $[P]/P_0$ representing the fraction of the clamp-switch in the folded and unfolded state can thus be expressed as a function of K_B^{urea} as follow:

$$\frac{[PT]}{P_0} = \frac{[T_{tot}]}{K_B^{urea}+[T_{tot}]} \qquad (5)$$

$$\frac{[P]}{P_0} = \frac{K_B^{urea}}{K_B^{urea}+[T_{tot}]}$$

Substituting these two equations into Eq. (1), we can express the fluorescence of the clamp-switch solution in equilibrium with its target as:

$$F = F_F \cdot \frac{[T_{tot}]}{K_B^{urea}+[T_{tot}]} + F_{UN} \cdot \frac{K_B^{urea}}{K_B^{urea}+[T_{tot}]} \qquad (6)$$

The fluorescence intensities of the clamp-switch in the folded (F_F) and unfolded (F_{UN}) states vary linearly as a function of the urea concentration[2,3] ($[U]$) according to:

$$F_F = F_F^{\circ} + \sigma_F \cdot [U] \qquad F_{UN} = F_{UN}^{\circ} + \sigma_{UN} \cdot [U] \qquad (7)$$

Where $F°_F$ and $F°_{UN}$ represent the fluorescence signal in absence of urea for the folded and unfolded state, respectively. Similarly, σ_F and σ_{UN} represent the dependence of the fluorescence signal of the folded and unfolded states, respectively, on urea concentration.

Equation (6) can be rearranged and combined with Equation (4) to give:

$$K_B^{urea} = \frac{[T_{tot}]\cdot(F_F - F)}{(F - F_{UN})} = e^{-\left(\Delta G_B^{\circ urea}/RT\right)} \qquad (8)$$

The binding free energy between the clamp-switch and its 6-base DNA target ($\Delta G°_B(H_2O)$) is also known[2,3] to vary linearly with urea concentration as shown in Eq.(9):

$$\Delta G_B^{\circ urea} = \Delta G_B^{\circ}(H_2O) - m\cdot[U] \qquad (9)$$

Where $\Delta G°_B(H_2O)$ is the binding free energy in absence of urea, and *m* represents the effect of urea on $\Delta G°_B{}^{urea}$. Combining Eq. (8-9) and Eq. (7) we obtain Eq. (10), which enables us to extrapolate the binding free energy between the DNA strands directly from the fluorescence data recorded at different urea concentrations (Figure 7) using $F°_F$, $F°_{UN}$, σ_F, σ_U, *m*, $\Delta G°_B(H_2O)$ as fitting parameters and $[T_{tot}]$ as constant value.

$$F = \frac{[T_{tot}]\cdot\left(F_F^{\circ} + \sigma_F[U]\right) + \left(F_{UN}^{\circ} + \sigma_{UN}[U]\right)\cdot e^{-\frac{\left(\Delta G_B^{\circ}(H_2O) - m\cdot[U]\right)}{RT}}}{\left([T_{tot}] + e^{-\frac{\left(\Delta G_B^{\circ}(H_2O) - m\cdot[U]\right)}{RT}}\right)} \qquad (10)$$

We fitted the urea denaturation curves of the clamp-switch/DNA target complex using Eq. (10), and the modeling enables to estimate the binding free energy of the triple helix. Depending on the denaturation curve profile we fixed either σ_F or σ_{UN} to improve the fitting performance and to

provide a smaller precision on the $\Delta G°_{B}(H_2O)$ values (see urea titration curves section of Experimental section).

Because at both pH 5.0 and 8.0 the duplex state is stable, the urea titration curves obtained with the clamp-switch and its 12-base DNA target (Figure S10) were fitted using a two-state dissociation model.[1-3]

2) Supporting figures and tables

Nucleotide 1	Nucleotide 2	Percentage of existence
C-4-CLAMP	G-21-CLAMP	8.16%
C-30-CLAMP	T-33-CLAMP	11.85%
T-19-CLAMP	C-24-CLAMP	15.09%
T-13-CLAMP	T-15-CLAMP	18.61%
T-7-CLAMP	C-30-CLAMP	19.10%
A-41-CLAMP	A-43-CLAMP	23.78%
T-5-CLAMP	C-37-CLAMP	26.50%
C-30-CLAMP	T-32-CLAMP	27.34%
C-2-CLAMP	T-5-CLAMP	29.54%
C-37-CLAMP	G-38-TARGET	31.29%
C-2-CLAMP	C-4-CLAMP	41.32%
T-20-CLAMP	C-22-CLAMP	45.60%
T-10-CLAMP	C-29-CLAMP	46.22%
T-15-CLAMP	G-41-TARGET	59.00%
T-20-CLAMP	T-25-CLAMP	59.56%
T-18-CLAMP	C-26-CLAMP	61.50%
C-12-CLAMP	T-27-CLAMP	62.82%
T-6-CLAMP	G-41-TARGET	66.84%
G-21-CLAMP	C-24-CLAMP	67.94%
T-11-CLAMP	T-28-CLAMP	78.31%
C-1-CLAMP	T-3-CLAMP	82.76%
T-33-CLAMP	C-30-CLAMP	85.34%
C-34-CLAMP	T-35-CLAMP	94.59%

