DNA-based nanodevices controlled by purely entropic linker domains

DNA-based nanodevices controlled by purely entropic linker domains

Journal:	*Journal of the American Chemical Society*
Manuscript ID	ja-2018-07640d.R1
Manuscript Type:	Article
Date Submitted by the Author:	14-Sep-2018
Complete List of Authors:	Mariottini, Davide; University of Rome Tor Vergata, Chemical Science and Technology Idili, Andrea; University of Rome Tor Vergata, Chemical Science and Technology Nijenhuis, Minke; Technical University Eindhoven, SMO de Greef, Tom; Technical University Eindhoven, SMO Ricci, Francesco; University of Rome Tor Vergata, Department of Chemistry

DNA-based nanodevices controlled by purely entropic linker domains

Davide Mariottini [1], Andrea Idili [1], Minke A.D. Nijenhuis [2], Tom de Greef [2,*] and Francesco Ricci [1,*]

[1] Chemistry Department, University of Rome, Tor Vergata, Via della Ricerca Scientifica, 00133, Rome, Italy; [2] Institute for Complex Molecular Systems, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands;

*Corresponding authors: t.f.a.d.greef@tue.nl; francesco.ricci@uniroma2.it

ABSTRACT:

We demonstrate here the rational design of purely entropic domains as a versatile approach to achieve control of the input/output response of synthetic molecular receptors. To do so and to highlight the versatility and generality of this approach we have rationally re-engineered two model DNA-based receptors: a clamp-like DNA-based switch that recognizes a specific DNA sequence and an ATP-binding aptamer. We show that, by varying the length of the linker domain that connects the two recognition portions of these receptors, it is possible to finely control their affinity for their specific ligand. Through mathematical modelling and thermodynamic characterization we also demonstrate for both systems that entropy changes associated with changes in linker length are responsible for affinity modulation and that the linker we have designed behaves as a disordered random coil polymer. The approach also allows us to regulate the ligand concentration range at which the receptors respond and show optimal specificity. Given these attributes, the use of purely entropic domains appears as a versatile and general approach to finely control the activity of synthetic receptors in a highly predictable and controlled fashion.

INTRODUCTION

Man-made synthetic molecular recognition systems and devices that can bind and recognize a ligand in a specific and selective way have become key tools in several fields including diagnostic, drug-delivery and therapeutic[1-6]. A fine and predictable regulation of the activity of such synthetic molecular receptors would allow to better control these tools and thus represents a highly relevant, yet challenging, objective[7-11]. Usually, modulation of the activity of a synthetic molecular receptor or nanodevice can be achieved by careful thermodynamic optimization of the recognition events involved [12-14], for example by directly modifying the recognition domains that take part in the ligand binding. However, this approach mostly affects the enthalpic contribution of binding and is not without limitation: first, this does not allow to finely control the affinity of a synthetic receptor with precision especially when less predictable interactions are in play; second, tuning the enthalpic contribution of the recognition event might affect also the affinity towards non-specific molecular targets thus ultimately affecting the specificity of the interaction. Finally, enthalpy-based approaches are prone to environmental effects such as changes of pH or of ionic strength. Finding new ways to overcome these limitations and to rationally control and modulate in a predictable fashion the activity of synthetic receptors thus represents an important goal with significant implications in several fields of research.

Nature has faced the same challenging goal: how to modulate the activity of biomolecular receptors, like proteins and enzymes, in a highly controllable way? Obviously many strategies, like heterotropic allostery, are well known, but recently, the discovery that many proteins contain intrinsically disordered domains without an apparent specific function[15-19] has provided an additional possible answer to this question challenging the original dogma that disorder plays against functional activity and that proteins require well-folded domains to function properly. Many proteins, in fact, employ conformational entropic contribution of thermodynamically different domains that are not directly involved in the recognition event to better control their activity (Figure 1)[20-23]. Such dynamic and purely entropic allostery represents a hallmark of many key proteins especially those involved in signalling pathways and transcription regulation, thus suggesting that this property allows a fine regulation of proteins response and activity in a very versatile and precise way[24]. Recreating this purely entropic control mechanism in man-made synthetic devices would allow an unprecedented fine regulation of their response and input/output behaviour.

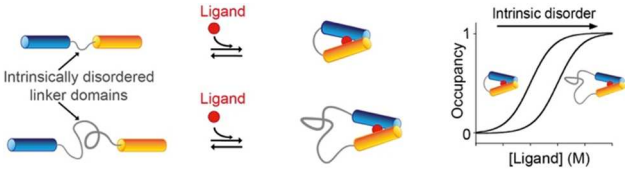


Figure 1. The affinity of a biomolecular receptor towards a specific ligand is strictly related to the presence of intrinsically disordered regions. With a higher entropic cost associated to the intrinsically disordered linker that connects two binding domains, for example, a poorer affinity of a protein for the same ligand will be observed (see binding curves). This mechanism is employed by Nature to dynamically control proteins function through the modulation of the entropy associated to intrinsically disordered domains distal from the binding site[20-23].

In response to the above considerations we report here a convenient and versatile approach to control the activity and response behaviour of synthetic molecular recognition systems by rationally designing intrinsically disordered domains. We demonstrate that, similarly to intrinsically disordered proteins, such approach allows to finely modulate the affinity of synthetic receptors through a purely entropic contribution in a highly versatile way without requiring any detailed thermodynamic design. To do so we have taken advantage of the high programmability of DNA interactions that allows the rational design of nanoscale synthetic DNA-based devices and structures with programmable features[25-28].

RESULTS

As a first system we have employed a synthetic DNA-based conformation switching receptor, named clamp-switch, that contains a pair of 10-nucleotide (nt) recognition domains (blue and orange, Figure 2a) joined by a poly(T) DNA linker (grey, Figure 2a)[29-30]. The first recognition domain (orange, Figure 2a) binds a specific DNA sequence through classic Watson-Crick interactions forming a duplex which is subsequently recognized by the second recognition domain (blue, Figure 2a) through intramolecular Hoogsteen interactions to form a DNA triplex

structure. The clamp-switch receptor is labelled at the two ends with a fluorophore/quencher pair to allow easy real-time detection of the ligand/receptor interaction. Formation of the triplex structure brings the fluorophore and quencher in close proximity thus leading to a suppression of the observed fluorescence signal. Because loop closure of short single-stranded poly(T) linkers is purely entropic and involves no additional enthalpic terms due to intramolecular base-stacking[31] we can finely control the overall affinity of our receptor by rationally modulating the length, and thus intrinsic disorder, of the linker domain. To demonstrate this, we have designed a library of DNA clamp-switch receptors sharing the same recognition domains and with a poly(T) linker of different lengths.

Poly(T) linkers are well described as random-coil polymers

To verify the binding mode of the clamp-switch receptor and quantify the contribution of the disordered poly(T) linker to the overall observed affinity towards a ssDNA ligand we performed titration experiments at increasing concentrations of a complementary 10-nt DNA sequence for all the clamp-switch receptor variants (Figure 2b). The data reveal that the observed affinity of the ligand decreases upon increasing the length of the poly(T) linker. More specifically, by analysing the data using a Langmuir isotherm (Figure SI1) we observe that varying the length of the linker domain from 10 to 38 nucleotides allows to tune the dissociation constant between the receptor and the target ($K_d^{Langmuir}$) from 7 ± 1 nM to 137 ± 22 nM, respectively.

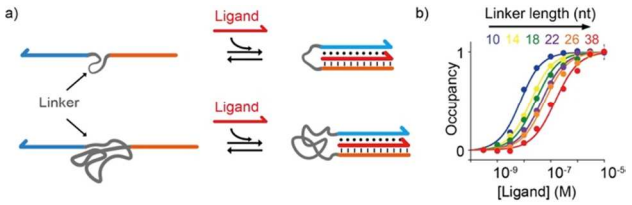


Figure 2. (a) The purely entropic regulation mechanism employed by proteins (Fig. 1) can be re-engineered into a synthetic triplex-forming clamp-like DNA-based receptor. By modulating the length of the linker domain that connects the two binding domains the observed affinity of the DNA-based clamp-switch for its target can be precisely tuned. (b) Binding curves obtained with a set of DNA clamp-switches sharing the same recognition elements (domains orange and blue) and varying lengths of the linker domain. Solid lines represent fits obtained by non-linear least square optimization with the equilibrium model outlined in Fig. 3. The titration experiments were performed in 50 mM TrisHCl, 10 mM $MgCl_2$ at pH 7.4, 37 °C at a concentration of clamp-switch of 3 nM and adding increasing concentrations of the 10-nt target strand. For a matter of clarity in these binding curves and in those in the following figures, error bars have been depicted for only one point on each curve and represent the average and standard deviations of measurements performed on at least three independent measurements.

