nature chemistry

Article

https://doi.org/10.1038/s41557-022-01127-4

Functional advantages of building nanosystems using multiple molecular components

Received: 26 January 2022

Accepted: 15 December 2022

Published online: 09 February 2023

Check for updates

D. Lauzon & A. Vallée-Bélisle

Over half of all the natural nanomachines in living organisms are multimeric and likely exploit the self-assembly of their components to provide functional benefits. However, the advantages and disadvantages of building nanosystems using multiple molecular components remain relatively unexplored at the thermodynamic, kinetic and functional levels. In this study we used theory and a simple DNA-based model that forms the same nanostructures with different numbers of components to advance our knowledge in this area. Despite its lower assembly rate, we found that a system built with three components may undergo a more cooperative assembly transition from less preorganized components, which facilitates the emergence of functionalities. Using simple variations of its components, we also found that trimeric nanosystems display a much higher level of programmability than their dimeric counterparts because they can assemble with various levels of cooperativity, self-inhibition and time-dependent properties. We show here how two simple strategies (for example, cutting and adding components) can be employed to efficiently programme the regulatory function of a more complex, artificially selected, RNA-cleaving catalytic nanosystem.

Finely regulated self-assembled molecular systems, or nanosystems, are central to life and are increasingly important in nanotechnology[1,2]. In living organisms, nanosystems have evolved to respond precisely to specific variations in stimuli, for example, temperature, pressure, light, pH, osmolarity, small molecules and macromolecules[3–5]. These nanosystems typically self-assemble through the formation of multiple non-covalent interactions, by either intramolecular folding or the intermolecular association of two or more molecular components[6,7]. Of all the proteins characterized so far in living organisms, more than 55% are multimeric and likely exploit molecular assembly to provide functional benefits[7,8]. The tetrameric protein haemoglobin represents a good example of functional adaptation through multimerization. While the monomeric myoglobin is better suited for oxygen storage due to its higher affinity, the tetrameric haemoglobin transports four O_2 molecules using a highly cooperative load-and-release mechanism that can be allosterically controlled through pH and other molecular effectors[9,10]. Inspired by such sophisticated nanosystems, chemists and engineers aspire to develop similar self-regulated systems for various nanotechnological applications, including biosensing, drug delivery and chemical computing[11–13]. However, the advantages or limitations of building such self-regulated molecular systems using either one or multiple molecular components are often overlooked by biologists, engineers and chemists.

Potential advantages and limitations of building self-regulated molecular systems using multiple molecular components have been proposed. These hypotheses, however, remain quite challenging to test by comparing natural monomeric and multimeric proteins given their differentiation throughout evolution[10,14,15]. An obvious advantage

Laboratoire de Biosenseurs & Nanomachines, Département de Chimie, Université de Montréal, Montréal, Québec, Canada.
e-mail: a.vallee-belisle@umontreal.ca

of unimolecular systems is their fast assembly due to typically high intramolecular folding rates[16]. However, large unimolecular systems increase the complexity inherent in their covalent synthesis (multistep, error rate, low yield and so on) and are more likely to misfold, thereby reducing their activity and even leading to harmful misfolded intermediates[17]. Proteins made from longer polypeptide chains, for example, typically require the assistance of complex molecular chaperones to limit misfolding during assembly[17]. This problem can be avoided, in principle, by reducing the sizes and complexities of these large individual components by employing several smaller self-assembling components[18]. Having more molecular components could also increase the regulatory potential of the assembly by rendering its function dependent on the concentration of its components[18–21], but could also lead to unwanted kinetic traps, thus slowing down its assembly[21]. These challenges are also typically observed when assembling artificial nanosystems made of DNA, RNA and proteins, which require long and optimized assembly protocols with often complex temperature annealing or dialysis strategies[22–24].

One strategy to understand the functional advantages of building nanosystems using one or multiple molecular components could consist of exploring the effect of fragmenting a functional unimolecular structure into multiple components. In the past decades, researchers have found that natural proteins and artificial nucleic acid-based systems can readily support fragmentation into two or multiple components (for example, protein-fragment complementary assay and split aptamers)[25–28]. However, so far, it remains unclear whether fragmenting a self-assembled system into more components can enable it to artificially evolve and, therefore, improve its functionality. In this study we explored this hypothesis by studying the thermodynamic and kinetic impacts of fragmenting a self-assembled nanosystem into multiple components that reassemble into the same exact nanostructure. Using theoretical simulations and experiments on two nucleic acid model systems, we have demonstrated how self-assembled nanosystems can be readily programmed to acquire complex regulation mechanisms through a simple fragmentation strategy.

Results

A DNA-based model

To explore whether fragmenting a self-assembled nanosystem into more components could enable the programming of novel regulatory functionalities, we designed a simple DNA three-way junction system that contains programmable interacting surfaces (Fig. 1a). We used DNA to recreate a multicomponent biological assembly instead of proteins due to the high predictability and programmability of its interactions (affinity and kinetic)[29,30], its ease of synthesis[31], its chemical modifications (for example, the addition of fluorophores and quenchers)[32] and its ability to create specific, well-defined and highly soluble structures[22]. The three-way junction, for example, is a widely occurring nucleic acid motif in all kingdoms of life (for example, rRNA, riboswitches, artificial ribozymes and DNAzymes) and an important building block in DNA nanotechnology[33–36]. This motif contains three arms, each of which is separated by a short two-thymine spacer. Each arm is made of ten base pairs (five AT and five GC) and has a predicted free energy of −10 kcal mol^{-1} (dissociation constant, $K_D \approx 100$ nM at 37 °C)[37], which enables assembly within a range of concentrations ideally suited for fluorescence measurements.

Assembly of three-way junctions from one, two or three components

We first characterized the assembly thermodynamics of three-way junctions built with one, two or three components by determining the difference in energy between their assembled and disassembled conformations, ΔG_{Ass}, using chemical and thermal denaturation procedures. Urea denaturation curves typically provide two important parameters: (1) an estimation of ΔG_{Ass} in the absence of urea ($\Delta G°_{Ass}$) and (2) the m value, which correlates with the amount of surface area exposed to solvent upon disassembly (Fig. 1b, top)[38,39]. The one-component system (1C) displays a free energy of assembly of −6.4 ± 0.4 kcal mol^{-1}, which remains substantially lower than the predicted free energy of assembly of the complete system (−22.9 kcal mol^{-1}, Extended Data Fig. 1). This suggests that part of the disassembled state of the 1C system may remain assembled even at high urea concentrations. To verify this hypothesis, we characterized the isolated hairpins and found that they indeed remained assembled even at urea concentrations above 10 M (Supplementary Figs. 1 and 2). The small $\Delta G°_{Ass}$ measured for the 1C nanosystem is also consistent with the value predicted (−5.7 kcal mol^{-1}) for a simple closure of the three-way junction with preorganized hairpins (Fig. 1a and Extended Data Fig. 1). We then cut the 1C system into two components by removing one of the hairpin loops, generating the two-component system (2C). Its $\Delta G°_{Ass}$ (−15.8 ± 0.2 kcal mol^{-1}) and m value (1.61 ± 0.05 kcal mol^{-1} M^{-1}) were approximately twice the values measured for the 1C system (Fig. 1b and Supplementary Table 1). This indicates that the 2C system forms approximately two times more interactions while burying two times more surfaces than the 1C system during its assembly (Supplementary Fig. 3). A similar trend was also observed when fragmenting the nanosystem into three components, with its assembly transition involving nearly three times more interactions ($\Delta G°_{Ass}$ = −21.8 ± 0.9 kcal mol^{-1} M^{-1}) and the burial of three times more surfaces (m value = 2.1 ± 0.5 kcal mol^{-1} M^{-1}) than the 1C system (Fig. 1b). Thermal denaturation experiments (Extended Data Fig. 2) also supported these observations and further revealed that the smaller $\Delta G°_{Ass}$ of the 1C system can be attributed to reductions in both enthalpy (54 ± 9%) and entropy (49 ± 10%) compared with those of the three-component (3C) system (Supplementary Table 1). A structural analysis of these systems using the NUPACK software[40] yielded results that were consistent with these experimentally measured free energies and this hypothesis (Extended Data Fig. 1).

We then determined the assembly kinetics of these nanosystems and found, as expected, that their rates of assembly decreased drastically as the number of components increased (Fig. 1c, top). For example, while 50% of the 1C system folded within 1 ms in the absence of urea (half-life, $t_{1/2}$ = 0.5 ± 0.1 ms), the 2C ($t_{1/2}$ = 32 ± 9 s) and 3C ($t_{1/2}$ = 1,287 ± 378 s) systems assembled five and seven orders of magnitude more slowly at similar concentration, respectively (Fig. 1c, top). This is consistent with the fact that unimolecular reactions are typically faster than multimolecular reactions and that many interactions are already formed in the 1C and 2C nanosystems (that is, preorganized hairpins) before the observed assembly transition takes place. We also compared the disassembly rate of all the three-way junctions at 10 M urea and found that they all disassembled at a similar rate ($t_{1/2}$ = 24 ± 10 ms; Fig. 1c, bottom). This suggests that fragmentation did not substantially impact the disassembly mechanism and half-life of each nanosystem. As noted above, however, the disassembly of the 1C and 2C systems was only partial as their hairpins remained organized above 10 M urea (see Extended Data Fig. 3 for more details). The $\Delta G°_{Ass}$ and m values estimated from the assembly and disassembly kinetics further support the presence of a preorganized structure given that they are within the experimental error of the values determined by equilibrium experiments (Supplementary Table 1)[41].

Taken together, these thermodynamic and kinetic results suggest that fragmenting a nanosystem into multiple components may help to reduce the amount of preorganized structures in the individual components, thus ensuring that the assembly transition maximizes the number of newly formed interactions. More specifically, our results demonstrate that the 1C DNA three-way junction assembles rapidly in a sequential manner through the preorganization of two hairpins followed by closure of the last arm, while the slower assembly of the 3C system displays a more global cooperative assembly, with all its interactions being formed simultaneously in a one-step process. Besides the variation in assembly rate, we further explored whether the presence

https://doi.org/10.1038/s41557-022-01127-4

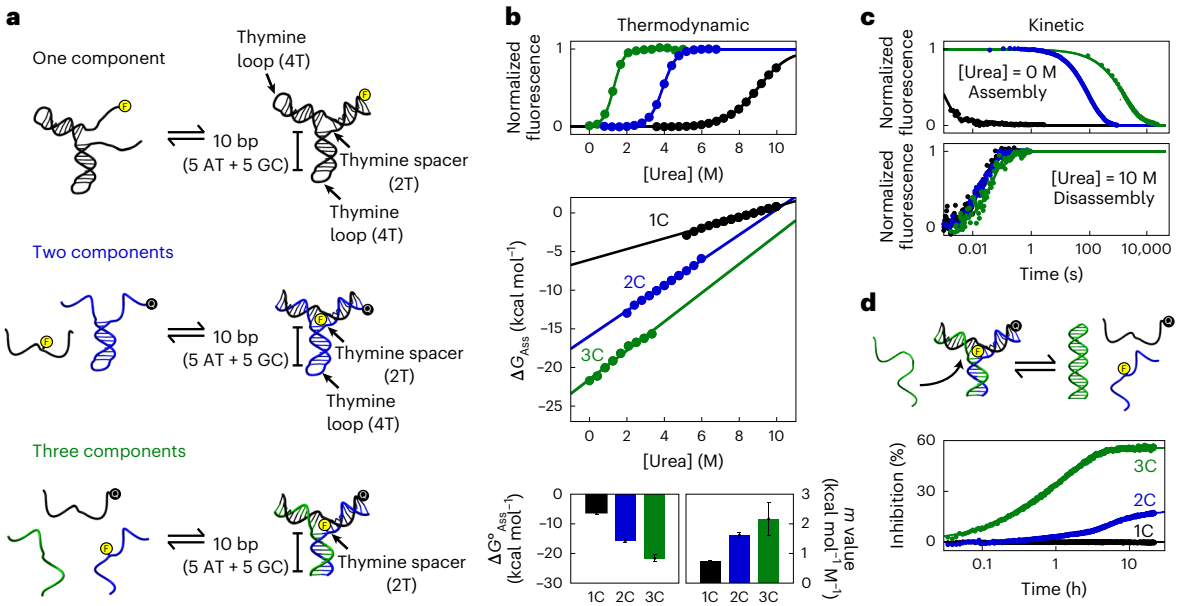


Fig. 1 | Thermodynamic and kinetic characterization of a nanosystem built with one, two or three molecular components. a, A simple DNA-based self-assembled 'three-way junction' nanosystem containing three arms made of ten base pairs (bp) consisting of five AT and five GC. The 2C and 3C systems were obtained through fragmentation by removing either one (2C) or two (3C) thymine loops. All assemblies were monitored using fluorescently labelled DNA strands (the yellow circle represents the fluorophore moiety FAM (fluorescein, F) and the black circle represents the quencher moiety BHQ-1 (black-hole quencher-1, Q)), and the data were normalized accordingly (see Methods). **b**, Thermodynamic analysis of nanosystem assembly using urea equilibrium denaturation curves (see also Extended Data Fig. 2 for the thermal denaturation curves). The free energy of assembly ($\Delta G°_{Ass}$) and the *m* values, which reflect the amount of surface area exposed to solvent upon disassembly, increase with the number of components (see Supplementary Figs. 1 and 2 for raw data and Table 1 for the values). The lower *m* values and energies of assembly of the 1C and 2C systems can be attributed to preorganized structures (hairpins) that remain in the disassembled state (see **a** and Supplementary Fig. 3). The data and errors are presented as the values obtained from a nonlinear regression ($n = 1$). The data for 1C, 2C and 3C nanosystems are indicated in black, blue and green, respectively, in all panels. **c**, Increasing the number of components reduces the rate of assembly (top) but does not affect the rate of disassembly at 10 M urea (bottom). Of note, the assembly parameters ($\Delta G°_{Ass}$ and *m* values) extracted from the urea kinetic traces (Extended Data Fig. 3) are within the experimental error of the values determined by equilibrium experiments (Supplementary Table 1). **d**, Increasing the number of components reduces the presence of preorganized structures, which facilitates the creation of a regulation mechanism via competitive inhibitors.

of preorganized structures in the disassembled state of a nanosystem could impact the regulation of its assembly. As preorganized structures reduce the availability of some interactions, we hypothesized that they could also prevent, or enable, the interaction or binding with other external molecules. To highlight this feature, we designed a complementary DNA strand that binds specifically both hairpin domains (Fig. 1d). We found, indeed, that the 1C system cannot be inhibited by this classic complementary DNA inhibitor due to the low accessibility of the nucleotides locked in the preorganized hairpins. In contrast, the 2C and 3C systems are increasingly more sensitive to the inhibitor because they have more exposed surfaces in their disassembled states (higher *m* value), thus displaying a better ability to develop more regulatory mechanisms.