Table S1. List of unspecific stacking interactions detected for the 25-base loop clamp-switch. For each nucleotide couple the percentage of existence, calculated over the last 100 ns of simulation time in which the unfolded conformations can be observed at pH 8.0 (see Figure 6), has been reported. The notation indicates the base (G, C, T or A) and its location in the clamp or target sequence (starting from the 5'-end).

Nucleotide 1	Nucleotide 2	Percentage of existence
C-14-CLAMP	A-22-TARGET	6.62%
T-5-CLAMP	T-13-CLAMP	22.83%
T-6-CLAMP	A-23-TARGET	31.11%
C-1-CLAMP	A-22-TARGET	38.80%
T-3-CLAMP	A-23-TARGET	41.71%
C-14-CLAMP	A-23-TARGET	50.68%
T-8-CLAMP	T-10-CLAMP	58.91%
T-10-CLAMP	T-12-CLAMP	87.61%

Table S2. List of unspecific stacking interactions detected for the 5-base loop clamp-switch. For each nucleotide couple the percentage of existence, calculated over the last 100 ns of simulation time in which the unfolded conformations can be observed at pH 8.0 (see Figure S5), has been reported. The notation indicates the base (G, C, T or A) and its location in the clamp or target sequence (starting from the 5'-end).

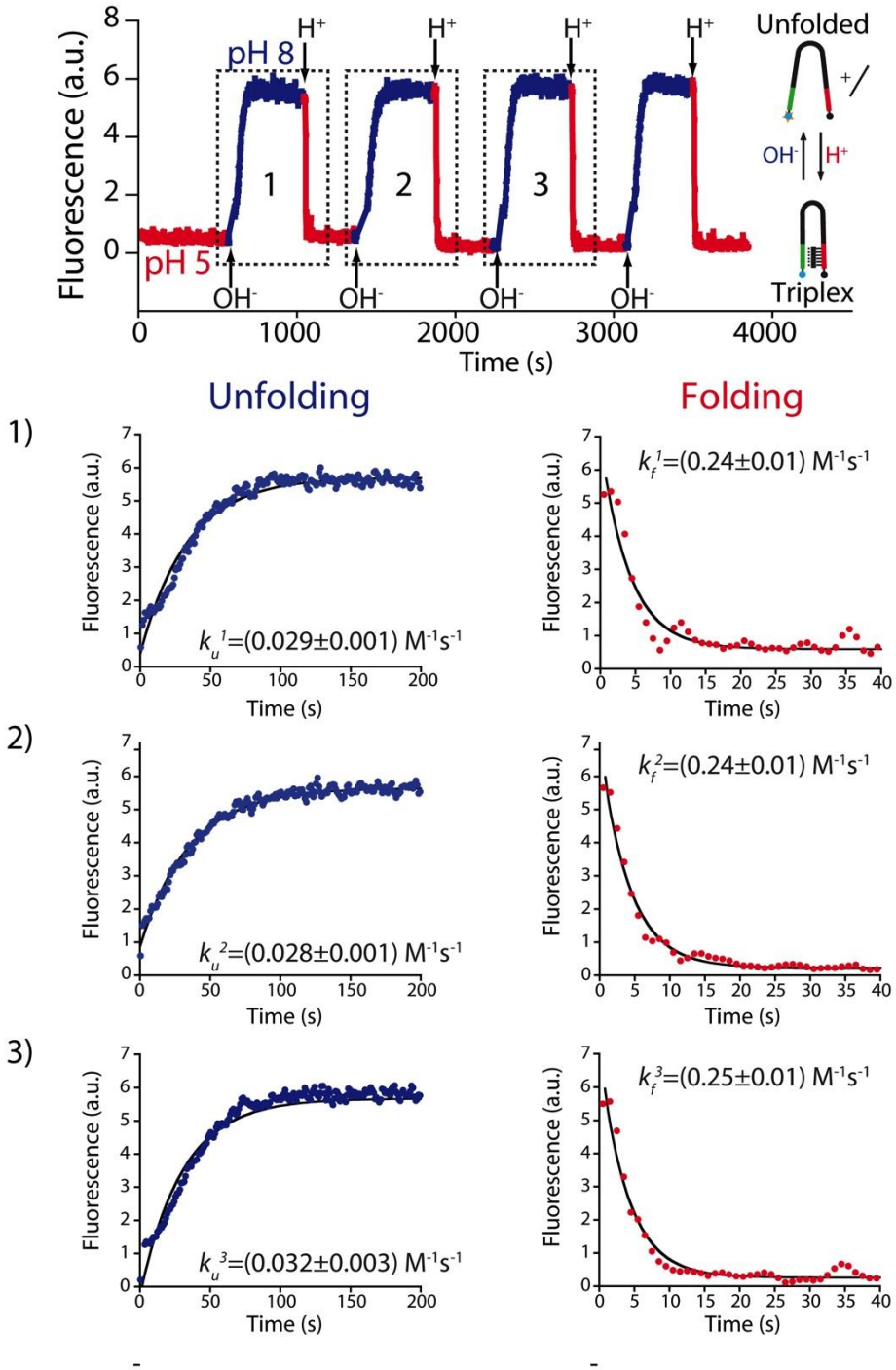

Figure S1: The folding rate of our clamp-switch is faster compared to the unfolding rate. Here we have analyzed the three pH cycles shown in Figure 2C and found that the folding rate is significantly faster than the unfolding