Entropy-Based Rational Modulation of the pKₐ of a Synthetic pH-Dependent Nanoswitch

Davide Mariottini, †‡ Andrea Idilli, † Minke A. D. Nijenhuis, ‡ Gianfranco Ercolani, † and Francesco Ricci†,*

†Chemistry Department, University of Rome, Tor Vergata, Via della Ricerca Scientifica, 00133 Rome, Italy
‡Interdisciplinary Nanoscience Center (iNANO), Aarhus University, DK-8000 Aarhus C, Denmark

Supporting Information

ABSTRACT: The rational regulation of the pKₐ of an ionizable group in a synthetic device could be achieved by controlling the entropy of the linker connecting the hydrogen bond forming domains. We demonstrate this by designing a set of pH-responsive synthetic DNA-based nanoswitches that share the same hydrogen bond forming domains but differ in the length of the linker. The observed acidic constant (pKₐ) of these pH-dependent nanoswitches is linearly dependent on the entropic cost associated with loop formation and is gradually shifted to more basic pH values when the length of the linker domain is reduced. Through mathematical modeling and thermodynamic characterization we demonstrate that the modulation of the observed pKₐ is due to a purely entropic contribution. This approach represents a very versatile strategy to rationally modulate the pKₐ of synthetic devices in a highly predictable and accurate way.

Hydrogen bonds are ubiquitous in nature and represent one of the most important interactions in life. Among the many advantages associated with hydrogen bonding such as reversibility, synthetic accessibility, and directionality, another feature makes this chemical interaction extremely appealing: its dependence on pH is highly tunable. The pKₐ of an ionizable group in a protein, for example, can shift by several pH units due to changes in the surrounding microenvironment.

For the above characteristics, hydrogen bonds are also among the most utilized noncovalent interactions used to design and build supramolecular assemblies, polymers, multifunctional materials, mechanically interlocked molecular architectures, and pH-responsive devices. Despite this extensive use, however, the possibility to tune the pKₐ of ionizable groups in synthetic systems with the same efficiency found in proteins has proven extremely difficult. Examples have demonstrated the possibility to control the pKₐ of supramolecular systems by photochemical or electrochemical inputs or by complexation, but the modulation of pKₐ remains difficult to predict and to achieve in a rational way.

Motivated by the above arguments, here we demonstrate that rational regulation of the pKₐ of an ionizable group in a synthetic device could be achieved through the design of intrinsically disordered regions connecting the hydrogen-bond forming domains (Figure 1). By doing so we show that through a purely entropic contribution it is possible to modulate the pKₐ of ionizable groups in a highly versatile and predictable way.

Figure 1. Observed acidic constant (pKₐ) of a pH-dependent synthetic switch can be rationally tailored by modulating the length of the linker domain that connects the portions undergoing hydrogen bonds (dotted line). When a lower entropic cost is associated with loop formation, the observed pKₐ of the switch is shifted to more basic pH values.

To achieve this goal, we took advantage of the high programmability of DNA interactions and designed a pH-responsive nucleic acid nanoswitch (Figure 2) that can form an intramolecular triplex structure through hydrogen bonds (Hoogsteen interactions) between a hairpin duplex domain and a single-strand triplex-forming domain. More specifically, we have designed a set of nanoswitches displaying the same pH-responsive sequence containing 2 protonation centers (i.e., two cytosines) but differing in the length of the linker connecting the duplex portion to the triplex-forming domain (Figure 2a). The nanoswitches are labeled with a pH-insensitive fluorophore/quencher pair to follow pH-dependent folding/unfolding (Figure S1). Experiments performed at a fixed concentration of the nanoswitch and at different pH values show that the pH at which the triplex structure folds/unfolds is strongly correlated to the linker length. By shortening the length of the linker domain we observe a gradual increase of the pH value at which the triplex structure is destabilized (Figure 2b).

Received: April 28, 2019
Published: July 11, 2019

DOI: 10.1021/jacs.9b04168
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX
To better understand these experimental data, we developed a thermodynamic model that describes the pH-dependent triplex formation as a multistep process in which the n cytosines present in the single-stranded triplex-forming portion (S) are first protonated with an intrinsic acidity constant \( K_{C(S)} \) presumably equal to that of free cytidine in aqueous solution (pKₐ = 4.08) followed by the intramolecular triplex (THₛ) formation characterized by the dimensionless intramolecular dissociation constant, \( K_{d}^{\text{intra}} \). As a result, it is evidenced that the HSH instability of the complex, we employed the effective molarity, EM, defined as the ratio between the intermolecular dissociation constant between the triplex-forming portion and the duplex when the linker is absent, \( K_{d}^{\text{inter}} \), and the intramolecular dissociation constant, \( K_{d}^{\text{intra}} \), defined above:

\[
\frac{K_{d}^{\text{inter}}}{K_{d}^{\text{intra}}} = EM
\]

The EM is a measure of the propensity for an intramolecular process to occur relatively to the corresponding intermolecular process. Thus, while the EM only depends on the structural features of the linker, \( K_{d}^{\text{inter}} \) can vary with the sequence of the duplex portion and the single-stranded triplex-forming portion (i.e., number of cytosines). Substitution of eq 5 into eq 3 gives

\[
pK_a = \frac{1}{n} \log EM + \frac{1}{n} pK_{d}^{\text{inter}} + pK_{C(S)}
\]

In the event of entropy control of the effective molarity, a linear log-log correlation with a negative slope between EM and the linker length (number of nucleotides, nt) is expected. For a series of n-protic switches with the same n value, since \( K_{d}^{\text{inter}} \) is constant, such a linear dependence would translate, according to eq 6, into a linear correlation with a negative slope between the pKₐ of the switch and the log of linker length (straight line marked TH₂, Figure 2e), thus confirming that the modulation of the pKₐ is regulated solely by entropy through the EM parameter.

The modularity of the system employed in this work offers the opportunity to study how the modulation of pKₐ is affected by the number of protonation centers in the triplex-forming portion of the switch. To this end, we have designed a new set of switches with the same linker lengths investigated above but with a single-stranded triplex-forming portion (S) containing four cytosines instead of the two employed before (Figure 2c). A linker dependent modulation of the pH value at which folding/unfolding of the triplex structure occurs is also observed in this case (see pKₐ values in Table S12. However, from visual comparison of the pH titrations of the diprotic and tetraprotic switches (Figures 2b,d), some differences are

\[
K_{C(T)}^n = K_d^{\text{intra}} K_{C(S)}^n = \frac{[S][H^+]^n}{[TH_n]}
\]

Since at half-titration \( pK_{C(T)} = \text{pH} \), the value of \( pK_{C(T)} \) can be referred in common parlance as the pKₐ of the switch:

\[
pK_a = pK_{C(T)} = \frac{1}{n} pK_{d}^{\text{intra}} + pK_{C(S)}
\]

Under a normalized fluorescent condition (0-1) in which only the duplex state is fluorescent, whereas the triplex state does not provide any signal, the following equation can be used to fit the experimental data by nonlinear least-squares:

\[
F = \frac{K_a^n}{K_a^n + [H^+]^n}
\]

Figure 2. pH nanoswitches with rationally tailored pKₐ. pH titration curves obtained with a set of diprotic, TH₂ (a,b) and tetraprotic, TH₄ (c,d) pH nanoswitches sharing the same pH-dependent triplex-forming domains and varying lengths of the linker domain. Solid lines in all pH curves represent fits obtained by nonlinear least-squares optimization with the equilibrium model outlined in the text. (e) pKₐ vs log(linker length, nt) for both the diprotic (TH₂) and tetraprotic (TH₄) nanoswitches. (f) ΔpKₐ between the shortest and longest linkers (i.e., 5 and 35 nt) for di-(TH₄) nanoswitches. (f) pH nanoswitches with rationally tailored pKₐ.
immediately apparent. As predicted by eq 4, the steepness of the titration curves increases on increasing the number n of available cytosines from 2 to 4. Moreover, the range of pK\textsubscript{a} modulation upon increasing the linker length appears to be more restricted in the case of n = 4 compared to n = 2. This trend is clearly illustrated by the linear correlations of pK\textsubscript{a} values against log(linker length, nt) for the diprotic and tetraprotic switches (Figure 2c) which show a slope of −2.7 and −1.4, respectively. The different slope is easily rationalized by the model embodied in eq 6, predicting that an increase of the number of protonation centers from n = 2 to n = 4, should decrease the slope by a factor of 2.

To further test the reliability of the proposed thermodynamic model, we have extended the study to the nanoswitches with n = 6 and n = 8, for the shortest (nt = 5) and longest (nt = 35) linkers (Table S13). These data allow to investigate how the ΔpK\textsubscript{a} between the shortest and longest linkers (i.e., 5 and 35 nt), which represents the range of pK\textsubscript{a} modulation, is affected by the number n of cytosines (n = 2, 4, 6, and 8). According to eq 6, for two nanoswitches of the same n value but different linker length, ΔpK\textsubscript{a} is equal to (1/n)Δ(log EM). As a result, a plot of ΔpK\textsubscript{a} vs 1/n should give a straight line passing through the origin, whose slope is Δ(log EM). The experimental data fit this prediction remarkably well (Figure 2f), thus strengthening our confidence in the proposed model.