Effect of number of nanosystem components on assembly mechanism

We then determined the impact of varying the number of nanosystem components on their assembly mechanisms. The assembly mechanism of unimolecular nanosystems, such as the folding of DNA[39] or protein[42] structures, can take place in milliseconds to hours and is generally hard to programme as it mostly depends on the complexity of their structures[43]. Some proteins, for example, start folding co-translationally as soon as they reach the ribosomal exit tunnel, while other unimolecular systems remain under kinetic control[44–47]. Unimolecular nanosystems, nonetheless, need to fold into their native structure to fulfil their functions[47] and their activity generally varies linearly with their concentration (that is, a higher yield of synthesis leads to a higher activity). Typical strategies for regulating these systems generally require the help of external molecules, such as allosteric effectors that will reduce (inhibitor) or enhance (activator) their activity[48]. In contrast, the assembly of the dimeric 2C system can be regulated by tuning the concentration of one of its components (herein called A), resulting in a classic dose–response behaviour (Fig. 2a) with two programmable parameters: (1) the midpoint or $[A]_{50\%}$, that is, the concentration at which 50% of the system is assembled, and (2) the cooperativity of the response or dynamic range (DR), that is, the broadness of the transition, defined as the change in [A] required to provide a change in response from 10% to 90% ($DR = [A]_{90\%}/[A]_{10\%}$). For example, when the 2C system assembles in the presence of a low concentration of one of its components (for example, $[B] = 1\ nM < K_D^{AB}$, the dissociation constant of AB; Fig. 2a, lightest blue curve), its $[A]_{50\%}$ remains constant and equal to the dissociation constant of the system ($[A]_{50\%} = K_D^{AB} = 5.6 \pm 0.5$ nM). In such cases, the DR of the assembly is approximately 81-fold (DR = 69 ± 12-fold). On the other hand, in a saturation regime, that is, when the concentration of component B exceeds K_D^{AB}, $[A]_{50\%}$ increases linearly with [B] ($0.5 \times [B]$), indicating a 1:1 binding regime (Fig. 2c, left)[49]. In such situations, the observed transition becomes more 'cooperative', and the DR is reduced by up to ninefold (Fig. 2c).

The assembly of the trimeric 3C system displays even more programmability by enabling assembly over a wider range of concentrations. For instance, varying the concentration of components B and C at levels lower or higher than the dissociation constant of the dimer BC ($K_D^{BC} = 41 \pm 23$ nM; Supplementary Fig. 7) allows their level of preorganization to be tuned (that is, [BC]). For example, when the

https://doi.org/10.1038/s41557-022-01127-4

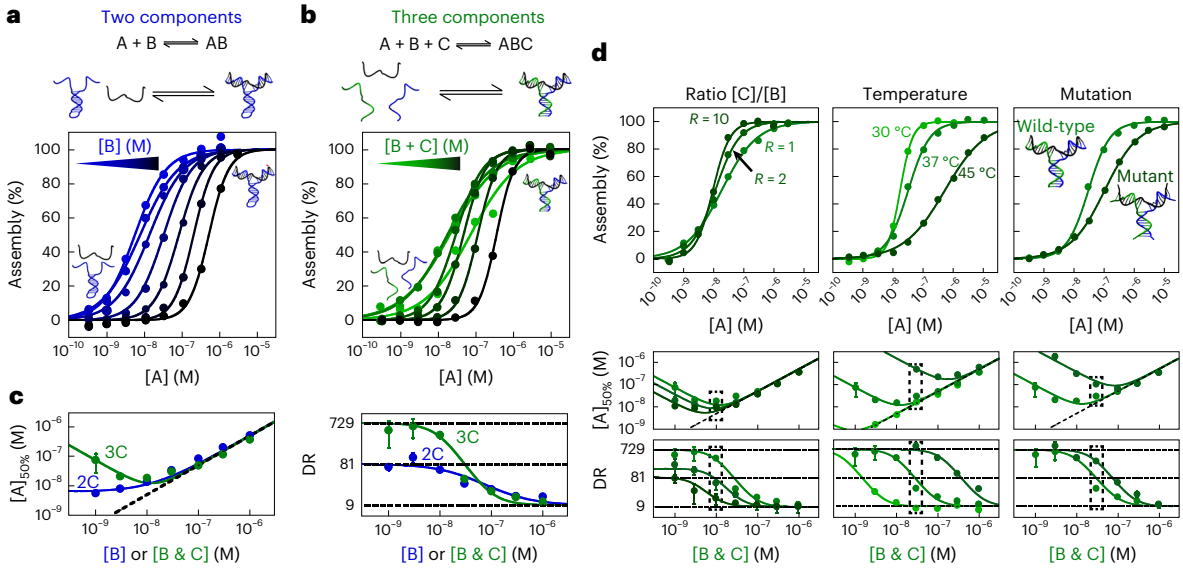


Fig. 2 | The assembly of a nanosystem becomes more programmable (regulable) as its number of components increases. a,**b**, Programming the assembly ($[A]_{50\%}$ and DR) of the 2C (**a**) and 3C (**b**) systems by increasing [A]. Assembly was performed at different fixed concentrations of strand B (2C) or strands B and C (3C). **c**, For the 2C system, as the concentration of strand B increases, the $[A]_{50\%}$ of the 2C system also increases from 5.6 to 516 nM (left), while its DR decreases from 81- to 9-fold (right). For the 3C system, as the concentrations of both strands B and C increase similarly, the $[A]_{50\%}$ of the 3C system shows a more complex relationship, changing from 92 to 18 to 372 nM (left), and its DR decreases from 729- to 9-fold (right). **d**, The assembly profile of the 3C nanosystem can be further programmed by varying the concentration ratio of components B and C (left), changing the temperature (middle) or introducing destabilizing mutations into the duplex arms (right). The $[A]_{50\%}$ and DR data shown in the dotted boxes in the lower plots are extracted from the experimental curves presented above (see Supplementary Figs. 9–11 for the rest of the data). All dose–response curves were fitted using the Hill equation, while the $[A]_{50\%}$ and DR data were fitted using equations derived from the 2C and 3C models. The experimental conditions were optimized to capture the entire assembly process over the experimentally allowed five orders of magnitude of [A]. The data and errors in **c** and **d** are presented as the values obtained from the nonlinear fitting ($n = 1$). The dashed line in panels **c** and **d** of the $[A]_{50\%}$ versus concentration plot represents half the concentration of fixed component used in the assay ($0.5 \times$ [B]) while the dashed line in panels **c** and **d** of the DR versus concentration plot represents the upper and lower limits of the 2C system (81 and 9) and the 3C system (729 and 9), respectively.

concentrations of B and C are higher than K_D^{BC}, the dimer preforms in a state similar to that of the 2C system, and the 3C system consequently behaves similarly to the 2C system (that is, a cooperative DR of ninefold with $[A]_{50\%}$ equal to $0.5 \times$ [B]; Fig. 2a,b, darker curves, and Fig. 2c). In contrast, for concentrations of B and C lower than K_D^{BC}, the system is not preorganized (that is, B and C remain dissociated) and the assembly displays a much higher $[A]_{50\%}$ and a broader DR because the assembly proceeds directly from the monomers to the trimer without forming dimeric intermediates (Supplementary Fig. 8). For example, at 1 nM B and 1 nM C (Fig. 2b, lightest green curve), the 3C system displays a $[A]_{50\%}$ of 76 ± 48 nM (compared with 5.6 ± 0.5 nM for 2C) with a broad anticooperative DR close to 729-fold (compared with 81-fold for 2C; Fig. 2c). These assembly behaviours of the 2C and 3C systems are also well modelled by numerical simulations (Extended Data Fig. 4).

Including more components in the assembly process helps to optimize the assembly profile and therefore conveniently tunes $[A]_{50\%}$ and DR. For example, simply increasing the ratio between components B and C (R = [C]/[B]) shifts the assembly profile from anticooperative behaviour to a typical 81-fold DR (for example, from DR = 383 ± 62 to DR = 57 ± 25) without affecting $[A]_{50\%}$ (Fig. 2d, left). This provides a useful strategy for specifically programming either $[A]_{50\%}$ or the DR independently. Alternatively, the assembly profile of a 3C system can be optimized by tuning the component B and C preorganization levels (that is, [BC]), which can be readily achieved by changing the temperature (Fig. 2d, middle) or through specific mutations (Fig. 2d, right). For example, increasing the temperature to 45 °C destabilizes the preorganized BC dimer by one order of magnitude (K_D^{BC} increases from 41 ± 23 to 547 ± 254 nM; Supplementary Fig. 7), thereby shifting $[A]_{50\%}$ and the DR profiles to the same extent. Decreasing the temperature to 30 °C has the opposite effect (K_D^{BC} decreases from 41 ± 23 to 3.8 ± 2.4 nM; Supplementary Fig. 7). Mutations can also be used to achieve a similar level of programmability. For example, when destabilizing the duplex hairpin by inserting one mismatch at each extremity (K_D^{BC} increases from 41 ± 23 nm to 3.3 ± 2.3 μM; Supplementary Fig. 7), the $[A]_{50\%}$ and DR profiles also shift to higher concentrations. Of note, the high preorganization of components B and C can help to mitigate the effects of perturbing mutations. For example, when using B and C at 1,000 nM, the wild-type and mutant show similar $[A]_{50\%}$ values (372 ± 22 versus 349 ± 42 nM) and DR (10 ± 1 versus 10 ± 2). We also used our numerical simulations to gain insights into a system that can be potentially divided into more than three molecular components. For example, when having four molecular components, the $[A]_{50\%}$ of the nanosystem follows a steeper bell-shaped trend than the 3C system, while the DR shifts from 9, at high concentration of components, to a much broader DR of 6,561 at low concentration of components (see Supplementary Fig. 12 for more details). Overall, these results demonstrate that the process of assembling nanostructures made from more than two components can be programmed to have substantially more diverse regulatory profiles than a structurally similar 1C or 2C system.

Effect of mutations on assembly

Specific mutations can have an even more drastic impact on the assembly profile of a 3C system. While the assembly of a 2C system consists of the simple dimerization of two monomers, that of a 3C system involves the assembly of multiple dimeric intermediates before the formation of the trimeric structure (Fig. 3a, top). We thus explored whether mutations that specifically affect the stability of the trimeric three-way junction ($\Delta G°_{Tri}$), without affecting the stability of the dimeric intermediates

https://doi.org/10.1038/s41557-022-01127-4

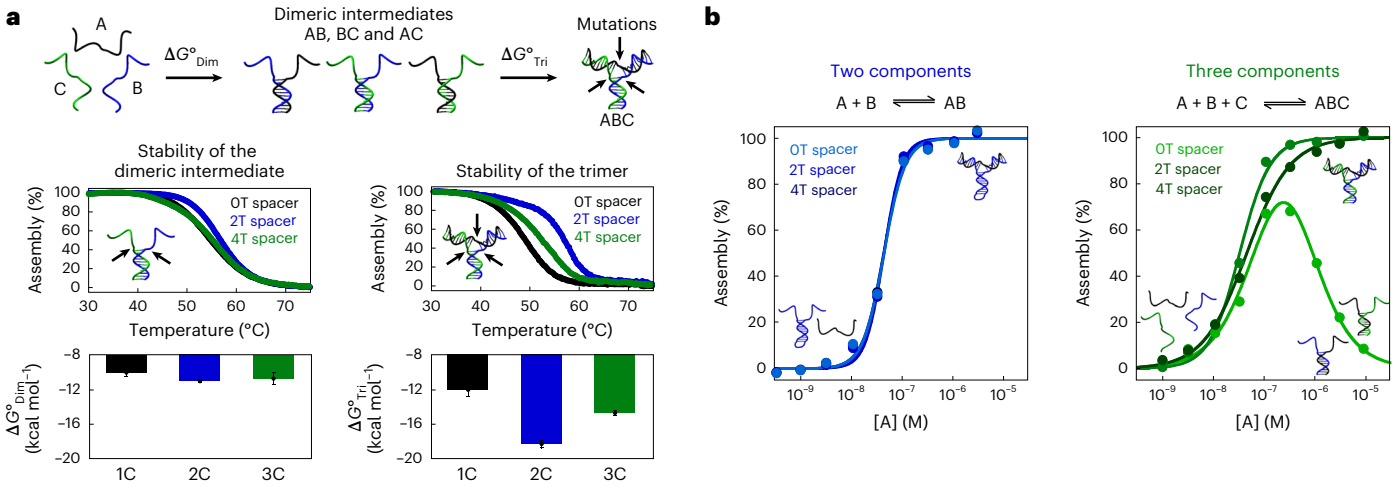


Fig. 3 | Specific mutations on trimeric assemblies are more impactful than on their dimeric counterparts and can lead to novel regulatory mechanisms. **a**, The assembly of a trimeric nanosystem is achieved via dimeric intermediates (top). To programme the assembly profile of a trimeric nanosystem, we introduced insertion or deletion mutations into all spacers connecting the duplex arms by using either 0, 2 or 4 thymines (0T, 2T and 4T). Thermal denaturation confirms that these mutations specifically affect the trimeric stability ($\Delta G°_{Tri}$) without destabilizing the dimeric intermediates ($\Delta G°_{Dim}$; bottom). The data and errors are presented as the values obtained from linear regression ($n = 1$; see Extended Data Fig. 5). **b**, The assembly of the 2C system is insensitive to spacer length mutations (left), whereas the same mutations in the 3C system result in new assembly profiles (right). For example, a low $\Delta G°_{Tri}$ (0T) creates a self-inhibited trimeric system that disassembles into two dimers (AB and AC) at a higher [A]. The inhibition of the trimeric assembly, leading to dimers, was confirmed by polyacrylamide gel electrophoresis (see Extended Data Fig. 7).

($\Delta G°_{Dim}$), could provide a novel avenue for diversifying functionality. We introduced simple insertion or deletion mutations at the junction between the duplex arms by either decreasing the spacer length from two thymines (2T, wild-type) to none (0T, deletion) or by increasing it to four thymines (4T, insertion)[50,51]. Using thermal denaturation (Fig. 3a, top, and Extended Data Fig. 5), we confirmed that these mutations do not affect the stability of the dimeric intermediates, but decrease the trimeric stability by 3.6 ± 0.6 kcal mol^{-1} (4T) and 6.3 ± 1.1 kcal mol^{-1} (0T). These mutations drastically impact the assembly profile of the 3C system, despite having no effect on the assembly of the structurally similar 2C system (Fig. 3b, left). While the most stable 2T trimeric system assembles in a highly cooperative manner (DR = 21 ± 5), the mildly destabilized 4T system displays less cooperative assembly (DR = 73 ± 18), whereas the most destabilized 0T system assembles via a novel self-inhibited profile ('none-all-none' mechanism[52]; Fig. 3b, right). In this self-inhibitory profile, the percentage of assembled trimer increases from 10% to 70% when [A] is increased from 5 to 200 nM and then decreases to 10% when [A] is further increased to 10 μM (DR = 2,000-fold, see also how to programme this self-inhibitory DR in Extended Data Fig. 6). This mechanism occurs when the stability of the trimeric system is similar to that of the dimeric intermediate (for example, for the 0T system, $\Delta G°_{Tri} - \Delta G°_{Dim} = -2.0 \pm 1.1$ kcal mol^{-1}). In such cases, the formation of dimeric intermediates becomes thermodynamically favoured at higher [A] (see Extended Data Fig. 7 for proof of the mechanism). In contrast, the same mutations have no effect on the assembly profile of the 2C system because they do not sufficiently destabilize the 2C system and its K_D to affect its assembly profile. Taken together, these results demonstrate that in addition to permitting novel regulation mechanisms, trimeric assemblies are also much more sensitive to small perturbations or mutations than their dimeric counterparts.