rate. This is probably due to non-specific interactions between the target and the 25-base loop that leads to a slower unfolding process. Here unfolding and folding were obtained by using a solution containing the clamp-switch (10 nM) with the 6-base

DNA target (1 μM), and cyclically changing the pH of the solution by adding small aliquots of 3 M NaOH or HCl in 40 mM Tris buffer, 12.6 mM $MgCl_2$ at 25°C (see also Figure 2C).

450-500	400-450	300-400	0-300	pH 8	0-300	300-400	400-450	450-500
-	26%	25%	44%	T — A ● T	69%	31%	31%	2%
-	-	-	72%	T — A ● T	67%	-	28%	2%
-	35%	100%	100%	C — G ● C	-	-	-	-
-	34%	67%	100%	T — A ● T	-	-	-	-
-	1%	60%	98%	C — G ● C	1%	-	-	-
70%	61%	57%	33%	C — G ● C	1%	-	-	-
				3' 5' 5'				

450-500	400-450	300-400	0-300	pH 5	0-300	300-400	400-450	450-500
100%	100%	100%	63%	T — A ● T	95%	98%	98%	98%
100%	100%	100%	99%	T — A ● T	99%	100%	100%	99%
100%	100%	100%	100%	C — G ● C	100%	100%	100%	100%
100%	100%	100%	100%	T — A ● T	100%	100%	100%	100%
100%	100%	100%	100%	C — G ● C	100%	100%	100%	94%
97%	100%	100%	100%	C — G ● C	1%	2%	1%	3%
				3' 5' 5'				

Figure S2: Analysis of the hydrogen bonds persistence evaluated at pH 8.0 (upper panel) and pH 5.0 (lower panel) within the double helix (red and black sequences) and between the double helix and the triplex-forming strand (green sequences). For each time window the calculated percentages of persistence are reported.