To understand this data quantitatively, we developed a thermodynamic binding model that describes the formation of the intramolecularly stabilised complex as a two-step process in which the ssDNA ligand binds the first domain with dissociation constant K_d^{target} (M) followed by intramolecular binding of the second domain characterized by the dimensionless intramolecular dissociation constant, K_d^{intra} (Figure 3a). The overall dissociation constant between the clamp-switch receptor and the ligand is then given by:

$$K_d = K_d^{target} K_d^{intra} \qquad \text{Eq. 1}$$

To isolate the effect of the linker length on the overall stability of the complex, we employ the effective molarity, EM:

$$\mathrm{EM} = \frac{K_d^{inter}}{K_d^{intra}} \qquad \text{Eq. 2}$$

with K_d^{inter}, the intermolecular dissociation constant of the appropriate reference reaction (Figure 3a) in which the linker is absent. To experimentally obtain K_d^{target} (dissociation constant of the first recognition event, i.e. duplex formation) we designed a control DNA receptor containing the first recognition domain and a second domain with a random sequence not able to form a triplex structure (Figure SI2). The dissociation constant of the first recognition event is, as expected, independent on the length of the linker domain ($K_d^{target\_10}$= 1.1 ± 0.1 µM for 10-nt linker and $K_d^{target\_38}$ = 1.3 ± 0.1 µM for 38-nt linker) (Figure SI2). Similarly, we measured the dissociation constant between a pre-formed hairpin duplex with the same sequence of the clamp receptor and a separate 10-nt triplex-forming strand (Figure SI3) to experimentally obtain K_d^{inter} (11.5 ± 0.2 µM). To extract an EM value for each clamp-switch receptor, we performed a global non-linear least-square optimization of the titration data using the thermodynamic binding model (Supporting Information and Table SI1). During the optimization, values of K_d^{inter} and K_d^{target} were fixed to their experimentally determined values. The thermodynamic equilibrium model is able to describe the titration data very well for all linker lengths (Figure 2b, solid lines). The match between model and experiment indicates that the affinity modulation is regulated solely through the EM parameter. To further verify the model, we calculated for each receptor the overall dissociation constant (K_d) from the estimated values of K_d^{target}, K_d^{inter} and EM and compared it to the dissociation constant values ($K_d^{Langmuir}$) obtained by fitting the experimentally determined data of the titration curves with a Langmuir isotherm (Figure SI1). The values correlate very well indicating the consistency of the binding model (Figure 3b). As expected for a disordered linker of sufficient length, the EM values monotonically decrease as a function of the linker length (Figure 3c).

The effective molarity, EM, is closely related to the effective concentration C_{eff}, a theoretical parameter that allows the equilibrium for intramolecular reactions to be estimated by assuming that the linker between the two associating end groups can be approximated as a random coil polymer[32-33]. Because the effective concentration of a random coil polymer has a purely entropic origin, correlations between EM and C_{eff} values would thus indicate that the linker has an entropic contribution to the overall binding process of the DNA receptor. The effective concentration, C_{eff}, of a random-coil polymer in which two end-groups are a distance d apart is given by:

$$C_{\mathrm{eff}}(d) = \frac{p}{N_A (C\sqrt{N})^3}\left(\frac{3}{2\pi}\right)^{3/2} \cdot e^{-\frac{3d^2}{2(C\sqrt{N})^2}} \qquad \text{Eq. 3}$$

In which the number of segments N in the linker is equal to the number of nucleotides in the linker, C is the length of each nucleotide along the direction of the chain, the parameter p takes into account excluded volume effects experienced by the linker when the two chains ends are associated with each other and d is the end-to-end distance[34-35]. We have performed non-linear least square regression of the EM data using Eq. 3 for C_{eff} (Supplementary Information, Table SI2). During this procedure,

the *C* parameter (nucleotide length) is fixed at 0.63 nm[36] while the parameters *p* and *d* were allowed to vary over a realistic range. As can be observed from Figure 3c, the experimental variation of EM upon increasing the length of the poly(T) linker can be described by Eq. 3, indicating that the linker behaves as a random coil polymer and thus has a purely entropic contribution to the stabilisation of the complex. The slight deviation between the experimentally determined EM and C_{eff} values is probably to be ascribed to non-specific base pair interactions between the loop and the stem. Previous works on hairpin loops have in fact shown that short poly(dT) loops in DNA hairpins are slightly more stable compared to what is expected from a purely entropic contribution[37]. The value of the end-to-end distance, *d* (0.1 ± 1.37 nm) could not be estimated accurately, however it is well within the expected range (Supporting Information). In contrast, the value of *p* (0.016 ± 0.009) could be estimated more reliably. The value is significantly lower compared to the theoretical value of 2, proposed by Lees and co-workers[34]. Previously, Whitesides and co-workers found a value of 0.12 for an intramolecular protein-ligand system[35]. We speculate that the low value of *p* in our synthetic DNA receptor is caused by the extended base-pairing between the two domains in the triplex state which results in exclusion of the poly(T) linker from this volume, which would lower *p* significantly.

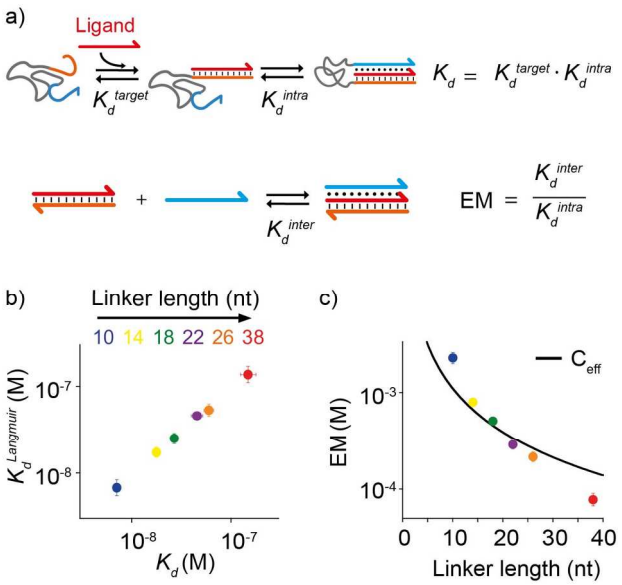


Figure 3. (a) The clamp-switch observed dissociation constant (K_d) can be described as the product of the binding dissociation constant of duplex formation (K_d^{target}) and the intramolecular binding dissociation constant for triplex formation (K_d^{intra}). K_d^{intra} can be defined as the ratio of the dissociation constant of the appropriate intermolecular reference reaction (K_d^{inter}) and the effective molarity (EM), which only depends on the length and flexibility of the linker. (b) Calculated values of K_d, obtained from the K_d^{inter} , K_d^{target}, and the estimated EM values, correlate well to the values obtained by analyzing the titration data using a Langmuir isotherm ($K_d^{Langmuir}$). (c) Variation of the effective concentration (C_{eff}) and effective molarity (EM) with linker length. The points are the estimated values of EM obtained by analysis of the titration curves in Fig. 2b while the solid curve is a fit to the data using the definition of C_{eff} for a random-coil linker. The $K_d^{Langmuir}$ values represent mean±s.d. of three separate measurements. The EM markers represent the estimated values±estimated s.d. (see SI). The K_d values represent the calculated values±calculated s.d. based on error propogation of the EM estimations.

Thermodynamic characterization

To better understand the role of the linker length on the binding activity of the clamp-switch receptor, we have experimentally determined the entropic contribution for the different linker domains through thermal melting curves and van't Hoff analysis[38]. To do so, we have designed a control unimolecular variant of the DNA clamp-switch receptor where the first recognition element and the 10-nt target are connected by a 5-nt domain thus leading to the formation of the duplex DNA in a concentration-independent fashion (Figure 4a, SI4-5). Such unimolecular clamp-switch receptor is labelled with a fluorophore/quencher pair in order to monitor triplex/duplex transition (Figure 4a). Formation of the triplex structure will bring the fluorophore and quencher in close proximity thus leading to a suppression of the fluorescence signal. This set of unimolecular clamp-switch receptors will allow to characterize only the folding/unfolding of the second recognition domain and will thus give a measure of the total free energy of the second recognition event. Because the enthalpic contribution of such event is likely not affected by the length of the linker domains, as this is not involved in the recognition event, the observed difference in free energy values between clamp-switch receptors with different linker lengths can be solely ascribed to the different entropic contribution associated to the linker domain (ΔS_{linker}). The entropy values of each clamp-switch variant calculated through van't Hoff analysis of the melting thermal curves (Figure 4b, SI4-8)[38] contain the entropic contribution of the second binding domain and that associated to the linker loop portion. By subtracting these values from the entropy obtained for the variant with the shortest linker length (and thus the lowest ΔS_{linker}), we can obtain an estimation of the entropic contribution of the linker for the other variants. These values are in good agreement with entropy values previously obtained by others using different DNA and RNA systems of comparable length[39]. As expected for a random-coil polymer, ΔS_{linker} scales with the natural logarithm of the number of monomers (nucleotides) in the linker (Figure 4c)[40] thus further supporting the model proposed. Moreover, the observed dissociation constant of clamp-switch receptors ($K_d^{Langmuir}$) is, as expected, strongly dependent on the total entropy of each variant (Figure 4d).

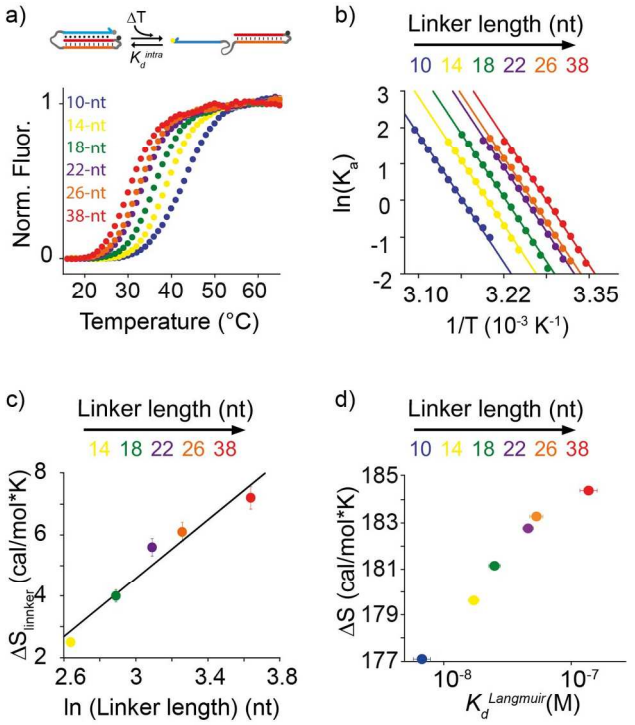


Figure 4. (a) Using an unimolecular control switch we have performed melting curve experiments and (b) obtained van't Hoff plots to measure the entropy associated to the linker domain (i.e. triplex formation). (c) The entropy scales linearly with the ln(linker length) (R^2 = 0.946) and d) with the dissociation constant values obtained from fitting the experimental values to a Langmuir isotherm ($K_d^{Langmuir}$). Melting curve experiments were performed at a concentration of control switch of 50 nM at a rate of 0.4 °C·min^{-1}. The experimental values represent mean±s.d. of three separate measurements.