Kinetic control over assembly

While the regulatory mechanisms discussed above all take place at equilibrium, the assembly of nanosystems with multiple components can also be controlled by kinetics[21,53,54]. As mentioned previously, unimolecular systems (1C) typically fold rapidly unless misfolding occurs[16]. Increasing the number of components to two can easily programme the assembly rate by exploiting the law of mass action. For example, increasing the concentration of A from 100 to 600 nM with a low concentration of B (10 nM) increases the observed assembly rate (k_{obs}) sixfold (Fig. 4a, left). For the 3C system, however, mixing an excess of component A (100–600 nM) with lower concentrations of B and C (10 nM) traps these two components in the non-functional dimeric intermediates AB and AC (Fig. 4a, right). Dissociation of these dimers is then required to enable trimer formation by the slow association of the formerly sequestered components B and C, thus resulting in biphasic kinetics. We confirmed that the fastest phase, k_{Dim}, represents the formation of dimers AB and AC, given their linear dependency on the concentration of monomer A. In contrast, the slowest phase, k_{Tri}, represents the formation of the trimer and is rate-limited by the formation of the productive dimer BC, thus explaining its insensitivity to the concentration of component A.

The kinetic analysis (Fig. 4a) and numerical simulations (Supplementary Figs. 13 and 14 and Extended Data Fig. 8) of trimeric assembly through the formation of dimeric kinetic traps provide strategies for precisely programming the assembly rate of nanosystems. For instance, simply increasing [A] can increase the rate of formation (activation) and the percentage of the transient intermediates AB and AC without affecting the trimer assembly rate (Fig. 4b, left). In contrast, increasing [B] and [C], and therefore their level of preorganization, can increase the trimer assembly rate (Fig. 4b, middle). Notably, increasing the rate of trimer formation also decreases the percentages of the transiently formed dimeric intermediates AB and AC, and specific mutations have similar effects (Fig. 4b, right, and Supplementary Fig. 17). For example, a mutation specifically destabilizing the transient intermediates AB and AC prevents the formation of these dimers without substantially affecting the rate of trimer formation, and k_{Tri} is reduced by only fivefold (Fig. 4b, right, top). On the other hand, a mutation specifically destabilizing the productive BC intermediate traps the nanosystem into the AB and AC dimers and prevents trimer formation (Fig. 4b, right, bottom). Similar trends and regulatory opportunities were also observed for nanosystems built with more than three molecular components. Numerical simulations of a nanosystem built with four components, for

https://doi.org/10.1038/s41557-022-01127-4

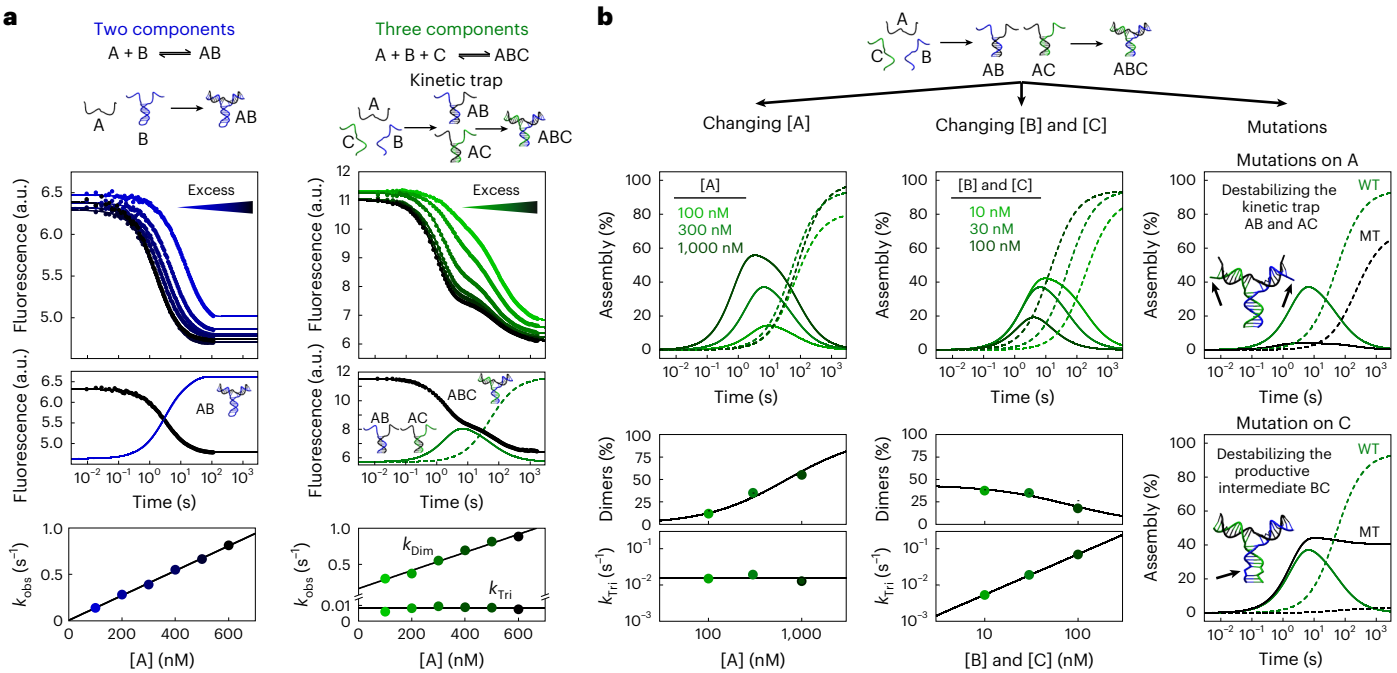


Fig. 4 | Programming the assembly of trimeric nanosystems over time by exploiting dimeric kinetic traps. a, For the 2C system, at a low concentration of monomer B (10 nM), the addition of an excess of monomer A (100 to 600 nM) results in increasingly faster single-phase kinetics (left). For the 3C system, at low concentrations of monomers B and C (10 nM, mainly unbound), the addition of an excess of monomer A (100–600 nM) results in slower biphasic kinetics due to the sequestration of monomers B and C in the dimeric intermediates AB and AC. Trimer formation then proceeds through a slow component-exchange mechanism limited by the availability of free components B and C (right). **b**, Tuning the concentration of A can programme the rate of formation (activation) and the percentage of dimers formed without affecting the rate of trimeric assembly (left). Increasing the concentrations of B and C, and thereby increasing their level of preorganization, can increase the rate of trimer assembly and decrease the percentage of transiently formed dimers (middle). Mutations specifically destabilizing the dimeric intermediates AB and AC (top) or the productive intermediate BC (bottom) prevent dimeric intermediate and trimer formation, respectively (right). WT, wild-type; MT, mutant. See Supplementary Figs. 15–17 for the raw data for **b**.

example, revealed more assembly pathways, leading to a greater variety of programmable time-dependent activation/deactivation profiles (see Supplementary Fig. 18 for more details). Having more components, however, can also increase the opportunities to form unwanted and harmful intermediate assemblies. All these results exemplify that building a nanosystem with multiple molecular components can generate programmable time-dependent nanosystems by exploiting the law of mass action and simple kinetic traps. Such time-dependent nanosystems are present in various biochemical processes, such as signal transduction, protein synthesis and the cell cycle, and will likely have promising applications in future self-regulated nanotechnologies[55–57].

Effect of fragmentation on functionality of a catalytic nanosystem

In this study, our three-way junction served as a convenient synthetic toolkit to quantitatively evaluate the impact of fragmentation and the number of components on the functionality of a nanosystem, and the extent to which this impact can be harnessed to create novel regulatory mechanisms. To test the generality and predictability of these findings, we used these rules to programme the catalytic activity of NaA43, a sodium-specific RNA-cleaving DNAzyme previously used as a sensor to monitor the sodium concentration inside cells[58]. We first measured the apparent activity of NaA43 and determined an apparent Michaelis constant, K_M^{App}, of 11 ± 1 nM, corresponding to approximately half of the DNAzyme concentration used in our assay (30 nM), and a DR of 11 ± 4 (Fig. 5a, left). We also estimated $K_D^{DNAzyme}$, the dissociation constant for the substrate and DNAzyme, to be in the femtomolar range (Supplementary Fig. 19), which confirms why this DNAzyme/substrate system operates in the saturation regime ([DNAzyme] > $K_D^{DNAzyme}$; see also Fig. 5a, left). We then enabled this catalytic nanosystem to evolve artificially by splitting the DNAzyme into two halves at the extremity of its loop (Fig. 5a, right). This modification still yields a functional DNAzyme despite a 36 ± 2% reduction in catalytic activity (Supplementary Fig. 20). We first explored the effect of DNAzyme concentration on the activity of this fragmented DNAzyme. At a high concentration of DNAzyme components (1 μM), which ensures high preorganization, $[A]_{50\%}$ remains close to half the concentration of the DNAzyme (778 ± 53 nM) with a 'cooperative' DR of 12 ± 3). In contrast, as predicted by our model (Fig. 4b), using a low concentration of DNAzyme components (that is, 30 nM) produces a kinetic trap, which can be used to programme a substrate inhibition regulatory mechanism (that is, 'none-all-none' regulation), where both DNAzyme components become sequestered into non-functional dimers at high substrate concentrations (Supplementary Fig. 21).

The function of this fragmented catalytic system can be further optimized through single point mutations. For example, when implementing a stabilizing mutation in the loop of the fragmented DNAzyme that increases its preorganization level (A → C, blue strand Fig. 5b), we increased the activity level by 25 ± 1% while maintaining the self-inhibition mechanism. In contrast, when implementing a mutation that destabilizes the kinetic trap (G → C, green strand, Fig. 5b), we disrupted the self-inhibition mechanism at the cost of reduced activity (53 ± 2% slower). On combining these two mutations, we obtained a nanosystem that displays both high activity and no self-inhibition mechanism. Interestingly, these two point mutations alone were able to shift the K_M value of the trimeric system back to the value of the original uncut DNAzyme (14 ± 2 versus 11 ± 1 nM) while displaying a broader DR (78 ± 33 versus 12 ± 3). Interestingly, none of these mutations impacted the catalytic activity of the native 2C system (Fig. 5b, left). These experiments illustrate how fragmented nanosystems are more affected by

https://doi.org/10.1038/s41557-022-01127-4

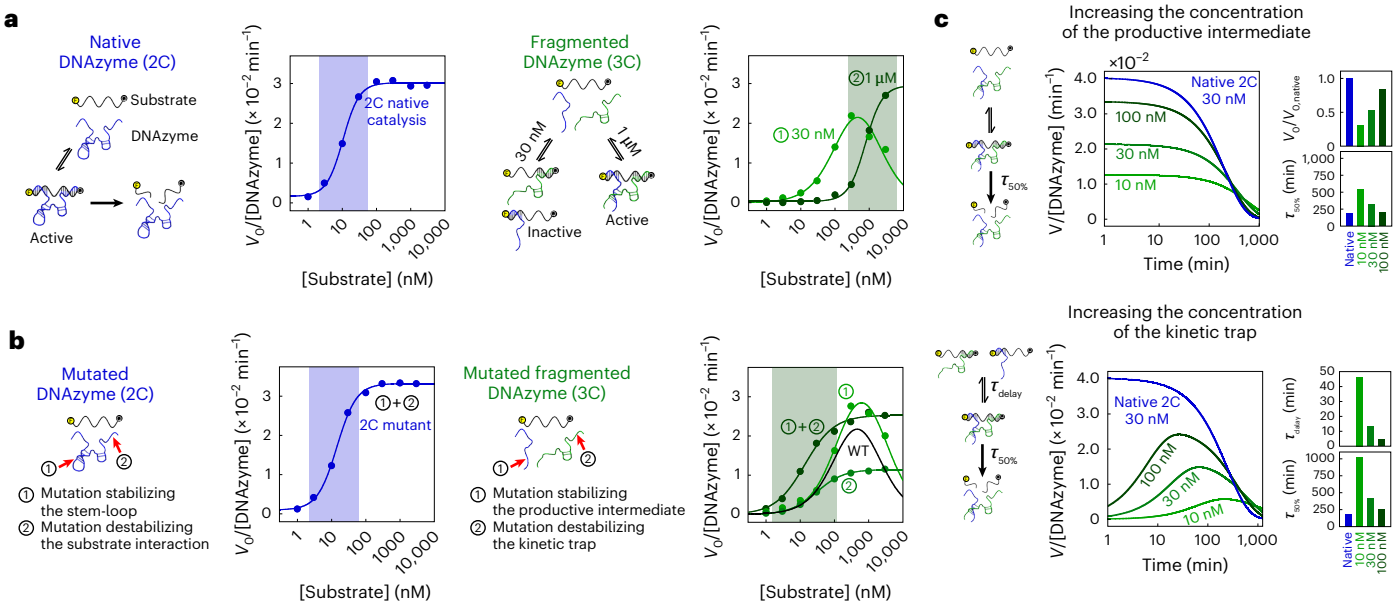


Fig. 5 | Programming novel regulatory mechanisms in a complex catalytic nanosystem through fragmentation. a, The NaA43 cleaving DNAzyme[58] displays a $[A]_{50\%}$ of 11 ± 1 nM at 30 nM DNAzyme and a DR of 11 ± 4, consistent with a saturation regime (left). Creating a trimeric assembly by fragmenting the DNAzyme can create a kinetic trap at a low concentration of DNAzyme (30 nM), leading to high sequestration of dimer intermediates and lower activity. However, preorganization of the productive dimer using a higher concentration of its component (1 μM) results in the system displaying a normal saturation regime profile ($[A]_{50\%}$ = 778 ± 53 nM and DR = 12 ± 3; right). **b**, Single point mutations that do not affect the catalytic activity of the native 2C system (left) have a drastic impact on the catalytic activity of the 3C system (right). A mutation stabilizing the productive dimer (mutation 1) enables the catalytic activity to be increased, while a mutation destabilizing the dimeric trap (mutation 2) enables self-inhibition to be disrupted at the cost of reduced catalytic activity. **c**, The catalytic activity of this trimeric nanosystem can be programmed by exploiting dimeric kinetic traps that allows precise control of the rate of the reaction (V) over time. Increasing the concentration of the productive intermediate of the 3C system enables the initial rate of the reaction (V_0) to be programmed (top). Increasing the concentration of the kinetic trap components enables the activation of the catalytic function to be delayed by a certain amount of time (τ_{delay}) which also affects the half-life of the catalysis ($\tau_{50\%}$) (bottom). The catalytic rates were obtained using the derivatives of the fitted kinetic traces of product generation. See Supplementary Fig. 22 for the raw data for **a** and **b**, and Supplementary Fig. 23 for the raw data for **c**.

single point mutations or chemical modifications and, therefore, are more programmable than their unfragmented counterparts.