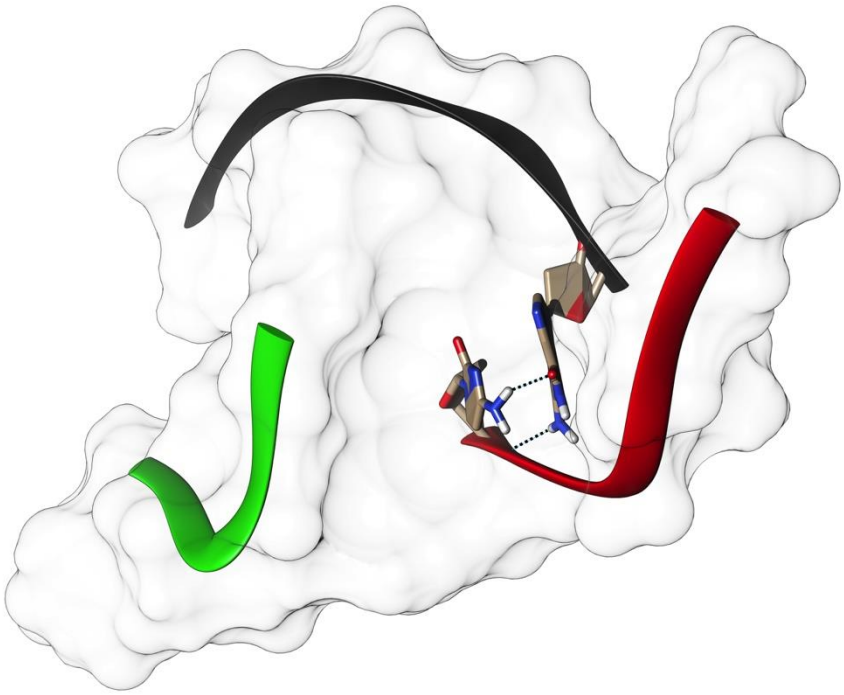

Figure S3: Ribbon view of the triple helix switch surrounded by molecular surface. The different ribbon colors indicate the different strands forming the triple helix. The stick models indicate the 5' guanine of the target oligo (black ribbon) and its base pair partner of the switch oligo (red ribbon) that engage the unusual stacking and hydrogen bond interactions.

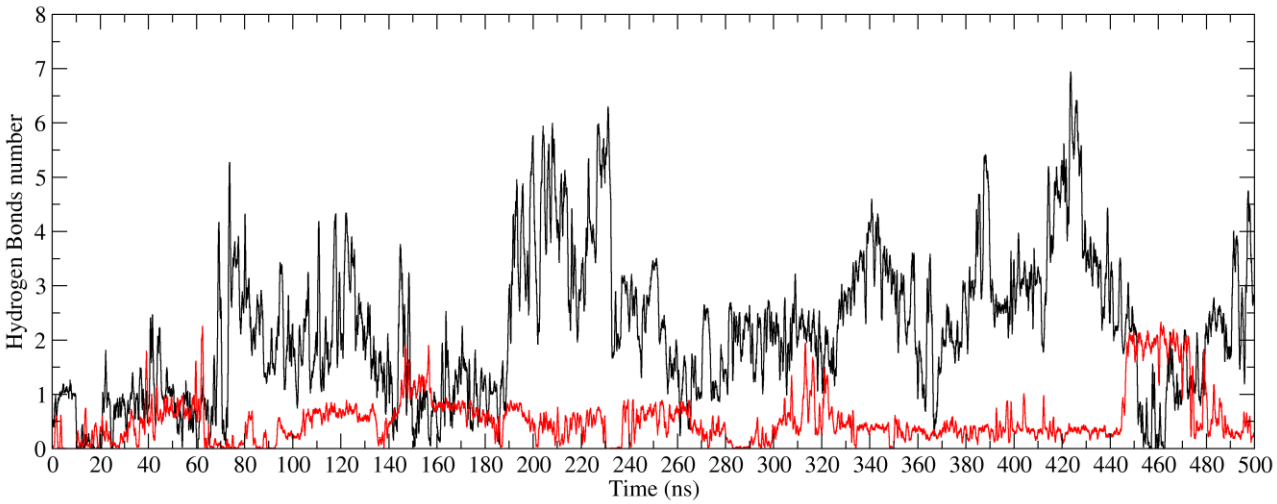

Figure S4. Evolution of the hydrogen bonds established between the loop and the target/switch complex at pH 8.0 for the 25-base loop (black line) and the 5-base loop (red line) clamp-switches, respectively. The increased number of non-specific hydrogen bond interactions for the 25-base loop switch compared to the 5-base loop one could explain the slower unfolding kinetic experimentally observed with the 25-base loop switch.