Purely entropic modulation of specificity and dynamic range of DNA-based receptors

Because the entropic cost associated with the linker domain affects the overall affinity of the receptor for its ligand without changing the recognition domains, this represents a means to modulate the target concentration window at which the receptor has optimal specificity. To demonstrate this we have selected two receptors with linkers of different lengths and performed titration experiments at increasing concentrations of a perfect match target and a target containing a single nucleotide mismatch (Figure 5a-b). The specificity window of our receptors can be graphically depicted by showing the difference between the signal obtained with a perfect match and that obtained with the same concentration of a mismatch target. We found that for the DNA receptor with the shorter linker such specificity window (here defined as the concentration range where the difference between the relative signal obtained with the perfect match and that obtained with the mismatch target is higher than 0.25) is centred at 100 nM and spans a 240±22-fold of target concentration (Figure 5c, blue). For the DNA receptor with the longer linker, the specificity window spans a similar width (i.e. 234±18-fold) but is centred at 480 nM of target concentration (Figure 5c, red). Through pure entropic contribution and without changing the recognition domains it is thus possible to shift the specificity window without altering its width.

By combining variants with different linker domains we can rationally tune the dynamic response of DNA-based nanodevices. To do this we have employed four triplex-forming DNA clamp-switches with different linker lengths (8, 16, 24 and 32 nucleotides) targeting an 11-nt perfect match target. As it is expected for receptors with a single binding site, the dynamic range of each of these receptors spans ca. 2 orders of magnitude (81-fold) of target concentration characteristic of a Langmuir-type isotherm[41] (Figure 5d). By mixing two receptors with different linker lengths it is possible to extend the observed dynamic range over almost 3 orders of magnitude of target concentration (Figure SI9). Similarly, by mixing 4 different receptors with 4 different linker lengths allows rational extension of the observed dynamic range over 3000 fold (Figure 5e).

Using the same library of receptors with different dissociation constants for the same target we can also narrow the dynamic range. To do this we have employed in the same solution two clamp-switch receptors with different linker lengths (8-nt and 36-nt) and thus different dissociation constants for the target ($K_d^{Langmuir\_8}$ = 3.4 ± 0.1 nM; $K_d^{Langmuir\_36}$ = 0.8 ± 0.1 □M). For this specific experiment the receptor with the lowest dissociation constant does not contain the fluorophore/quencher pair and thus the binding of the target to this receptor will not result in any measureable signal. Under these experimental conditions the signalling incompetent receptor will sequester the ligand until it saturates. Only when the ligand concentration surpasses the concentration of the signalling incompetent receptor, the signalling competent receptor is activated. As a result the observed dynamic range of the receptor mixture spans a much narrower range of target concentration (3-fold) (Figure 5f).

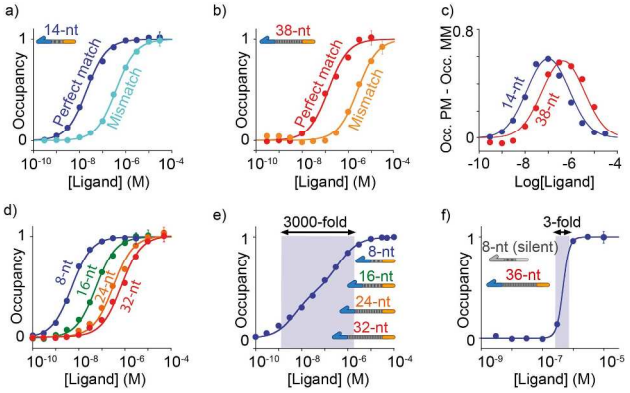


Figure 5. Tuning the specificity window and dynamic range of synthetic DNA-based clamp-switches through intrinsic disorder. (a) Binding curves of a 14-nt linker clamp-switch (Figure 2) with 10-nt perfect match (red) and mismatch (orange) targets. (b) Binding curves of a 38-nt linker clamp-switch with the same 10-nt perfect match (red) and mismatch (orange) targets. (c) Specificity windows for the two clamp-switches obtained by subtracting the values obtained in the presence of the perfect match from those obtained with the mismatch target. (d) Mixing in the same solution different clamp-switch variants with different linker lengths[42] allows to broaden (e) and narrow (f) the dynamic range of the clamp-switch receptor towards a perfect match target. Extended dynamic range experiments (d-e) shown here were performed in 50 mM Tris HCl at pH 6.5, 37 °C. For binding curves with a single clamp-switch a concentration of 3 nM was used. To broaden the dynamic range (panel e) we have employed a mixture of four clamp-switch variants at the following concentrations ([8-nt]= 0.85 nM; [16-nt]= 0.85 nM; [24-nt]= 0.85 nM; [32-nt]= 0.45 nM. To narrow the dynamic we used a signaling incompetent version of the 8-nt linker clamp-switch (at a 10 nM concentration) together with a 36-nt linker clamp-switch (at a 1 μM concentration). All solid lines in a, b and d

are Langmuir-type fits while those in e and f are Hill-type fits (see SI). Solid lines in c are only meant to guide the eye. The experimental values represent mean±s.d. of three separate measurements.

Rational design of DNA-based aptamers controlled by disorder

Disorder can be also used to control the affinity and response behaviour of DNA-based aptamers. To demonstrate this we have selected a DNA aptamer able to bind ATP[43-44]. We have split this aptamer into two fragments that are connected by a poly-T domain of varying length (Figure 6a). The ATP-binding split-aptamer is labelled at the two ends with a fluorophore/quencher pair to allow easy real-time detection of the ATP binding. Formation of the aptamer/ATP complex brings the fluorophore and quencher in close proximity thus leading to a suppression of the observed fluorescence signal. We show here that the length of such poly-T linker domain, and thus its associated disorder, allows to finely control the affinity towards ATP. To demonstrate this we performed titration experiments at increasing concentrations of ATP for all the ATP-binding split-aptamer variants (Figure 6b). The data reveal that the observed affinity for ATP decreases upon increasing the length of the poly(T) linker. More specifically, by analysing the data using a Langmuir isotherm (Figure SI10) we observe that varying the length of the linker domain from 4 to 70 nucleotides allows to tune the dissociation constant between the aptamer and ATP ($K_d^{Langmuir}$) from 0.050 ± 0.009 mM to 1.0 ± 0.2 mM, respectively.

The response of this aptamer receptor can also be modelled in a similar way as the clamp-switch system using the same thermodynamic binding model that describes the formation of the ATP-aptamer stabilised complex as a two-step process in which the ATP binds the first split domain with dissociation constant K_d^{target} (M) followed by intramolecular binding of the second split domain characterized by the dimensionless intramolecular dissociation constant, K_d^{intra} (Figure 6a). Also in this case the overall observed dissociation constant is given by Eq. 1 and to isolate the effect of the linker length we again employ the effective molarity, EM, given by Eq. 2. In this case, K_d^{target} (M) is the dissociation constant between the first split aptamer domain and ATP and K_d^{inter} is the intermolecular dissociation constant between the second split aptamer domain and the complex formed between ATP and the first split domain in a reference reaction where the linker domain is not present. For a similar split ATP-binding aptamer a K_d^{inter} in the high µM range was reported[45] and thus for our analysis we fixed this value to 100 µM. Conversely, because K_d^{target} is too high and cannot be experimentally derived, and considering that the binding mechanism suggests this should be higher than K_d^{inter} (see SI) we fixed this value to 5 mM. To extract an EM value for each ATP-binding aptamer variant, we performed nonlinear least-squares optimization with the Levenberg-Marquardt algorithm of the titration data (see Supporting Information and Figures SI11-13 for a detailed description of the fitting procedure). The thermodynamic equilibrium model is once again able to describe the titration data very well (Figure 6b, solid lines). The match between model and experiment indicates that also for this system the affinity regulation is achieved solely through the EM parameter. Because in this system, the precise value of K_d^{target} and K_d^{inter} cannot be determined experimentally, we repeated the non-linear least square analysis of the titration data using a broad range of K_d^{target} and K_d^{inter} values (Figures SI11). The analysis reveals that the binding model is able to accurately fit the titration data for values of K_d^{target} > 1 mM independent of K_d^{inter} indicating that our hypothesized value for K_d^{target} is in the right order of magnitude. Comparison of the dissociation constant values calculated from our model (K_d) using the hypothesized values of K_d^{target}, K_d^{inter} and EM with the dissociation constant values ($K_d^{Langmuir}$) obtained by fitting the experimentally determined data of the titration curves with a Langmuir isotherm (Figure SI10) demonstrate the consistency of the binding model (Figure 6c). Also, in this case the EM values monotonically decrease as a function of the linker length (Figure 6d). The optimized EM values were then fitted with the random coil model (Eq. 3). During the fitting the C parameter (nucleotide length) is fixed at 0.63 nm, and the volume exclusion parameter p and the end-to-end distance parameter d are optimized while being constrained to real numbers. There is a good fit found between the calculated C_{eff} and EM values (Figure 6d, Table SI3-4) indicating that the linker behaves as a random coil. However, because the precise value of K_d^{target} and K_d^{inter} cannot be determined experimentally, we investigated how changes in these two parameters would affect this conclusion (Figures SI12). Our analysis reveals that the EM and calculated C_{eff} values based on Eq. 3 correlate very well for a broad range of K_d^{target} and K_d^{inter} values.