Fragmentation of this catalytic system into a 3C nanosystem also enabled us to programme its activity over time by either increasing the amount of productive dimer or by favouring the kinetic trap. The activity of the fragmented DNAzyme, for example, can be easily tuned by increasing the concentration of its components (that is, favouring the productive dimer; Fig. 5c, top). At a high level of the productive dimer, this catalytic 3C system shows a fast kinetic profile that behaves similarly to the 2C system. Its activity can be reduced by decreasing its preorganization level using a lower concentration of DNAzyme components (Fig. 5c, top, light green). The catalytic activity of the nanosystem can also be delayed by increasing the concentration of the dimeric trap (Fig. 5c, bottom). By sequestering both DNAzyme components in a complex with the substrate (that is, each DNAzyme fragment is first put in contact with the substrate), we can delay the catalytic activity by up to 46 min. Such kinetically programmed functional assemblies are also often employed in biological assembly. For example, compartmentalization can be used to organize the stepwise interaction of components to avoid misassembly[59], while kinetic traps are employed to delay the recruitment of protein units to control specific functions over time[60]. It is also interesting to note that a similar kinetics profile and level of programmability can be achieved by increasing the number of components through the addition of an extra DNA strand (Supplementary Fig. 24 and Extended Data Fig. 9)[61]. Overall, these results exemplify the simplicity and efficiency of the fragmentation strategy to programme and optimize the function of complex nanosystems by increasing the number of components involved in their assembly.

Discussion

Here, we have demonstrated how the functionalities of nanosystems can simply be optimized by dividing their structures into multiple components. We first demonstrated that despite its lower assembly rate, our model 3C system undergoes a more global and cooperative assembly transition with fewer preorganized structures, which facilitates the creation of novel functionalities (for example, a complementary DNA inhibitor; Fig. 1d). Another advantage of systems with more components, such as the 3C nanosystem, is their ability to be programmed to permit assembly using both a 'cooperative' and an anticooperative process (Fig. 2). We also showed that 3C systems, in contrast to 2C systems, are more sensitive to mutations and can be readily tuned to exhibit self-inhibition mechanisms (Fig. 3) as well as time-dependent activation/deactivation mechanisms (Fig. 4). All these complex regulatory profiles (at equilibrium or over time) are readily accessible by simply fragmenting an existing self-assembled nanosystem into multiple components. We further demonstrated and validated the simplicity and usefulness of this strategy by artificially evolving the functionality of a more complex catalytic nanosystem that possesses RNA-cleaving activity (Fig. 5).

The fragmentation strategy can be easily implemented in nucleic acid-based nanosystems and could also be employed to create novel functionalities in proteins given their similar self-assembled characteristics[39,62]. Experts in DNA/RNA nanotechnology, for example, have exploited split aptamers to improve the properties of DNA/RNA-based sensors and therapeutics[26,63]. Similarly, proteins can also sustain fragmentation into multiple molecular components by folding back into their native conformation[27,28,64–67]. Indeed, multiple protein-fragment

complementation assays have been developed using such artificial dimeric proteins, and even a trimeric protein[66], to study and control biological systems[28]. Additionally, computational design strategies have recently been developed to identify the cutting site in proteins and to optimize the reconstitution of various fragments through specific mutations[27,67,68]. Although these studies demonstrated the flexibility of self-assembled protein nanosystems to sustain fragmentation, no studies, to the best of our knowledge, have highlighted, or even suggested that a fragmented nanosystem displays more ability to acquire novel regulation mechanisms. We believe that this fragmentation strategy, which simply consists of cutting an already existing nanosystem into multiple fragments, is also conceptually easier to perform than designing, from scratch, additional components that can interact and regulate the function of a nanosystem. Given the apparent simplicity of fragmenting a nanosystem, we anticipate that this strategy may represent an efficient alternative to engineering allosteric regulatory mechanisms to programme and create novel regulatory functions in self-assembled systems[69].

Engineering complex self-regulatory mechanisms using multiple components provides a programmable and quantitative chemical strategy for developing and optimizing the function of any self-assembled nanosystem with applications ranging from biosensing to chemical computing and drug delivery. For example, current strategies for narrowing the DR of sensors typically employ allosteric mechanisms[70,71], while strategies for extending the DR combine two or multiple sensors with different affinities[72–76]. In contrast, here, we illustrated how a DNA sensing nanosystem can be programmed with a narrow (9-fold) or extended (729-fold) DR by simply building it with three components. This ability to tune the DR is also useful for programming and optimizing the responses of molecular logic gates such as all-or-none or none-all-none responses (Fig. 3b, right)[77], a regulatory mechanism observed in many cellular functions[78]. A 3C system may also help to maintain drug concentrations within a specific therapeutic window. Nature, for example, employs various substrate inhibition strategies analogous to that presented in Fig. 5a to maintain the levels of crucial product metabolites despite large variations in substrate concentration[79,80]. Finally, a 3C system with programmed kinetic traps enables the time-specific activation and deactivation of various active biomolecules, leading to flexible and custom disease treatment strategies (Fig. 5c)[81,82].

In addition to providing new strategies to optimize complex self-regulated nanosystems, we believe that the thermodynamic and kinetic principles described herein will contribute to better understand the advantages of natural protein complexes. Indeed, it has been demonstrated that natural evolution has exploited fragmentation strategies to evolve the function of its various nanosystems[83–85]. Our findings now demonstrate that, in addition to being regulated through different independent promoters (for example, an 'AND' logic gate)[86], protein assemblies can additionally be regulated through the creation and optimization of various mechanisms (for example, cooperativity, anticooperativity, self-inhibition and molecular timer). Given that more than 55% of all proteins in living organisms are multimeric[8], it will be interesting to explore whether the functional gains of these protein complexes have simply emerged from the advantages derived from being built with more components[87]. Answering this question remains challenging, however, given that multimeric proteins have evolved and diverged over billions of years, making their direct comparison quite difficult[8,87,88]. In perspective, the ability of fragmented nanosystems to self-assemble through a myriad of weak intermolecular forces and acquire new functionalities remains surprising, if not amazing. This may well represent a unique feature of self-assembled nanosystems because such a strategy is unlikely to be useful in macroscale human-designed technologies, where components are typically held together by screws, glue or by welding.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-022-01127-4.

References

1. Grzybowski, B. A. & Huck, W. T. S. The nanotechnology of life-inspired systems. *Nat. Nanotechnol.* **11**, 585–592 (2016).
2. Murphy, K. G. & Bloom, S. R. Gut hormones and the regulation of energy homeostasis. *Nature* **444**, 854–859 (2006).
3. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390 (2014).
4. Benner, S. A. & Sismour, A. M. Synthetic biology. *Nat. Rev. Genet.* **6**, 533–542 (2005).
5. Camilli, A. & Bassler, B. L. Bacterial small-molecule signaling pathways. *Science* **311**, 1113–1116 (2006).
6. Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).
7. Marsh, J. A. & Teichmann, S. A. Structure, dynamics, assembly, and evolution of protein complexes. *Annu. Rev. Biochem.* **84**, 551–575 (2015).
8. Lynch, M. The evolution of multimeric protein assemblages. *Mol. Biol. Evol.* **29**, 1353–1366 (2012).
9. Thomas, C. & Lumb, A. B. Physiology of haemoglobin. *Continuing Education in Anaesthesia Critical Care & Pain* **12**, 251–256 (2012).
10. Pillai, A. S. et al. Origin of complexity in haemoglobin evolution. *Nature* **581**, 480–485 (2020).
11. Erbas-Cakmak, S. et al. Molecular logic gates: the past, present and future. *Chem. Soc. Rev.* **47**, 2228–2248 (2018).
12. Galés, C. et al. Real-time monitoring of receptor and G-protein interactions in living cells. *Nat. Methods* **2**, 177–184 (2005).
13. Pakulska, M. M., Miersch, S. & Shoichet, M. S. Designer protein delivery: from natural to engineered affinity-controlled release systems. *Science* **351**, aac4750 (2016).
14. Kinch, L. N. & Grishin, N. V. Evolution of protein structures and functions. *Curr. Opin. Struct. Biol.* **12**, 400–408 (2002).
15. Nooren, I. M. A. & Thornton, J. M. Diversity of protein–protein interactions. *EMBO J.* **22**, 3486–3492 (2003).
16. Plaxco, K. W., Simons, K. T., Ruczinski, I. & Baker, D. Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry* **39**, 11177–11183 (2000).
17. Grantcharova, V., Alm, E. J., Baker, D. & Horwich, A. L. Mechanisms of protein folding. *Curr. Opin. Struct. Biol.* **11**, 70–82 (2001).
18. Marianayagam, N. J., Sunde, M. & Matthews, J. M. The power of two: protein dimerization in biology. *Trends Biochem. Sci.* **29**, 618–625 (2004).
19. Whitty, A. Cooperativity and biological complexity. *Nat. Chem. Biol.* **4**, 435–439 (2008).
20. Williamson, J. R. Cooperativity in macromolecular assembly. *Nat. Chem. Biol.* **4**, 458–465 (2008).
21. Mattia, E. & Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **10**, 111–119 (2015).
22. Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nat. Rev. Mater.* **3**, 17068 (2017).
23. Li, M. et al. In vivo production of RNA nanostructures via programmed folding of single-stranded RNAs. *Nat. Commun.* **9**, 2196 (2018).
24. Gradišar, H. et al. Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat. Chem. Biol.* **9**, 362–366 (2013).
25. Shekhawat, S. S. & Ghosh, I. Split-protein systems: beyond binary protein–protein interactions. *Curr. Opin. Chem. Biol.* **15**, 789–797 (2011).

26. Debiais, M., Lelievre, A., Smietana, M. & Müller, S. Splitting aptamers and nucleic acid enzymes for the development of advanced biosensors. *Nucleic Acids Res.* **48**, 3400–3422 (2020).
27. Dolberg, T. B. et al. Computation-guided optimization of split protein systems. *Nat. Chem. Biol.* **17**, 531–539 (2021).
28. Michnick, S. W., Ear, P. H., Manderson, E. N., Remy, I. & Stefan, E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat. Rev. Drug Discov.* **6**, 569–582 (2007).
29. Cisse, I. I., Kim, H. & Ha, T. A rule of seven in Watson–Crick base-pairing of mismatched sequences. *Nat. Struct. Mol. Biol.* **19**, 623–627 (2012).
30. Vallée-Bélisle, A., Ricci, F. & Plaxco, K. W. Thermodynamic basis for the optimization of binding-induced biomolecular switches and structure-switching biosensors. *Proc. Natl Acad. Sci. USA* **106**, 13802–13807 (2009).
31. Hughes, R. A. & Ellington, A. D. Synthetic DNA Synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* **9**, a023812 (2017).
32. Madsen, M. & Gothelf, K. V. Chemistries for DNA nanotechnology. *Chem. Rev.* **119**, 6384–6458 (2019).
33. Sadowski, J. P., Calvert, C. R., Zhang, D. Y., Pierce, N. A. & Yin, P. Developmental self-assembly of a DNA tetrahedron. *ACS Nano* **8**, 3251–3259 (2014).
34. Hao, C. et al. Construction of RNA nanocages by re-engineering the packaging RNA of Phi29 bacteriophage. *Nat. Commun.* **5**, 3890 (2014).
35. Silverman, S. K. Catalytic DNA: scope, applications, and biochemistry of deoxyribozymes. *Trends Biochem. Sci.* **41**, 595–609 (2016).
36. Peña, M. D. L., Dufour, D. & Gallego, J. Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold. *RNA* **15**, 1949–1964 (2009).
37. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
38. Idili, A., Ricci, F. & Vallée-Bélisle, A. Determining the folding and binding free energy of DNA-based nanodevices and nanoswitches using urea titration curves. *Nucleic Acids Res.* **45**, 7571–7580 (2017).
39. Lawrence, C. et al. A comparison of the folding kinetics of a small, artificially selected DNA aptamer with those of equivalently simple naturally occurring proteins. *Protein Sci.* **23**, 56–66 (2014).
40. Zadeh, J. N. et al. NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
41. Jackson, S. E. & Fersht, A. R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428–10435 (1991).
42. Maxwell, K. L. et al. Protein folding: defining a 'standard' set of experimental conditions and a preliminary kinetic data set of two-state proteins. *Protein Sci.* **14**, 602–616 (2005).
43. Plaxco, K. W., Simons, K. T. & Baker, D. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994 (1998).
44. Sanchez-Ruiz, J. M. Protein kinetic stability. *Biophys. Chem.* **148**, 1–15 (2010).
45. Baskakov, I. V., Legname, G., Prusiner, S. B. & Cohen, F. E. Folding of prion protein to its native α-helical conformation is under kinetic control. *J. Biol. Chem.* **276**, 19687–19690 (2001).
46. Waudby, C. A., Dobson, C. M. & Christodoulou, J. Nature and regulation of protein folding on the ribosome. *Trends Biochem. Sci* **44**, 914–926 (2019).
47. Thommen, M., Holtkamp, W. & Rodnina, M. V. Co-translational protein folding: progress and methods. *Curr. Opin. Struct. Biol.* **42**, 83–89 (2017).
48. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009 (2007).
49. Hulme, E. C. & Trevethick, M. A. Ligand binding assays at equilibrium: validation and interpretation. *Br. J. Pharmacol.* **161**, 1219–1237 (2010).
50. Tian, L. & Heyduk, T. Bivalent ligands with long nanometer-scale flexible linkers. *Biochemistry* **48**, 264–275 (2009).
51. Lescoute, A. & Westhof, E. Topology of three-way junctions in folded RNAs. *RNA* **12**, 83–93 (2006).
52. Ercolani, G. & Schiaffino, L. Allosteric, chelate, and interannular cooperativity: a mise au point. *Angew. Chem. Int. Ed.* **50**, 1762–1768 (2011).
53. Korevaar, P. A. et al. Pathway complexity in supramolecular polymerization. *Nature* **481**, 492–496 (2012).
54. Wang, J., Liu, K., Xing, R. & Yan, X. Peptide self-assembly: thermodynamics and kinetics. *Chem. Soc. Rev.* **45**, 5589–5604 (2016).
55. Li, G. & Zhang, X. C. GTP hydrolysis mechanism of Ras-like GTPases. *J. Mol. Biol.* **340**, 921–932 (2004).
56. Lu, K. P., Finn, G., Lee, T. H. & Nicholson, L. K. Prolyl *cis-trans* isomerization as a molecular timer. *Nat. Chem. Biol.* **3**, 619–629 (2007).
57. Hou, W.-S. & Van Parijs, L. A Bcl-2-dependent molecular timer regulates the lifespan and immunogenicity of dendritic cells. *Nat. Immunol.* **5**, 583–589 (2004).
58. Torabi, S.-F. et al. In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing. *Proc. Natl Acad. Sci. USA* **112**, 5903–5908 (2015).
59. Maity, T. S. & Weeks, K. M. A threefold RNA–protein interface in the signal recognition particle gates native complex assembly. *J. Mol. Biol.* **369**, 512–524 (2007).
60. Yi, J., Balagopalan, L., Nguyen, T., McIntire, K. M. & Samelson, L. E. TCR microclusters form spatially segregated domains and sequentially assemble in calcium-dependent kinetic steps. *Nat. Commun.* **10**, 277 (2019).
61. Hunter, C. A. & Anderson, H. L. What is cooperativity? *Angew. Chem. Int. Ed.* **48**, 7488–7499 (2009).
62. Robustelli, P., Piana, S. & Shaw, D. E. Mechanism of coupled folding-upon-binding of an intrinsically disordered protein. *J. Am. Chem. Soc.* **142**, 11092–11101 (2020).
63. Chen, A., Yan, M. & Yang, S. Split aptamers and their applications in sandwich aptasensors. *Trends Analyt. Chem.* **80**, 581–593 (2016).
64. Müller, J. & Johnsson, N. Split-ubiquitin and the split-protein sensors: chessman for the endgame. *ChemBioChem* **9**, 2029–2038 (2008).
65. Shiba, K. & Schimmel, P. Functional assembly of a randomly cleaved protein. *Proc. Natl Acad. Sci. USA* **89**, 1880 (1992).
66. Cabantous, S. et al. A new protein-protein interaction sensor based on tripartite split-GFP association. *Sci. Rep.* **3**, 2854 (2013).
67. Rohl, C. A., Strauss, C. E. M., Misura, K. M. S. & Baker, D. in *Methods in Enzymology*, Vol. 383 66–93 (Academic Press, 2004).
68. Dagliyan, O. et al. Computational design of chemogenetic and optogenetic split proteins. *Nat. Commun.* **9**, 4042 (2018).
69. Ricci, F., Vallée-Bélisle, A., Porchetta, A. & Plaxco, K. W. Rational design of allosteric inhibitors and activators using the population-shift model: in vitro validation and application to an artificial biosensor. *J. Am. Chem. Soc.* **134**, 15177–15180 (2012).
70. Simon, A. J., Vallée-Bélisle, A., Ricci, F. & Plaxco, K. W. Intrinsic disorder as a generalizable strategy for the rational design of highly responsive, allosterically cooperative receptors. *Proc. Natl Acad. Sci. USA* **111**, 15048–15053 (2014).
71. Simon, A. J., Vallée-Bélisle, A., Ricci, F., Watkins, H. M. & Plaxco, K. W. Using the population-shift mechanism to rationally introduce 'Hill-type' cooperativity into a normally non-cooperative receptor. *Angew. Chem. Int. Ed.* **53**, 9471–9475 (2014).