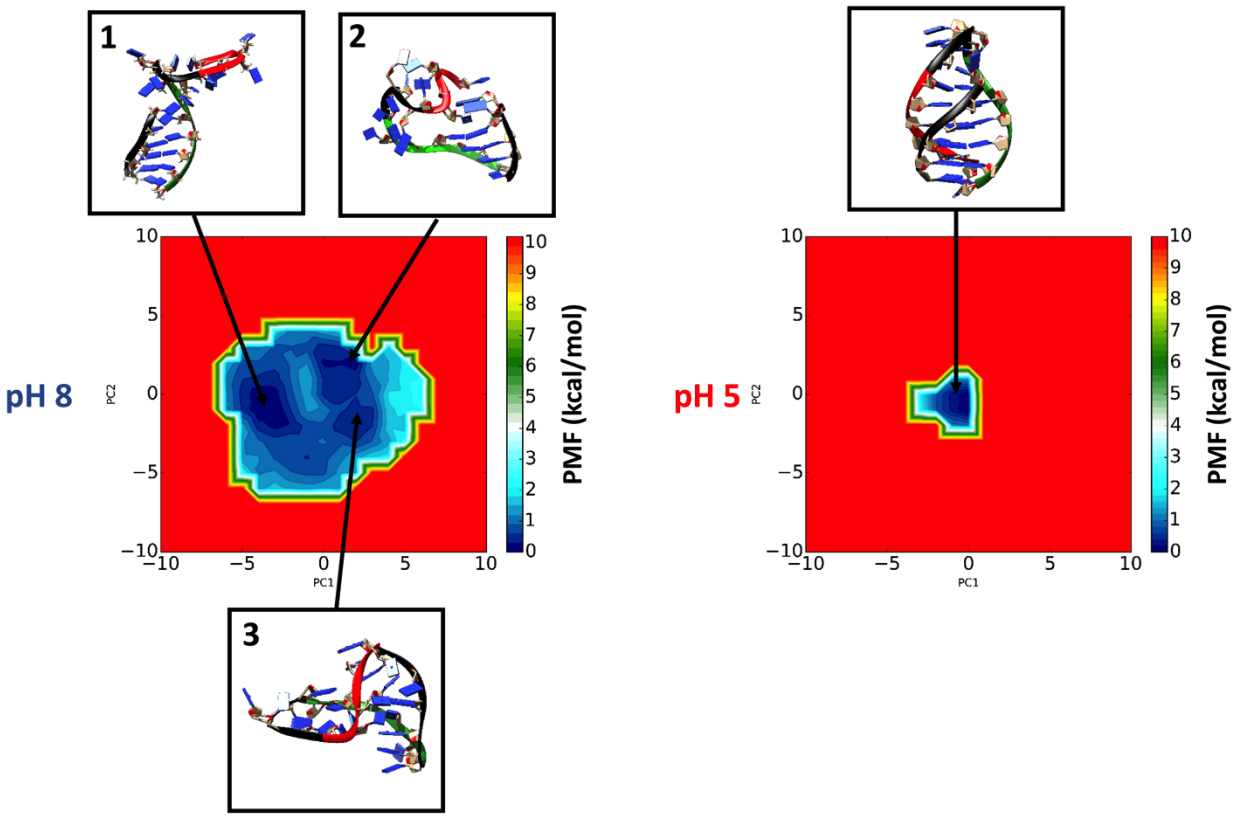

Figure S5: Free energy principal component projection of the 5-base loop system with the 6-base target simulated at pH 8.0 (left) and pH 5.0 (right). 1, 2 and 3 identify the representative three-dimensional structures corresponding to the low energy long-lived conformations. These conformations are characterized by the absence of strong interactions between the loop and the triplex region.