The optimized d value is 1.20 ± 0.07 nm, which is well within the expectations considering that the linker ends are attached to a portion of the aptamer that is reported as a WC stem[46], with an expected C1-C1 distance of 1.05 nm. The value of p is found to be 0.19 ± 0.02, which would suggest a high amount of excluded volume for the linker end (a value of p=2 should indicate a full sphere of access, whereas p=1 only a hemisphere). Because we could not measure accurate values of K_d^{inter} and K_d^{target} any inaccuracies in these parameters will be compensated through the p parameter during the optimization procedure (Figure SI13), which further underlines that the value of p requires a loose interpretation. These results thus suggest that also for the re-engineered split ATP-binding aptamer the affinity modulation is indeed a purely entropic process.

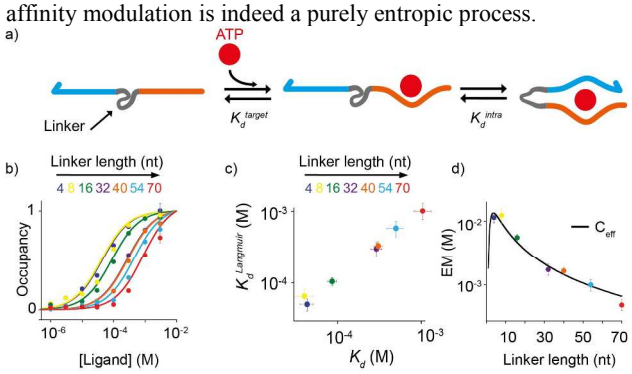


Figure 6. We have re-engineered an ATP-binding aptamer by splitting it into two portions connected by a poly(T) linker domain. By varying the length of the linker domain the observed dissociation constant of the split ATP-binding aptamer can be modulated. (b) Binding curves obtained with a set of split ATP-binding aptamers with varying lengths of the linker domain. Solid lines represent fits obtained by non-linear least square optimization with the equilibrium model outlined in panel a. (c) Calculated values of K_d, obtained from the K_d^{inter}, K_d^{target}, and the estimated EM values correlate well to the values obtained by fitting the experimental data to a Langmuir isotherm ($K_d^{Langmuir}$) (Figure SI10). (d) Variation of the effective concentration (C_{eff}) and effective molarity (EM) with linker length. The points are estimated values of EM obtained by analysis of the titration curves in b) while the solid curve is a fit to the data using the definition of C_{eff} for a random-coil linker (see text). The titration experiments were performed in 100 mM Tris HCl, 10 mM $MgCl_2$, pH 6.5 at 37 °C at a concentration of ATP-

binding aptamer of 50 nM and adding increasing concentrations of ATP. The $K_d^{Langmuir}$ values represent mean±s.d. of three separate measurements. The EM markers represent the estimated values±estimated s.d. (see SI). The K_d values represent the calculated values±calculated s.d. based on error propogation of the EM estimations.

To elucidate the role of the linker length on the binding activity of the ATP-binding split-aptamer, we have experimentally determined the entropic contribution for the different linker domains. To do so we have first estimated the free energy of the different ATP-binding aptamer variants by performing urea denaturation experiments in the absence and presence of saturating concentration of ATP (Figure 7a, SI14)[47]. This method has been recently proven efficient to determine thermodynamic free energy of aptamer/target interactions[47]. The free energy values obtained are in good agreement with the binding free energies obtained from the binding curves (Figure SI15). By increasing the linker length we can modulate the binding free energy of the ATP-binding aptamer from -6.2 ± 0.2 kcal/mol (4-nt linker) to -4.3 ± 0.2 kcal/mol (70-nt linker). By assuming that the linker does not take part in the ATP binding event we can estimate the entropic contribution due to the poly-T linker domain (ΔS_{linker}) for each ATP-binding split-aptamer. To do so we have used the split-aptamer variant with the shortest linker (i.e. 4 nucleotides) as our reference and we have subtracted its binding free energies from the binding energies estimated for the other aptamers variants (see SI). The estimated ΔS_{linker} values of the aptamer variants are once again linearly correlated with the natural logarithm of the number of monomers (nucleotides) in the linker as expected for a random-coil polymer (Figure 7b)[40]. The observed dissociation constant values of the different receptors, $K_d^{Langmuir}$, show also in this case a linear relation with the entropy associated with each linker (Figure 7c). Also in this case, the possibility to have a set of ATP-binding aptamers with finely modulated dissociation constants can be employed to tune the dynamic range of the aptamer in a very versatile way and to broaden the range of ATP concentration at which the aptamer can give valuable information (Figure SI16).

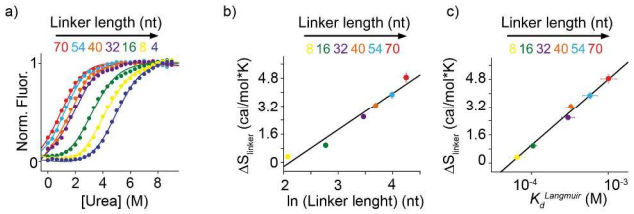


Figure 7. The entropy associated to the linker domain of the ATP-binding aptamer shown in Figure 6a, measured by urea denaturation experiments (a) scales linearly with ln(linker length) (R^2 = 0.964) (b) and correlates with the observed dissociation constant values ($K_d^{Langmuir}$) (R^2 = 0.994) (c). Urea denaturation experiments were performed in 100 mM Tris HCl, 10 mM $MgCl_2$, pH 6.5 at 37 °C at a concentration of ATP-binding aptamer of 50 nM. The experimental values represent mean±s.d. of three separate measurements.

DISCUSSION

Disordered domains are fundamental units of protein function and regulation[48]. It has been demonstrated that the control of entropic cost associated with disordered domains is employed by proteins to finely modulate their function, binding affinity and activity[49,50]. The ubiquitous nature of such mechanism and the efficiency with which it is employed by proteins and biomolecular receptors suggests that the same principle could be recreated in-vitro to finely control the activity and response behaviour of re-engineered proteins or synthetic receptors. Recently, for example, light and ligand sensitive domains have been used to modulate the structural disorder and thus activity of different proteins[51] or of protein-based sensors[52]. Motivated by the above arguments we have demonstrated here an unprecedented approach to finely modulate the activity of different synthetic DNA-based nanodevices by controlling the entropy of domains that are not directly involved in the recognition event.

The high versatility of this approach allows to adapt it to different DNA-based nanodevices. We have demonstrated this by employing the same principle to control the affinity of a clamp-like DNA-based receptor that recognizes a specific DNA sequence and an ATP-binding aptamer. We have shown that, by varying the length of the linker domain that connects the two recognition domains of these receptors, it is possible to finely control their affinity for their specific target. Through modelling and thermodynamic characterization we have demonstrated that entropy changes associated with changes in linker length are responsible for affinity modulation and that the linker we have designed behaves as a disordered random coil polymer. We have also demonstrated that the possibility of designing a library of intrinsically disordered nanodevices with controlled affinity provides a means of extending and narrowing the dynamic range at which they respond to the target.

Several methods have been employed to date to achieve control on the activity and tune the dynamic range of synthetic receptors and switches[41, 53-56] including the use of strategies inspired by allostery[57,58]. However, the majority of these methods are either based on the modulation of the enthalpic contribution of the binding event, for example by mutating the recognition domain itself, or on the rational design of a conformational state alternative to the binding-state. Both these approaches, however, are not easy to predict and are often based on a trial and error process. As a result, the fine modulation of a synthetic receptor in a straightforward and predictable fashion has proven a difficult and challenging task. The possibility to rationally design intrinsically disordered domains to control the activity and fine tune the dynamic range of synthetic receptors and switches thus appears extremely advantageous. First, as we demonstrated here, the entropic behaviour of a sequence of consecutive thymines could be quite easily predicted thus making the control of the disorder associated to a poly-T linker a straightforward task. Second, the level of accuracy that can be achieved by modulating only the entropic contribution of a linker domain cannot be achieved with other *enthalpy-based* approaches to modulate target affinity. For example, a 1-nucleotide modification in the sequence of a DNA-based receptor that affects a W-C base-pair interaction in the binding domain or in a non-binding conformation will cause a change in the binding free energy for a target that is comprised between 0.6 and 2.3 kcal/mol[59, 60]. Conversely, we have shown here that our purely entropic approach allows to modulate the binding free energy of a DNA-based receptor in a more controllable and precise way. For example, using the clamp-switch receptor and a linker length shorter than 22 nucleotides (a range for which the binding free energy vs linker length dependence is linear), the addition of a single thymine to the linker changes the binding free energy towards the 10-nt target of an average value of 0.14 ± 0.02 kcal/mol.