72. Kang, D., Vallée-Bélisle, A., Porchetta, A., Plaxco, K. W. & Ricci, F. Re-engineering electrochemical biosensors to narrow or extend their useful dynamic range. *Angew. Chem. Int. Ed.* **51**, 6717–6721 (2012).
73. Porchetta, A., Vallée-Bélisle, A., Plaxco, K. W. & Ricci, F. Allosterically tunable, DNA-based switches triggered by heavy metals. *J. Am. Chem. Soc.* **135**, 13238–13241 (2013).
74. Harroun, S. G. et al. Programmable DNA switches and their applications. *Nanoscale* **10**, 4607–4641 (2018).
75. Vallée-Bélisle, A., Ricci, F. & Plaxco, K. W. Engineering biosensors with extended, narrowed, or arbitrarily edited dynamic range. *J. Am. Chem. Soc.* **134**, 2876–2879 (2012).
76. Ricci, F., Vallée-Bélisle, A., Simon, A. J., Porchetta, A. & Plaxco, K. W. Using nature's 'tricks' to rationally tune the binding properties of biomolecular receptors. *Acc. Chem. Res.* **49**, 1884–1892 (2016).
77. Lauzon, D., Zhu, G. & Vallée-Bélisle, A. in *DNA- and RNA-Based Computing Systems* 105–124 (Wiley-VCH, 2021).
78. Gao, X. J., Chong, L. S., Kim, M. S. & Elowitz, M. B. Programmable protein circuits in living cells. *Science* **361**, 1252–1258 (2018).
79. Best, J. A., Nijhout, H. F. & Reed, M. C. Homeostatic mechanisms in dopamine synthesis and release: a mathematical model. *Theor. Biol. Med. Model.* **6**, 21 (2009).
80. Reed, M., Best, J., Golubitsky, M., Stewart, I. & Nijhout, H. F. Analysis of homeostatic mechanisms in biochemical networks. *Bull. Math. Biol.* **79**, 2534–2557 (2017).
81. Patra, J. K. et al. Nano-based drug delivery systems: recent developments and future prospects. *J. Nanobiotechnol.* **16**, 71 (2018).
82. Lu, H. et al. Recent progress on nanostructures for drug delivery applications. *J. Nanomater.* 5762431 (2016).
83. Pasek, S., Risler, J.-L. & Brézellec, P. Gene fusion/fission is a major contributor to evolution of multi-domain bacterial proteins. *Bioinformatics* **22**, 1418–1423 (2006).
84. Snel, B., Bork, P. & Huynen, M. Genome evolution: gene fusion versus gene fission. *Trends Genet.* **16**, 9–11 (2000).
85. Kuriyan, J. & Eisenberg, D. The origin of protein interactions and allostery in colocalization. *Nature* **450**, 983–990 (2007).
86. Pandey, N., Nobles, C. L., Zechiedrich, L., Maresso, A. W. & Silberg, J. J. Combining random gene fission and rational gene fusion to discover near-infrared fluorescent protein fragments that report on protein–protein interactions. *ACS Synth. Biol.* **4**, 615–624 (2015).
87. Hagner, K., Setayeshgar, S. & Lynch, M. Stochastic protein multimerization, activity, and fitness. *Phys. Rev. E* **98**, 062401 (2018).
88. Hashimoto, K., Nishi, H., Bryant, S. & Panchenko, A. R. Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization. *Phys. Biol.* **8**, 035007 (2011).

https://doi.org/10.1038/s41557-022-01127-4

Methods

Fluorescent experiments

Urea titration curves. Urea titration curves were measured following a method developed by our laboratory[38]. We first prepared a 900 µl solution of the DNA-based system of interest in 10 M urea-buffered solution (10 mM NaH_2PO_4, 40 mM NaCl, pH 7.00). We then sequentially diluted this solution with a buffered solution containing the same concentration of the DNA-based system but without the urea. Each sample was equilibrated for 2 min before recording the fluorescence (Cary Eclipse, Agilent). The unimolecular system ($A_{Diss} \rightleftharpoons A_{Ass}$) was titrated using a concentration of 10 nM of fluorescent DNA strand and fitted using equation (1). The bimolecular system ($A + B \rightleftharpoons AB$) was titrated using a concentration of 10 nM of fluorescent DNA strand and 100 nM of the quencher DNA strand (tenfold excess) and fitted using equation (2). The trimolecular system ($A + B + C \rightleftharpoons ABC$) was titrated using a concentration of 10 nM of fluorescent DNA strand and 100 nM of the quencher and unlabelled DNA strands (tenfold excess) and fitted using equation (3). In equations (1)–(3), F is the fluorescence signal, $F°_{Ass}$ and $F°_{Diss}$ are the intrinsic fluorescence of the assembled and disassembled states, σ_{Ass} and σ_{Diss} are the urea dependency of the fluorescence signals of the assembled and disassembled states, [U] is the concentration of urea, $\Delta G°_{Ass}$ is the Gibbs free energy of assembly in the absence of urea, m is the m-value, $[D_Q]$ and $[D_{Un}]$ are the concentrations of the quencher and unlabelled DNA strands (which are in tenfold excess), T is the temperature and R is the gas constant.

$$F = \frac{\left(F°_{Ass} + \sigma_{Ass}[U]\right) + \left(F°_{Diss} + \sigma_{Diss}[U]\right) \times e^{-\frac{\left(\Delta G°_{Ass} - m[U]\right)}{RT}}}{\left(1 + e^{-\frac{\Delta G°_{Ass} - m[U]}{RT}}\right)} \quad (1)$$

$$F = \frac{[D_Q] \times \left(F°_{Ass} + \sigma_{Ass}[U]\right) + \left(F°_{Diss} + \sigma_{Diss}[U]\right) \times e^{-\frac{\left(\Delta G°_{Ass} - m[U]\right)}{RT}}}{\left([D_Q] + e^{-\frac{\Delta G°_{Ass} - m[U]}{RT}}\right)} \quad (2)$$

$$F = \frac{[D_Q] \times [D_{Un}] \times \left(F°_{Ass} + \sigma_{Ass}[U]\right) + \left(F°_{Diss} + \sigma_{Diss}[U]\right) \times e^{-\frac{\left(\Delta G°_{Ass} - m[U]\right)}{RT}}}{\left([D_Q] \times [D_{Un}] + e^{-\frac{\Delta G°_{Ass} - m[U]}{RT}}\right)} \quad (3)$$

Temperature melting curves. Thermodynamic constants were calculated from temperature melting curves according to Owczarzy and co-workers[89]. The melting curves of all complexes were measured at different concentrations using a fluorimeter (Cary Eclipse, Agilent). All samples were heated to 90 °C at 10 °C min^{-1}, held for 2 min, then cooled to 20 °C at 10 °C min^{-1} and held for 10 min before recording the melting curve from 20 to 90 °C at 0.8 °C min^{-1}. The melting temperature, T_m, was extracted from the derivative of the melting curve (dF/dT), which displays a 'bell-shape' curve. The maximum of the dF/dT plot was evaluated from Gaussian distribution fitting. The change in enthalpy, ΔH, and change in entropy, ΔS, were determined by linear regression using equation (4), enabling the calculation of ΔG. For all experiments, the concentration of the monomer with a fluorophore (denoted A) was kept constant at 1 µM for $\Delta G°_{Dim}$ measurement and at 0.1 µM for $\Delta G°_{Tri}$ measurement, while the concentration of the other monomer containing the quencher (denoted B) was varied between 5 and 200 µM for $\Delta G°_{Dim}$ measurement and between 0.5 and 20 µM for $\Delta G°_{Tri}$ measurement. In equation (4), $[A]_T$ is the total concentration of strand A and $[B]_T$ is the total concentration of strand B.

$$\frac{1}{T_m} = \frac{R}{\Delta H} \ln\left([B]_T - \frac{[A]_T}{2}\right) + \frac{\Delta S}{\Delta H} \quad (4)$$

Binding curves. To measure the binding curves, appropriate dilutions were made such that the volume of the probe solution (strand containing the fluorophore) was always 900 µl. The probe solution contained either one strand (for the 2C system, denoted B) or two strands (for the 3C system, denoted B and C). To this solution, a small volume of titrant (strand with a quencher, denoted A) was added and the fluorescence was recorded with a fluorimeter (Cary Eclipse, Agilent) after 20 min of equilibration. The total volume of titrant added to the probe solution was kept below 5% to avoid high dilution of the probe. We could then assume that the concentration of the probe remained approximately constant over the titration. The titration curve was then fitted using the Hill equation (5), where K_{obs} is the observed dissociation constant (or the concentration at which 50% of the probe was bound, also denoted $[A]_{50\%}$), n_H is the Hill factor (an indicator of the cooperativity enabling the calculation of the DR[70]), F is the fluorescence signal, Bsl is the baseline fluorescence and Amp is the amplitude, or change in the fluorescence upon the binding of strand A. The dependency of K_{obs} and the DR on the concentration of the probe is described in detail in the Supplementary Information.

$$F = \mathrm{Bsl} + \mathrm{Amp}\frac{[A]^{n_H}}{K_{obs}^{n_H} + [A]^{n_H}} \quad (5)$$

Kinetic experiments. For fast kinetics (<30 min), appropriate dilutions were made and solutions of A and B were rapidly mixed using a stopped-flow instrument coupled with a fluorimeter (SX20, Applied Photophysics). Slower kinetics (>30 min) were recorded using a standard fluorimeter (Cary Eclipse, Agilent). A small volume of solution A was added with a pipette to a cuvette containing solution B and manually mixed using the same pipette. Oil was deposited on top of the solution to avoid water evaporation over the long kinetic measurement. DNA hybridization follows second-order kinetics: $A + B \rightleftharpoons AB$, where A and B are single-stranded DNA and AB is the duplex formed from these two strands[90]. Pseudo-first-order kinetics were achieved using at least a tenfold excess of strand A. The kinetic traces were then fitted using exponential equation (6), where F is the fluorescence signal, Bsl is the baseline fluorescence, Amp is the amplitude, or change in fluorescence upon binding, t is the time and k_{obs} is the observed first-order rate constant. The relationship between k_{obs} and the concentration of A was linear, enabling the measurement of the rate constants for association (k_{ass}) and dissociation (k_{diss}). Second-order kinetics were achieved using equimolar concentrations of A and B ($[A]_T = [B]_T$). This system acts like a dimerization reaction (that is, $2A \rightleftharpoons A_2$) and can thus be fitted using equation (7), where $[A]_T$ is the total concentration of strand A and k_{obs} is now the observed second-order rate constant.

$$F = \mathrm{Bsl} + \mathrm{Amp} \times e^{-k_{obs}t} \quad (6)$$

$$F = \mathrm{Bsl} + \mathrm{Amp} \times \frac{1}{1 + 2[A]_T k_{obs} t} \quad (7)$$

The trimeric association follows the reaction $A + B + C \rightleftharpoons ABC$, where A, B and C are single-stranded DNA and ABC is the three-way junction formed from these strands. When measuring the assembly, we used at least a tenfold excess of molecule A in buffer. which was rapidly mixed using a stopped-flow instrument coupled with a fluorimeter (SX20, Applied Photophysics) with a solution of B and C in water (to avoid pre-association of B and C). Data were fitted using a combination of pseudo-first-order and second-order kinetics (equation (8)), where Amp_1 is the amplitude of the first phase (that is, pseudo-first-order kinetic), Amp_2 is the amplitude of second phase (that is, second-order kinetic), $k_{obs,1}$ is the observed

rate constant of the first phase and $k_{obs,2}$ is the observed rate constant of the second phase.

$$F = \mathrm{Bsl} + \mathrm{Amp}_1 \times e^{-k_{obs,1}t} + \mathrm{Amp}_2 \times \frac{1}{1 + 2\,[A]_T\,k_{obs,2}t} \quad (8)$$

Native polyacrylamide gel electrophoresis experiments

Appropriate dilutions of unlabelled DNA solutions were made such that the concentration of strands B and C was kept at 1 µM and the concentration of strand A was varied from 30 nM to 100 µM. Solutions were then mixed in a 5:1 ratio with the 6× loading buffer (2.5 mg ml^{-1} bromothymol blue, 2.5 mg ml^{-1} xylene cyanol FF and 30% glycerol in water). A 15% polyacrylamide gel was hand cast following the Bio-Rad protocol and incubated in the running buffer (0.5× Tris/Borate/EDTA buffer (TBE) containing 5 mM $MgCl_2$) for 1 h. Then, 10 µl of samples were run for 90 min at 120 V using the Mini-PROTEAN Tetra cell electrophoresis unit (Bio-Rad) and the Bio-Rad PowerPac Basic power supply. Gels were stained with a 0.5× solution of GelRed (Biotium) for 10 min and analysed on a ChemiDoc XRS+ imaging system (Bio-Rad). The band intensity was then integrated to evaluate the amount of assembled DNA-based systems.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available within the paper and the Supplementary Information. The datasets generated and/or analysed during the study are also available from the corresponding author upon reasonable request. Source data are provided with this paper.

