ACCOUNTS of chemical research

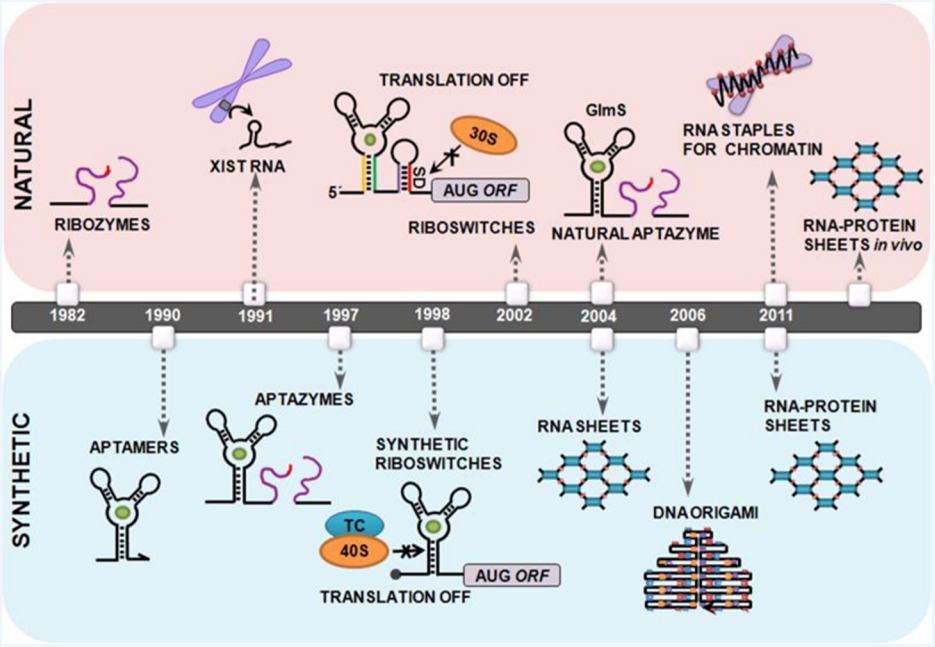
Article

pubs.acs.org/accounts

# The Predictive Power of Synthetic Nucleic Acid Technologies in RNA Biology

Saikat Chakraborty, Shabana Mehtab, and Yamuna Krishnan*

National Centre for Biological Sciences, TIFR, GKVK Bellary Road, Bangalore, 560065 Karnataka, India




**CONSPECTUS:** The impact of nucleic acid nanotechnology in terms of transforming motifs from biology in synthetic and translational ways is widely appreciated. But it is also emerging that the thinking and vision behind nucleic acids as construction material has broader implications, not just in nanotechnology or even synthetic biology, but can feed back into our understanding of biology itself. Physicists have treated nucleic acids as polymers and connected physical principles to biology by abstracting out the molecular interactions. In contrast, biologists delineate molecular players and pathways related to nucleic acids and how they may be networked. But *in vitro* nucleic acid nanotechnology has provided a valuable framework for nucleic acids by connecting its biomolecular interactions with its materials properties and thereby superarchitecture ultramanipulation that on multiple occasions has pre-empted the elucidation of how living cells themselves are exploiting these same structural concepts.

This Account seeks to showcase the larger implications of certain architectural principles that have arisen from the field of structural DNA/RNA nanotechnology in biology. Here we draw connections between these principles and particular molecular phenomena within living systems that have fed in to our understanding of how the cell uses nucleic acids as construction material to achieve different functions. We illustrate this by considering a few exciting and emerging examples in biology in the context of both switchable systems and scaffolding type systems. Due to the scope of this Account, we will focus our discussion on examples of the RNA scaffold as summarized.

In the context of switchable RNA architectures, the synthetic demonstration of small molecules blocking RNA translation preceded the discovery of riboswitches. In another example, it was after the description of aptazymes that the first allosteric ribozyme, glmS, was discovered. In the context of RNA architectures as structural scaffolds, there are clear parallels between DNA origami and the recently emerging molecular mechanism of heterochromatin formation by Xist RNA. Further, following the construction of well-defined 2D DNA−protein architectures, the striking observation of remarkably sculpted 2D RNA−protein hydrogel sheets in *Caenorhabditis elegans* speaks to the *in vivo* relevance of designer nucleic acid architectures. It is noteworthy that discoveries of properties in synthetic space seem to precede the uncovering of similar phenomena *in vivo*.

## INTRODUCTION

Synthetic or *in vitro* nucleic acid technologies occupy a unique and powerful space in that they exploit chemical insight to

**Special Issue:** Nucleic Acid Nanotechnology

**Received:** December 31, 2013
**Published:** April 8, 2014

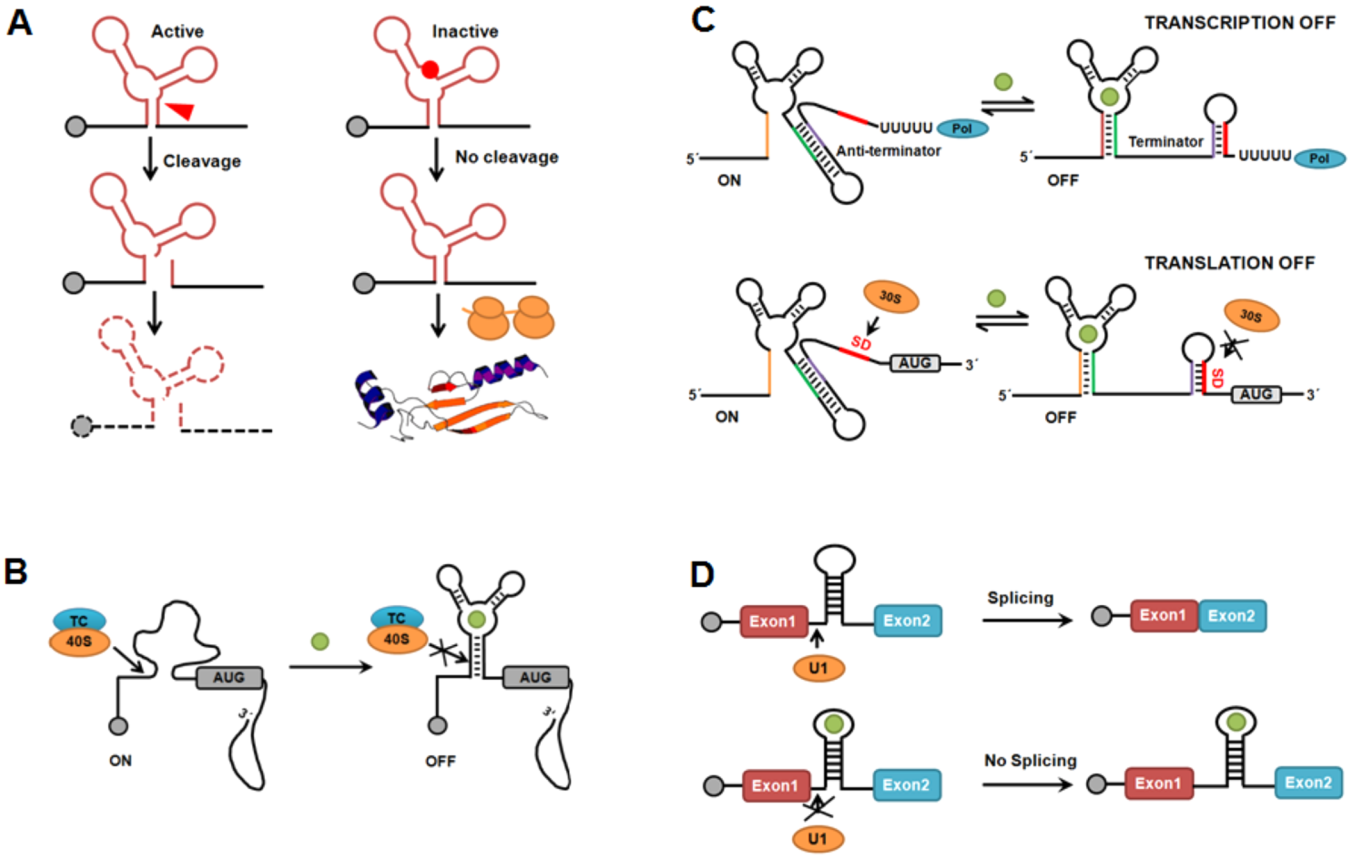



**Figure 1.** (A) Ribozymes can modulate translation by $Mg^{2+}$ dependent self-cleavage (red arrowhead) at a well-defined location. (B) Small molecule mediated inhibition of eukaryotic translation. (C) Natural riboswitches can terminate transcription (top) or inhibit translation (bottom). (D) Synthetic riboswitches can modulate splicing. Small molecule effectors are indicated as green circles.