The use of entropy or disorder to control DNA-based reactions or assembly has seen only limited applications. Examples in this direction include the demonstration of DNA-based reactions controlled by the entropy gain of the released molecules rather than enthalpy gain of base pair formation[61] or

the control of the hybridization kinetics through the use of disordered spacers in the recognition domain[62]. In a similar approach, Sleiman and co-workers have also shown that the intrinsic disorder of single stranded DNA sequences can be used to change the conformation bending of DNA nanostructures[63]. Despite these inspiring examples, however, entropy-based strategies have received limited attention and their use to achieve control and predictable optimization of the input/output response of bio-sensing DNA-based devices in a versatile, precise and tunable fashion has not been yet demonstrated.

While we use here DNA as material to build our nanodevices we also note that the same approach could in principle be easily applied to other synthetic biosensors including those based on peptides and other biopolymers[64-66]. And, exactly as it occurs with several naturally-occurring biomolecular receptors, it would in principle be straightforward to dynamically modulate the entropic contribution of intrinsically disordered domains with external triggers (i.e. pH, temperature, ionic strength, etc) to allow a further level of control of the nanodevice activity. Finally, we believe that the recreation of such mechanism in simple and versatile synthetic systems could also represent an ideal way to disclose and better understand the function and the mechanism with which intrinsically disorder regions control the activity of proteins.

METHODS

Chemicals

All reagent-grade chemicals, including Tris HCl, $MgCl_2$ NaCl, Phosphate buffered saline, Urea and ATP (Adenosine 5'-triphosfate) (all from Sigma-Aldrich, St. Louis, Missouri), were used as received.

Oligonucleotides and DNA-based receptors

HPLC purified oligonucleotides were purchased from IBA, (Gottingen, Germany) or Biosearch Technologies (Risskov, Denmark). The DNA probes were modified with Alexa Fluor 488 (A-488), Alexa Fluor 647 (A-647) or Alexa Fluor 680 (A-680) at 5' end, and Black Hole Quencher 1 (BHQ-1) or Black Hole Quencher 2 (BHQ-2) at 3' end or internally to the sequence. All oligonucleotides were dissolved in TE buffer (100 mM Tris buffer, 10 mM $MgCl_2$, pH 7.8) at a concentration of 100 μM and frozen at -20 °C. The final concentration of the oligonucleotides was confirmed using Tecan Infinite M200pro (Männedorf, Switzerland) through NanoQuant Plate™ measuring the relative absorbance at 260 nm. The sequences and the relative modifications are reported below.

Clamp-switch receptors and targets.

For titration experiments the following clamp-switch receptors and ligand targets were employed:

8nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** CTTTTTTC *CTC TCT TTC C* T(BHQ-1)-3'

10nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** CTTTTTTTTC *CTC TCT TTC C* T(BHQ-1)-3'

12nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** CTTT TTT TTT TC *CTC TCT TTC C* T(BHQ-1)-3'

16nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** CTTT TTT TTT TTT TTC *CTC TCT TTC C* T(BHQ-1)-3'

18nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** CTTT TTT TTT TTT TTT TC *CTC TCT TTC C* T(BHQ-1)-3'

20nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT C *CTC TCT TTC C* T(BHQ-1)-3'

22nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTC *CTC TCT TTC C* T(BHQ-1)-3'

24nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTT TC *CTC TCT TTC C* T(BHQ-1)-3'

26nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTT TTT C *CTC TCT TTC C* T(BHQ-1)-3'

28nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTT TTT TTC *CTC TCT TTC C* T(BHQ-1)-3'

32nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT C *CTC TCT TTC C* T(BHQ-1)-3'

36nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TC *CTC TCT TTC C* T(BHQ-1)-3'

38nt-linker clamp-switch: 5'-(A-488) T **CCT TTC TCT C** C TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT C *CTC TCT TTC C* T(BHQ-1)-3'

For the above sequences the portion in bold represents the triplex-forming domain, the portion in italics the duplex forming domain and the underlined portion the linker domain.

10-nt perfect match ligand target: 5'-GGA AAG AGA G-3'

10-nt mismatch ligand target: 5'- GGA AAT AGA G-3'

11-nt perfect match ligand target: 5'-AGGAAAGAGAG-3'

For thermodynamic characterization (Figure 4) the following unimolecular switches were employed:

8nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TC *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

12nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TTT TTC *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

16nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TTT TTT TTT C *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

22nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TTT TTT TTT TTT TTT C *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

24nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TTT TTT TTT TTT TTT TTC *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

28nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TTT TTT TTT TTT TTT TTT TTT C *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

32nt-linker unimolecular switch: 5'-(A-488)T **CCT TTC TCT C** CTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TC *CTC TCT TTC C* TTTTTT(BHQ-1)A *GGA AAG AGA* G-3'

For the above sequences the portion in bold represents the triplex-forming domain, the two portions in italics are complementary to each other and form the duplex and are separated by a 7-nt loop and the underlined portion represents the linker domain.

To estimate K_d^{target} the following control switches were employed:

8nt-linker control switch (K_d^{target}): 5'- **GTT TGT TGT TT** C TTT TTT(A-488)C *CTC TCT TTC C* T-(BHQ-1)-3'

36nt-linker control switch (K_d^{target}): 5'- **GTT TGT TGT TT** C TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT T (A-488)C *CTC TCT TTC C* T-(BHQ-1)-3'

For the above sequences the portion in bold represents a random domain, the portion in italics is the duplex-forming domain and the underlined portion represents the linker domain.

To estimate K_d^{inter} the following hairpin duplex probe and ligand strand were employed:

Hairpin duplex probe (K_d^{inter}): 5'-(A-488)TAG GAA AGA GAG GTT TTT CCT CTC TTT CCT T-3'

Target K_d^{inter} 10-nt: 5'-(A-647)TC CTT TCT CTC T-3'

6nt-linker clamp-switch incompetent receptor: 5'-TCC TTT CTC TCC TTT TTT CCT CTC TTT CCT-3'

ATP-binding aptamers:

For titration experiments the following ATP-binding aptamers were employed:

4nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT T TTT **TGC GGA GGA AGG A**(BHQ-1)-3'

8nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT T TTT TTT T **TG CGG AGG AAG GA**(BHQ-1)-3'

16nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT T TTT TTT TTT TTT TTT **TGC GGA GGA AGG A**(BHQ-1)-3'

32nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT T TTT TTT TTT TTT TTT TTT TTT TTT TTT T **TG CGG AGG AAG GA**(BHQ-1)-3'

40nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT T TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT **TGC GGA GGA AGG A**(BHQ-1)-3'

54nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT **TG CGG AGG AAG GA**(BHQ-1)-3'

70nt-linker ATP aptamer: 5'-(A-488) ACC TGG GGG AGT AT T TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT **TGC GGA GGA AGG A**(BHQ-1)-3'

For the above sequences the portions underlined and in bold represent the two split binding domains of the ATP-binding aptamer separated by the poly-T linker domain.

Fluorescent experiments

Titration experiments with clamp-switch receptors were conducted in 50 mM Tris HCl buffer, or in 50 mM Tris HCl and 10 mM $MgCl_2$ buffer at 37ºC in a 800 µL cuvette using 3 nM or 10 nM of clamp-switch receptor at the indicated pH. Equilibrium fluorescence measurements were obtained using a Cary Eclipse Fluorimeter with excitation at 490 (± 5) nm and acquisition at 517 (± 5) nm (for DNA strands labeled with A-488 and BHQ-1) or with excitation at 468 (± 5) nm and acquisition at 665 (± 5) nm (for DNA strands labeled with the FRET pair A-488/A-647). For each concentration the fluorescence signal was recorded every 10 minutes until it reached equilibrium.

For the titration experiments the observed fluorescence data, $F_{[T]}$, were fitted with the following simplified "Langmuir-type" single site binding equation:

$$F_{[T]} = F_0 + \frac{[\text{Target}](F_B - F_0)}{[\text{Target}] + K_d^{Langmuir}} \qquad \text{Eq. 6}$$

Where [Target] = target concentration; F_B = fluorescence in the presence of saturating concentration of target; $F_{[T]}$ = fluorescence in the presence of different concentration of target; F_0 = background fluorescence and $K_d^{Langmuir}$ is the equilibrium target concentration at half-maximum signal. This model is not necessarily physically relevant, but it does a good (empirical) job of fitting effectively bi-linear binding curves such as those we obtain for most of our nanoswitches, providing a convenient and accurate means of estimating the observed dissociation constant.

Unimolecolar Triplex melting curves were conducted at pH 6.5 in 10 mM PBS, 200 mM NaCl, and 10 mM $MgCl_2$ buffer using 50 nM of the unimolecular control clamp-switch in an 800 µL cuvette. The gradient was fixed at 0.4 °C/min.

ATP binding curves were conducted at pH 6.5 in 100 mM Tris HCl and 10 mM $MgCl_2$ buffer at 37 °C using 50 nM of the clamp-switch in an 800 µL cuvette.

ATP urea denaturation curves were conducted at pH 6.5 in 100 mM Tris HCl, 10 mM $MgCl_2$ and 10 M Urea at 37 °C using 50 nM of the ATP-binding aptamer switch in an 800 µL cuvette.

Thermal melting curves

Fluorescence versus temperature profiles (thermal melting curves) were obtained using a Cary Eclipse Fluorimeter (Agilent Technologies) with an excitation wavelength at 490 (± 5) nm and an acquisition wavelength at 517 (± 5) nm. Melting curves were performed by heating from 15 to 95 °C at a rate of 0.4 °C·min^{-1} using a total reaction volume of 800 µl in a quartz cuvette. To limit the evaporation of the sample during the experiment a thin layer of mineral oil to the top of the solution was added. The stock solution of the triplex clamp-switch unimolecular receptors were diluted in 10 mM phosphate buffer + 200 mM NaCl + 10 mM $MgCl_2$ at pH 6.5 at a final concentration of 50 nM. Before the experiment the solutions were heated to 95 °C for 5 min and then allowed to cool to room temperature for 1 h.