Code availability

The equations and examples of the codes used for the simulations are described in the Supplementary Information. The MATLAB codes used to perform the numerical simulations of the equilibrium binding experiments and the kinetic traces are also available from the corresponding author upon reasonable request.

References

89. You, Y., Tataurov, A. V. & Owczarzy, R. Measuring thermodynamic details of DNA hybridization using fluorescence. *Biopolymers* **95**, 472–486 (2011).
90. Galau, G. A., Britten, R. J. & Davidson, E. H. Studies on nucleic acid reassociation kinetics: rate of hybridization of excess RNA with DNA, compared to the rate of DNA renaturation. *Proc. Natl Acad. Sci. USA* **74**, 1020–1023 (1977).

Acknowledgements

This research was conducted through Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants (RGPIN-2020-06975, A.V.-B.). A.V.-B. is Canada Research Chair in Bioengineering and Bionanotechnology, Tier II. D.L. acknowledges a Canada graduate scholarship master (CGS M) from NSERC and a third cycle scholarship from the Fonds de recherche du Québec—Nature et technologies (FRQNT). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors would like to thank J. A. Marsh, J. N. Pelletier, L. Pedro, K. Nemčeková and S. G. Harroun for their helpful discussions and comments on the paper. The authors would also like to thank all members of the Quebec Network for Research on Protein Function, Engineering, and Applications (PROTEO) for helpful discussion, more specifically C. R. Landry. The authors would also like to thank the Département de Biochimie et Médecine Moléculaire de l'Université de Montréal for providing us access to their instruments.

Author contributions

D.L. and A.V.-B. designed the experiments and D.L. performed all the experiments. D.L. and A.V.-B. designed the figures and wrote the paper. Both authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

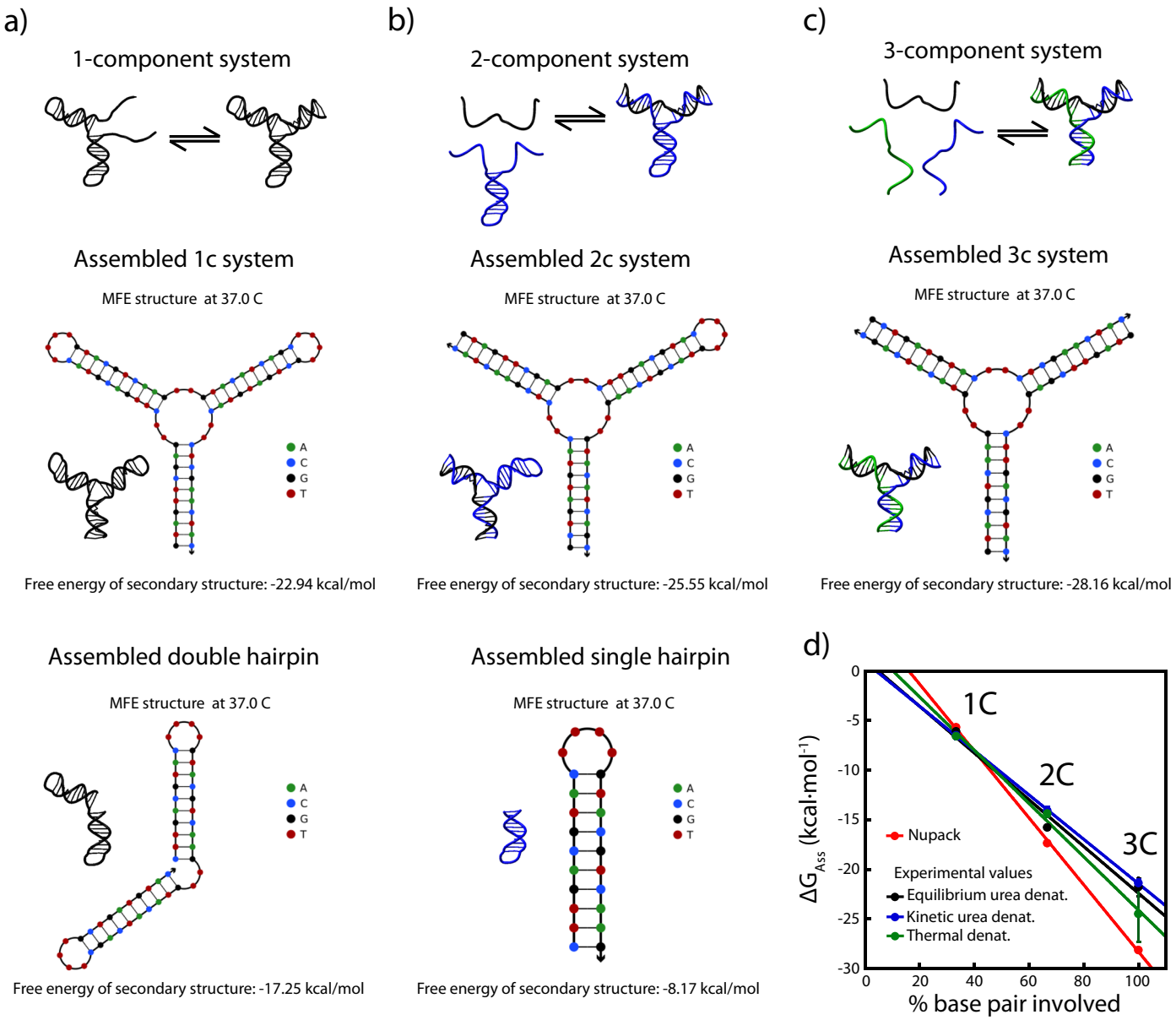
Extended data is available for this paper at https://doi.org/10.1038/s41557-022-01127-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41557-022-01127-4.

Correspondence and requests for materials should be addressed to A. Vallée-Bélisle.

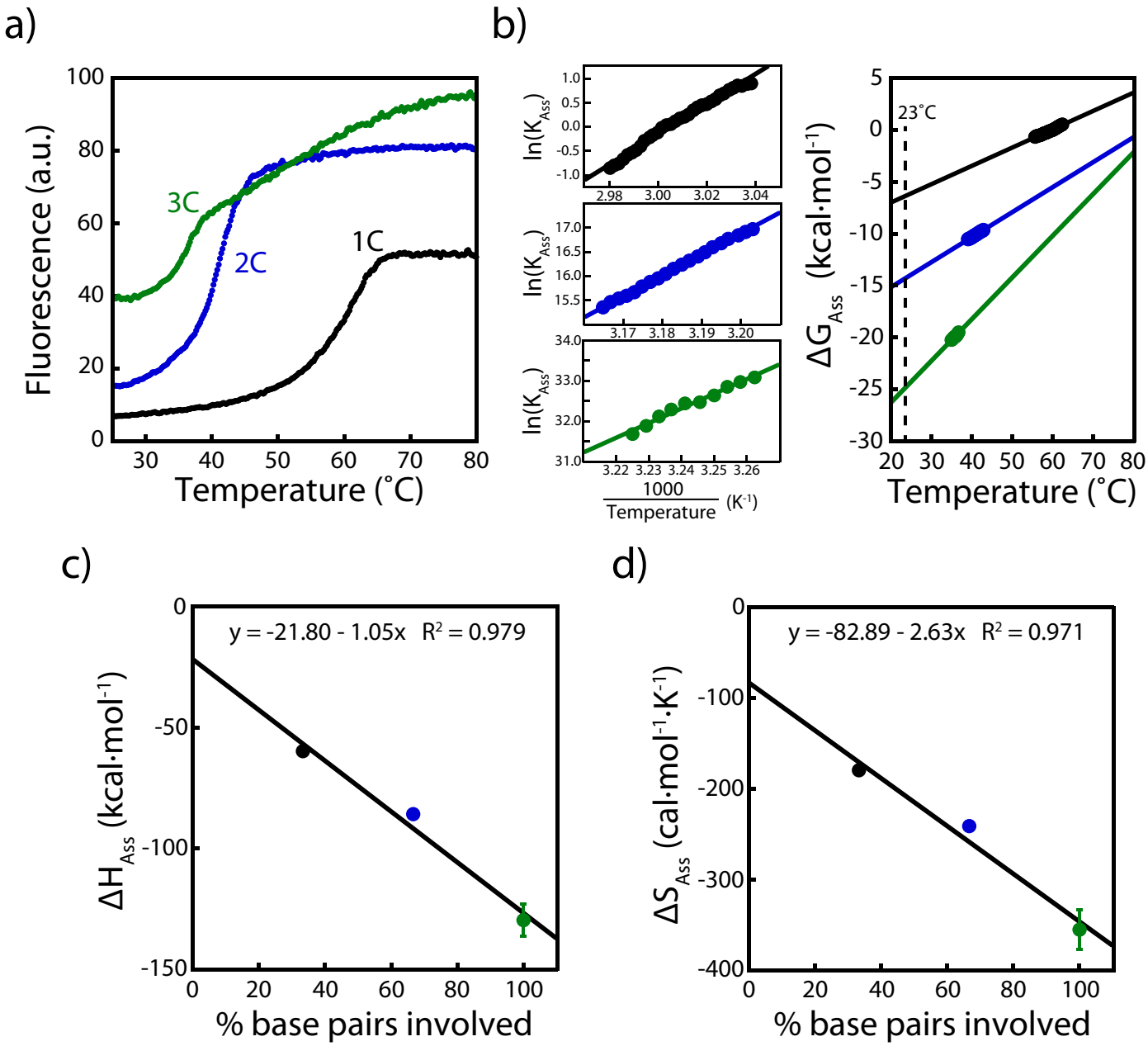
Peer review information *Nature Chemistry* thanks Aleksei Aksimentiev, James Carothers and Guillaume Gines for their contribution to the peer review of this work.

https://doi.org/10.1038/s41557-022-01127-4



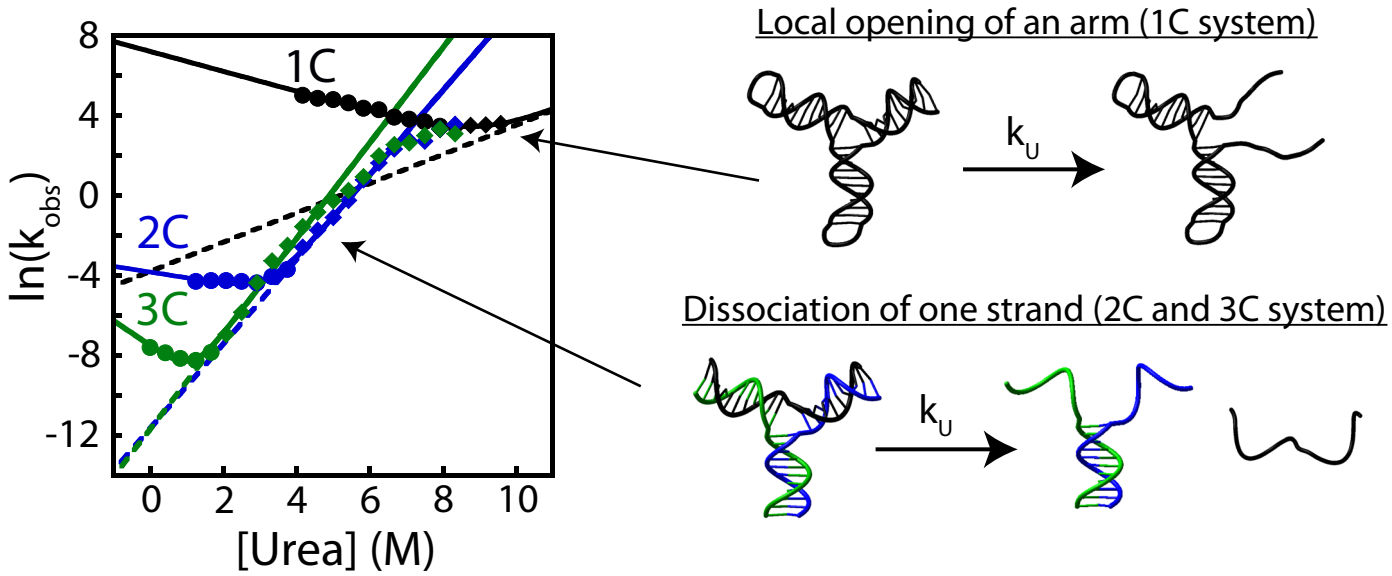
Extended Data Fig. 1 | Nupack analysis of our DNA based nanostructures. (**a**) The 1C system displays a predicted free energy ($\Delta G°_{Ass}$) of −22.94 kcal·mol^{-1} while its preorganized loops, which cannot be unfolded using urea (Fig. S2), show a predicted $\Delta G°_{Ass}$ of −17.25 kcal·mol^{-1}. The subtraction of both leads to a free energy of assembly ($\Delta G°_{Ass}$) of −5.69 kcal·mol^{-1}. (**b**) The 2C system displays a predicted $\Delta G°_{Ass}$ of −25.55 kcal·mol^{-1} while its preorganized loop has a predicted $\Delta G°_{Ass}$ of −8.17 kcal·mol^{-1}. The subtraction of both leads to a $\Delta G°_{Ass}$ of −17.38 kcal·mol^{-1}. (**c**) The 3C system displays a $\Delta G°_{Ass}$ of −28.16 kcal·mol^{-1}. (**d**) All $\Delta G°_{Ass}$ predicted by NUPACK are in good agreement with the experimentally derived values and correlate with the number of base pairs involved in the transition. Of note, NUPACK seems to overestimate the $\Delta G°_{Ass}$ of the 2C system and the 3C system. This is not unusual as similar discrepancies were observed for the urea denaturation of DNA-DNA complexes[38]. Data and errors are presented as the values obtained from the non-linear regression of the denaturation experiments (see Fig. S1, Extended Data Fig. 2 and Extended Data Fig. 3) (n = 1).