manipulate biomolecular interactions and simultaneously straddle the mechanophysical properties of nucleic acid filaments that are central to its use as a material. This has generated not only the ability to control nucleic acid superstructure to near-angstrom level precision as manifested in superhierarchical, rigid architectures but also the ability to achieve controlled, ultramanipulation of functional nucleic acid structures as seen in the realization of various dynamic or switchable architectures.[1−3] The general perception of directed nucleic acid assembly has evolved from one of fascinating yet esoteric molecular gymnastics to one of creative exploitation for neat applications. In this Account, we discuss that beyond all of this, its importance lies in it being a hotbed to uncover engineering principles unique to this molecular scaffold that can feed back into a better understanding of how Nature herself exploits nucleic acids. Using selected examples related to both rigid scaffolds and dynamic devices, we illustrate how organizational principles gleaned from *in vitro* nucleic acid structure manipulation have pre-empted landmark discoveries in biology where the same design principles have been exploited by Nature.

## ■ DYNAMIC DEVICES: FUNCTIONAL NUCLEIC ACIDS

The discovery of catalytic RNAs in the early 1980s opened new vistas in RNA function, whose importance formerly lay confined to protein production. RNA sequences with enzymatic function, termed "ribozymes", were first described in *Tetrahymena* where nuclear pre-mRNA of the 23S rRNA excised an intervening sequence from adjacent exons without protein assistance.[4] All identified naturally occurring ribozymes may be functionally classified as either cleaving or splicing ribozymes. The general mechanism involves a nucleophilic attack of a polarized water molecule on an adjacent phosphate in the RNA backbone, resulting in well-defined cleavage products. However unlike ribonucleases, ribozymes cleave at a unique location, determined by base-pairing and tertiary interactions mediated by divalent cations, particularly $Mg^{2+}$, to form an "active conformation" crucial for cleavage (Figure 1A). Despite the small number of naturally occurring ribozymes, each new discovery has brought considerable excitement.[5] Given their small size, ease of chemical synthesis, and facile reconstitution under diverse conditions, ribozymes have proved versatile tools for extrinsic control of gene expression in living systems. For instance, one can inactivate mRNA transcripts by incorporating self-cleaving ribozymes as demonstrated by the inclusion of the Sm1 ribozyme within a lacZ reporter in HEK293 cells.[6] Similarly, one can selectively destabilize pre-mRNA transcripts by inhibiting their effective splicing through the incorporation of the hammerhead ribozyme (HHR) within the exon of the $\beta$-globin gene.[7] Given that splicing exploits kinetic control, it was shown that the incorporation of ribozymes with different cleavage rates, such as the HHR or the faster hepatitis delta virus (HDV) ribozyme, in the same location of a given pre-mRNA transcript could change the processing fate of the pre-mRNA.[8]

While many naturally occurring, functional RNA motifs have been harnessed to modulate gene expression, *in vitro* selection strategies have allowed researchers to access a palette of


dx.doi.org/10.1021/ar400323d | *Acc. Chem. Res.* 2014, 47, 1710−1719

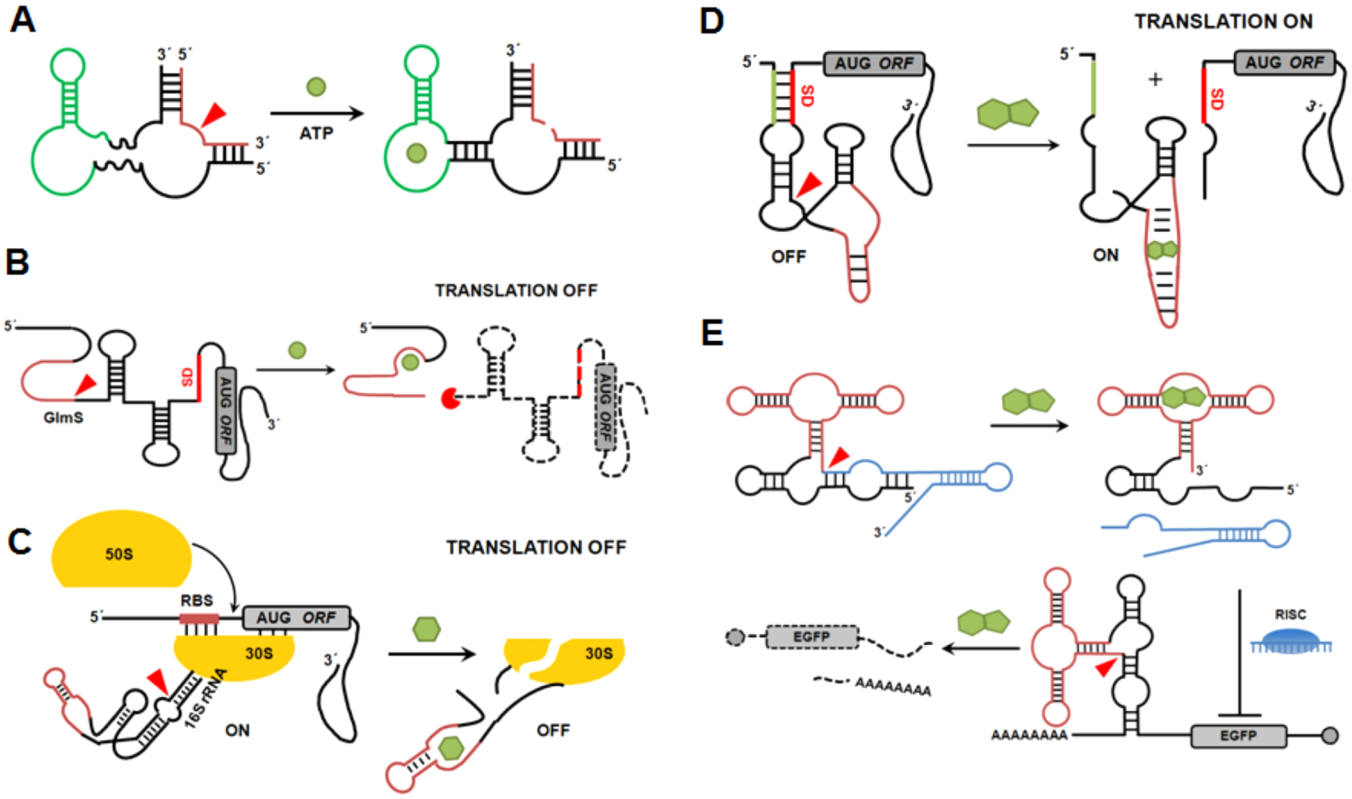



**Figure 2.** (A) Design concept of allosteric ribozymes or "aptazymes". (B) Naturally occurring allosteric ribozyme, *glmS* ribozyme. Effector-induced mRNA cleavage can switch translation (C) OFF or (D) ON or (E) activate RNA processing and poly(A) tail cleavage. Small molecule effectors are indicated as green circles or polygons. Cleavage sites are indicated by red arrowheads.

orthogonal chemical toggles of gene expression. *In vitro* selection of a binder from a library of nucleic acid sequences (SELEX) against a target molecule has provided a new paradigm to obtain DNA and RNA molecules with novel functions.[9] SELEX affords "aptamers", which are single strands of DNA or RNA that form stable structures in complex with a specific ligand. A landmark demonstration of such novel function in 1998, was small-molecule mediated control over mRNA translation *in vivo* that was achieved synthetically by the integration of an aptamer into an mRNA (Figure 1B). An aptamer to a Hoechst dye was incorporated into the 5′ untranslated region (UTR) of the $\beta$-galactosidase gene, which was transfected into Chinese hamster ovary (CHO) cells. In the presence of Hoechst, the resultant mRNAs formed stable Hoechst−aptamer complexes, which blocked the scanning of 43S subunit on the mRNA and thereby inhibited its translation.[10]