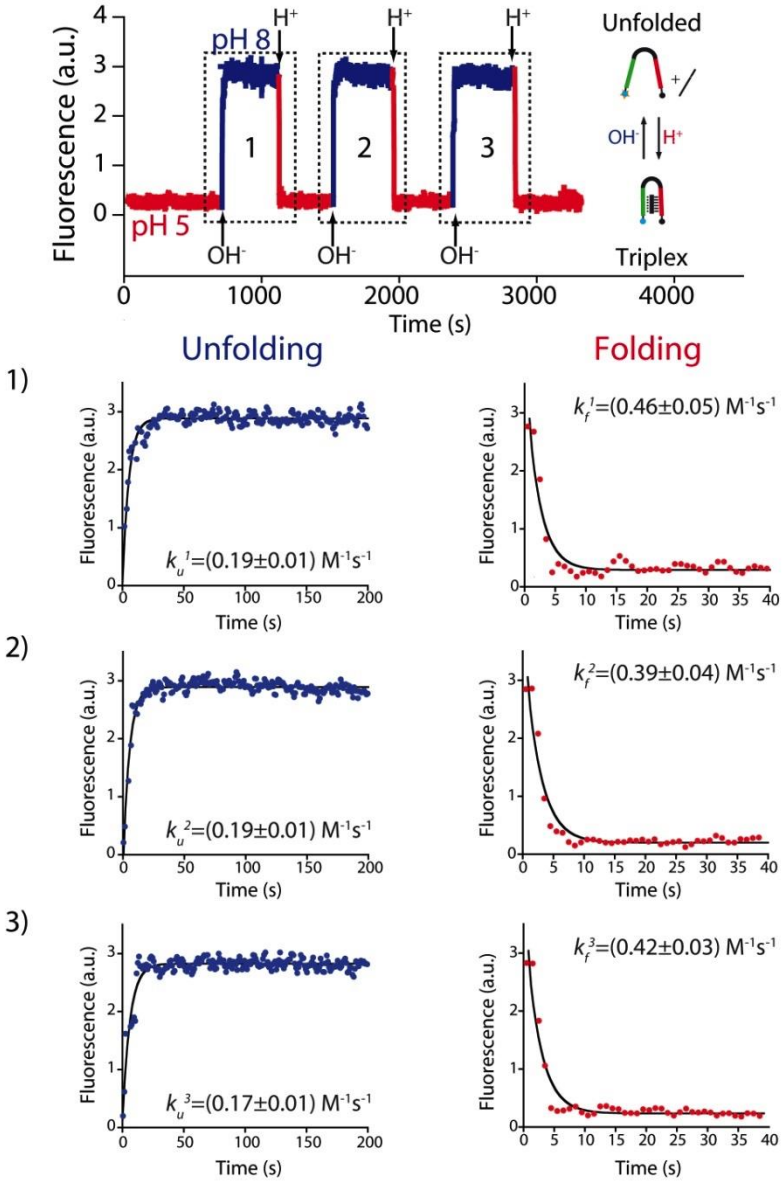

Figure S6: The unfolding rate of the clamp-switch containing only 5 bases in the loop and recognizing the same 6-base target of the 25-base loop clamp-switch (see Figure 2) is similar to the folding rate. This is probably due to the fact that in this clamp-switch non-specific interactions between the 5-base loop and the target are significantly reduced (see Table S1 and S2). Here unfolding and folding were obtained by using a solution containing the 5-base loop clamp-switch (10 nM) with the 6-base DNA target (1 μM), and cyclically changing the pH of the solution by

adding small aliquots of 3 M NaOH or HCl in 40 mM Tris buffer, 12.6 mM $MgCl_2$ at 25°C (see also Figure 2C).

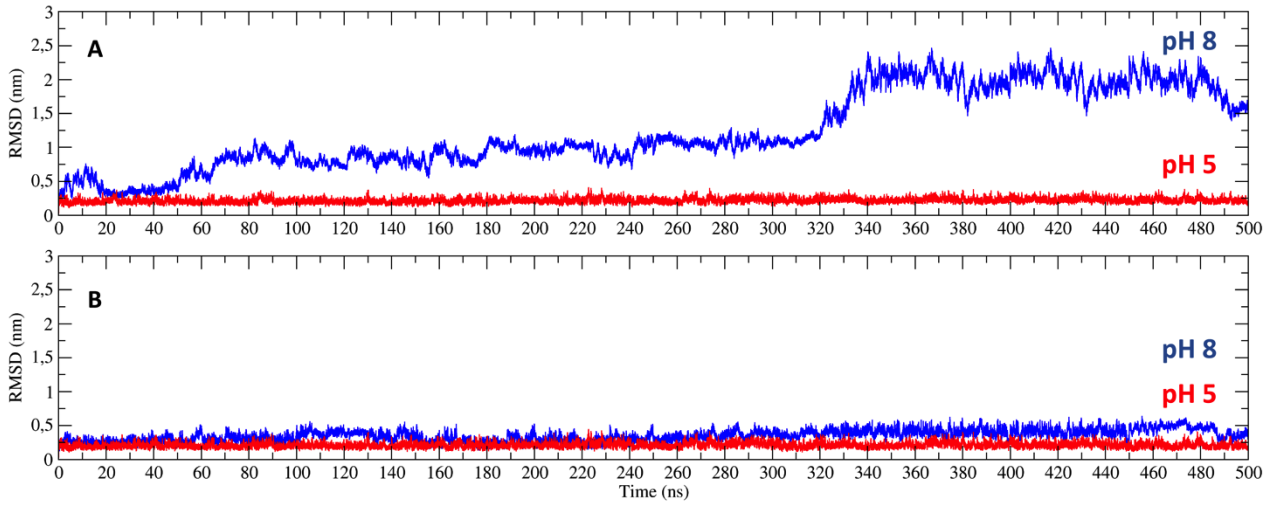

Figure S7: Time dependent evolution of RMSD calculated for the clamp-switch designed with a 12-base target. (A) RMSD of triple helix region (upper panel) and of the double helix region (lower panel) at pH 5.0 (red line) and 8.0 (blue line), respectively. The RMSD values indicate the presence of a stable double helix at both pH 8.0 and 5.0, while the triplex region is fully disrupted at pH 8.0.