https://doi.org/10.1038/s41557-022-01127-4



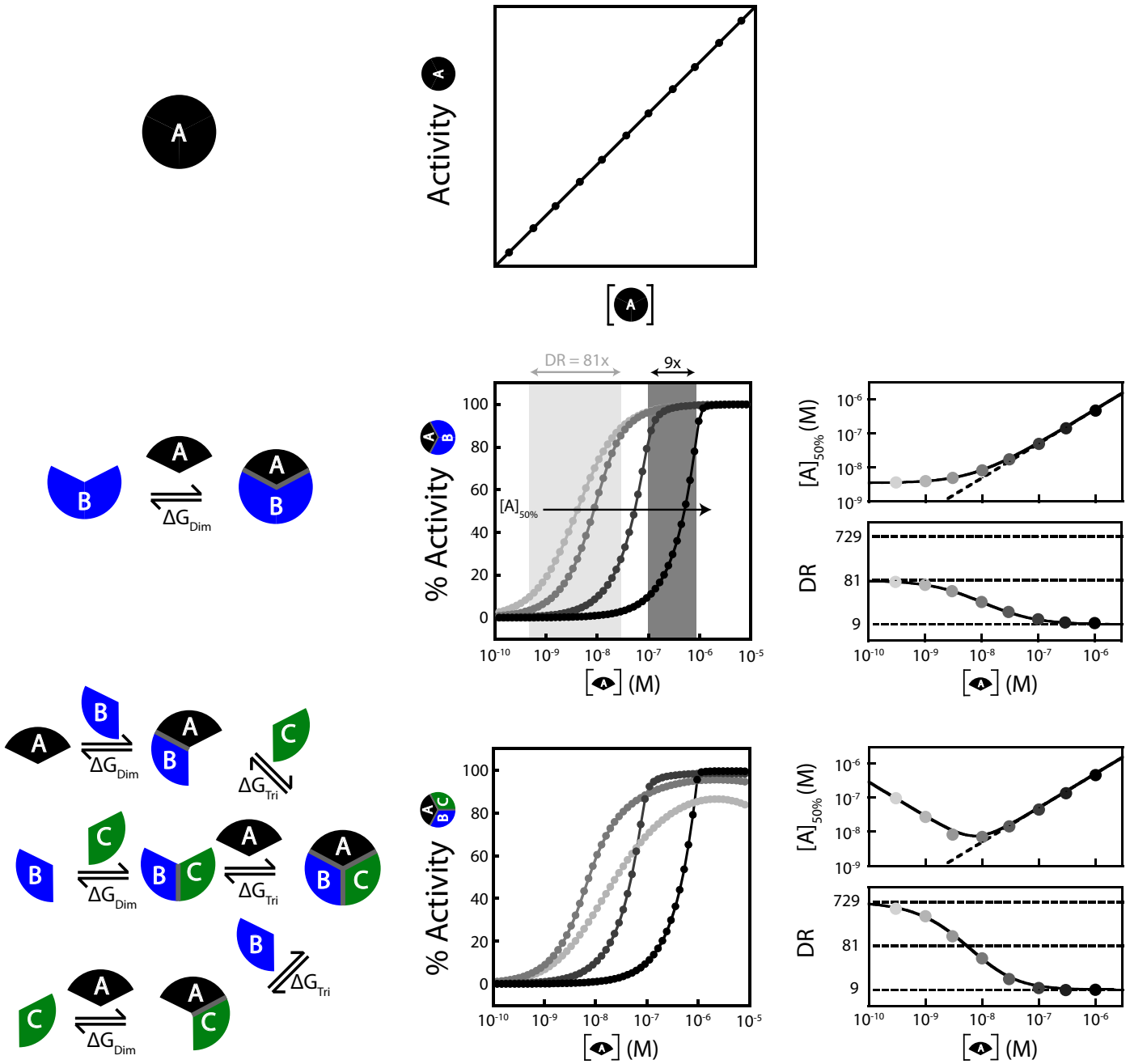
Extended Data Fig. 2 | Thermal denaturation analyses of the 1C, 2C and 3C systems. (**a**) Thermal denaturation profiles of the 1-component (black), 2-component (blue) and 3-component systems (green). (**b**) *Left*: Van't Hoff analysis of the thermal denaturation curves enables the extraction of the thermodynamic parameters $\Delta H°_{Ass}$ and $\Delta S°_{Ass}$ (see panels c) and d), respectively). *Right*: extrapolation of the $\Delta G°_{Ass}$ at 23˚C (right side) are in good agreement with the values obtained using urea denaturation. (**c**, **d**) As observed for the *m*-values and $\Delta G°_{Ass}$ (see Fig. S3), $\Delta H°_{Ass}$ and $\Delta S°_{Ass}$ are also linearly dependant on the number of base pairs broken/formed in the assembly/disassembly transition. $\Delta H°_{Ass} = -60 \pm 1$ kcal·mol^{-1}, -85.8 ± 0.8 kcal·mol^{-1}, and -130 ± 7 kcal·mol^{-1}, for the 1C, 2C, and 3C system, respectively. $\Delta S°_{Ass} = -179 \pm 4$ cal·mol^{-1}·K^{-1}, -241 ± 3 cal·mol^{-1}·K^{-1}, and -355 ± 22 cal·mol^{-1}·K^{-1} for the 1C, 2C, and 3C system, respectively. All experiments are performed in PBS buffer (10 mM NaH_2PO_4, 40 mM NaCl, pH = 7.00). Data and errors are presented as the values obtained from the Van't Hoff linear regression (n = 1).

https://doi.org/10.1038/s41557-022-01127-4



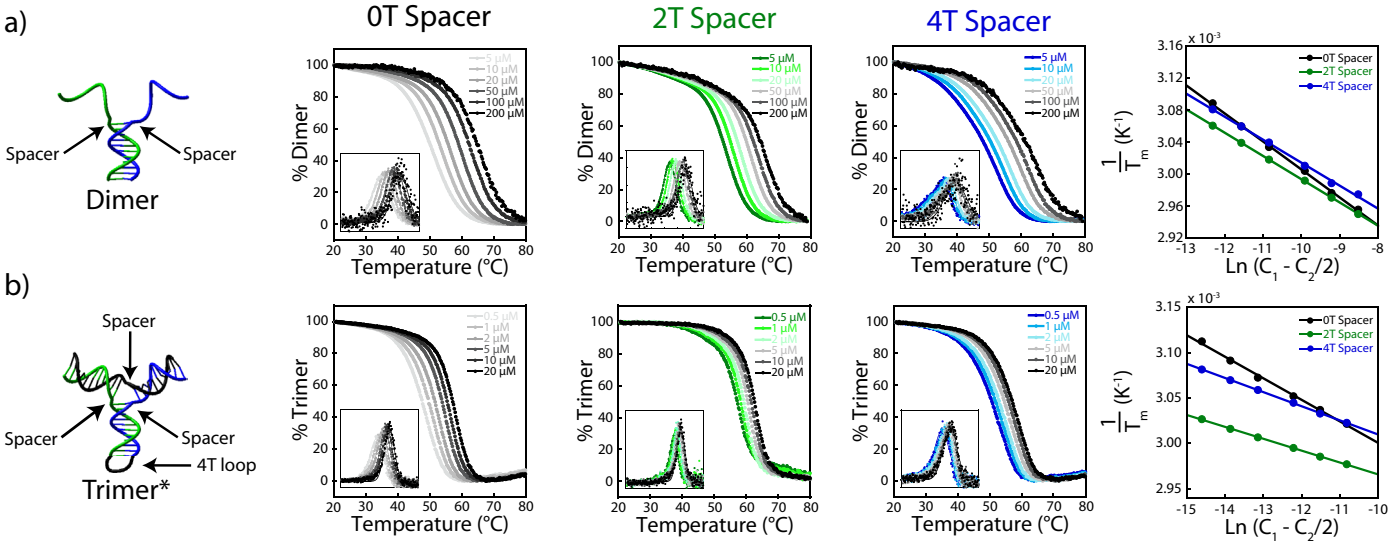
Extended Data Fig. 3 | Kinetic analysis (chevron plot) of the 1C, 2C, and 3C systems. The kinetics of the association/folding (circles) and dissociation/unfolding (diamonds) of all nanosystems under various urea concentrations (chevron plot) reveal a two-state association/dissociation mechanism around their denaturation transition (see Table S1). The 2C and 3C systems display similar unfolding transitions ($k_{U,2C} = 8.6 \pm 3.1 \times 10^{-6}$ s^{-1} and $k_{U,3C} = 8.3 \pm 3.2 \times 10^{-6}$ s^{-1}; $m_{U,2C} = 1.24 \pm 0.04$ kcal·mol^{-1}·M^{-1} and $m_{U,3C} = 1.41 \pm 0.06$ kcal·mol^{-1}·M^{-1}) suggesting that they dissociate via the same mechanism (dissociation of a complete strand). In contrast, the 1C system displays an unfolding rate that is 3000-time faster ($k_{U,1C} = 2.2 \pm 0.9 \times 10^{-2}$ s^{-1}) with a 3-time smaller urea dependency ($m_{U,1C} = 0.43 \pm 0.01$ kcal·mol^{-1}·M^{-1}). This is consistent with the smaller local unfolding expected to take place for the 1C system, the unfolding slope (m_U) is proportional to the surface area made accessible upon dissociation/unfolding. The slower dissociation kinetics of the 2C and 3C systems, extrapolated in absence of urea, is also consistent with the fact their dissociation requires the disruption of 20 base pairs compared to the disruption of 10 base pairs for the 1C system (see cartoon). Interestingly, the dissociation mechanism for the 2C and 3C systems becomes non-linear and similar to the 1C system at a high concentration of urea (>7 M) suggesting that their disassembly becomes only rate limited by the opening of one arm. For raw kinetic traces see Fig. S4 to Fig. S6.

https://doi.org/10.1038/s41557-022-01127-4



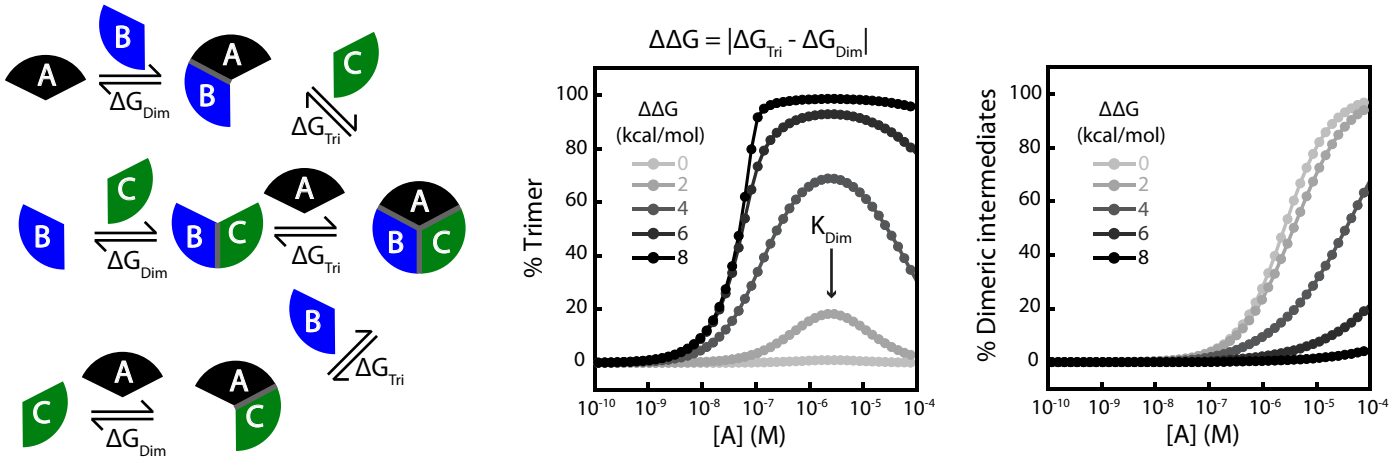
Extended Data Fig. 4 | The assembly behaviour of the 2C and 3C systems (Fig. 2) are well predicted by numerical simulations. *Top*. Because the one-component system already folds into its active conformation, its activity remains linear with its concentration. *Middle*. The assembly of the two-components system induced by an increase of component A can be triggered at different $[A]_{50\%}$ by increasing the concentration of the limiting component (here B). While $[A]_{50\%}$ shifts towards higher concentration, the dynamic range (DR) shifts from 81-fold to 9-fold. *Bottom*. The assembly of the three-components system induced by an increase of component A can also be triggered at different $[A]_{50\%}$ by increasing the concentration of the limiting components (here B and C). Interestingly, the $[A]_{50\%}$ shifts from high to low concentration and shifts back to high concentration, while the dynamic range (DR) shifts from 729-fold to 9-fold. Of note, at a low concentration of components B and C, the overall yield of the 3-component system decreases in the presence of large excess of component A. This is because the system favours the formation of the two dimeric intermediates (here AB and AC) instead of the trimer (see also Fig. 3 and Extended Data Fig. 6). Numerical simulations are done using MATLAB® with a dimeric stability of −8 kcal·mol^{-1} ($\Delta G°_{Dim}$) and a trimeric stability of −16 kcal·mol^{-1} ($\Delta G°_{Tri}$) leading to an overall assembly stability of −24 kcal·mol^{-1} ($\Delta G°_{Ass} = \Delta G°_{Dim} + \Delta G°_{Tri}$).