Subsequently in 2002, "riboswitches" were codiscovered independently in two different contexts. Breaker and colleagues showed that the 5′ leader sequences of *Escherichia coli* btu mRNAs bind a small molecule, namely, the coenzyme B12. This complex sequesters the ribosome binding site (RBS) of btu mRNA, which then suppresses its translation.[11] Simultaneously Nudler et al. showed that a feedback regulation of riboflavin and thiamin genes occurs by the complexation of small molecules, that is, flavin mononucleotide or thiamine pyrophosphate (TPP), with the leader region on their corresponding RNAs, which in turn attenuated transcription.[12] Riboswitches are thus regulatory domains on mRNAs that comprise a naturally evolved aptamer to a small molecule, such as a metabolite, that functions as an effector molecule. Here the aptamer domain is interfaced to an expression platform that, in prokaryotes, modulates gene expression either at the transcriptional or at the translational level[13] (Figure 1C). In eukaryotes such as *Neurospora crassa*, a TPP-specific riboswitch within the intron of NMT1 RNA was shown to regulate alternative splicing of the host transcript.[14] Synthetic control of gene expression using a tetracycline aptamer that controlled pre-mRNA splicing was almost immediately realized thereafter[15] (Figure 1D).

With the discovery of riboswitches in 2002, it became obvious that Nature used essentially the same synthetic design concept outlined by Green et al. in 1998, where small-molecule complexation by mRNA domains could prevent translation and thereby regulate gene expression quickly and efficiently. Thus the molecular recognition capabilities of RNA to diverse ligands has been expanded to construct a panoply of synthetic molecular switches, toggled by the specific binding of diverse effector molecules.[16]

Small molecule effectors can bind to the allosteric sites in enzymes and alter catalytic function via structural changes to bring about allosteric regulation.[17] Although RNAs were known to be catalytic and also recognize ligands, no natural ribozyme was known to operate as a true allosteric enzyme. In 1997, a synthetic "aptazyme" was realized *in vitro* by Breaker et al. that brought the catalytic activity of HHR under the control of ATP as an effector molecule by fusing a synthetic ATP aptamer to the HHR thus realizing a synthetic allosteric ribozyme[18](Figure 2A). It was only thereafter that the *glmS* ribozyme, a natural allosteric ribozyme governing the synthesis of glucosamine-6-phosphate (GlcN6P) in Gram-positive bacteria, was discovered. The *glmS* ribozyme undergoes self-cleavage upon binding of


dx.doi.org/10.1021/ar400323d | *Acc. Chem. Res.* 2014, 47, 1710−1719

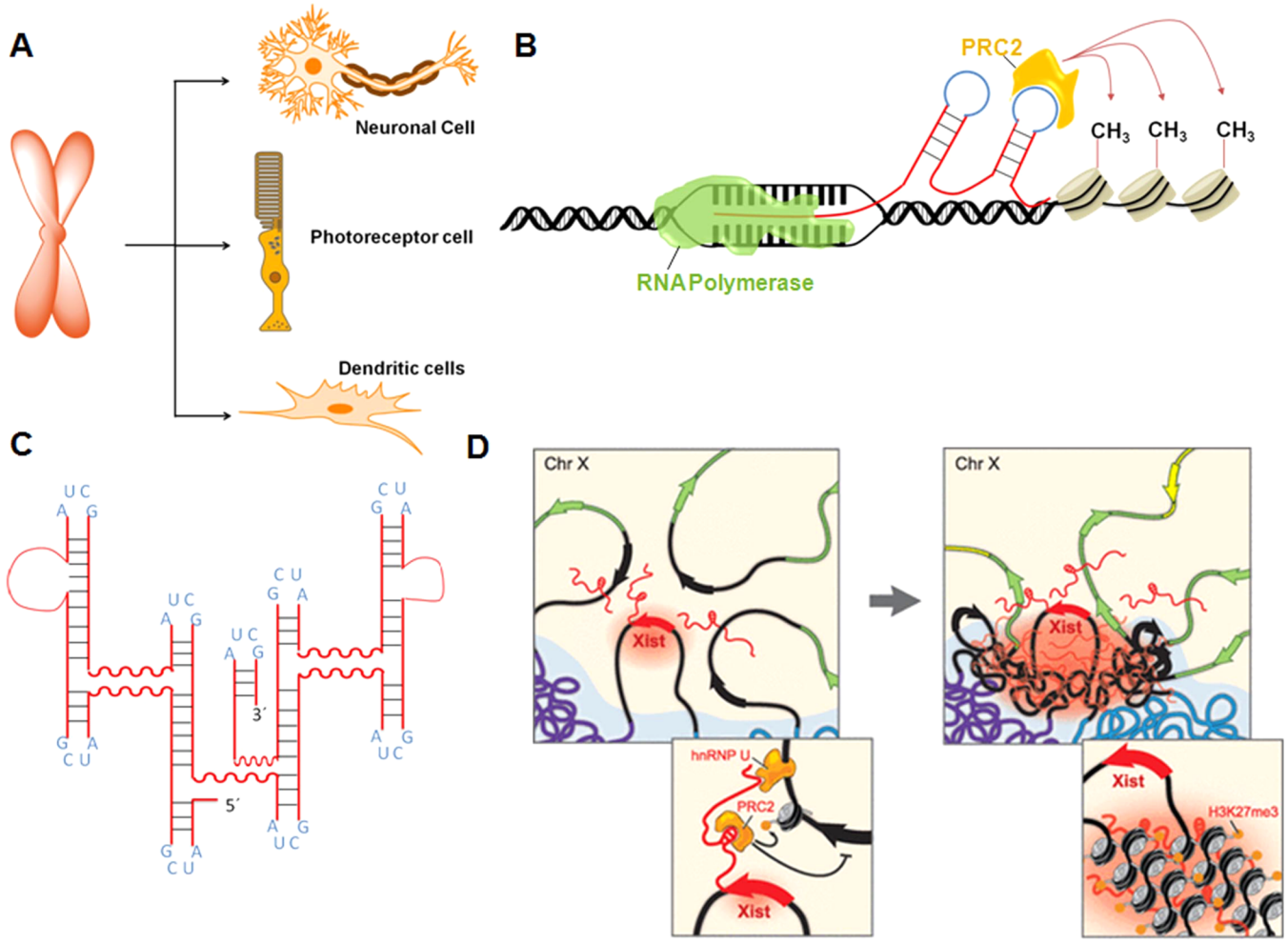



**Figure 3.** Cellular pseudo-origami. (A) Differential accessibility of same genomic DNA yields different cellular architectures. (B) Chromatin tethered Xist RNA-mediated, site-specific histone methylation. (C) Secondary structure of Xist 'A' repeats displaying AUCG tetraloops. (D) Xist alters 3D genomic architecture by "stapling" remote sites on Xi chromatin. Panel D reproduced with permission from ref 49. Copyright 2013 AAAS.