The meaning of the plot is straightforward: on increasing the number of cytosines (n) the pK\textsubscript{a} of switches of different linker lengths converge to the same value. In other words, the entropy modulation achieved by varying the length of the linker is more and more dissipated by an increase of the number of protonation centers (i.e., cytosines). These same data also offer the possibility to understand how the number of cytosines might affect the intermolecular dissociation constant between the triplex-forming portion and the duplex when the linker is absent (i.e., pK\textsubscript{a}inter). As expected, the change in pK\textsubscript{a}inter observed by increasing the number of cytosines in the switch is indistinguishable, within error, between the switches with a 5 and 35 nt linker (see SI and Figure S12). Moreover, the increase in pK\textsubscript{a}inter is remarkable (>11 units on passing from n = 2 to n = 8) further demonstrating the importance of the protonation of cytosines in the triplex formation.34

To better understand the role of the linker length on the pK\textsubscript{a} modulation of the switch, we have experimentally determined the entropic contribution for the different linker lengths through thermal melting curves and van’t Hoff analysis at a fixed pH (5.5) for both the diprotic (Figure 3a,b) and tetraprotic (Figure 3c,d) switches.35 We found that the melting temperature of all our nanoswitches scales linearly with the linker length (Figures S13 and S14). The length of the linker domains does not affect the enthalpic contribution of hydrogen bonding formation in triplex folding (see S1). The free energy values calculated through van’t Hoff analysis with nanoswitches of different linker lengths can thus be used to obtain an estimate of the entropic contribution associated with the linker domain (ΔS\textsubscript{linker}) by subtracting the ΔS value of the shortest linker (used as reference) to that of the other switches. As expected for a purely entropic controlled phenomenon, ΔS\textsubscript{linker} scales linearly with the natural logarithm of the number of monomers (nucleotides) in the linker (Figure 3e),32 as well as with the pK\textsubscript{a} of the corresponding switch (Figure 3f).

Finally, we observe a trend in the values of folding rate constants (k\textsubscript{f}, s\textsuperscript{-1}) of the switches with different linker lengths with slower folding kinetics as the linker increases (Figures S15 and S16). Conversely, because the unfolding process is not affected by the linker length, we do not observe any significant difference in the rate constants with different linker lengths (Figure S17).

In conclusion, we have designed a family of multiprotic DNA-based pH-dependent nanoswitches that share the same hydrogen bond-forming domains but differ in the length of the linker connecting such domains. We have demonstrated that the modulation of the observed pK\textsubscript{a} is due to a purely entropic contribution of the linker domain, which can be dissipated on increasing the number of protonation centers in the single-stranded triplex-forming portion.

This approach represents a highly versatile and predictable strategy to rationally modulate the pK\textsubscript{a} of synthetic devices and might be used for a wide range of different applications, including the design of molecular machines and noncovalent assemblies with a finely tunable pH dependence. Compared to other strategies in which modulation of the pK\textsubscript{a} of pH-dependent DNA-based nanoswitches has been achieved by modifying the pH-dependent domain36,38,39 or through allosteric control,40 our approach appears advantageous because of the ease with which the entropy of single-stranded
DNA sequences can be gradually regulated. We also note that a similar strategy might be employed with other more complex biopolymers, including for example peptide sequences, thus allowing the rational design of synthetic protein-like biomolecular systems and assemblies with tunable pKₐ values.

**ASSOCIATED CONTENT**

**Supporting Information**
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b04168.

Experimental, thermodynamic model and data analysis, thermodynamic analysis of thermal melting curves, supporting figures (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**
*francesco.ricci@uniroma2.it*

**ORCID**

Davide Mariottini: 0000-0001-6211-6428
Andrea Idili: 0000-0002-6004-270X
Gianfranco Ercolani: 0000-0003-2437-3429
Francesco Ricci: 0000-0003-4941-8646

**Notes**
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This work was supported by the European Research Council, ERC (project no. 336493, FR), by Associazione Italiana per la Ricerca sul Cancro, AIRC (project no. 21965, FR) and by the Italian Ministry of Health (project no. GR-2013-02356714).

**REFERENCES**


(33) Error bars in the titration curves are depicted for only one point on each curve and represent the typical standard deviations obtained on at least three independent measurements.