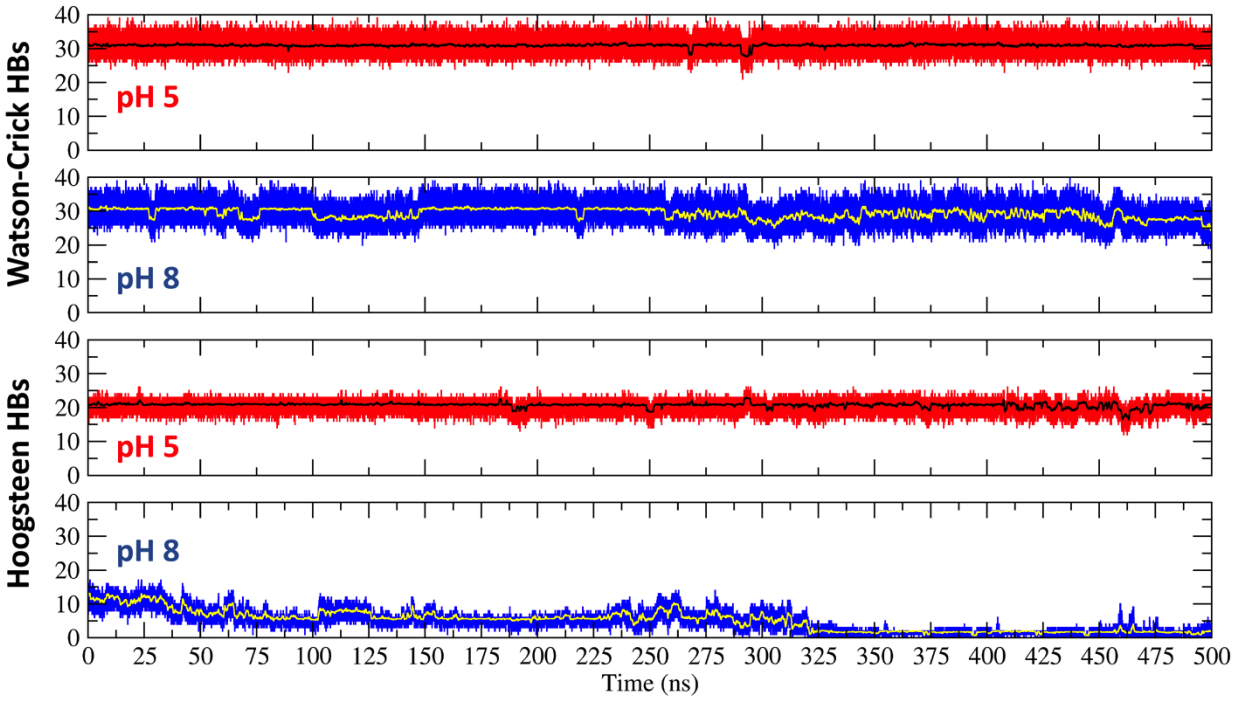

Figure S8: Evolution of the hydrogen bonds established within the double helix (Watson-Crick panel) and between the double helix and the triplex-forming strand (Hoogsteen panel) at pH 5.0 (red line) and 8.0 (blue line) of the clamp-switch designed with a 12-base target. The number of hydrogen bonds, averaged every 100 ps and highlighted by the black and yellow lines, are in agreement with the RMSD results.

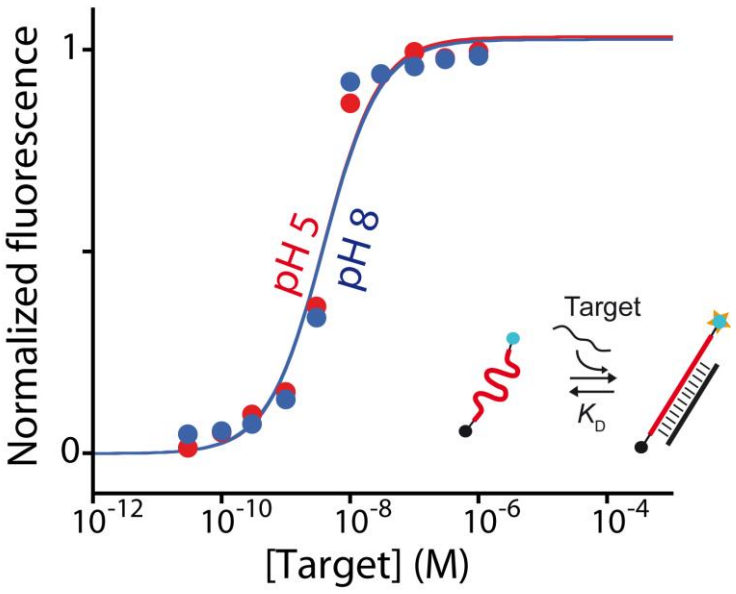

Figure S9: Binding curve experiments using a control switch that recognizes the same 12-base target of the clamp-switch but lacks the triplex-forming portion. The experiment suggests that the duplex state is equally stable at both pH 5.0 and pH 8.0.

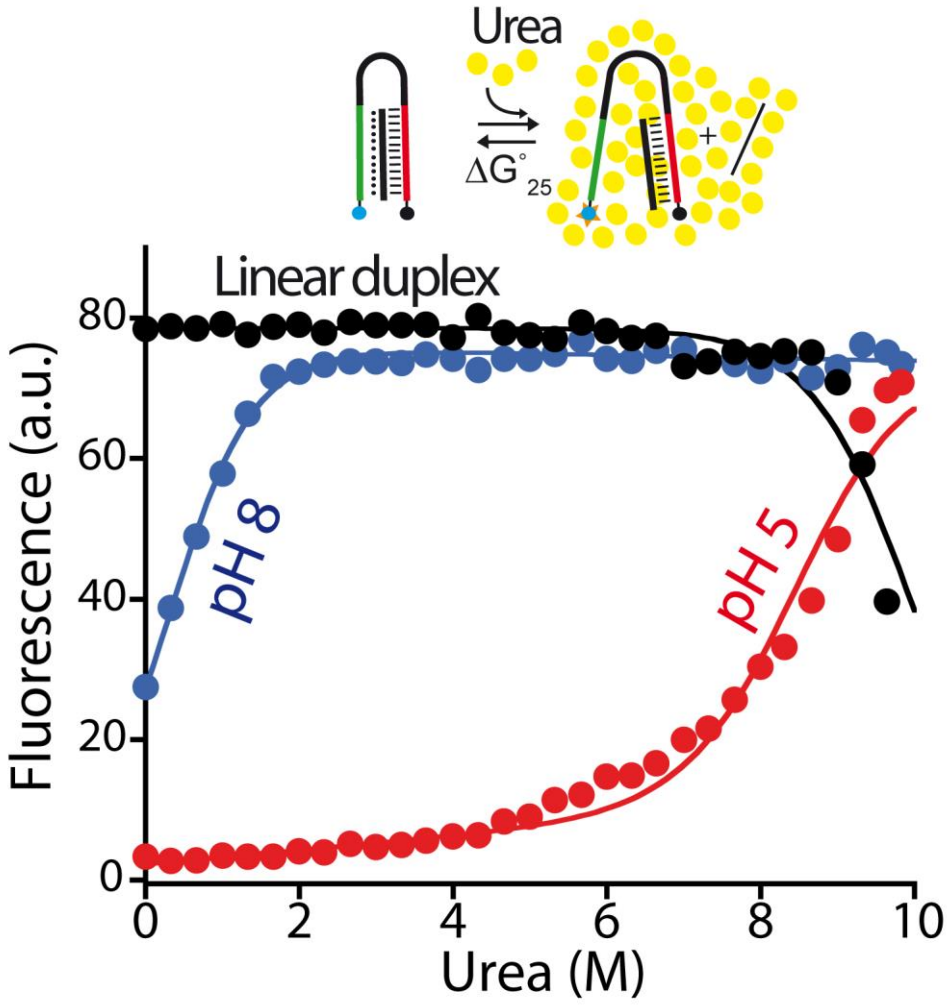

Figure S10: Urea denaturation curves performed with the clamp-switch recognizing a 12-base target. As expected, the unfolding/dissociation of the clamp-switch at pH 5.0 (red curve) occurs at higher urea concentrations than that at pH 8.0 (blue curve), thus confirming the higher stability of the triplex motif at acid pH. Of note, contrarily to what observed with the clamp-switch recognizing a 5-base target, the duplex is highly stable under these experimental conditions (see black curve at pH 8.0). Using this approach we found free energy values of the Hoogsteen interactions of 7.3 + 0.3 kcal/mol and 2.9 + 0.2 kcal/mol at pH 5.0 and 8.0 respectively. Here the urea denaturation curves were obtained in 40 mM Tris buffer, 12.6 mM $MgCl_2$ at the indicated pHs and at 25°C.

7) References:

(1) Idili, A.; Plaxco, K. W.; Vallée-Bélisle, A.; Ricci, F. *ACS Nano* **2013**, *7*, 10863.

(2) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063.

(3) Pace, C. N.; Grimsley, G. R.; Scholtz, J. M. in Buchner, J.; Kiefhaber, T. *Protein Folding Handbook*; Wiley-VCH Verlag: Weinheim, Germany, **2005**, *1*, 45.