All the reported melting curves have been normalized through the use of the interpolation model[38] that allows to estimate the melting temperature (T_m) for each experiment. Two baselines (upper and lower) have been chosen as straight lines fitting the fluorescence signal before and after the melting transition. Such baselines correspond to the unfolded (duplex) and folded (triplex) states, respectively. Trough the average of the estimated baselines it is possible to calculate a median line. Such median line will be drawn within the two baselines crossing the experimental curve in the melting transition region. The T_m will correspond to the crossing point between the experimental curve and the median line, and its uncertainty is estimated at ± 0.5 °C [38].

ACKNOWLEDGEMENTS

This work was supported by the European Research Council, ERC (project n. 336493, Nature Nanodevices) (FR), (project n. 677313 BioCircuit) (T.d.G.), by Associazione Italiana per la Ricerca sul Cancro, AIRC (project n. 14420) (FR), an NWO-VIDI grant from the Netherlands Organization for Scientific Research (NWO, 723.016.003) (T.d.G.), and funding from the Ministry of Education, Culture and Science (Gravity programs, 024.001.035 & 024.003.013) (T.d.G.).

SUPPORTING INFORMATION AVAILABLE

Supporting Information. Thermodynamic analysis of thermal and urea melting curves, Thermodynamic binding model and data analysis, estimation of the effective molarity, supporting figures. This material is available free of charge at http://pubs.acs.org.

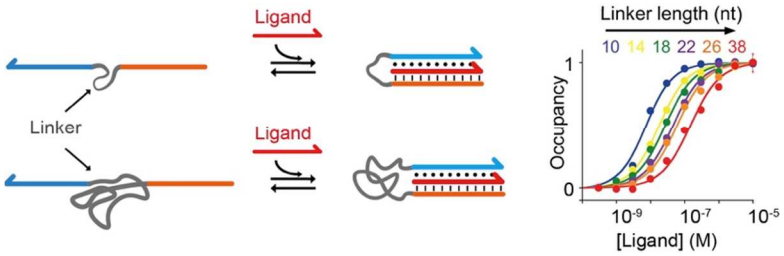
REFERENCES

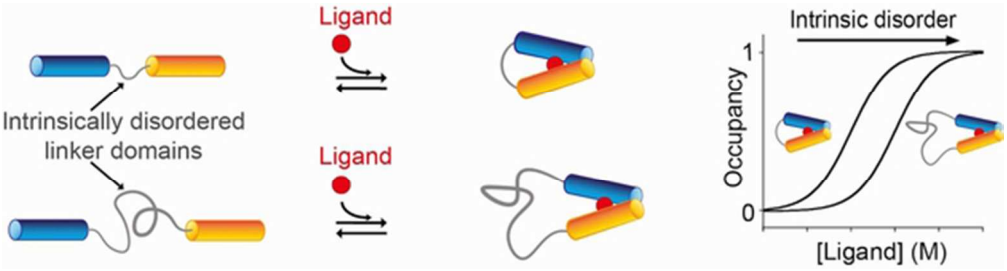
(1) Kinbara, K.; Aida, T. *Chem. Rev.* **2005**, *105*, 1377-1400.

(2) Krishnan, Y.; Simmel, F.C. *Angew. Chem., Int. Ed.* **2011**, *50*, 124-3156.

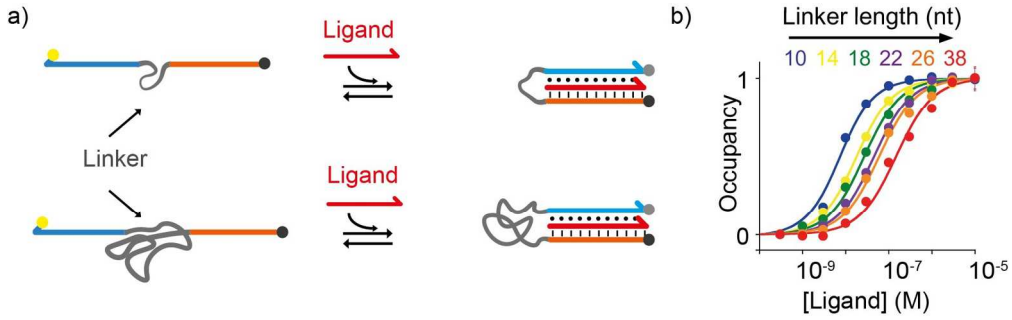
(3) Yang, Y.-W.; Sun, Y.-L.; Song, N. *Acc. Chem. Res.* **2014**, *47*, 1950-1960.
(4) Wang, J. *ACS Nano*. **2009**, *3*, 4-9.
(5) Lutolf, M. P.; Hubbell, J. A. *Nature Biotechnol*. **2005**, *23*, 47-55.
(6) Haupt, K.; Mosbach, K. *Chem. Rev*. **2000**, *100*, 2495-2504.
(7) Looger, L. L.; Dwyer, M. A.; Smith, J. J.; Hellinga, H. W. *Nature*. **2003**, *423*, 185-190.
(8) Levskaya, A.; Weiner, O. D.; Lim, W. A.; Voigt, C. A. *Nature*. **2009**, *461*, 997-1001.
(9) Volgraf, M.; Gorostiza, P.; Numano, R.; Kramer, R. H.; Isacoff, E.Y.; Trauner, D. *Nat. Chem. Biol*. **2006**, *2*, 47-52.
(10) Kurishita, Y.; Kohira, T.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc*. **2010**, *132*, 13290-13299.
(11) Van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. *J. Am. Chem. Soc*. **2017**, *139*, 13960-13968.
(12) Hancock, R. D. *Chem. Soc. Rev*. **2013**, *42*, 1500-1524.
(13) Jeong, B.; Gutowska, A. *Trends Biotechnol*. **2002**, *20*, 305-311.
(14) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc*. **2013**, *135*, 9222-9239.
(15) Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N; Obradovic, Z. *J. Prot. Res*. **2007**, *6*, 1882-1898.
(16) Pelka, P.; Ablack, J. N. G.; Fonseca, G. J.; Yousef, A. F.; Mymryk, J. S. *J. Virol*. **2008**, *82*, 7252-7263.
(17) Dunker, A. K.; Brown, C.J.; Lawson, J. D.; Iakoucheva, L. M.; Obradović, Z. *Biochemistry*. **2002**, *41*, 6573-6582.
(18) Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.; Obradovic, Z.; Dunker, A. K. *Nucleic Ac. Res*. **2004**, *32*, 1037-1049.
(19) Ferreon, A. C.; Moran, C. R.; Gambin, Y.; Deniz, A. A. *Methods Enzymol*. **2010**, *472*, 179-204.
(20) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. *Nature*. **2014**, *508*, 331-339.
(21) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol*. **2009**, *5*, 789-796.
(22) Tzeng, S.-R.; Kalodimos, C. G. *Nature*. **2012**, *488*, 236-240.
(23) Hilser, V. J. *Nature*. **2013**, *498*, 308-310.
(24) Hilser, V. J.; Thompson, E. B. *Proc. Natl. Acad. Sci. U.S.A*. **2007**, *104*, 8311-8315.
(25) Ricci, F.; Vallée-Bélisle, A.; Simon, A. J.; Porchetta, A.; Plaxco, K. W. *Acc. Chem. Res*. **2016**, *49*, 1884-1892.
(26) Lin, M.; Song, P.; Zhou, G.; Zuo, X.; Aldalbahi, A.; Lou, X.; Shi, J.; Fan, C. *Nat. Protocols*. **2016**, *11*, 1244-1263.
(27) a) Vallée-Bélisle, A.; Plaxco, K. W. *Curr. Opin. Struct. Biol*. **2010**, *20*, 518-526; b) Pardee, K.; Green, A. A.; Takahashi, M. K.; Braff, D.; Lambert, G.; Lee, J. W.; Ferrante, T.; Ma, D.; Donghia, N.; Fan, M.; Daringer, N. M.; Bosch, I.; Dudley, D. M.; O'Connor, D. H.; Gehrke, L.; Collins, J. J10. *Cell*. **2016**, *165*, 1255-1266.
(28) Li, D.; Song, S.; Fan, C. *Acc. Chem. Res*. **2010**, *43*, 631-641.
(29) Idili, A.; Plaxco, K. W.; Vallée-Bélisle, A.; Ricci, F. *ACS Nano*. **2013**, *7*, 10863-10869.
(30) Del Grosso, E.; Idili, A.; Porchetta, A.; Ricci, F. *Nanoscale*. **2016**, *8*, 18057-18061.
(31) Goddard, N. L.; Bonnet, G.; Krichevsky, O.; Libchaber, A. *Phys. Rev. Lett*. **2000**, *85*, 2400-2403.
(32) Mandolini, L. *In Advances in Physical Organic Chemistry*; Gold, V.; Bethell, D.; Eds.; *Academic Press: London*, **1986**, *22*, 1-111.
(33) Winnik, M. A. *Chem. Rev*. **1981**, *81*, 491-524.
(34) Gargano, J. M.; Ngo, T.; Kim, J. Y., Acheson, D. W. K.; Lees, W. J. *J. Am. Chem. Soc*. **2001**, *123*, 12909-12910.
(35) Krishnamurthy, V. M.; Semetey, V.; Bracher, P. J.; Shen, N.; Whitesides, G. M. *J. Am. Chem. Soc*. **2007**, *129*, 1312-1320.
(36) Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. *Biophys. J*. **2004,** *86*, 2530-2537.
(37) Kuznetsov, S. V.; Shen, Y.; Benight, A. S.; Ansari, A. *Biophys. J*. **2001,** *81*, 2864-2875.
(38) Mergny, J.-L.; Lacroix, L. *Oligonucleotides*. **2003**, *13*, 515-537.
(39) a) Mak, C.H.; Phan, E.N.H. *Biophys. J*. **2018,** *8*, 2059-2071; b) Marenduzzo, D., Micheletti, C., Cook, P.R. *Biophys. J*. **2006**, *90*, 3712-3721; c) Kuznetsov, S.V., Shen, Y., Benight, A.S., Ansari, A. *Biophys. J*. **2001**, *81*, 2864-2875.
(40) Jacobson, H.; Stockmayer, W. H. *J. Chem. Phys*. **1950**, *18*, 1600-1606.
(41) Ricci, F.; Vallée-Bélisle, A.; Simon, A. J.; Porchetta, A.; Plaxco, K. W. *Acc. Chem. Res*. **2016**, *49*, 1884-1892.
(42) Idili, A.; Vallée-Bélisle, A.; Ricci, F. *J. Am. Chem. Soc*. **2014**, *136*, 5836–5839.
(43) Huizenga, D. E.; Szostak, J. W. *Biochemistry*. **1995**, *34*, 656-665.
(44) Mo, R.; Jiang, T.; DiSanto, R.; Tai W.; Gu, Z. *Nat. Commun*. **2014**, *5*, 3364.
(45) Zuo, X.; Xiao, Y.; Plaxco, K.W. *J. Am. Chem. Soc.* **2009**, *131*, 6944-6945.
(46) Xia, T.; Yuan, J.; Fang, X. *J. Phys. Chem. B.* **2013**, *117*, 14994-15003
(47) Idili, A.; Ricci, F.; A. Vallée-Bélisle. *Nucleic Ac. Res*. **2017**, *27*, 7571-7580.
(48) Wright, P. E.; Dyson, H. J. *Nat. Rev. Mol. Cell Biol*. **2015**, *16*, 18-29.
(49) Flock, T.; Weatheritt, R. J.; Latysheva, N. S.; Babu, M. M. *Curr. Opin. Struct. Biol*. **2014**, *26*, 62-72.
(50) Ferreon, A. C. M.; Ferreon, J. C.; Wright, P. E.; Deniz, A. A. *Nature*. **2013**, *498*, 390-394.
(51) Dagliyan, O.; Tarnawski, M.; Chu, P.-H.; Shirvanyants, D.; Schlichting, I.; Dokholyan, N. V.; Hahn, K. M. *Science*. **2016**, *354*, 1441-1444.
(52)Elisabeth, M. W; van Dongen, M.; Evers, T. E.; Dekkers, L. M.; Meijer, E. W.; Klomp, L. W. J.; Merkx M. *J. Am. Chem. Soc.* **2007**, *129*, 3494.
(53) Xia, Y.; DiPrimio, N.; Keppel, T. R.; Vo, B.; Keith Fraser, K.; Battaile, K. P.; Egan, C.; Bystroff, C.; Lovell, S.; Weis, D. D.; J. Anderson, C.; Karanicolas, J. *J. Am. Chem. Soc*. **2013**, *135*, 18840-18849.
(54) Ha, J.-H.; Loh, S.N. *Chemistry Eur. J*. **2012**, *18*, 7984-7999.
(55) Marvin, J. S.; Corcoran, E. E.; Hattangadi, N. A.; Zhang, J. V.; Gere, S. A.; Hellinga, H. W. *Proc. Natl. Acad. Sci. U.S.A*. **1997**, *94*, 4366-4371.
(56) Yamazaki, T.; Kojima, K.; Sode, K. *Anal. Chem*. **2000**, *72*, 4689-4693.
(57) Song, P.; Li, M.; Shen, J.; Pei, H.; Chao, J.; Su, S.; Aldalbahi, A.; Wang, L.; Shi, J.; Song, S.; Wang, L.; Fan, C.; Zuo, X. *Anal. Chem*. **2016**, *88*, 8043-8049.
(58) Centola, M.; Valero, J.; Famulok, M. *J. Am. Chem. Soc.* **2017**, *139*, 16044-16047.
(59) SantaLucia Jr. J. *Proc. Natl. Acad. Sci. U.S.A*. **1998**, *95*, 1460-1465.
(60) SantaLucia Jr. J.; Allawi, H. T.; Seneviratne, P. A. *Biochemistry*. **1996**, *35*, 3555-3562.
(61) Zhang, D. Y.; Turberfield, A. J.; Yurke, B.; Winfree, E. *Science*. **2007**, *318*, 1121-1125.
(62) Machinek, R. R. F.; Ouldridge, T. E.; Haley, N. E. C.; Bath, J.; Turberfield, A. J. *Nat. Commun.* **2014**, *5*, 5324.
(63) Aldaye, F. A.; Lo, P. K.; Karam, P.; Mclaughlin, C. K.; Cosa, G.; Sleiman, H. F. *Nat. Nanotechnol.* **2009**, *4*, 349−352.
(64) Choi, J. H.; Laurent, A. H.; Hilser, V. J.; Ostermeier, M. *Nat. Commun*. **2015**, *6*, 6968.
(65) Lim, W.A. *Nat. Rev. Mol. Cell Biol*. **2010**, *11*, 393-403.
(66) Dueber, J. E.; Yeh, B. J.; Chak, K.; Lim, W. A. *Science*. **2003**, *301*, 1904-1908.

TOC Graphic

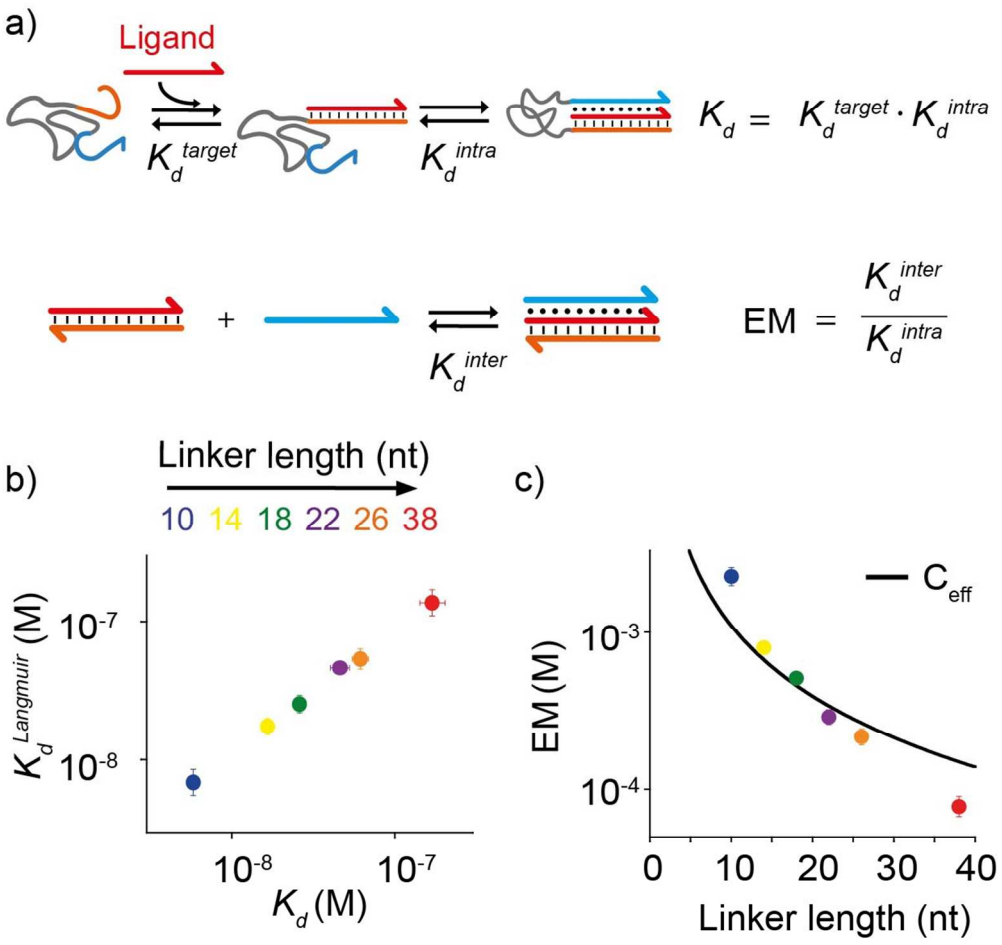

Ligand
Linker
Ligand
Linker length (nt)
10 14 18 22 26 38
Occupancy
1
0
10-9
10-7
10-5
[Ligand] (M)



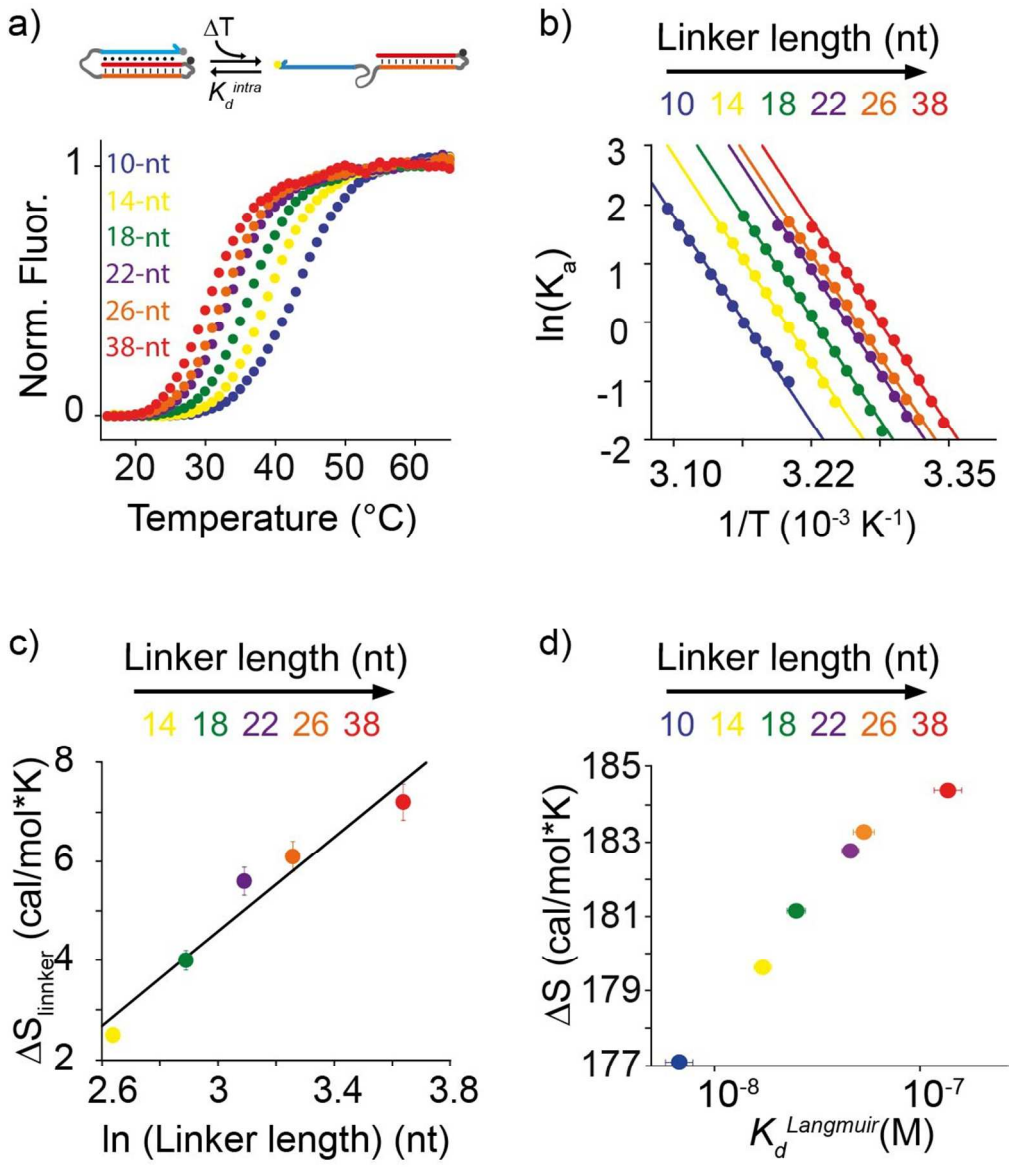
53x14mm (300 x 300 DPI)



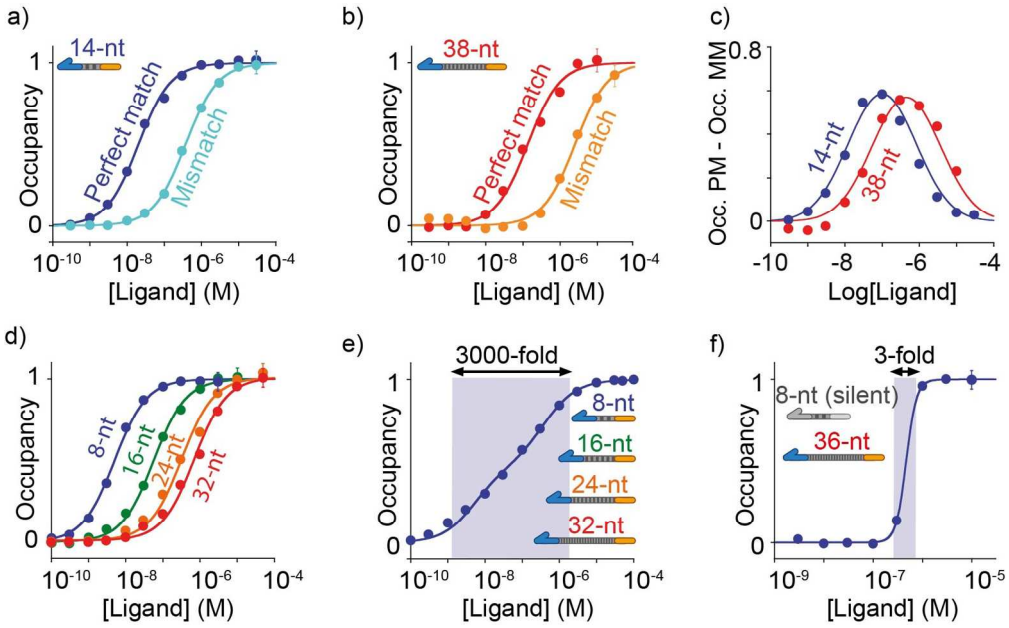
666x209mm (72 x 72 DPI)



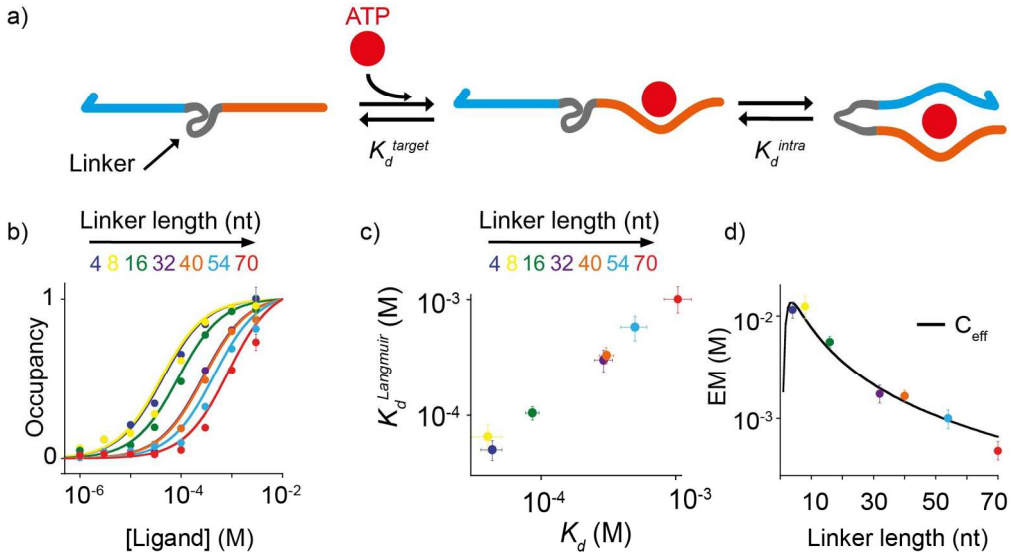
437x410mm (72 x 72 DPI)



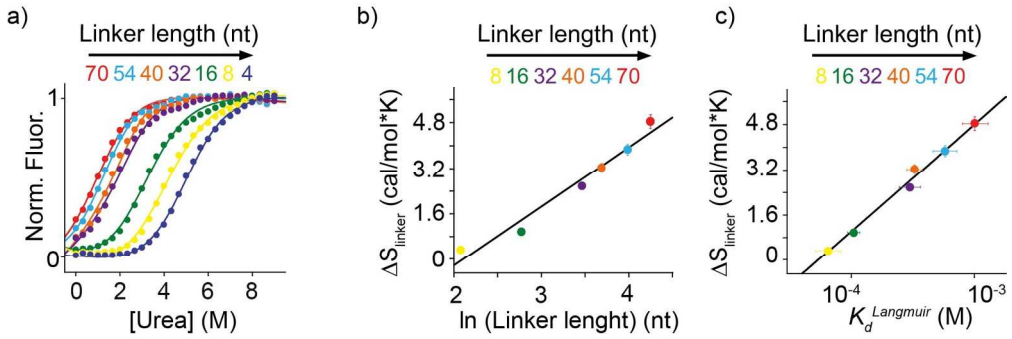
420x490mm (72 x 72 DPI)



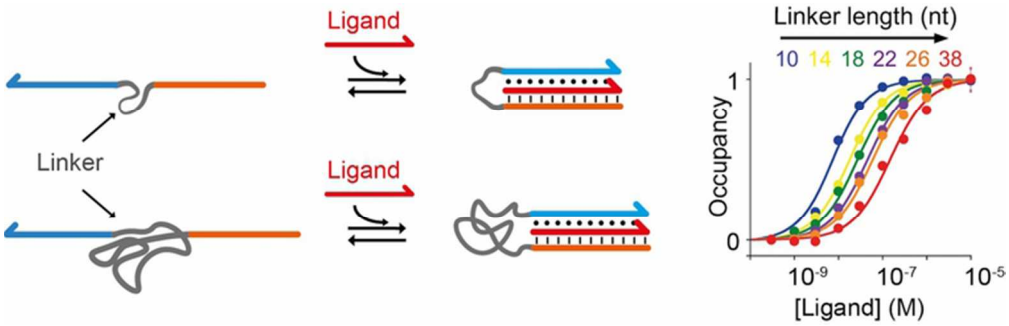
644x399mm (72 x 72 DPI)



659x366mm (72 x 72 DPI)



664x224mm (72 x 72 DPI)



TOC Figure

64x20mm (300 x 300 DPI)