GlcN6P leading to the degradation of its host mRNA[19](Figure 2B). More recently a natural allosteric group I ribozyme was identified in *Clostridium difficile* where self-splicing is mediated by a distinct riboswitch class that senses cyclic-di-GMP.[20]

The mechanistic elucidation of *in vivo* allosteric ribozymes has in turn spurred the diversity of allosteric control of gene expression using synthetic RNA modulators.[21] Gene regulation has been achieved in prokaryotes by different mechanisms, for example, coupling a theophylline-responsive aptamer and HHR to an mRNA domain that is complementary to the ribosome binding site (RBS). Addition of theophylline activates HHR, which in turn liberates the RBS, switching translation ON[22] (Figure 2D). Alternatively regulation can be achieved by inserting a thiamine-responsive aptazyme into the 16SrRNA of an orthogonal 16SrRNA−mRNA pair in *E. coli*. Addition of thiamine leads to cleavage of 16SrRNA switching translation OFF[23] (Figure 2C). The RNAi machinery in mammalian cells can be toggled using small molecules. Recently, a guanine-responsive aptazyme (GRA) module was used to achieve dual action on the same gene in cells. Addition of guanine activated the GRA to cleave and release a pri-miRNA transcript that underwent further processing to release RNAi that specifically silences a given mRNA. The mRNA on the other hand also contained a GRA at the poly(A) tail, which underwent simultaneous cleavage from the mRNA (Figure 2E). Both in concert silence the reporter gene expression with enhanced efficiency.[24]

## RIGID SCAFFOLDS: NUCLEIC ACID ORIGAMI

The remarkable physicochemical properties of DNA lie at the heart of the success of DNA origami. It consists of folding long, single-stranded scaffold DNA (ssDNA), typically a phage genome, into literally any desired shape and topology with nanoscale precision, with the aid of shorter DNA strands called staples.[25] These DNA staples direct the overall 2D or 3D shape adopted by the long scaffold DNA strand and rely on sequence complementarity with the long scaffold.[26,27] This is now the predominant route to organize DNA into diverse, intricately sculpted nanoscale architectures. Since the short DNA staples can be integrated to functional modules through a rich repertoire of chemistries,[28] one can uniquely position desired functional groups via the staples on a mesoscale architecture with nanoscale precision. DNA is thus fast becoming a versatile medium with which to organize nanomaterials in programmable and addressable ways. Along richer veins, RNA is also emerging as a powerful substrate for programmable self-assembly despite being chemically more labile than DNA.[29] RNA's structural and therefore functional space far eclipses that of DNA while still retaining the powerful properties of antisense recognition and modularity, plainly evident from


dx.doi.org/10.1021/ar400323d | *Acc. Chem. Res.* 2014, 47, 1710−1719

the explosion of layers of RNA-based gene regulation in biology.[30] For example, naturally occurring RNA tertiary motifs such as K-turns, kissing-loop interactions, tetraloop-receptor interactions, or tetramolecular motifs offer a plethora of association modes for RNA building blocks, whereas DNA architectures overwhelmingly rely on sticky-end cohesions or *de novo* designed, laboriously realized artificial building blocks.[31,32] Such tertiary motifs can be directly implemented in synthetic "tectoRNA" assemblies to realize designer RNA scaffolds.[33] Further, the principles of DNA origami can be directly co-opted into the RNA domain where an *in vitro* transcribed scaffold RNA strand can be sculpted into superarchitectures by short DNA staples.[34]

## CELLULAR RNA PSEUDO-ORIGAMI

The delineation of DNA origami in 2006 and the remarkable precision to which the same DNA strand can be differentially compacted in 2D and 3D reveal the following fundamental organizational principle inherent in filamentous DNA:[25,27] irrespective of the primary sequence of the long strand, the capacity to proximally attach pairs of selected distant locations uniquely in space, identified and addressed by their sequence, can lead to distinct superstructures.[35] The combinatorics of such pairwise attachments on a single filament leads to the diversity of superarchitectures. From a series of important findings in a seemingly unrelated field, it is recently emerging that Nature seems to be exploiting this powerful organizational principle to control nuclear architecture of genomic DNA. Within a given multicellular organism, a neuron, a photoreceptor cell, or a dendritic cell all have the same content of genomic DNA within their nuclei (Figure 3A). Yet the morphologies and functions of the resultant cells are completely different. Since 1979, it has been known that the difference in cellular architecture and function is due to distinct transcription programs that arise from the differential 3D organization of genomic DNA within their nuclei.[36,37] Meter-long genomic DNA is packaged by histone proteins as chromatin and condensed within a micrometer-sized cell nucleus. In order to transcribe this highly packaged DNA within the cell, sequence information needs to be accessed in a finely orchestrated manner, through this highly condensed chromatin. Based on transcription, chromatin is broadly divided into active, relatively loosely packaged "euchromatin" and inactive, highly packaged "heterochromatin". Heterochromatin has been shown to function as crucial nodes that impart structural integrity to the 3D nuclear architecture.[38] Thus differential 3D packaging of the same genomic DNA within the space of the cell nucleus leads to different transcription programs and therefore different cell types. One of the first steps in the packaging of genomic DNA is the formation and establishment of heterochromatin from loosely packed euchromatin.

"Active" and "inactive" chromatin domains are demarcated by signature histone and DNA modifications.[39] For instance, histone H3 lysine 4 (H3K4) on its $\varepsilon$-$NH_2$ can be monomethylated, dimethylated, or trimethylated, which demarcate enhancers, active genes, and promoters, respectively, while H3K36 trimethylation marks elongation, that is, actively transcribed regions.[39] These modifications are deposited by histone methyltransferases, for example, MLL proteins in mammals. A long-standing enigma was how such enzymes, with little or no DNA sequence specificity, were able to find their correct target regions dispersed throughout genomic DNA and achieve exquisite target-specific action. As in origami, this is also a problem where nanoscale precision is critical to mesoscale ordering. A flurry of exciting recent findings have revealed that heterochromatin is established at distinct sites on chromatin by long noncoding RNAs (lincRNAs), for example, HOTAIR, Xist, and TERRA, where the general mechanism of lincRNA action bears a palpable resemblance to the short staple strands and where chromatin acts as the long scaffold DNA strand.[40] The interaction of chromatin with these lincRNAs changes its compaction state and thereby transcriptional activity. These findings reveal that lincRNAs bind designated sites on chromatin and importantly target chromatin through DNA sequence specificity.[41] LincRNAs thus precisely position multiprotein complexes on the chromatin scaffold site-specifically to chemically modify the chromatin and bring about targeted heterochromatinization. Thus, the lincRNA "staples" and the chromatin "scaffold strand" that sculpt 3D genomic DNA architecture can be considered cellular "pseudo-origami". We describe the emerging parallel between cellular pseudo-origami and *in vitro* DNA origami.

Paul and Duerksen observed in 1975 that biochemically pure chromatin contained twice as much RNA as DNA, implying RNA's intimate connection with chromatin structure.[42] Modern sequencing now reveals that 90% of the genome is pervasively transcribed into noncoding RNAs. A subset of these, known as long intergenic RNAs (lincRNAs, >200 nt), is associated with chromatin. They shape histone and DNA modifications and are implicated in processes such as dosage compensation, imprinting, pluripotency, development, cancer metastasis, and DNA repair.[40] Most well described lincRNAs thus far tend to control gene expression by targeting chromatin-modifying complexes site-specifically onto chromatin to locally induce heterochromatin formation given that the latter do not possess DNA sequence specificity. Mechanisms by which lincRNAs guide the chromatin-modifying complexes remain elusive although evidence for RNA/DNA hybrids, RNA/dsDNA triplexes, or specific RNA binding protein adapters are emerging.[41] Using Xist RNA as an example, we illustrate how lincRNAs initiate the "stapling" of chromatin at specific sites to form heterochromatin.

In mammals females possess two X-chromosomes and therefore require one of these to be silenced by heterochromatinization for dosage compensation between XY males and XX females. This X chromosome inactivation (XCI) is mediated by a 17 kb lincRNA, called Xist RNA.[43] It is responsible for nucleating and spreading heterochromatin on the X chromosome for XCI in female somatic cells. Despite being extensively studied since 1991, the exact mechanism of Xist-mediated XCI is emerging only now.[43] XCI begins with the two X chromosomes transiently pairing to redistribute transcriptional activators leading to a "symmetry breaking" between the two X alleles. This results in one of the X chromosomes expressing Tsix RNA that is antisense to Xist and originates from the complementary DNA strand. Tsix RNA and its activator Xite are indispensable to pair or "staple" the two X alleles together.[44] Post-symmetry breaking, Tsix RNA transcripts are lost by the X allele that is ultimately inactivated (Xi), and it starts transcribing Xist RNA that remains tethered to its site of transcription on the chromatin via an RNA/DNA/RNAPol II ternary complex (Figure 3B). This Xist RNA then has to nucleate at and spread to remote sites on the rest of Xi for its complete silencing. A transcription factor YY1 is also required to tether Xist RNA at the nucleation site on Xi.[45] YY1

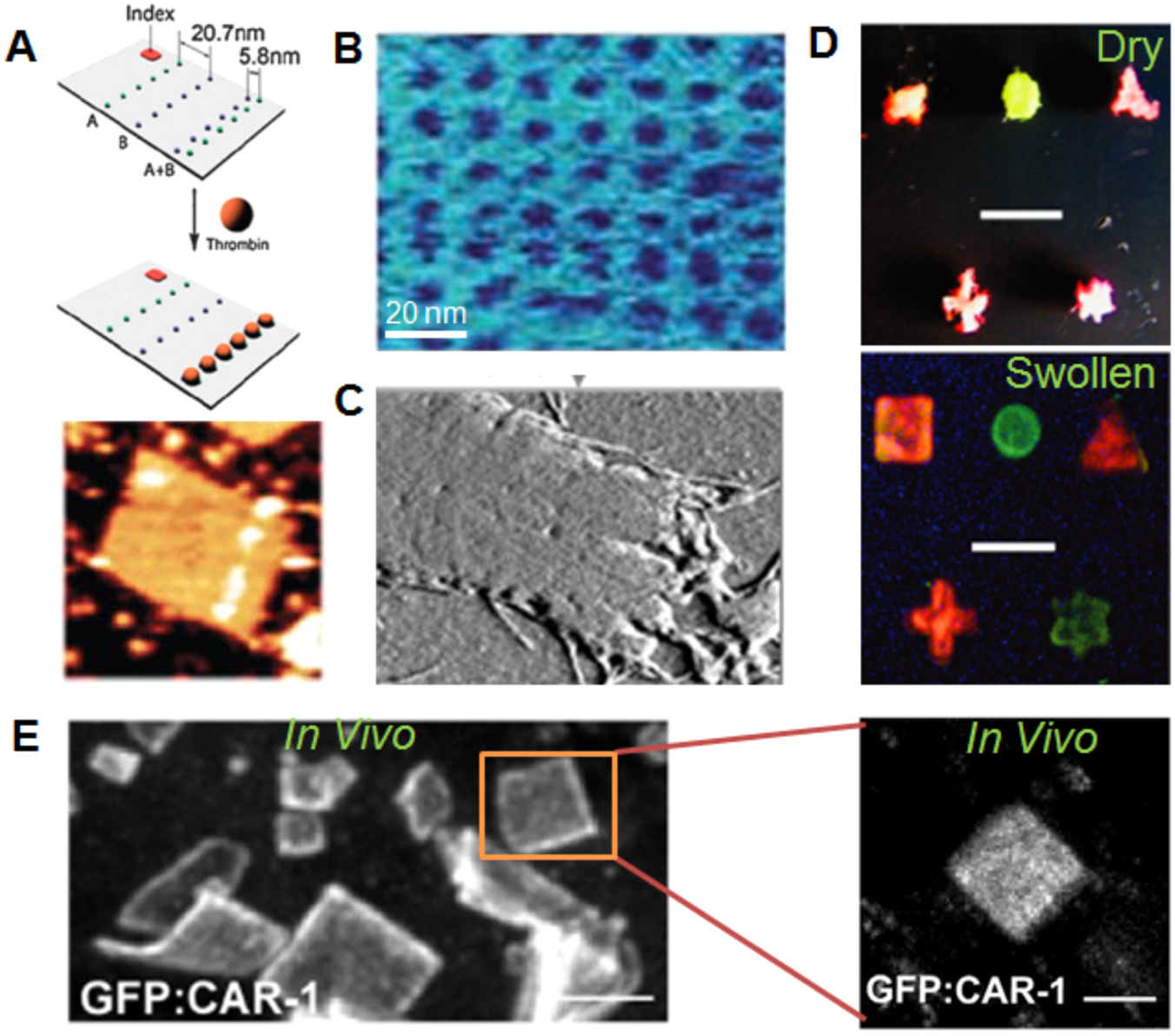



**Figure 4.** Designer *in vitro* and *in vivo* nucleic acid scaffolding. (A) Thrombin positioned on DNA origami; (B) 2D RNA sheet; (C) synthetic RNA–protein origami inside bacteria; (D) dry (top) and swollen (bottom) hydrogels from X-DNA (scale bars, 1 cm); (E) natural RNP hydrogels in *C. elegans* oocytes (scale bars, 5 $\mu$m). Panel A reproduced with permission from ref 51, copyright 2008 NPG. Panel B reproduced with permission from ref 29, copyright 2004 AAAS. Panel C reproduced with permission from ref 53, copyright 2011 AAAS. Panel D reproduced with permission from ref 61, copyright 2006 Nature Publishing Group. Panel E reproduced with permission from ref 59, copyright 2013 Cell Press.

contains a DNA binding domain and an RNA binding domain acting as a bifunctional adaptor to site-specifically lock Xist RNA on the target chromatin.

The 5′-end of human Xist RNA contains ~8 repeats called "Repeat A", which binds the PRC2 complex and targets it site-specifically on chromatin where it deposits H3K27 methylation marks, which is the first step in heterochromatin formation.[46] "Repeat A" is 26 nt, it is highly conserved, and its *in vitro* structure is partially established[47] with the more probable structures being those that display a conserved AUCG tetraloop important for PRC2 binding[48] (Figure 3C). This AUCG tetraloop is indispensable for Xist-mediated heterochromatin formation.[48] The display of multiple A-repeats on Xi either from single or multiple Xist RNAs creates a high local concentration of AUCG tetraloops in close spatial proximity. The presentation of multiple AUCG motifs by inter-repeat duplexation would increase the binding avidity of key RNA binding proteins that would otherwise have low binding affinities to a single AUCG motif and thus help the spreading of heterochromatin on Xi.

At the start of XCI, Xist from the as yet unsilenced Xi needs to seek out its remote target sites on the Xi chromosome.[49] The transcribing Xist RNA locks onto those transcriptionally active genes of Xi that are spatially closely positioned and silences these loci first. This through-space transfer of Xist to chromatin sites that are remote from its site of transcription is mediated by its interaction with hnRNP U, a nuclear matrix protein (Figure 3D). Being more compact than euchromatin, heterochromatin formation changes the 3D chromatin structure on Xi thereby repositioning adjacent euchromatin regions.[49] During the course of this repositioning, Xist reels in new regions of euchromatin closer to the Xist locus and allows spreading by


dx.doi.org/10.1021/ar400323d | *Acc. Chem. Res.* 2014, 47, 1710–1719

proximity transfer. Intermolecular duplexation between two or multiple Xist RNAs via the "Repeat A" motif could function to bridge these interactions acting as staples on a longer chromatin scaffold strand.

## ■ *IN VIVO* NUCLEIC ACID SELF-ASSEMBLY

Scaffolding is a universal engineering principle also exploited by Nature. For instance, multiprotein signaling complexes are often scaffolded into spatially organized clusters.[50] Such organization helps channel substrates between interacting enzymes, limits cross-talk and increases the yields of sequential metabolic reactions. Primarily, scaffolds exert their effects by simply tethering molecular partners to increase the effective concentrations of enzymes or substrates. Thus the spatial organization of biomolecules in 3D to control signaling or multienzyme cascades remains one of the Holy Grails of nucleic acid nanotechnology[51] (Figure 4A). As early as 2009, it was shown that when glucose oxidase and horseradish peroxidase were perfectly positioned on a DNA strip, it allowed the product of the first enzyme to act as the substrate for second, resulting in an enzyme cascade of enhanced efficiency.[52] However, such applications of DNA-based scaffolding have remained *in vitro* due to challenges associated with producing ssDNA inside cells. For applications related to *in vivo* nucleic acid scaffolding, RNA presents a potent medium because it is genetically encodable (Figure 4B,C). Subsequently in 2011, a 2D synthetic RNA sheet[53] that site-specifically displayed protein-binding aptamers was assembled within bacteria such that protein players could localize with controllable orientation and proximity (Figure 4C). By positioning of ferredoxin and bacterial [Fe−Fe]-hydrogenase, two hydrogen synthesis enzymes, on these RNA sheets, hydrogen production was enhanced in bacteria. Thus a compartment-free spatial sequestration of a biochemical process could be achieved within the crowded cellular milieu.[53] It is noteworthy that although membrane compartmentalization is used for specific reactions to occur in isolation from the rest of the cell, several intracellular bodies that are not membrane delimited can also achieve highly efficient, localized intracellular chemistry. A large and important class of such intracellular assemblies are RNA−protein (RNP) complexes or RNP granules, for example, nucleoli, Cajal bodies, speckles, P-bodies, and P-granules.[54] Located in the nucleoplasm or cytoplasm, they spatiotemporally separate biochemical processes from the rest of the cell and play key roles in growth, development, and homeostasis. Changes in their composition and structure lead to pathological conditions such as Huntington's disease or spinal muscular dystrophy.[55]

RNP granules are typically highly dynamic at the molecular level, exchanging subunits with bulk nucleoplasm on typical time scales of tens of seconds. A physical description of such assemblies started to emerge in 2012 where they were considered as a viscoelastic hydrogel-like state formed by an intracellular phase transition involving liquid−liquid demixing.[56] RNP granule formation and disintegration presumes that reversible multivalent interactions between individual RNPs and other molecular partners lead to phase transitions between diffuse (gas-like), liquid, or solid states.[56] Multivalency can be provided by RNAs displaying multiple binding sites for either RNAs or proteins, since many RNP-related proteins can induce sol−gel transitions or liquid−liquid demixing to form dynamic hydrogels in reconstituted reaction mixtures *in vitro*.[57] Liquid-like condensation is strongly reflected in germ granules found in *Caenorhabditis elegans* embryos and nucleoli in *Xenopus* oocytes.[58] However, RNPs can also polymerize into well-defined, sculpted solid architectures. In 2013, the Evans group showed that RNPs that form germ line P-bodies (grPBs) in *C. elegans* oocytes can phase separate from a diffuse gas-like phase to a liquid-like phase and these states are regulated by early developmental cues *in vivo*.[59] Translational mRNA regulators induce specific conserved RNP ($\eta \approx 1$ Pa·s) components to coassemble into large, viscoelastic, semiliquid ($\eta \approx 1000$ Pa·s) grPB granules.[59] Such a high viscosity of RNP droplets restricts repressed RNPs from the cytosol where active translation occurs. An RNA helicase CGH-1 maintains the liquid-like state of the RNP granule preventing a phase transition into the solid-like state[59] (Figure 4E). Dramatically, the loss of this RNA helicase leads to the formation of solid-like structures with astonishingly well-defined geometry within the *C. elegans* oocyte. These RNA sheet-based structures have lengths up to 10 $\mu$m, >300 nm depth, and cornices with ∼90° angles.[59] Despite minor length variation, these sheets maintain a length/width ratio of 0.99. Such regular, geometrically constrained growth is suggestive of specific interactions between RNP components.[59] Solid RNP aggregates are often associated with neurological disorders.[55] These suggest that RNPs can assemble into a variety of supramolecular states that are carefully regulated toward specific functions.

Recently both in mouse brain and in human cell extracts, proteins with low-complexity (LC) sequence domains (regions with low amino acid diversity) separate into a different phase together with RNA during liquid−liquid demixing.[57,60] These studies suggest a model in which RNAs bind to RNA binding proteins, which in turn phase separate using their LC domains. LC domains, abundantly found in RNA binding proteins, are now thought to promote a hydrogel-like state by liquid−liquid demixing. This directly recapitulates hydrogel systems described in 2006 when Dan Luo's group realized a DNA-based hydrogel using the X-DNA motif as a scaffold.[61] A branched multivalent DNA structure, X-DNA contains palindromic sticky ends that aid gelation and that are further stabilized by ligation (Figure 4D). Recently a ssDNA template and a polymerase were shown to amplify and noncovalently weave DNA into a hydrogel.[62] The resulting hydrogel, called a meta-hydrogel, has liquid-like properties when taken out of water and solid-like properties when in water. Moreover after completely disintegrating when out of water, the addition of water causes the hydrogel to reform where it exhibits shape-memory, similar to solid RNP aggregates formed in *C. elegans*. The importance of such systems is not restricted to materials or biological applications; they serve as pioneering examples that could provide a framework to understanding similar complex phenomena *in vivo*.

## ■ CONCLUSIONS AND OUTLOOK

Using selected examples, we have tried to illustrate the predictive power of synthetic nucleic acid technologies in identifying new functions of endogenous RNAs in biology. Although the original vision of the field was to position matter in 3D with maximal precision using nucleic acids, it is now proving to be a fertile hub to discover fundamental organizational principles for nucleic acids that Nature uses for similar functions *in vivo*. For example, hierarchical self-assembly of DNA nanostructures into defined polyhedral architectures could in principle shed light on viral capsid assembly.[63,27,64] Viral capsids that are icosahedral are an outstanding example of

natural self-assembly where modular protein units hierarchically assemble through diffusion, reproducibly, into well-defined polyhedra devoid of any kinetic and thermodynamic traps.[65] Understanding the rules underlying such robust supramolecular assembly could be important for the development of synthetic assemblies for applications in gene therapy, drug delivery, and vaccine development.

The strong parallels between the operating principles of DNA origami and chromatin compaction could inspire the expansion of the existing repertoire of origami-based architectures. To date, DNA and RNA origami have yielded largely static architectures. Yet, chromatin states occupy specific ordered 3D organizations, and these ordered 3D states in the nucleus (i) change progressively during the process of differentiation or (ii) can switch between specific organizations reversibly, say during cell division. It is now well recognized that the appearance or removal of lincRNAs is involved in altering chromatin compaction. The charged chromatin filament is heavily compacted by chemical modifications on DNA and histones brought about by lincRNA guided enzymes. On the other hand, the removal of the RNA could induce the decompaction of the chromatin filament into a well-defined structure exploiting the entropic regime. With respect to DNA origami, very rare examples do exist of static architectures that irreversibly convert to another distinct static state upon the addition of molecular cues. These include DNA architectures such as the clam-shell,[66] the icosahedron,[67] and the box[68] that may all be "opened" by the addition of chemical triggers as well as specific nucleic acid sequences. However, there exists a single example by way of Turberfield's tetrahedron, which though not origami may be expanded and contracted reversibly between specific static shapes by the introduction of a nucleic acid sequence.[69] This suggests that it might be possible to realize a new class of origami-based architectures that can flip reversibly between specific structured states by the introduction or removal of shorter staple strands. Such shape-changing origami could be reductionist systems to understand principles behind chromatin state change that are still a mystery.

DNA and RNA origami could be exploited in the context of bacterial genome architecture, which is much simpler than complex eukaryotic genomes. New technological developments, especially sequencing, propel core molecular assays such as HiC, which in conjunction with novel computational methods to analyze "big data" could prove game-changing in understanding and manipulating genome architecture. Chromosome conformation capture coupled with deep sequencing (HiC) and high resolution fluorescence *in situ* hybridization (FISH) technologies show that the bacterial genome of *Caulobacter crescentus* adopts a bottle-brush like structure composed of arrays of supercoiled plectonemes.[70] Sequencing methods[71] integrated to powerful computational approaches[72] could provide a potent future combination to realize synthetic plectonemic architectures that could function as reductionist synthetic genomes.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: yamuna@ncbs.res.in.

### Notes

The authors declare no competing financial interest.

### Biographies

**Saikat Chakraborty** received his Bachelors in Chemistry from Presidency College, Kolkata, in 2003 and M.Sc. in Biochemistry from University of Calcutta in 2005. Currently, he is a Ph.D. student at NCBS. His current research interests are in structural biology of miRNA biogenesis.

**Shabana Mehtab** obtained her Bachelors and Masters in Biochemistry from Sri Krishnadevaraya University, Anantapur, in 2004. Currently she is a Ph.D. student at NCBS and her research interest is in RNA splicing.

**Yamuna Krishnan** started her independent laboratory at NCBS, Bangalore, in 2005 and is a Wellcome Trust-DBT India Alliance for a Senior Research Fellowship. Her research interests have evolved from self-assembly and supramolecular chemistry to the applications of nucleic acids in probing and programming living systems.

## ACKNOWLEDGMENTS

Y.K. thanks Prof. Eric Westhof for a critical reading of the manuscript and the Wellcome Trust-DBT India Alliance.

## REFERENCES

(1) Krishnan, Y.; Bathe, M. Designer Nucleic Acids to Probe and Program the Cell. *Trends Cell Biol.* **2012**, *22*, 624−633.
(2) Bhatia, D.; Sharma, S.; Krishnan, Y. Synthetic, Biofunctional Nucleic Acid-Based Molecular Devices. *Curr. Opin. Biotechnol.* **2011**, *22*, 475−484.
(3) Krishnan, Y.; Simmel, F. C. Nucleic Acid Based Molecular Devices. *Angew. Chem., Int. Ed.* **2011**, *50*, 3124−3156.
(4) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. Self-Splicing RNA: Autoexcision and Autocyclization of the Ribosomal RNA Intervening Sequence of Tetrahymena. *Cell* **1982**, *31*, 147−157.
(5) Doherty, E.; Doudna, J. Ribozyme Structures and Mechanisms. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 457−475.
(6) Yen, L.; Svendsen, J.; Lee, J.; Gray, J. T. Exogenous Control of Mammalian Gene Expression through Modulation of RNA Self-Cleavage. *Nature* **2004**, *431*, 471−476.
(7) Dye, M. J.; Gromak, N.; Proudfoot, N. J. Exon Tethering in Transcription by RNA Polymerase II. *Mol. Cell* **2006**, *21*, 849−859.
(8) Fong, N.; Ohman, M.; Bentley, D. L. Fast Ribozyme Cleavage Releases Transcripts from RNA Polymerase II and Aborts Co-Transcriptional Pre-mRNA Processing. *Nat. Struct. Mol. Biol.* **2009**, *16*, 916−922.
(9) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346*, 818−822.
(10) Werstuck, G.; Green, M. R. Controlling Gene Expression in Living Cells through Small Molecule-RNA Interactions. *Science* **1998**, *282*, 296−298.
(11) Nahvi, A.; Sudarsan, N.; Ebert, M. S.; Zou, X.; Brown, K. L.; Breaker, R. R. Genetic Control by a Metabolite Binding mRNA. *Chem. Biol.* **2002**, *9*, 1043−1049.
(12) Mironov, A. S.; Gusarov, I.; Rafikov, R.; Lopez, L. E.; Shatalin, K.; Kreneva, R. A.; Perumov, D. A.; Nudler, E. Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria. *Cell* **2002**, *111*, 747−756.
(13) Serganov, A.; Nudler, E. A Decade of Riboswitches. *Cell* **2013**, *152*, 17−24.
(14) Cheah, M. T.; Wachter, A.; Sudarsan, N.; Breaker, R. R. Control of Alternative RNA Splicing and Gene Expression by Eukaryotic Riboswitches. *Nature* **2007**, *447*, 497−500.
(15) Weigand, J. E.; Suess, B. Tetracycline Aptamer-Controlled Regulation of Pre-mRNA Splicing in Yeast. *Nucleic Acids Res.* **2007**, *35*, 4179−4185.

(16) Wittmann, A.; Suess, B. Engineered Riboswitches: Expanding Researchers' Toolbox with Synthetic RNA Regulators. *FEBS Lett.* **2012**, *586*, 2076−2083.
(17) Goodey, N. M.; Benkovic, S. J. Allosteric Regulation and Catalysis Emerge via a Common Route. *Nat. Chem. Biol.* **2008**, *4*, 474−482.
(18) Tang, J.; Breaker, R. R. Rational Design of Allosteric Ribozymes. *Chem. Biol.* **1997**, *4*, 453−459.
(19) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. Control of Gene Expression by a Natural Metabolite-Responsive Ribozyme. *Nature* **2004**, *428*, 281−286.
(20) Lee, E. R.; Baker, J. L.; Weinberg, Z.; Sudarsan, N.; Breaker, R. R. An Allosteric Self-Splicing Ribozyme Triggered by a Bacterial Second Messenger. *Science* **2010**, *329*, 845−848.
(21) Vinkenborg, J. L.; Karnowski, N.; Famulok, M. Aptamers for Allosteric Regulation. *Nat. Chem. Biol.* **2011**, *7*, 519−527.
(22) Wieland, M.; Hartig, J. S. Improved Aptazyme Design and in Vivo Screening Enable Riboswitching in Bacteria. *Angew. Chem., Int. Ed.* **2008**, *47*, 2604−2607.
(23) Wieland, M.; Berschneider, B.; Erlacher, M. D.; Hartig, J. S. Aptazyme-Mediated Regulation of 16S Ribosomal RNA. *Chem. Biol.* **2010**, *17*, 236−242.
(24) Nomura, Y.; Kumar, D.; Yokobayashi, Y. Synthetic Mammalian Riboswitches Based on Guanine Aptazyme. *Chem. Commun.* **2012**, *48*, 7215−7217.
(25) Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 297−302.
(26) Saccà, B.; Niemeyer, C. M. DNA Origami: The Art of Folding DNA. *Angew. Chem., Int. Ed.* **2012**, *51*, 58−66.
(27) Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M. Self-Assembly of DNA into Nanoscale Three-Dimensional Shapes. *Nature* **2009**, *459*, 414−418.
(28) Ganesh, K. N.; Krishnan, Y. Nucleic Acids - Chemistry and Applications. *J. Org. Chem.* **2013**, *78*, 12283−12287.
(29) Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L. Building Programmable Jigsaw Puzzles with RNA. *Science* **2004**, *306*, 2068−2072.
(30) Guo, P. The Emerging Field of RNA Nanotechnology. *Nat. Nanotechnol.* **2010**, *5*, 833−842.
(31) Westhof, E.; Masquida, B.; Jaeger, L. RNA Tectonics: Towards RNA Design. *Folding Des.* **1996**, *1*, R78−88.
(32) Chakraborty, S.; Modi, S.; Krishnan, Y. The RNA2-PNA2 Hybrid I-Motif-a Novel RNA-Based Building Block. *Chem. Commun.* **2008**, 70−72.
(33) Jaeger, L.; Chworos, A. The Architectonics of Programmable RNA and DNA Nanostructures. *Curr. Opin. Struct. Biol.* **2006**, *16*, 531−543.
(34) Endo, M.; Yamamoto, S.; Tatsumi, K.; Emura, T.; Hidaka, K.; Sugiyama, H. RNA-Templated DNA Origami Structures. *Chem. Commun.* **2013**, *49*, 2879−2881.
(35) Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. A Primer to Scaffolded DNA Origami. *Nat. Methods* **2011**, *8*, 221−229.
(36) Wu, C.; Bingham, P. M.; Livak, K. J.; Holmgren, R.; Elgin, S. C. The Chromatin Structure of Specific Genes: I. Evidence for Higher Order Domains of Defined DNA Sequence. *Cell* **1979**, *16*, 797−806.
(37) Parada, L. a; McQueen, P. G.; Misteli, T. Tissue-Specific Spatial Organization of Genomes. *Genome Biol.* **2004**, *5*, No. R44.
(38) Mazumder, A.; Shivashankar, G. V Gold-Nanoparticle-Assisted Laser Perturbation of Chromatin Assembly Reveals Unusual Aspects of Nuclear Architecture within Living Cells. *Biophys. J.* **2007**, *93*, 2209−2216.
(39) Mikkelsen, T. S.; Ku, M.; Jaffe, D. B.; Issac, B.; Lieberman, E.; Giannoukos, G.; Alvarez, P.; Brockman, W.; Kim, T.-K.; Koche, R. P.; Lee, W.; Mendenhall, E.; O'Donovan, A.; Presser, A.; Russ, C.; Xie, X.; Meissner, A.; Wernig, M.; Jaenisch, R.; Nusbaum, C.; Lander, E. S.; Bernstein, B. E. Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells. *Nature* **2007**, *448*, 553−560.
(40) Rinn, J. L.; Chang, H. Y. Genome Regulation by Long Noncoding RNAs. *Annu. Rev. Biochem.* **2012**, *81*, 145−166.
(41) Chu, C.; Qu, K.; Zhong, F. L.; Artandi, S. E.; Chang, H. Y. Resource Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions. *Mol. Cell* **2011**, *44*, 667−678.
(42) Paul, J.; Duerksen, J. D. Chromatin-Associated RNA Content of Heterochromatin and Euchromatin. *Mol. Cell. Biochem.* **1975**, *9*, 9−16.
(43) Brown, C. J.; Ballabio, A.; Rupert, J. L.; Lafreniere, R. G.; Grompe, M.; Tonlorenzi, R.; Willard, H. A Gene from the Region of the Human X Inactivation Centre Is Expressed Exclusively from the Inactive X Chromosome. *Nature* **1991**, *349*, 38−44.
(44) Jeon, Y.; Sarma, K.; Lee, J. T. New and Xisting Regulatory Mechanisms of X Chromosome Inactivation. *Curr. Opin. Genet. Dev.* **2012**, *22*, 62−71.
(45) Jeon, Y.; Lee, J. T. YY1 Tethers Xist RNA to the Inactive X Nucleation Center. *Cell* **2011**, *146*, 119−133.
(46) Wutz, A.; Rasmussen, T. P.; Jaenisch, R. Chromosomal Silencing and Localization Are Mediated by Different Domains of Xist RNA. *Nat. Genet.* **2002**, *30*, 167−74.
(47) Maenner, S.; Blaud, M.; Fouillen, L.; Savoye, A.; Marchand, V.; Dubois, A.; Sanglier-Cianférani, S.; Van Dorsselaer, A.; Clerc, P.; Avner, P.; Visvikis, A.; Branlant, C. 2-D Structure of the A Region of Xist RNA and Its Implication for PRC2 Association. *PLoS Biol.* **2010**, *8*, No. e1000276.
(48) Duszczyk, M. M.; Wutz, A.; Rybin, V.; Sattler, M. The Xist RNA A-Repeat Comprises a Novel AUCG Tetraloop Fold and a Platform for Multimerization. *RNA* **2011**, *17*, 1973−1982.
(49) Engreitz, J. M.; Pandya-Jones, A.; McDonel, P.; Shishkin, A.; Sirokman, K.; Surka, C.; Kadri, S.; Xing, J.; Goren, A.; Lander, E. S.; Plath, K.; Guttman, M. The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread across the X Chromosome. *Science* **2013**, *341*, No. 1237973.
(50) Burack, W. R.; Shaw, A. S. Signal Transduction: Hanging on a Scaffold. *Curr. Opin. Cell Biol.* **2000**, *12*, 211−216.
(51) Rinker, S.; Ke, Y.; Liu, Y. A. N.; Chhabra, R.; Yan, H. A. O. Self-Assembled DNA Nanostructures for Distance-Dependent Multivalent Ligand − Protein Binding. *Nat. Nanotechnol.* **2008**, *3*, 418−422.
(52) Wilner, O. I.; Weizmann, Y.; Gill, R.; Lioubashevski, O.; Freeman, R.; Willner, I. Enzyme Cascades Activated on Topologically Programmed DNA Scaffolds. *Nat. Nanotechnol.* **2009**, *4*, 249−254.
(53) Delebecque, C. J.; Lindner, A. B.; Silver, P. A.; Aldaye, F. A. Organization of Intracellular Reactions with Rationally Designed RNA Assemblies. *Science* **2011**, *333*, 470−474.
(54) Anderson, P.; Kedersha, N. RNA Granules. *J. Cell Biol.* **2006**, *172*, 803−808.
(55) Ramaswami, M.; Taylor, J. P.; Parker, R. Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders. *Cell* **2013**, *154*, 727−736.
(56) Weber, S. C.; Brangwynne, C. P. Getting RNA and Protein in Phase. *Cell* **2012**, *149*, 1188−1191.
(57) Kato, M.; Han, T. W.; Xie, S.; Shi, K.; Du, X.; Wu, L. C.; Mirzaei, H.; Goldsmith, E. J.; Longgood, J.; Pei, J.; Grishin, N. V; Frantz, D. E.; Schneider, J. W.; Chen, S.; Li, L.; Sawaya, M. R.; Eisenberg, D.; Tycko, R.; McKnight, S. L. Cell-Free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* **2012**, *149*, 753−767.
(58) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Jülicher, F.; Hyman, A. A. Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* **2009**, *324*, 1729−1732.
(59) Hubstenberger, A.; Noble, S. L.; Cameron, C.; Evans, T. C. Translation Repressors, an RNA Helicase, and Developmental Cues Control RNP Phase Transitions during Early Development. *Dev. Cell* **2013**, *27*, 161−173.
(60) Han, T. W.; Kato, M.; Xie, S.; Wu, L. C.; Mirzaei, H.; Pei, J.; Chen, M.; Xie, Y.; Allen, J.; Xiao, G.; McKnight, S. L. Cell-Free Formation of RNA Granules: Bound RNAs Identify Features and Components of Cellular Assemblies. *Cell* **2012**, *149*, 768−779.

(61) Um, S. H.; Lee, J. B.; Park, N.; Kwon, S. Y.; Umbach, C. C.; Luo, D. Enzyme-Catalysed Assembly of DNA Hydrogel. *Nat. Mater.* **2006**, *5*, 797−801.
(62) Lee, J. B.; Peng, S.; Yang, D.; Roh, Y. H.; Funabashi, H.; Park, N.; Rice, E. J.; Chen, L.; Long, R.; Wu, M.; Luo, D. A Mechanical Metamaterial Made from a DNA Hydrogel. *Nat. Nanotechnol.* **2012**, *7*, 816−820.
(63) Bhatia, D.; Mehtab, S.; Krishnan, R.; Indi, S. S.; Basu, A.; Krishnan, Y. Icosahedral DNA Nanocapsules by Modular Assembly. *Angew. Chem., Int. Ed.* **2009**, *48*, 4134−4137.
(64) He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. Hierarchical Self-Assembly of DNA into Symmetric Supramolecular Polyhedra. *Nature* **2008**, *452*, 198−201.
(65) Hagan, M. F.; Chandler, D. Dynamic Pathways for Viral Capsid Assembly. *Biophys. J.* **2006**, *91*, 42−54.
(66) Douglas, S. M.; Bachelet, I.; Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **2012**, *335*, 831−834.
(67) Banerjee, A.; Bhatia, D.; Saminathan, A.; Chakraborty, S.; Kar, S.; Krishnan, Y. Controlled Release of Encapsulated Cargo from a DNA Icosahedron Using a Chemical Trigger. *Angew. Chem., Int. Ed.* **2013**, *52*, 6854−6857.
(68) Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L. P.; Pedersen, J. S.; Birkedal, V.; Besenbacher, F.; Gothelf, K. V; Kjems, J. Self-Assembly of a Nanoscale DNA Box with a Controllable Lid. *Nature* **2009**, *459*, 73−76.
(69) Goodman, R. P.; Heilemann, M.; Doose, S.; Erben, C. M.; Kapanidis, A. N.; Turberfield, A. J. Reconfigurable, Braced, Three-Dimensional DNA Nanostructures. *Nat. Nanotechnol.* **2008**, *3*, 93−96.
(70) Umbarger, M. A.; Toro, E.; Wright, M. A.; Porreca, G. J.; Baù, D.; Hong, S.-H.; Fero, M. J.; Zhu, L. J.; Marti-Renom, M. A.; McAdams, H. H.; Shapiro, L.; Dekker, J.; Church, G. M. The Three-Dimensional Architecture of a Bacterial Genome and Its Alteration by Genetic Perturbation. *Mol. Cell* **2011**, *44*, 252−264.
(71) Balasubramanian, S. Sequencing Nucleic Acids: From Chemistry to Medicine. *Chem. Commun.* **2011**, *47*, 7281−7286.
(72) Kim, D.-N.; Kilchherr, F.; Dietz, H.; Bathe, M. Quantitative Prediction of 3d Solution Shape and Flexibility of Nucleic Acid Nanostructures. *Nucleic Acids Res.* **2012**, *40*, 2862−2868.