https://doi.org/10.1038/s41557-022-01127-4



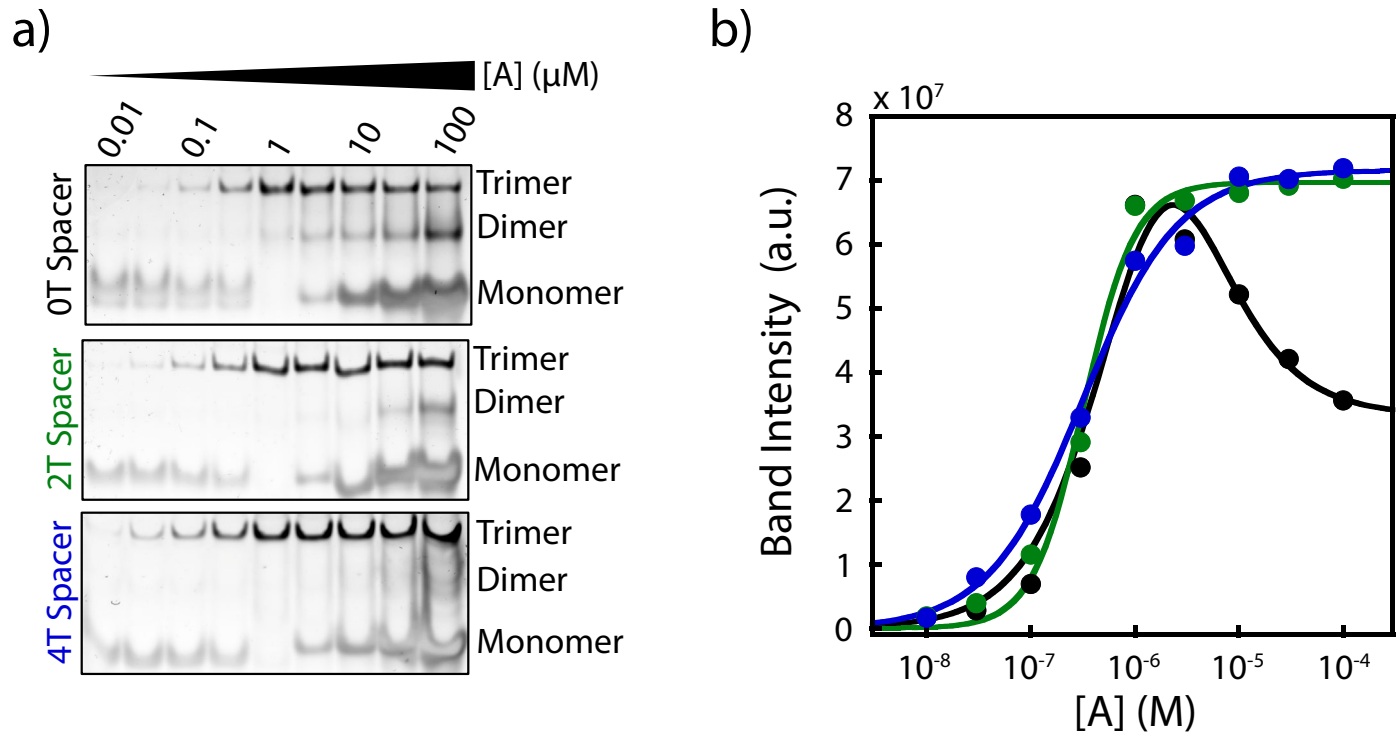
Extended Data Fig. 5 | Effect of the thymine spacer length between each arm on the 3-way junction stability. (**a**) Melting curve analysis of the dimeric equilibrium reveals that, as expected, the spacer length (0 T, 2 T, and 4 T) does not substantially change the dimeric affinity ($\Delta G°_{Dim}$). (**b**) Melting curve analysis of the trimeric equilibrium reveals a strong dependence of the spacer length on the trimeric affinity ($\Delta G°_{Tri}$). *Of note, to measure the trimeric affinity, $\Delta G°_{Tri}$, and to exclude the dimeric energy, we 'locked' the dimer into a unimolecular system using a 4 T loop (black loop at the bottom of the blue-green arm). The DNA-hairpin, therefore, mimics the dimeric structure. Inset of panels (a) and (b) represent the derivative of the melting curve (dF/dT) fitted with a Gaussian distribution, which has been used to extract the T_m that are then fitted using Eq. 4 to extract thermodynamic parameters. All experiments have been done in PBS buffer (50 mM Na_2HPO_4, 400 mM NaCl, pH = 7.00).

https://doi.org/10.1038/s41557-022-01127-4



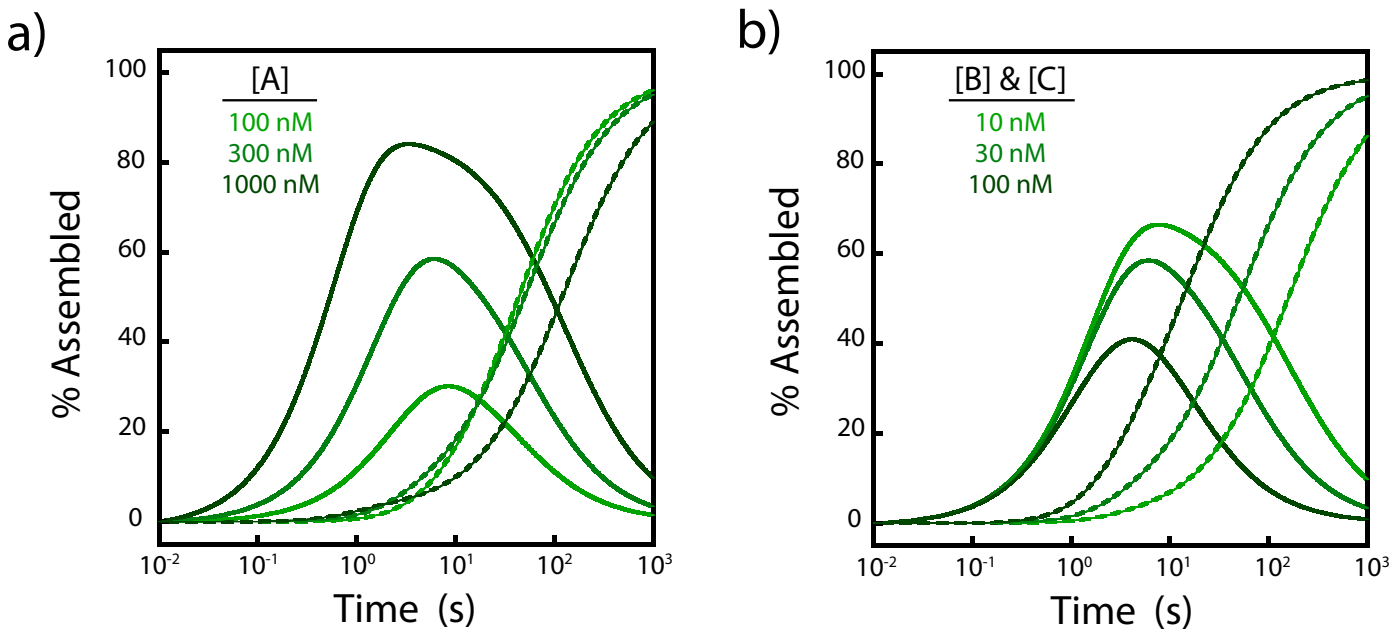
Extended Data Fig. 6 | Numerical simulations of the effect of destabilizing only $\Delta G°_{Tri}$ on the assembly of a 3-component system. When $\Delta G°_{Tri}$ and $\Delta G°_{Dim}$ are similar ($\Delta\Delta G°_{Tri\text{-}Dim}$ = 0 kcal·mol^{-1}) the assembly of the 3-component system is not favoured (lightest grey). Increasing the value of $\Delta\Delta G°_{Tri\text{-}Dim}$ leads to a better yield of assembly of the 3-component system. However, when component A becomes more concentrated than the dissociation constant of the dimeric intermediates (K_{Dim}), the assembly of the 3-component system becomes less favourable and the dimeric intermediates AB and AC form instead. A higher $\Delta\Delta G°_{Tri\text{-}Dim}$ enables to minimize this effect. Numerical simulations are done using MATLAB®. All $\Delta G°_{Dim}$ are fixed at −8 kcal·mol^{-1} and $\Delta G°_{Tri}$ is varied between −8 kcal·mol^{-1} and −16 kcal·mol^{-1}. Temperature is fixed at 37 °C (310.15 K), the concentration of components B and C are fixed at 100 nM and the concentration of component A is varied from 0.1 nM to 100 µM.

https://doi.org/10.1038/s41557-022-01127-4



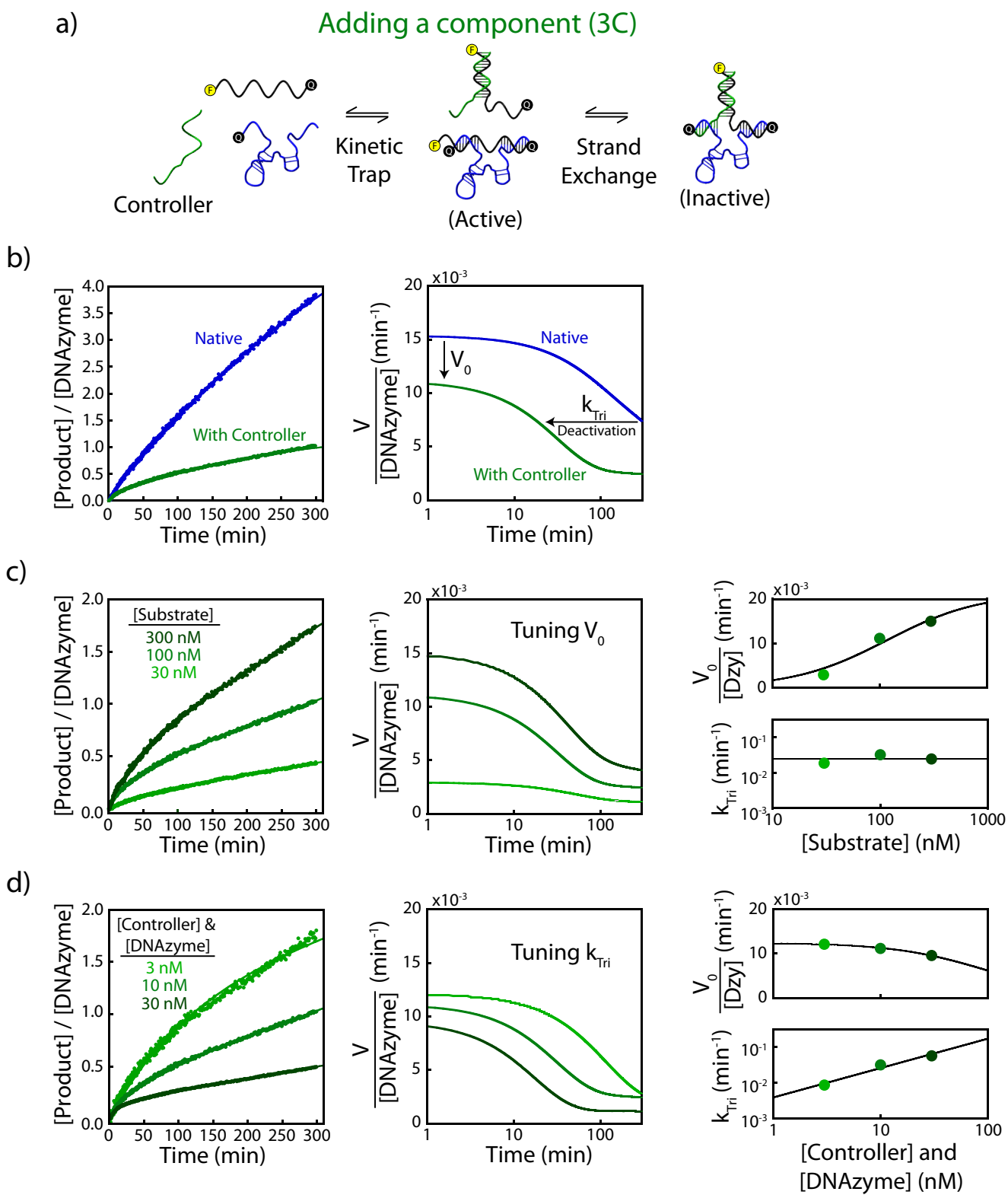
Extended Data Fig. 7 | Native PAGE titration. (**a**, **b**) Titrations analyzed by native polyacrylamide gel electrophoresis (native PAGE) support the dimers sequestration mechanism. Like our fluorescence titration experiments (Fig. 3b, *right*), the 2 T system displays more cooperativity (DR = 12 ± 5) than the 4 T system (DR = 100 ± 47). Furthermore, the 0 T system (black) shows a decrease in assembly when the concentration of component A is larger than the concentration of components B and C (1 μM) (see the reduction of trimer band and increase of dimer band, dimeric intermediates AB and AC). This phenomenon is also observed for the 2 T and 4 T systems, although requiring a much higher concentration of component A. This result is in good agreement with our numerical simulations (see Extended Data Fig. 6) Black = 0 T spacer, Green = 2 T spacer and Blue = 4 T spacer. *Of note, this experiment has been done at room temperature (~ 23 °C) whereas the fluorescence titration has been done at 37 °C. Also, because PAGE is a less sensitive technique compared to fluorescence titration, this experiment has been performed at 1 μM of components B and C compared to the 100 nM used in fluorescence titration. Nonetheless, both experiments demonstrate the same tendencies and are in good agreement with our numerical simulations. Binding curves of 2 T and 4 T spacer are fitted using the Hill equation (Eq. 5) while the binding curve of the 0 T spacer is fitted using a double dose-response curve.

https://doi.org/10.1038/s41557-022-01127-4



Extended Data Fig. 8 | Numerical simulation of the time-dependent assembly of a 3C system. Using the rate constants measured from Fig. S13 and Fig. S14, we can predict the kinetic traces of assembly of the dimer (plain line) and the trimer (dashed line) using numerical simulation. This enables us to test different concentration conditions before experimentation. (**a**) Changing [A] only affects dimer assembly while (**b**) changing [B] and [C] affects the formation of the trimer. This illustrates that the limiting step of trimer assembly, in these conditions, is the formation of the productive dimer BC. These results agree with our experimental data (Fig. 4b in main text and Fig. S16 and Fig. S17). Numerical simulations are performed using MATLAB®.

https://doi.org/10.1038/s41557-022-01127-4



Extended Data Fig. 9 | Adding a third component. (**a**) A trimeric assembly can also be created by introducing a new component, called the controller, that can interact with both the native DNAzyme and the substrate. (**b**) This controller strand can modulate the level of activity and the deactivation time via the formation of an inactive trimer. (**c**) In the presence of a 10 nM controller and DNAzyme, increasing the concentration of substrate increases the formation of the active dimer and the catalytic rate without substantially affecting the trimeric deactivation rate k_{Tri}. (**d**) At 100 nM substrate, increasing the concentration of controller and DNAzyme increases the rate of trimer formation and thus the rate of DNAzyme deactivation. Kinetic traces (*left* panel) are fitted using a single exponential (native) or a double exponential (with controller) to extract rate constants of dimeric and trimeric formation (that is, deactivation of the DNAzyme) (Eq. 6). The derivatives (*middle* panel) are plotted to better show the variation in initial rates and deactivation times. Fluorescence data were converted to the concentration of product using a calibration curve (data not shown) and normalized by the concentration of DNAzyme. All experiments have been done in PBS buffer (50 mM Na_2HPO_4, 400 mM NaCl, pH = 7.00) at 25 °C.

nature research

Corresponding author(s): Alexis Vallée-Bélisle

Last updated by author(s): Oct 6, 2022

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Fluorescence data were collected using the Cary Eclipse software (Agilent) and using the Applied Photophysics software. Gel analysis were performed using ChemiDoc XRS+ software.
Data analysis	Numerical simulations were performed using MatLab R2019b. Codes were created in house and are described in SI. All curves (binding curves, melting curves, urea denaturation and kinetic traces) were fitted using the KaleidaGraph software. All errors reported are the errors obtained from the best fit (n=1) or the average +/- standard deviation from a triplicat (n=3).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data supporting the findings of this study are available within the article, its Supplementary Information files, the Source Data file or from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Most data reported represent one single experiment (n=1) and the error reported represent the error of the fit. When triplicate (n=3) were made, the value represents the average +/- standard deviation.
Data exclusions	n/a
Replication	Most experiments (binding curves, kinetic trace, DNAzyme assay) were performs at least twice using completely independent samples to ensure that data are replicable (e.g. different batch of synthesized DNA, different stock solutions, different days, ...)
Randomization	n/a
Blinding	n/a

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging