Organic & Biomolecular Chemistry

Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-REV-09-2018-002436.R2
Article Type:	Review Article
Date Submitted by the Author:	16-Nov-2018
Complete List of Authors:	MacCulloch, Tara; Arizona State University College of Liberal Arts and Sciences, School of Molecular Sciences; Arizona State University Biodesign Institute, Center for Molecular Design and Biomimetics Buchberger, Alex; Arizona State University College of Liberal Arts and Sciences, School of Molecular Sciences; Arizona State University Biodesign Institute, Center for Molecular Design and Biomimetics Stephanopoulos, Nicholas; Arizona State University College of Liberal Arts and Sciences, School of Molecular Sciences; Arizona State University Biodesign Institute, Center for Molecular Design and Biomimetics

Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials

Tara MacCulloch[1,2]*, Alex Buchberger[1,2]*, Nicholas Stephanopoulos[1,2]**
1. School of Molecular Sciences, Arizona State University, Tempe AZ
2. Biodesign Center for Molecular Design and Biomimetics, Arizona State University, Tempe AZ
(* these authors contributed equally to this work; ** nstepha1@asu.edu)

Abstract: Peptide-oligonucleotide conjugates (POCs) are covalent constructs that link a molecule like DNA to a synthetic peptide sequences. These materials merge the programmable self-assembly of oligonucleotides with the bioactivity and chemical diversity of polypeptides. Recent years have seen the widespread use of POCs in a range of fields, driven the by relative advantages of each molecular type. In this review, we will present an overview of the synthesis and application of POCs, with an emphasis on emerging areas where these molecules will have a unique impact. We first discuss two main strategies for synthesizing POCs from synthetic monomers such as phosphoramidites and functionalized amino acids. We then describe four key fields of research in POCs: (1) biomaterials for interfacing with, and controlling the behavior of cells; (2) hybrid self-assembling systems that balance peptide and oligonucleotide intermolecular forces; (3) template-enhanced coupling of POCs into larger molecules; and (4) display of peptides on self-assembled oligonucleotide scaffolds. We also highlight several promising areas for future applications in each of these four directions, and anticipate ever increasing uses of POCs in interdisciplinary research.

Motivation for and synthesis of peptide-oligonucleotide conjugates (POCs)

Chemists have been fascinated with chemical structure and diverse functions of biological molecules like nucleic acids and polypeptides since their discovery. Solid-phase methods for synthesizing oligonucleotides from phosphoramidite monomers and polypeptides from protected amino acids have revolutionized structural and functional studies, and enabled the design of novel materials. In particular, these methods allow for incorporation of chemical moieties not found in nature, imbuing these materials with properties not possible through their biological production. Interestingly, although many examples exist in nature of supramolecular association between polypeptides and nucleic acids (e.g. transcription factors, the ribosome, chromatin), naturally occurring *covalent* conjugates between proteins and oligonucleotides are confined to specific biological systems that are hard to adapt synthetically[1]. Thus, research into peptide-oligonucleotide conjugates (POCs) has generally been the purview of synthetic chemists, who have developed a host of methods for building these intriguing molecules. Fig. 1a shows a schematic of a 20-nucleotide single-stranded DNA molecule conjugated to peptides of three different lengths, to demonstrate the relative sizes of the components, and Fig. 1b the chemical structure of DNA attached to a short peptide. The ultimate goal of POCs is to leverage the unique advantages of each biomolecule in a single system. For example, the oligonucleotide allows binding to complementary molecules inside of cells (e.g. mRNA), programmable polymerization on a complementary template, or integration with complex DNA nanostructures. The peptide component in turn enables targeting to cells of interest, membrane penetration, or biological signaling to cells.

In this review, we will discuss the great potential of POCs made through solid-phase organic synthesis. We will first briefly describe synthetic strategies for obtaining these molecules,

followed by four key areas of application where they have unique potential. The field of POCs is quite broad, and this review is not meant to be comprehensive, so we first will mention several research areas that space limitations will not allow us to cover. First, one of the most active and in-depth application of POCs is in antisense therapy: delivering either siRNA or nucleotides that can bind to and block the translation of mRNA. In these materials, the peptides play an important role for cell targeting, uptake, and endosomal escape or, in one case, catalytic cleavage of intracellular RNA.[2] However, the use of POCs in these applications have been extensively covered elsewhere, so we refer the interested reader to several excellent reviews on this topic.[3-7] We will also not discuss molecular evolution schemes, such as mRNA display,[8] which generally use biochemical machinery instead of synthetic reactions to generate POCs with appended DNA barcodes. Furthermore, although a wealth of elegant examples exist of oligonucleotides linked to recombinant proteins,[9] we will restrict this discussion to peptides made via solid-phase methods, and linked covalently to synthetic oligonucleotides. We will also primarily discuss papers from the past 10 years, since extensive reviews exist on the literature prior to that.[10]

There are two main conceptual approaches for synthesizing POCs (Fig. 2): (1) on-resin synthesis of the entire hybrid molecule, and (2) separate synthesis and purification of each component followed by their coupling in solution. The first approach essentially requires the coupling of amino acid monomers followed by nucleic acid monomers (or vice versa) to create a heteroblock copolymer (Fig. 2a). Traditionally, peptides and oligonucleotides are synthesized on solid support from suitably protected monomers—Fmoc- or Boc-protected amino acids, and phosphoramidites, respectively (Fig. 2b)—but the conditions for these methods are often incompatible with one another. Namely, the most common protocol for Fmoc-based peptide synthesis involves cleavage of the final molecule from the resin and side chain deprotection in strongly acidic conditions (95% trifluoroacetic acid) that will cleave or depurinate native oligonucleotides. One way to circumvent this limitation is through the use of peptide nucleic acid (PNA) for the oligonucleotide component,[11] which replaces the phosphate and sugar backbone of natural nucleic acids with an N-(2-aminoethyl)-glycine polymer with pendant nucleobases (Fig. 2c). PNA can be made using standard solid-phase peptide synthesis conditions, allowing for a continuous and seamless transition between the peptide and oligonucleotide component of a POC. In this approach, the peptide and PNA are seamlessly attached without an intervening linker (which can introduce unwanted flexibility for some applications) if desired. The downside is that PNA monomers, though they are commercially available, are significantly more expensive than most commercially available phosphoramidites. Furthermore, the lack of charge the backbone makes PNA quite hydrophobic, so the final PNA-peptide POC may be insoluble in water if the concatenated peptide is not hydrophilic enough to solubilize it. Another way to avoid the harsh acidic cleavage conditions used in peptide synthesis is to employ protecting groups on the peptide that can be removed by mild acid, or basic conditions. A full discussion of protecting group strategies and synthesis is beyond the scope of this work, but we refer the interested reader to several excellent reviews.[10, 12-14]

The second approach—in-solution stoichiometric coupling of peptides and oligonucleotides—is more commonly used because it sidesteps compatibility issues with solid phase synthesis protocols, and does not require exotic protecting group strategies. In this approach (Fig. 2d), each individual molecule is synthesized separately on solid phase bearing a unique reactive functional group. The two pieces are then linked either directly, or through a bifunctional linker, followed by subsequent purification of the desired POC. Common methods of

purification include chromatography (e.g. HPLC, anion exchange), elution from a denaturing polyacrylamide gel, or precipitation with ethanol. Because many linkers are commercially available, as are peptides and oligonucleotides bearing common reactive moieties (e.g. amines, thiols), the barrier for the non-expert is low.

Native oligonucleotides like DNA or RNA, or synthetic analogues like PNA, lack suitable nucleophiles or electrophiles for coupling them to peptides, so these are introduced through functionalized phosphoramidite monomers. Some of the most common reactive moieties include amines, thiols, azides/alkynes (for click chemistry), alkoxyamines or hydrazines, or aldehydes. On the peptide side, natural amino acids containing amines (lysine), carboxylic acids (glutamic acid, aspartic acid), or thiols (cysteine) are introduced. When there are multiple such residues—which would preclude site-specificity, or potentially modify a bioactive epitope—non-canonical amino acids (NCAAs) can be used through appropriately functionalized Fmoc-protected monomers. The most commonly used NCAAs for this purpose are azides (e.g. azidolysine, 4-azidophenylalanine) and alkynes (e.g. propargylglycine) for click chemistry (Fig. 2e). However, less widely known bioconjugation reactions, like the suite of oxidative coupling reactions developed by the Francis lab[15] can also be employed to synthesize POCs. Commonly homo- or hetero-bifunctional linkers are used in conjunction with certain standard reactive moieties on the two coupling partners. These have the advantage of being commercially available with spacers of varying length and flexibility, which can be important for the application at hand. Examples of linkers include (Fig. 2f): (1) linking two amines using N-hydroxysuccinimidyl esters (one on the oligonucleotide, and the other either a lysine or the N-terminus on the peptide); (2) a thiol to an amine (either one the DNA, or on the peptide as lysine/N-terminus or cysteine) using maleimides or disulfide exchange for thiol modification; (3) two thiols (one on each component); or (4) converting an amine or thiol into a new reactive handle (like an azide or alkyne). One downside of these commercial linkers is that they usually rely on common functional groups like amines or thiols, and if the peptide of interest has multiple such groups, site specificity is not possible.

Applications of POCs

The key advantage of POCs lies in their ability to seamlessly integrate oligonucleotide functionality with the biological, chemical, or structural behavior of the grafted peptide. In this section, we discuss four particularly exciting applications of these conjugates: (1) functional biomaterials for interfacing with cells, (2) hybrid self-assembling systems, (3) templated-enhanced coupling of POCs, and (4) scaffolding of peptides on DNA nanostructures. This list is not intended to be exhaustive, but rather to highlight several diverse and emerging fields where these hybrid biomolecules have found use, and propose directions for the future of the field. A common theme in all many of the examples presented is the use of oligonucleotides as a scaffold or linker for presenting the peptide, and imbuing it with added functionality or spatial control.

1) Functional biomaterials. One area where POCs have found increasing use in the past few years is to impart bioactivity to DNA-based biomaterials such as hydrogels, nanofibers, or functional surfaces designed to interact with cell and influence their behavior. The central goal of these materials is to mimic the properties of the extracellular matrix (ECM) in order to

enhance cell survival, guide differentiation of stem cells, spur regeneration of injured tissues, or study fundamental biological processes in a more native-like environment. Peptides have been employed extensively in biomaterials, in both functionalizing other scaffolds to impart them with bioactivity, as well as self-assembling scaffolds in their own right.[16-21] However, only a handful of examples exist of integrating these systems with oligonucleotides due to the higher cost of DNA and concerns about its immunogenicity and degradability. Conversely, oligonucleotides can imbue biomaterials with unique properties that are difficult to accomplish with other materials—such as dynamic control of multiple signals, or controlled nanoscale spacing between ligands—so we foresee increasing incorporation of DNA into ECM mimics in the future.

Although hydrogels comprised of branched DNA structures[22] or DNA-modified polymers[23] were first reported 10-15 years ago, peptides have only recently been incorporated. In 2015, the Liu and Shu laboratories demonstrated the "bioprinting" of a hydrogel comprised of two "inks": a polypeptide bearing DNA handles, and a complementary DNA crosslinker (Fig. 3a).[24] The peptide in this case served as the bulk hydrogel material, whereas the DNA mediated crosslinking and gelation via Watson Crick pairing with the handles attached to the peptide. In contrast to most of the POCs reported herein—which link a single peptide to a single oligonucleotide—this work designed a synthetic polymer of glutamic acid and propargyl-L-glutamate, followed by click coupling of azide-DNA onto the pendant terminal alkynes of the propargyl-L-glutamate. Each polypeptide (which contained on average 240 glutamate monomers and 20 alkynes) displayed 5-6 DNA handles. Mixing this POC with double stranded DNA linkers with complementary sticky ends crosslinked the peptide chains and resulted in self-supporting hydrogels within seconds. These hydrogels possessed physiologically relevant stiffness values (G' ~ 5 kPa), could be printed into various shapes using a co-extrusion process, and successfully encapsulated mammalian AtT20 and HEK-293 for long-term culture (Fig. 3b). Interestingly, the gels could be degraded by both proteases (which cleaved the polypeptide) and nucleases (which degraded the DNA crosslinks), and the incorporation of DNA hints at future applications with other oligonucleotide-specific mechanisms like strand displacement.[25] A concurrent report by the Liu laboratory with this hydrogel system demonstrated another advantage of DNA: facile functionalization through DNA-tethered ligands.[26] The authors designed an X-shaped DNA structure for both gelation of the polypeptide-DNA POC and to introduce pendant ssDNA handles throughout the gel (Fig. 3c). Following hydrogel formation, these handles could immobilize model ligands (in this case red and green fluorophores) with high specificity and orthogonality. Shortly after these reports, the Marx laboratory demonstrated that branched DNA hydrogels could be functionalized with cyclic-RGD peptides via copper-catalyzed click.[27] In this example, branched DNA structures were used as primers for PCR on the surface, generating the branched DNA hydrogel *in situ* (Fig. 3d), an interesting and novel DNA-specific mechanism for hydrogel synthesis.

All the hydrogels in the above work resulted in a hydrogel that lacked a defined morphology at the nanoscale, whereas the native ECM incorporates fibrous proteins like collagen. In 2015, the Stupp laboratory sought to mimic this nanoscale architecture by modifying a DNA nanostructure with bioactive signals through POCs.[28] Nanotubes constructed from the self-assembly of double-crossover "tile" building blocks were functionalized with the integrin-binding peptide RGDS (derived from fibronectin) by attaching it to one of the constituent DNA strands via copper-free click chemistry. Co-assembly of this POC with the other strands of the tile resulted in nanotubes with multivalent display of the peptide on the surface (Fig. 4a). Neural stem cells could adhere to surfaces coated with these peptide-DNA nanotubes, and the peptide

both suppressed astrocyte generation and enhanced differentiation into neurons. Interestingly, the nanotube morphology was critical for bioactivity. When the authors omitted one of the strands of the DNA tile (which prevented assembly of nanostructures), the remaining strands (including the POC) only formed random aggregates. Surfaces coated with this material did still contain the RGDS signal, so cells could adhere and differentiate readily; however, the desired increase in the fraction of cells that became neurons was not observed. The cell adhesive effect of the peptide signal was in effect decoupled from the nanostructure, something that is difficult if not impossible to achieve with self-assembling peptide fibers but relatively facile with DNA, as the nanostructure assembly is not perturbed by the presence of the much smaller peptide. In this way, the authors demonstrated that POCs could play an important role in determining structure-activity relationships in synthetic extracellular matrices, with a DNA nanostructure serving as a shape-programmable “skeleton” for a bioactive peptide “skin.”

The Stupp group demonstrated another powerful application of POCs in biomaterials in 2017, using DNA as a functional linker to attach peptides onto a surface coated with the non-bioactive polymer alginate (Fig. 4b).[29] Fibroblast cells plated on the surfaces adhered and spread in the presence of a DNA-tethered RGDS peptide but remained rounded on surfaces bearing the DNA alone, or displaying a control non-bioactive peptide. The key advantage of the DNA in this case was to allow for dynamic removal of the peptide through a process called toehold-mediated strand displacement,[25] whereby an externally added oligonucleotide out-competes the surface strand due to increased base-pairing. Stripping the RGDS signal from the surface resulted in cells rounding up and detaching; since this process regenerated the surface tether, a fresh batch of POC could be added to restore bioactivity. In this fashion, cells were able to spread and contract over multiple cycles, with complete reversibility (Fig. 4c). Furthermore, the added oligonucleotide DNA was an exceptionally mild “trigger” compared with other mechanisms for dynamic control of biomaterial signals (such as electrochemical potentials or UV light). The authors also used the DNA as a “molecular ruler” to immobilize two peptides—RGDS and a “synergy sequence” PHSRN, which is roughly 5 nm from RGDS on fibronectin—and demonstrated that when these signals were presented within ~5.5 nm of one another, cell bioactivity was enhanced.

The true power of using DNA in this work, however, was that it allowed the dynamic control of multiple signals through the use of sequence-specific displacement strands. The authors used independent DNA linkers to attach two peptide signals: the migration and differentiation signal IKVAV (derived from laminin) and a peptide mimicking the proliferative growth factor FGF-2 (Fig. 4d). Each peptide was controlled by a unique displacement strand, and the specificity of Watson-Crick pairing meant that one could be switched independently of the other, and over multiple cycles. Neural stem cells cultured on these surfaces could be coaxed to migrate and differentiate in response to the IKVAV, or proliferate in response to the FGF-2; interestingly, cells that had migrated could be induced to retract upon removing the IKVAV signal. Each signal could be turned on and off independently of the other, and over multiple cycles, with no cross-talk between the two peptides. Somewhat surprisingly, the DNA linkers remained functional over a week of cell culture, obviating fears of degradation by nucleases, or occlusion by secreted matrix proteins from the cells. Such an orthogonal control of multiple biological signals will be especially critical to recapitulating the complex extracellular milieu for enhancing biomaterial activity, and DNA—with its vast number of potential sequences—is an ideal material for accomplishing this goal.

DNA-linked peptide surfaces have also been used as sensors to probe the forces between cell receptors and the ECM.[30-32] In these approaches, a bioactive peptide (again, usually RGD) is tethered to a surface using a DNA linker, either a duplex or a hairpin structure. This linker contains a fluorophore-quencher pair, such that no fluorescent signal is seen until a cell receptor binds the peptide and exerts a force that breaks the DNA hybridization. Because the relationship between length of hybridized DNA or hairpin structure and the force required to break the interactions is well known, these constructs can serve as precisely tunable probes for the forces applied by cells. In one recent report by the Ke and Salaita laboratories, this approach was further extended by using more complex DNA nanostructures to bind multiple POCs and correlate multivalency with the force exerted.[33] Combining force sensors like those described above with multiple signals (and dynamic control thereof) has the potential to create nanomaterials for probing complex biological behavior with unprecedented precision in a multiplexed fashion.

2) Hybrid self-assembling systems. Self-assembly is one of the most promising methods for generating functional nanomaterials from simple molecular components. By balancing intermolecular forces—such as electrostatics, hydrogen bonds, hydrophobic collapse, dipole interactions, or van der Waals forces—diverse structure such as fibers, sheets, or spheres consisting of hundreds to thousands of molecule can be formed spontaneously. Peptides alone possess a rich catalogue of self-assembly motifs, often derived from natural proteins, such as α-helical coiled-coils, β-sheet fiber assemblies, or collagen-mimetic triple helices.[16] Grafting a self-assembling peptide motif to a charged oligonucleotide like DNA opens up new avenues for co-assembly of the resulting POCs. Furthermore, the DNA handle allows for the intermolecular forces to be tuned reversibly through hybridization to a complementary strand. Unlike DNA duplexes or more complex nanostructures, most of the hybrid systems in this section are not monodisperse, a result of relying on the balancing of only 2-3 forces rather than the highly specific Watson-Crick pairing. However, the ability of these materials to form extended structures that possess oligonucleotide-specific properties makes them attractive for a range of applications not possible with more defined assemblies. We also note that while some examples exist of self-assembled peptide structures modified with DNA or RNA,[34, 35] we will only discuss systems where both components play a role in the morphology of the final self-assembled structure.

In 2012 Vebert-Nardin and coworkers described the first POC linking a self-assembling peptide motif to a DNA strand.[36] They conjugated a 12-nt CT-rich sequence to one of the simplest peptides known to form spontaneous nanostructures: diphenylalanine (Fig. 5a). As opposed to the one-dimensional nanofibers (and hydrogels) formed by diphenylalanine alone, the large charged head group of the DNA instead gave spherical structures 200-300 nm in diameter. The authors demonstrated that these spheres consisted of a single POC bilayer surrounding a hollow core (Fig. 5b), and that the head groups could bind to their DNA complement. A dye could be incorporated into the center of these spheres as a model cargo and dropping the pH from 6.5 to 4.5 resulted in release of the dye due to protonation of the cytosine-rich oligonucleotides and disruption of the membrane. This mechanism could, in principle, be used to release encapsulated cargo follow uptake by cells and acidification of the environment in the endosome. Interestingly, the authors proposed that the spherical morphology resulted partly due to hydrogen-bonding (but *not* base-pairing) between the DNA bases, and adding urea to disrupt this effect while leaving the diphenylalanine interactions intact resulted in

an elongated morphology halfway between the spheres and nanofibers. This work thus demonstrated that DNA had the potential for rich self-assembly behavior not based on Watson-Crick pairing, which could be balanced in new and interesting ways with peptide self-assembly. In a follow-up study two years later, the Verbert-Nardin group explored the self-assembly of ditryptophan-DNA POCs, which gave spheres at low concentrations, but fibers at higher concentrations.[37] This change was attributed to the stronger π-π stacking forces between tryptophan side chains, which could be monitored by fluorescence spectroscopy. The simplicity of these peptide motifs belied the more complex assemblies possible, and hints at a rich diversity of structures possible with longer peptides.

One of the driving motivations for POC self-assembly is the generation of synthetic nanoparticles containing both protein and oligonucleotides, for applications in gene delivery; to create, in effect, artificial viruses. Many noncovalent approaches exist for incorporating proteins and peptides with DNA or RNA, but these tend to give polydisperse structures without defined internal structure. To address this limitation, Kye and Lim investigated covalently attaching DNA to a β-sheet forming peptide in order to generate an artificial “deoxyribonucleoprotein” (DNP) through synergistic assembly of each component (Fig. 5c).[38] In order to prevent peptide assembly in solution—which could dramatically reduce the DNA conjugation yield—the authors carried out the copper-click coupling of the peptide to protected DNA on-resin in organic solvent, a method that could be particularly useful in the future as longer and more hydrophobic peptides are attached to DNA handles. They also incorporated an RGD epitope in the peptide to enable cell internalization, and synthesized two POCs: one with the peptide conjugated to a 20-nt antisense DNA for GFP, and one with the peptide conjugated to the complement of this sequence. By tuning the conditions the authors were able to either first form the β-sheet prior to DNA hybridization (by assembling the peptide above the melting temperature of DNA and then cooling the solution), or vice versa (by hybridizing the DNA below the critical aggregation concentration of the peptide and then concentrating the solution). Interestingly, both pathways resulted in identical (and fairly homogeneous) toroidal nanostructures, compared with a more heterogeneous population from simply mixing without controlling the self-assembly pathways. These structures could be switching to a fibrous morphology through DNA strand exchange, and displayed enhanced GFP knockdown efficiency compared with the commercial standard transfection agent Lipofectamine. Once again, the careful control of multiple orthogonal self-assembly forces was able to yield interesting hybrid nanostructures with properties not possible with materials made of peptides or DNA alone. An open question going forward, however, is just how general and predictable this method is, and whether more complex nanostructures be designed by subtle tweaking of the peptide and DNA components.

Similar to the above work, very recently Ashkenasy and coworkers reported another self-assembling system comprised of a β-sheet peptide linked to either single- or double-stranded DNA (Fig. 5d).[39] In contrast to the previous works, the authors probed the change in morphology of nanostructures formed from pure peptides (not linked to DNA) upon co-assembly with increasing amounts of POC. Pure peptide-based nanofibers formed bundled fiber-like assemblies, as expected, and this morphology was not changed by the addition of low amounts of POC. As the POC concentration increased, however, the fiber morphology changed to be wider and more flexible, and when two POCs with complementary DNA strands were added, a new spherical structure with diameters 20-80 nm were observed. Closer examination of these spheres by electron microscopy revealed nested multi-lamellar structures, akin to the layers of an onion, and the authors proposed an assembly of alternating layers of DNA duplexes and

peptide fibers (Fig. 5e). The structures could be denatured using thermal melting of the DNA, or increasing the pH of the solution, and the structures could be loaded with hydrophobic guest molecules like doxorubicin for drug delivery applications. Much like the previous two examples, the emergent structure could not be predicted from general heuristic principles (like the Israelachvili packing parameter for simple amphiphiles),[40] so either a new theory will be required—likely aided by computational simulation methods—or systematic screening of peptides and DNA.

The relative paucity of works in this area highlights the great potential for self-assembling POCs that integrate multiple types of intermolecular forces. Going forward, additional peptide self-assembly motifs beyond the β-sheet (e.g. α-helical coiled-coils, collagen triple helices, peptide amphiphiles) can be integrated with DNA. Going beyond double-stranded DNA helices to complex nanostructures like those described in Section 4 below would provide additional functionality and access to larger length scales. Finally, the reversibility and stimulus-responsiveness of DNA—through displacement strands, as outlined in the Kye and Lim work—will further enhance the already rich world of peptide-based self-assembly.

3) Template-enhanced coupling of POCs. The ribosome is a marvel of evolution, able to synthesize long proteins from constituent amino acids through a combination of catalysis and templated enhancement of local concentration. The exquisite control over monomer sequence is achieved through the use of an oligonucleotide template (the mRNA) in conjunction with complementary adaptors bearing the amino acids (the charged tRNAs, which incidentally are one of the few examples of “natural” conjugates of amino acids with oligonucleotides). Researchers have recognized the value of a programmable template for increasing the effective concentration of the reactive partners, and facilitating their reaction into larger constructs that would be far more challenging to synthesize by solution-phase coupling alone. Indeed, pioneering work by scientists like Orgel, Letsinger, von Kiedrowski, Eschenmoser, and Kool[41-58] have demonstrated the potential for oliognucleotides to template the condensation of smaller fragments through promixity-based concentration enhancement. Much of this work has focused on elucidating plausible prebiotic chemistry that led to the origin of life, and the eventual emergence of complicated molecular machines like the ribosome. In this section, we will describe several examples that use a DNA or PNA template to bring together peptides (through the corresponding POCs) and couple them in a sequence-specific manner. Ultimate applications of these materials include full-length protein synthesis, or protein-like synthesis of synthetic polymers and hybrid materials.

One of the first examples of DNA-templated peptide ligation chemistry was reported by the Joyce group in 1996, for the synthesis of a POC via an amide bond.[59] A peptide thioester was first reacted with a thiol-terminated DNA strand to link the components via a new thioester. Next, a DNA template was used to bring this POC into close proximity with an amine-terminated DNA strand, resulting in amide bond formation enhanced by the higher effective concentration of the reactants (Fig. 6a). Notably, this reaction did not occur in the absence of the template, and the overall reaction was highly reminiscent of native chemical ligation (NCL) between two peptide fragments, a reaction that links one with a C-terminal thioester and the other with an N-terminal cysteine.[60] NCL is a powerful method for linking together unprotected peptide fragments into longer conjugates, and has been used to build full-length, functional proteins with completely control over their sequence. However, as molecules grow in length, the coupling yield drops off

due to inter- and intra-molecular aggregation. In 2011, the Seitz lab used a DNA template to bring two PNA-peptide conjugates into close proximity and enhance the NCL between them (Fig. 6b),[61] as a prelude to coupling longer fragments in a ribosome-mimetic fashion. The template enhanced the rate of reaction approximately 200-fold, and because its binding to the POCs was reversible it was able to act catalytically over multiple turnovers. In a follow-up study, Seitz and coworkers used PNA-containing POCs to demonstrate the NCL of a 14-residue cytotoxic peptide[62] and an apoptosis-inducing peptide[63] on RNA templates. The Diederichsen group reported an extension of this theme in 2017, linking short peptides to complementary PNA handles through *o*-nitrobenzyl ester photocleavable linkers.[64] Hybridization of the handles and NCL followed by irradiation with UV light liberated the full-length peptide free of the PNA handles, paving the way for a truly traceless ligation in the future. We highlight that the examples above were restricted to either model peptides or short bioactive sequences, whereas non-templated NCL can synthesize functional proteins ~200 residues in length.[65] Thus, a critical challenge in the future will be using peptide templation to access dramatically longer peptides/proteins that are not possible by in-solution coupling alone.

Although NCL has the distinct advantage of producing peptides with native amide bonds, the use of DNA templates has to date been limited to relatively short fragments and a single coupling step. In 2013, the Liu laboratory reported a conceptually different method for linking multiple molecular fragments together templated by DNA.[66] Taking direct inspiration from the ribosome, this approach used a set of PNA-based “adaptor” molecules to bring together short polymeric building blocks on a DNA template and co-localize their reactive termini. These termini were linked using copper-catalyzed click chemistry, followed by reductive cleavage of the disulfide bonds connecting them to the PNA adaptors and liberating a sequence-specific synthetic polymer (Fig. 6c,d). By “pre-synthesizing” the polymer-adaptor building blocks, longer molecules could be joined together after assembly on the DNA template, and the authors demonstrated efficient synthesis of polymers with up to six units. Furthermore, the order of these building blocks could be precisely controlled through the order of “codons” on the DNA template. Amide bond formation was unfortunately not efficient with this system, so native polypeptides could not be synthesized. However, this method was flexible enough to allow for diverse polymer backbones, including β-peptides and α-(D)-peptides up to 15 residues in length. Although the cyclic peptide-PNA building blocks used in this work were highly unusual POCs compared to the others in this review, they demonstrate the power of innovative molecular design and DNA-templated reactions for synthesizing novel materials in a highly biomimetic fashion. Extension of this method to amide bond formation (e.g. via NCL) has the potential to create a truly synthetic ribosome.

Yet another method for templated-directed ligation of multiple POCs was reported by the Hili lab in 2015. Rather than use a DNA template to link POCs at the peptide termini, the authors synthesized short (five-nucleotide) phosphorylated DNA fragments with oligopeptides linked to the third nucleobase.[67] These penta-nucleotide POCs could be assembled on a complementary DNA scaffold and ligated using T4 DNA ligase (Fig. 6e). In this fashion, a DNA polymer bearing up to eight pendant peptides could be made in a sequence-programmable manner, in a single step. In this case, the template was critical, since T4 ligase cannot link single-stranded DNA termini. Moreover, by using a natural enzyme instead of synthetic chemical reactions extremely high efficiency could be achieved, with the overall reaction yielding almost exclusively full-length molecules. Although the authors optimized the method with simple dipeptides protruding from the pentanucleotides, up to eight residues were tolerated. The method was applied to

generating DNA strands containing multiple pendant hexahistidine peptides, which could then be subjected to a metal-based affinity selection with the final enriched molecules decoded through the attached DNA barcode. Although conceptually similar to methods like mRNA or phage display, the use of fully synthetic peptides encoded by natural codons allowed for highly non-natural peptides to be used.[68] The authors posit that this system could be used to generate protein analogues, whereby the DNA provides a programmable (and readable) scaffold for the three-dimensional presentation of peptides much the way that a protein core positions surface residues in space.

4) Oligonucleotide nano-scaffolds for displaying peptides. The nanoscale assembly and presentation of biological signals is a critical property of natural systems, with importance in areas like signal transduction, mechanical force generation, and enzymatic cascades among others. Cells achieve the controlled assembly of these signals largely through protein-based self-assembly, which has spurred extensive efforts to mimic these processes in designed proteins.[69] These approaches, while highly promising, rely on an intricate understanding of protein self-assembly that can be challenging for a non-expert. An alternate approach is to use oligonucleotides to create programmable templates for displaying peptides, effectively decoupling the scaffold design from the biological signal. The greater simplicity of oligonucleotide design (namely the Watson-Crick pairing rules), combined with extensive advances in the field of DNA nanotechnology in the past 30 years,[70, 71] have resulted in a wealth of programmable oligonucleotide structures. In this section we will describe the merging of POCs with programmable scaffolds to control the nanoscale display of peptides, and highlight some exciting applications of these materials.

The simplest “scaffold” for displaying POCs is the DNA double helix, which has predictable nanoscale dimensions and helical pitch. In 2009, the Johnston and Chaput groups reported the use of double-stranded DNA to bind two peptides and create a synthetic antibody, or “synbody” (Fig. 7a).[72] A peptide array was first used to identify several short (12-mer) peptides with micromolar affinity to a target protein. These peptides were synthesized as POCs, and a DNA template used to create heterobivalent combinations of peptides that could be screened for enhanced binding to the target. The key hypothesis of the work was that different peptides would bind to different facets of the protein, so that some of the synbody constructs would “clamp” the target in a cooperative fashion, similar to antibody variable loops. Using the DNA-templated approach, a DNA-peptide synbody for the protein Gal80 was discovered with K_d of 5 nM, a roughly 1,000-fold increase in affinity compared with the individual peptides. The DNA scaffold served not only as a mere linker, however, but also allowed for rational tuning of the distance and angle (through the number of intervening base pairs) between the peptides to find an optimal combination for binding. Such scaffolds could, in the future, be used in conjunction with DNA barcoding approaches to create vast libraries of synbodies for panning against a target, followed by amplification and decoding. The power of this approach to answer fundamental biological questions was demonstrated by the Seitz lab, which attached two different POCs to a DNA template to control the spacing between them. This approach allowing them to probe the distance-dependent synergy between them in binding to the endocytic AP-2 complex.[73] In a follow-up report, spatial control of two peptides allowed discrimination between binding to spleen tyrosine kinase (Syk) and the ζ-chain-associated protein kinase (ZAP-70)[74], elucidating how spatial variations in the presentation of the same signal could have dramatically different biological effects. Two other works used a DNA scaffold as a molecular ruler—with

control over distance and flexibility—to present two identical ligands in order to probe and enhance their binding to protein[75] or antibody[76] targets in a bivalent fashion.

In 2012 Appella and coworkers extended this concept beyond homo- or hetero-dimers, and used a repeating DNA scaffold to present a much larger number of a single peptide (Fig. 7b).[77] The peptides were attached to short PNA oligomers via the amine on an L-lysine-γ-substituted monomer, allowing for rapid generation of multivalent constructs with tunable POC number (up to 45 units) and density. By attaching a cyclic-RGD peptide to the PNA building blocks, the authors were able to generate highly potent binding agents for $\alpha_v\beta_3$ integrins and block their activity in a melanoma model both *in vitro* and *in vivo*. The multivalency afforded by the DNA-scaffolded approach enhanced the inhibitory activity of these molecules by over two orders of magnitude compared with the monomeric peptide. Interestingly, increasing the number of ligands past 15 did not yield any increases in binding, which the authors attributed to saturation of integrin binding sites. It is important to highlight that the use of DNA hybridization allowed for rapid and quantitative generation of multiple systems from a limited set of building blocks, via thermal annealing of the monomers with a suitable template. By contrast, individually synthesizing each component via organic chemistry (e.g. on a polymeric or dendrimeric scaffold) would have been far more challenging and expensive, and would require separation of the polymer with the desired valence.

The above examples used DNA as a scaffold to display multiple copies of a peptide, but in 2016 Wengel, Jensen and coworkers described a new paradigm for DNA-templated POCs: using the oligonucleotide to nucleate a peptide self-assembly motif.[78] A 30-residue peptide that could assemble into a coiled-coil of two or three monomers in solution was linked to three distinct DNA strands using copper free click chemistry. The DNA strands could form a triplex structure, thereby bringing the three copies of the peptide into close proximity and “folding” them into the desired homotrimeric coiled-coil (Fig. 7c). The peptide assembly was stabilized against thermal denaturation by the DNA triplex relative to the free peptides, and showed enhanced alpha helicity by circular dichroism. In effect, the DNA triplex played a similar role to a protein core, concentrating and pre-arranging the tethered peptides to help enhance their assembly. A similar approach was reported by Ghosh and Hamilton in 2012, whereby a POC was synthesized consisting of a peptide flanked on both ends by a G-rich DNA strands.[79] This DNA-peptide-DNA triblock could fold into a G-quadruplex in the presence of potassium ions, thereby “pinning” the two ends of the peptide in a constrained loop morphology (Fig. 7d). Once again, the DNA assembly effectively folded the peptide into a restricted conformation, much the way that a protein scaffold like an antibody constrains peptide loops at the ends of the hypervariable region in order to reduce entropic freedom and enhance binding. The use of POCs wherein the peptide is flanked by *two* oligonucleotide handles (usually PNA) allowed the Seitz group to control the bioactivity of a peptide by changing its conformation using a complementary DNA strand. By switching between “stretched” and flexible forms of the peptide loop, the bioactivity (protein binding in this case) could be reversibly cycled.[80, 81] A similar system could be converted into a peptide-based “molecular beacon” if the PNA handles were complementary to one another and unfolded upon protein binding[82, 83], or cleaved upon protease activity[84], creating a fluorescent signal. Winssinger and colleagues used a similar hairpin-like approach to create a peptide loop with 25-fold enhanced phosphatase activity over the linear peptide[85], mimicking protein folding into an active site using a POC.

In addition to duplex or triplex structures, DNA can be used to construct highly complex two- and three-dimensional nanostructures through the co-assembly of multiple programmable

strands.[71] These structures can be functionalized with peptides to impart novel bioactivity, either through direct co-assembly of POCs with unfunctionalized strands, or through the incorporation of ssDNA handles that immobilize the POCs in a secondary hybridization step. In 2007, the Yan and Chaput labs reported a two-dimensional nano-array of peptides constructed using four different double-crossover DNA tiles (Fig. 8a).[86] One of the four tiles was modified with a capture strand to immobilize a POC bearing the myc peptide epitope. Co-assembly of the tiles resulted in two-dimensional sheets with peptides spaced ~64 nm apart; exposure of the arrays to an anti-myc antibody then yielded antibody arrays with the same spacing. These protein arrays could in principle be used as ultra-sensitive detection platforms for proteins, potentially down to the single-molecule level by using AFM as a readout. DNA nanostructures have also been investigated as three-dimensional carriers for targeted drug delivery to diseased cells in order to minimize off-target side effects. The Anderson group reported one such approach using a self-assembled DNA tetrahedron functionalized with up to six ligands as a nano-carrier for siRNA (Fig. 8b).[87] The unique DNA strands comprising the cage allowed for a systematic probing of ligand number and three-dimensional presentation to find the optimal arrangement for cell targeting. The authors screened a number of cell targeting peptides, attached to handles on the cage via complementary POCs, although in the end the folate molecule proved to be the most effective ligand for targeting cancer cells. However, decoupling the nanostructure assembly from the ligand functionalization allowed for rapid screening of 30 different ligands by simply attaching them to a DNA handle, greatly simplifying the screening of multiple valences and geometrical presentations. A few years later Sun, He, and coworkers demonstrated a similar approach, using a DNA tetrahedron modified with a tumor-penetrating peptide to deliver doxorubicin to glioblastoma cells (Fig. 8c).[88] Interestingly, the peptide-DNA conjugation chemistry (via copper click) was carried out on the fully intact nanostructure, not on a single constituent strand prior to assembly. Aside from drug delivery applications, PNA-peptide conjugates have also been used to tether peptides to the inside of a DNA nanocage, as a first step to reconstituting protein-like function inside a well-defined environment.[89] Future work immobilizing multiple, longer peptides in a confined nano-cage have the potential to even recapitulate enzyme activity by creating an “active site” inside a structurally defined cavity. Critical to such efforts, however—and a direction for future research—is the ability to control the orientation and co-assembly of two or more peptides within a DNA structure, which in turn requires judicious design of linkers and sites of modification on both the peptide and the oligonucleotide.

One of the most ambitious demonstrations of DNA-templated peptide assembly was demonstrated by the Spruijt and Bayley groups in 2018, and employed a circular DNA nano-scaffold to create a lipid membrane-spanning nanopore from 38-residue amphiphilic peptides.[90] The DNA scaffold contained twelve unique handles, allowing for precise control over the number of peptides attached through the corresponding POCs (Fig. 8d,e). Molecular modeling confirmed that the structure was flexible enough to allow all twelve peptides to face the same direction and co-assemble into the desired nanopore in the presence of a membrane. The peptide was derived from a naturally octameric *E. coli* transporter protein, and only DNA scaffolds containing eight peptides allowed for conductance through the membrane as measured by electrical current recording. Seven or fewer peptides did not result in stable nanopores, and more than eight failed to increase the conductance, suggesting that superfluous peptides were excluded from the assembly. The DNA scaffold was not only necessary to initiate pore formation, but also to maintain it, and cleaving the DNA from the peptides resulted in rapid disassembly of the octameric structure. Furthermore, the DNA could be used to tether a guest

polymer (such as polyethylene glycol or a C_{30} DNA strand) and thread it into the pore; strand displacement to break the tethering could then release that guest on-demand. This work thus demonstrated that the programmable properties of DNA assemblies could be used to great effect to template protein-like activity from synthetic peptides, and enable advances in sensing or the construction of nano-devices that incorporate the functionality of multiple molecular categories.

Conclusions and future applications of POCs

We hope that the discussion of the four topics mentioned here have provided a flavor for the wide range of fields where POCs have played a unique role. We specifically focused on areas where the functionality of the peptide was enhanced by using DNA to either tether it to other materials (like biomaterials) or display it on a nanostructure. In this way, POCs occupy an interesting middle ground between pure peptides and proteins, with oligonucleotides providing a programmability and accessibility that is still elusive in protein design for a non-expert. We foresee a bright future for POCs, both in the four areas mentioned above and in conjunction with areas we did not discuss, such as oligonucleotide tags for identification following selection. Below we discuss future applications of POCs in each of the four areas covered herein:

1) Simultaneously controlling multiple signals in and the morphology of the extracellular matrix. Processes like embryogenesis or regeneration following injury involve coordinated changes of many proteins (or other signaling molecules) in the ECM. Just as importantly, the mechanical and morphological properties of the matrix change as well. As demonstrated by Stupp and coworkers,[29] DNA can control two peptide epitopes readily; extending this system to three or more ligands should be straightforward in order to probe complicated biological processes. DNA can also be used as a reversible crosslinker to mediate matrix stiffness, or to direct the reversible polymerization and/or bundling of peptide nanofibers to tune morphology. Taken together, we expect that the next decade of research in biomaterials will incorporate POCs in increasingly elaborate ways, with the ultimate goal of creating materials that can both uncover fundamental biology, and serve as practical scaffolds for regenerative medicine. Animals like the axolotl can regenerate entire limbs from a severed stump, perhaps someday POC-containing hydrogels will do the same for humans.

2) Hierarchical self-assembly across multiple length scales. One advantage of more traditional self-assembling systems—like lipid bilayers, or one-dimensional peptide nanofibers—is that they can give extended assemblies across length scales ranging from the molecular to the macroscopic. By merging oligonucleotides with self-assembling peptides, these materials have the potential to dynamically, and reversibly, control these extended structures. DNA strand displacement, thermal melting, or enzymatic susceptibility can all be used to modulate the structure of materials in a complex manner. For example, self-assembling peptide fibers could be coaxed to assemble into novel structures by crosslinking them (reversibly) with DNA, with applications in both biomedical hydrogels and stimulus-responsive materials for artificial muscles and soft robotics. Hierarchically structured systems could also be used in interesting photonic applications, photovoltaics, and metamaterials through integration with nanoparticles, organic dyes, catalysts, and semiconducting polymers.

3) Synthesis of full-length, functional proteins or polymers from synthetic peptide components. To date, native chemical ligation has primarily focused on the synthesis of small

functional proteins. Using oligonucleotide templates to enhance the coupling of larger fragments has the potential to dramatically extend the size limit of this method, especially if multiple steps can be successfully templated in a sequential fashion. Full-length, functional proteins composed of artificial peptides would be transformational for studies of structure and function of these molecules, or their application in other materials. Synthetic peptides can contain arbitrarily many non-canonical amino acids, so novel catalytic functionality, tracking moieties (e.g. dyes, nanoparticles), stabilizing polymers, or biophysical tags (e.g. for NMR, EPR) can be incorporated. A DNA scaffold can also be used to create sequence-specific peptide polymers not linked by amide bonds, which could yield materials with resistance to enzymatic degradation or completely new folding motifs and functions. Highly charged DNA handles (which are often comparable in size to small-to-medium size proteins) will be particularly useful to solubilize the hydrophobic peptides necessary for the synthesis of membrane proteins, which are particularly prone to the aggregation that hinders NCL.

4) Artificial antibodies and active sites organized by DNA scaffolds. The majority of protein sequence in molecules like antibodies or enzymes is dedicated to correctly folding a handful of peptide loops or key residues that bind a target or catalyze a reaction. Although great strides have been made in *de novo* protein design,[69] POCs templated on DNA nano-scaffolds provide an attractive alternative thanks to the complexity and programmability of structures produced by DNA nanotechnology researchers. Researchers routinely create three-dimensional objects with complex surfaces or cavities,[91, 92] and position DNA handles with nanometer precision in order to immobilize POCs. Especially important will be the attachment of *multiple* peptides on these scaffolds, to bind different faces of a target protein, or pre-organize reactive amino acids to create a catalytic active site. The use of synthetic peptides allows for the incorporation of an arbitrary number of non-canonical amino acids, enabling functions far beyond those of natural proteins. While the rational design of these materials poses computational difficulties similar to those in protein design, one alternative is to create vast combinatorial libraries of nanostructures with three or more peptides bound, and select for those with function. For such an applications, tagging the peptide with a DNA barcode (as part of the POC) will be particularly useful for determining the active sequences after the selection.

Acknowledgements: N.S. acknowledges startup funds from Arizona State University. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-17-1-0053.

1. B. A. Juodka, *Nucleosides Nucleotides & Nucleic Acids*, 1984, **3**, 445-483.
2. O. A. Patutina, E. V. Bichenkova, S. K. Miroshnichenko, N. L. Mironova, L. T. Trivoluzzi, K. K. Burusco, R. A. Bryce, V. V. Vlassov and M. A. Zenkova, *Biomaterials*, 2017, **122**, 163-178.
3. M. Gooding, M. Malhotra, J. C. Evans, R. Darcy and C. M. O'Driscoll, *European Journal of Pharmaceutics and Biopharmaceutics*, 2016, **107**, 321-340.
4. B. Lebleu, H. M. Moulton, R. Abes, G. D. Ivanova, S. Abes, D. A. Stein, P. L. Iversen, A. A. Arzumanov and M. J. Gait, *Advanced Drug Delivery Reviews*, 2008, **60**, 517-529.
5. R. Juliano, M. R. Alam, V. Dixit and H. Kang, *Nucleic Acids Research*, 2008, **36**, 4158-4171.
6. M. Taskova, A. Mantsiou and K. Astakhova, *Chembiochem*, 2017, **18**, 1671-1682.

7. E. Valeur, S. M. Gueret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann and A. T. Plowright, *Angewandte Chemie-International Edition*, 2017, **56**, 10294-10323.
8. H. Wang and R. H. Liu, *Expert Review of Proteomics*, 2011, **8**, 335-346.
9. B. Sacca and C. M. Niemeyer, *Chemical Society Reviews*, 2011, **40**, 5910-5921.
10. N. Venkatesan and B. H. Kim, *Chemical Reviews*, 2006, **106**, 3712-3761.
11. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1500.
12. K. Lu, Q. P. Duan, L. Ma and D. X. Zhao, *Bioconjugate Chemistry*, 2010, **21**, 187-202.
13. H. Lonnberg, *Bioconjugate Chemistry*, 2009, **20**, 1065-1094.
14. A. Erben and O. Seitz, *Israel Journal of Chemistry*, 2011, **51**, 876-884.
15. A. M. ElSohly and M. B. Francis, *Accounts of Chemical Research*, 2015, **48**, 1971-1978.
16. N. Stephanopoulos, J. H. Ortony and S. I. Stupp, *Acta Materialia*, 2013, **61**, 912-930.
17. R. V. Ulijn and A. M. Smith, *Chemical Society Reviews*, 2008, **37**, 664-675.
18. D. N. Woolfson and Z. N. Mahmoud, *Chemical Society Reviews*, 2010, **39**, 3464-3479.
19. M. P. Lutolf and J. A. Hubbell, *Nature Biotechnology*, 2005, **23**, 47-55.
20. R. Pugliese and F. Gelain, *Trends in Biotechnology*, 2017, **35**, 145-158.
21. I. W. Hamley, *Chemical Reviews*, 2017, **117**, 14015-14041.
22. S. H. Um, J. B. Lee, N. Park, S. Y. Kwon, C. C. Umbach and D. Luo, *Nature Materials*, 2006, **5**, 797-801.
23. D. C. Lin, B. Yurke and N. A. Langrana, *Journal of Biomechanical Engineering-Transactions of the Asme*, 2004, **126**, 104-110.
24. C. Li, A. Faulkner-Jones, A. R. Dun, J. Jin, P. Chen, Y. Z. Xing, Z. Q. Yang, Z. B. Li, W. M. Shu, D. S. Liu and R. R. Duncan, *Angewandte Chemie-International Edition*, 2015, **54**, 3957-3961.
25. D. Y. Zhang and G. Seelig, *Nature Chemistry*, 2011, **3**, 103-113.
26. C. Li, P. Chen, Y. Shao, X. Zhou, Y. Z. Wu, Z. Q. Yang, Z. B. Li, T. Weil and D. S. Liu, *Small*, 2015, **11**, 1138-1143.
27. A. Finke, H. Busskamp, M. Manea and A. Marx, *Angewandte Chemie-International Edition*, 2016, **55**, 10136-10140.
28. N. Stephanopoulos, R. Freeman, H. A. North, S. Sur, S. J. Jeong, F. Tantakitti, J. A. Kessler and S. I. Stupp, *Nano Letters*, 2015, **15**, 603-609.
29. R. Freeman, N. Stephanopoulos, Z. Alvarez, J. A. Lewis, S. Sur, C. M. Serrano, J. Boekhoven, S. S. Lee and S. I. Stupp, *Nature Communications*, 2017, **8**.
30. X. F. Wang and T. Ha, *Science*, 2013, **340**, 991-994.
31. Y. Zhang, C. H. Ge, C. Zhu and K. Salaita, *Nature Communications*, 2014, **5**.
32. B. L. Blakely, C. E. Dumelin, B. Trappmann, L. M. McGregor, C. K. Choi, P. C. Anthony, V. Duesterberg, B. M. Baker, S. M. Block, D. R. Liu and C. S. Chen, *Nature Methods*, 2014, **11**, 1229-+.
33. P. K. Dutta, Y. Zhang, A. T. Blanchard, C. H. Ge, M. Rushdi, K. Weiss, C. Zhu, Y. G. Ke and K. Salaita, *Nano Letters*, 2018, **18**, 4803-4811.
34. M. O. Guler, J. K. Pokorski, D. H. Appella and S. I. Stupp, *Bioconjugate Chemistry*, 2005, **16**, 501-503.
35. P. J. S. King, A. Saiani, E. V. Bichenkova and A. F. Miller, *Chemical Communications*, 2016, **52**, 6697-6700.
36. N. Gour, D. Kedracki, I. Safir, K. X. Ngo and C. Vebert-Nardin, *Chemical Communications*, 2012, **48**, 5440-5442.
37. N. Gour, J. N. Abraham, M. Chami, A. Castillo, S. Verma and C. Vebert-Nardin, *Chemical Communications*, 2014, **50**, 6863-6865.
38. M. Kye and Y. B. Lim, *Angewandte Chemie-International Edition*, 2016, **55**, 12003-12007.

39. A. Chotera, H. Sadihov, R. Cohen-Luria, P. A. Monnard and G. Ashkenasy, *Chemistry-a European Journal*, 2018, **24**, 10128-10135.
40. J. N. Israelachvili, D. J. Mitchell and B. W. Ninham, *Journal of the Chemical Society-Faraday Transactions Ii*, 1976, **72**, 1525-1568.
41. L. E. Orgel, *Accounts of Chemical Research*, 1995, **28**, 109-118.
42. A. Luther, R. Brandsch and G. von Kiedrowski, *Nature*, 1998, **396**, 245-248.
43. C. Bohler, P. E. Nielsen and L. E. Orgel, *Nature*, 1995, **376**, 578-581.
44. M. K. Herrlein, J. S. Nelson and R. L. Letsinger, *Journal of the American Chemical Society*, 1995, **117**, 10151-10152.
45. Y. Z. Xu and E. T. Kool, *Journal of the American Chemical Society*, 2000, **122**, 9040-9041.
46. Y. Z. Xu, N. B. Karalkar and E. T. Kool, *Nature Biotechnology*, 2001, **19**, 148-152.
47. Y. Z. Xu and E. T. Kool, *Nucleic Acids Research*, 1999, **27**, 875-881.
48. S. Sando and E. T. Kool, *Journal of the American Chemical Society*, 2002, **124**, 9686-9687.
49. Y. Z. Xu and E. T. Kool, *Tetrahedron Letters*, 1997, **38**, 5595-5598.
50. Y. Z. Xu and E. T. Kool, *Nucleic Acids Research*, 1998, **26**, 3159-3164.
51. T. F. Wu and L. E. Orgel, *Journal of the American Chemical Society*, 1992, **114**, 5496-5501.
52. T. Wu and L. E. Orgel, *Journal of the American Chemical Society*, 1992, **114**, 7963-7969.
53. G. Vonkiedrowski, *Angewandte Chemie-International Edition in English*, 1986, **25**, 932-935.
54. K. U. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347-1351.
55. X. L. Wu, S. Guntha, M. Ferencic, R. Krishnamurthy and A. Eschenmoser, *Organic Letters*, 2002, **4**, 1279-1282.
56. X. L. Wu, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Organic Letters*, 2002, **4**, 1283-1286.
57. I. A. Kozlov, B. De Bouvere, A. Van Aerschot, P. Herdewijn and L. E. Orgel, *Journal of the American Chemical Society*, 1999, **121**, 5856-5859.
58. K. U. Schoning, P. Scholz, X. L. Wu, S. Guntha, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Helvetica Chimica Acta*, 2002, **85**, 4111-4153.
59. R. K. Bruick, P. E. Dawson, S. B. Kent, N. Usman and G. F. Joyce, *Chemistry & Biology*, 1996, **3**, 49-56.
60. P. E. Dawson, T. W. Muir, I. Clarklewis and S. B. H. Kent, *Science*, 1994, **266**, 776-779.
61. A. Erben, T. N. Grossmann and O. Seitz, *Angewandte Chemie-International Edition*, 2011, **50**, 2828-2832.
62. O. Vazquez and O. Seitz, *Chemical Science*, 2014, **5**, 2850-2854.
63. M. Di Pisa, A. Hauser and O. Seitz, *Chembiochem*, 2017, **18**, 872-879.
64. S. Middel, C. H. Panse, S. Nawratil and U. Diederichsen, *Chembiochem*, 2017, **18**, 2328-2332.
65. V. Y. Torbeev and S. B. H. Kent, *Angewandte Chemie-International Edition*, 2007, **46**, 1667-1670.
66. J. Niu, R. Hili and D. R. Liu, *Nature Chemistry*, 2013, **5**, 282-292.
67. C. Guo, C. P. Watkins and R. Hili, *Journal of the American Chemical Society*, 2015, **137**, 11191-11196.
68. D. H. Kong, Y. Lei, W. Yeung and R. Hili, *Angewandte Chemie-International Edition*, 2016, **55**, 13164-13168.
69. P. S. Huang, S. E. Boyken and D. Baker, *Nature*, 2016, **537**, 320-327.
70. N. C. Seeman and H. F. Sleiman, *Nature Reviews Materials*, 2018, **3**.

71. F. Hong, F. Zhang, Y. Liu and H. Yan, *Chemical Reviews*, 2017, **117**, 12584-12640.
72. B. A. R. Williams, C. W. Diehnelt, P. Belcher, M. Greving, N. W. Woodbury, S. A. Johnston and J. C. Chaput, *Journal of the American Chemical Society*, 2009, **131**, 17233-17241.
73. F. Diezmann, L. von Kleist, V. Hauckeb and O. Seitz, *Organic & Biomolecular Chemistry*, 2015, **13**, 8008-8015.
74. M. Marczynke, K. Groger and O. Seitz, *Bioconjugate Chemistry*, 2017, **28**, 2384-2392.
75. H. Eberhard, F. Diezmann and O. Seitz, *Angewandte Chemie-International Edition*, 2011, **50**, 4146-4150.
76. B. M. G. Janssen, E. H. M. Lempens, L. L. C. Olijve, I. K. Voets, J. L. J. van Dongen, T. F. A. de Greef and M. Merkx, *Chemical Science*, 2013, **4**, 1442-1450.
77. E. A. Englund, D. Y. Wang, H. Fujigaki, H. Sakai, C. M. Micklitsch, R. Ghirlando, G. Martin-Manso, M. L. Pendrak, D. D. Roberts, S. R. Durell and D. H. Appella, *Nature Communications*, 2012, **3**.
78. C. G. Lou, M. C. Martos-Maldonado, C. S. Madsen, R. P. Thomsen, S. R. Midtgaard, N. J. Christensen, J. Kjems, P. W. Thulstrup, J. Wengel and K. J. Jensen, *Nature Communications*, 2016, **7**.
79. P. S. Ghosh and A. D. Hamilton, *Journal of the American Chemical Society*, 2012, **134**, 13208-13211.
80. L. Roglin, F. Altenbrunn and O. Seitz, *Chembiochem*, 2009, **10**, 758-765.
81. L. Roglin, M. R. Ahmadian and O. Seitz, *Angewandte Chemie-International Edition*, 2007, **46**, 2704-2707.
82. K. J. Oh, K. J. Cash, A. A. Lubin and K. W. Plaxco, *Chemical Communications*, 2007, 4869-4871.
83. S. Thurley, L. Roglin and O. Seitz, *Journal of the American Chemical Society*, 2007, **129**, 12693-+.
84. M. Fischbach, U. Resch-Genger and O. Seitz, *Angewandte Chemie-International Edition*, 2014, **53**, 11955-11959.
85. T. Machida, S. Dutt and N. Winssinger, *Angewandte Chemie-International Edition*, 2016, **55**, 8595-8598.
86. B. A. R. Williams, K. Lund, Y. Liu, H. Yan and J. C. Chaput, *Angewandte Chemie-International Edition*, 2007, **46**, 3051-3054.
87. H. Lee, A. K. R. Lytton-Jean, Y. Chen, K. T. Love, A. I. Park, E. D. Karagiannis, A. Sehgal, W. Querbes, C. S. Zurenko, M. Jayaraman, C. G. Peng, K. Charisse, A. Borodovsky, M. Manoharan, J. S. Donahoe, J. Truelove, M. Nahrendorf, R. Langer and D. G. Anderson, *Nature Nanotechnology*, 2012, **7**, 389-393.
88. Z. W. Xia, P. Wang, X. W. Liu, T. Liu, Y. N. Yan, J. Yan, J. Zhong, G. Sun and D. N. He, *Biochemistry*, 2016, **55**, 1326-1331.
89. J. D. Flory, S. Shinde, S. Lin, Y. Liu, H. Yan, G. Ghirlanda and P. Fromme, *Journal of the American Chemical Society*, 2013, **135**, 6985-6993.
90. E. Spruijt, S. E. Tusk and H. Bayley, *Nature Nanotechnology*, 2018, **13**, 739-+.
91. S. M. Douglas, H. Dietz, T. Liedl, B. Hogberg, F. Graf and W. M. Shih, *Nature*, 2009, **459**, 1154-1154.
92. Y. G. Ke, L. L. Ong, W. M. Shih and P. Yin, *Science*, 2012, **338**, 1177-1183.

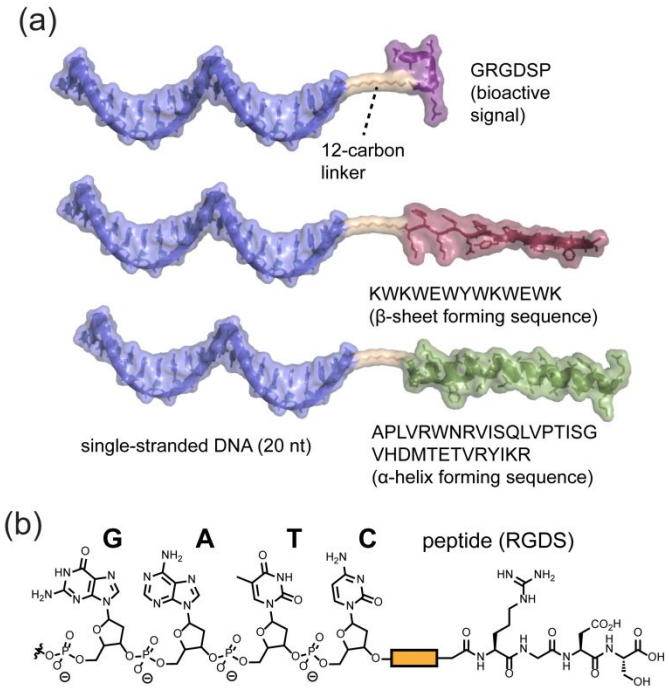


Figure 1: Structure of peptide-oligonucleotide conjugates. (a) To-scale molecular representations of single-stranded DNA linked to peptide sequences of varying length and structure. (b) Chemical structure of the four DNA bases linked to a short peptide, to demonstrate the relative size of each molecule. (a) modified from PDB accession codes 3ZE2 and 2J58

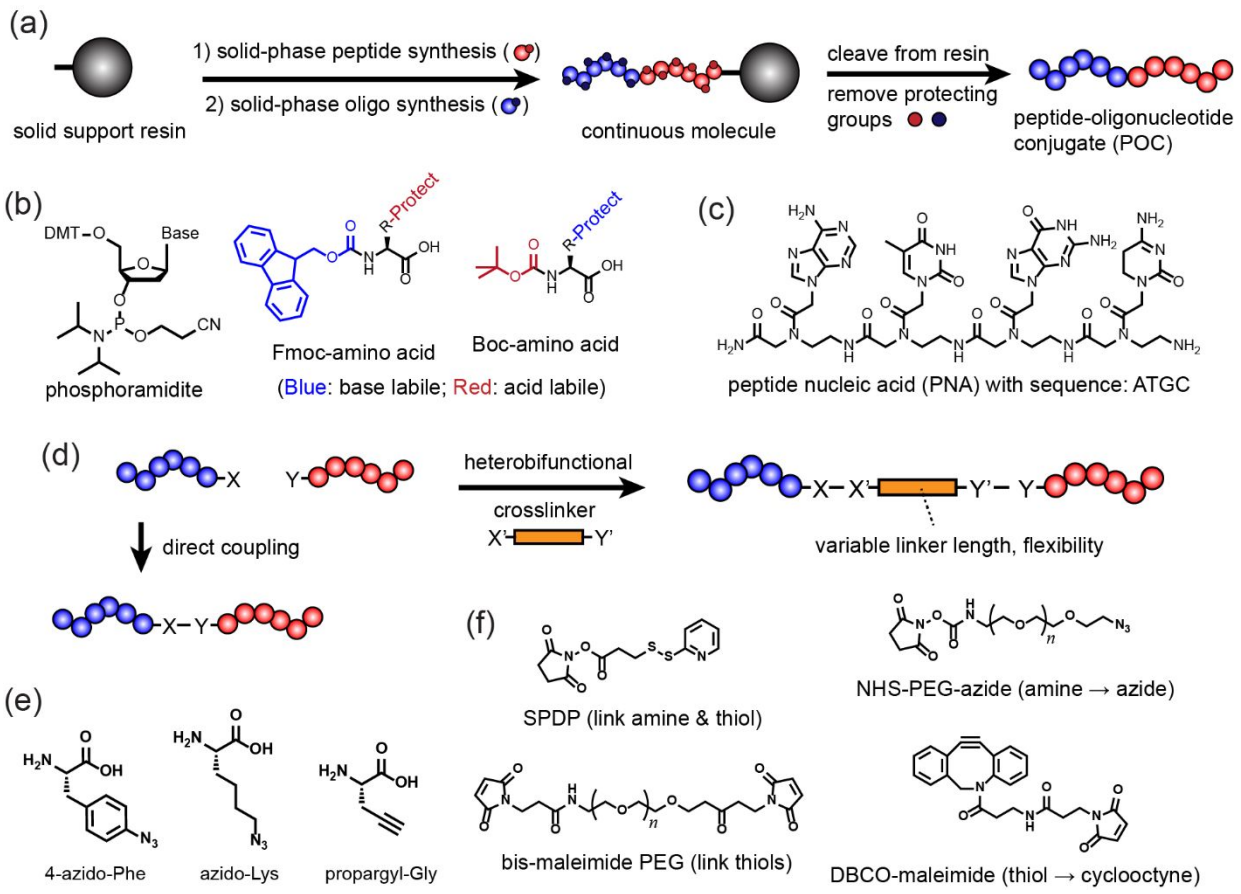


Figure 2: Chemical approaches for POC synthesis. (a) Continuous synthesis of a POC on-resin through the sequential coupling of protected peptide and oligonucleotide monomers. Cleavage from the resin and removal of the protecting groups generates a full-length POC. (b) Chemical structure of a generic phosphoramidite used in solid-phase oligonucleotide synthesis, and Fmoc- or Boc-protected amino acids used in solid-phase peptide synthesis. (c) Structure of peptide nucleic acid (PNA), depicting the backbone and pendant nucleobases. (d) In-solution coupling of peptides and oligonucleotides, either directly or through a bifunctional linker. (e) Structures of three non-canonical amino acids incorporated into peptides for subsequent click coupling to oligonucleotides. (f) Chemical structures of several homo- or hetero-bifunctional crosslinkers or functionalization reagents for solution-phase POC synthesis.

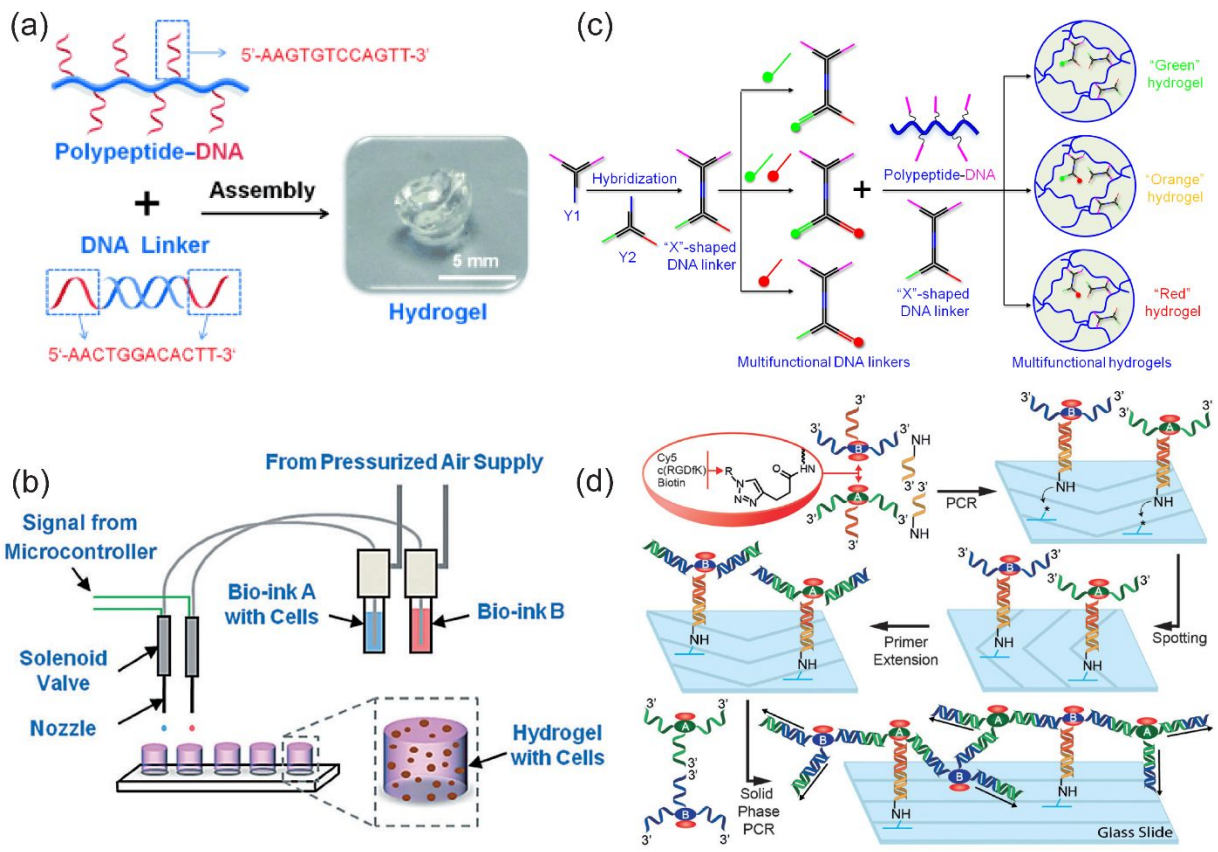


Figure 3: Branched POC biomaterials. (a) Polypeptides bearing ssDNA handles can be crosslinked with complementary DNA strands to create a self-supporting hydrogel.[24] (b) The two components can be extruded as "inks" from a bioprinter to generate shape-controlled hydrogels.[24] (c) Branched DNA crosslinkers containing pendant DNA handles can be used to create hydrogels with addressable ligand functionalization.[26] (d) DNA building blocks containing peptides can be used to polymerize a hydrogel on a surface through PCR.[27]

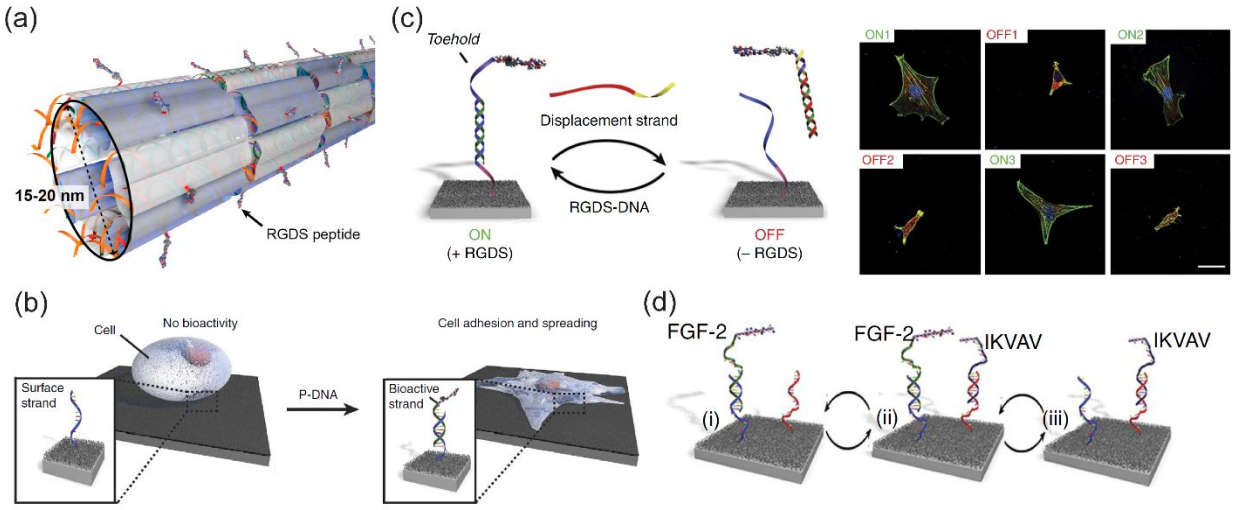


Figure 4: DNA biomaterial scaffolds for peptide presentation. (a) Schematic of a self-assembling DNA nanotube with multivalent display of the bioactive ligand RGDS.[28] (b) DNA as a functional linker to bind a POC and imbue an otherwise inert surface with bioactivity (e.g. cell adhesion and spreading).[29] (c) Removal of the POC from the surface through toehold-mediated strand displacement results in loss of bioactivity and contraction of cells; adding the POC to the surface restores bioactivity in a highly reversible manner.[29] (d) Two different peptide signals can be controlled independently through the use of orthogonal displacement strands.[29]

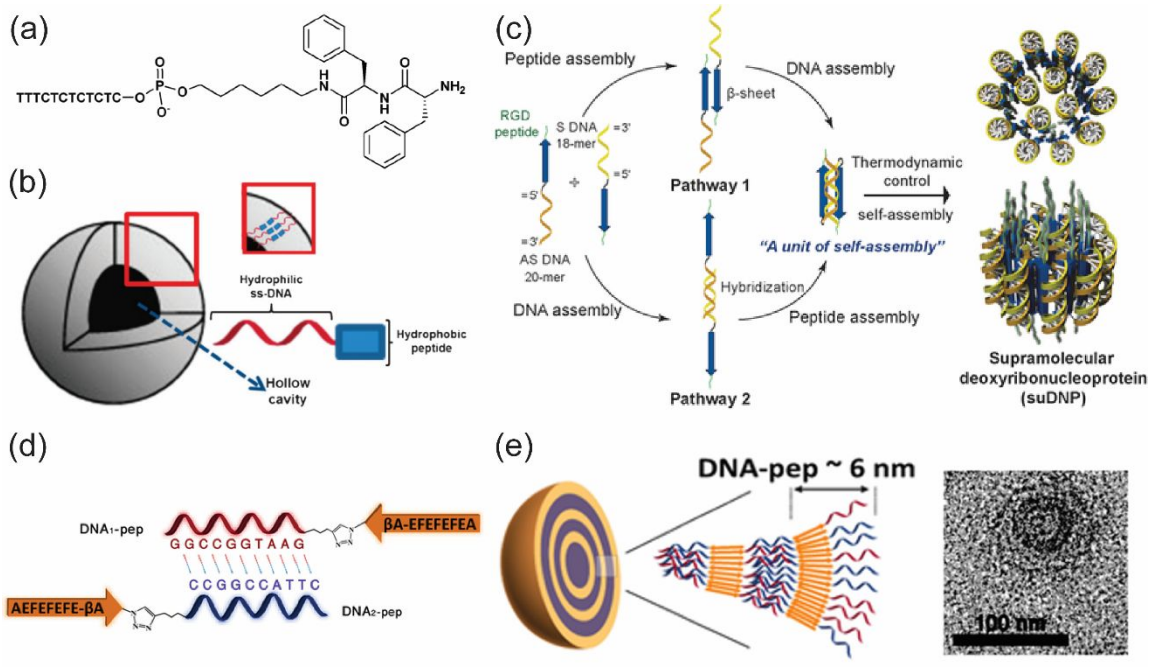


Figure 5: Hybrid self-assembling systems from peptides and DNA. (a) Structure of DNA appended to diphenylalanine, which self-assembles to form a hollow spherical structural with a POC bilayer (b).[36] (c) Beta-sheet forming peptides linked to DNA can form toroidal nanoparticles by balancing the forces of DNA hybridization and hydrogen bonding.[38] (d) DNA linked to a beta-sheet forming peptide results in multi-lamellar nanostructures of alternating peptide and oligonucleotide layers (e).[39]

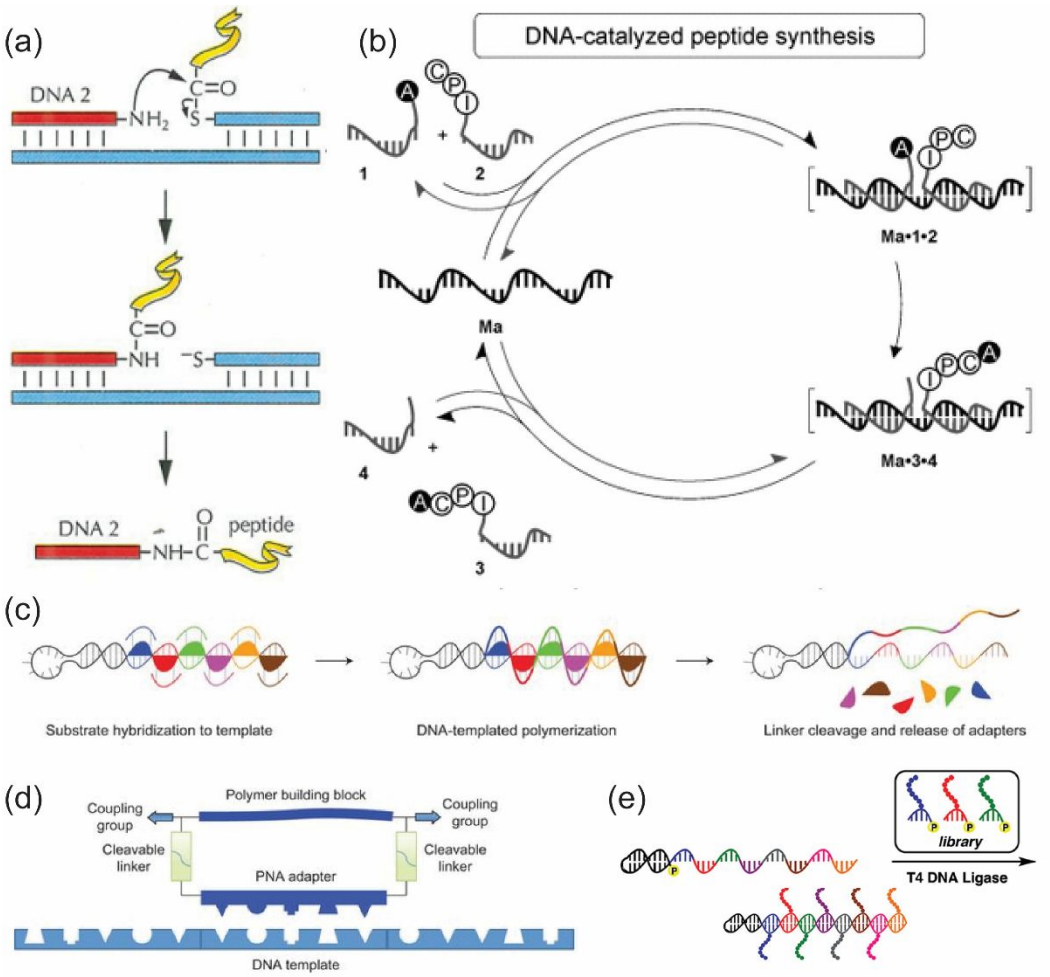


Figure 6: DNA-templated peptide coupling. (a) DNA was used to enhance amide bond formation between an amine-terminated oligonucleotide and a POC thioester.[59] (b) Example of native chemical ligation between two POCs enhanced by the high local concentration enforced by a complementary DNA template.[61] (c) Synthesis of sequence-specific polymers on a DNA template through the use of programmable PNA-polymer "adaptors" (d).[66] (e) Pentanucleotides bearing pendant peptides can be linked using T4 DNA ligase on a programmable oligonucleotide template.[67]

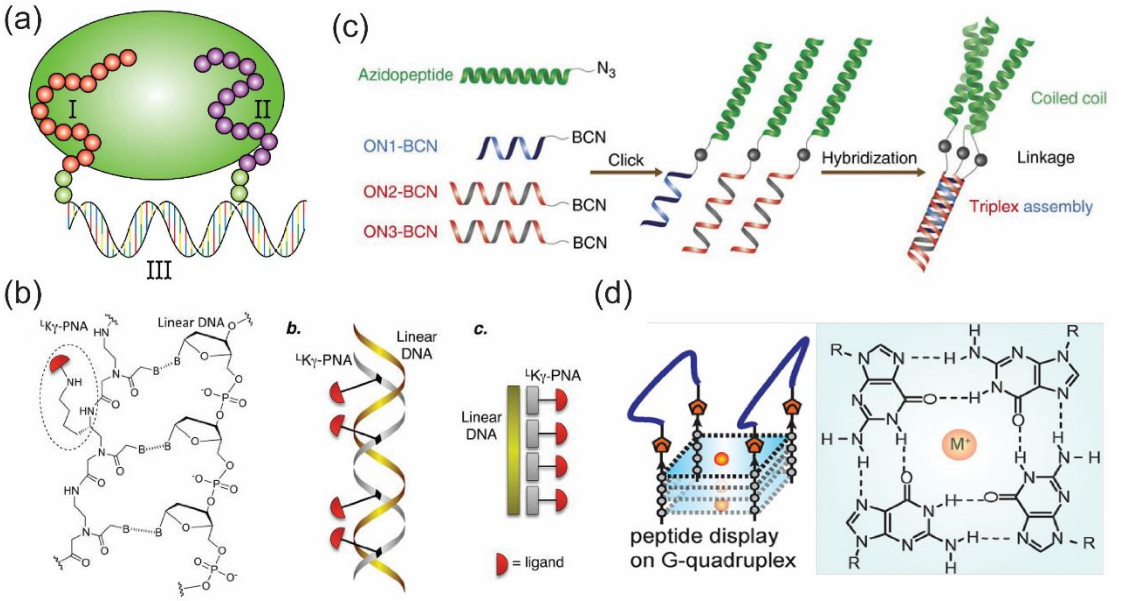


Figure 7: Peptides templated by simple DNA structures. (a) Duplex DNA can serve as a tunable linker to immobilize two peptides with weak affinity for a protein and enhance their binding to different faces of the target.[72] (b) PNA-POCs can be used to create a multivalent display of a bioactive peptide on a repetitive DNA template.[77] (c) DNA triplex formation can enhance the stability of a homotrimeric coiled-coil peptide.[78] (d) DNA-peptide-DNA triblock molecules constrain the peptides into loops when the oligonucleotides form a G-quadruplex.[79]

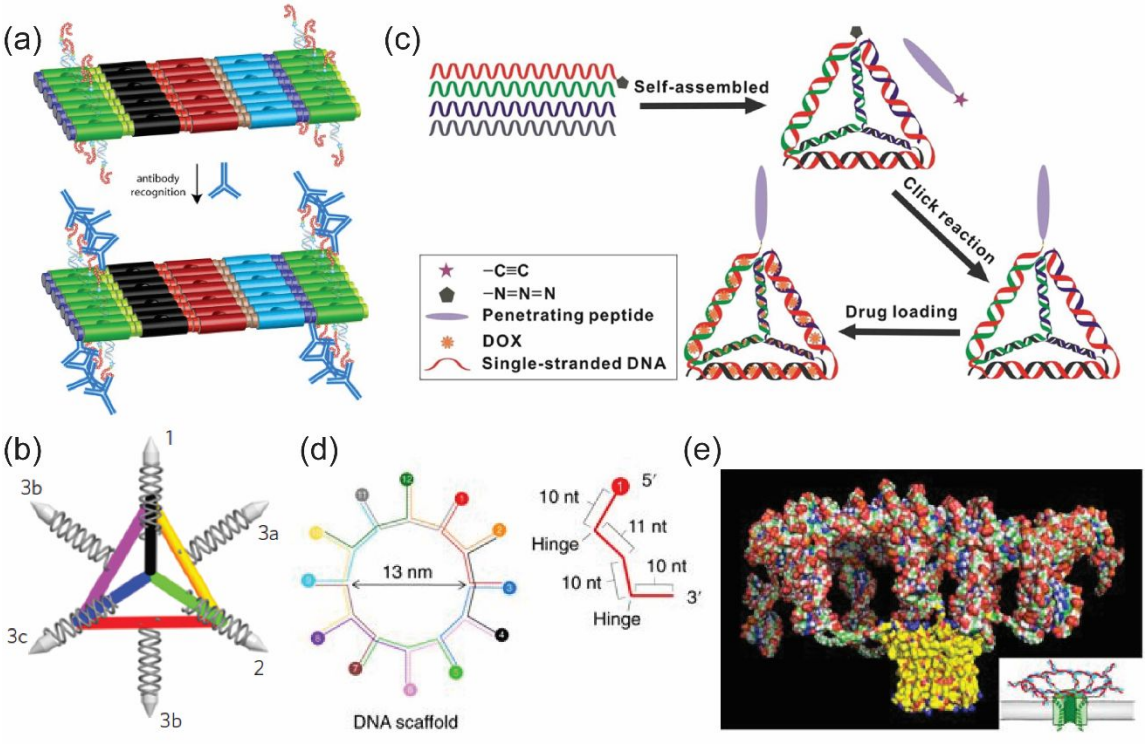
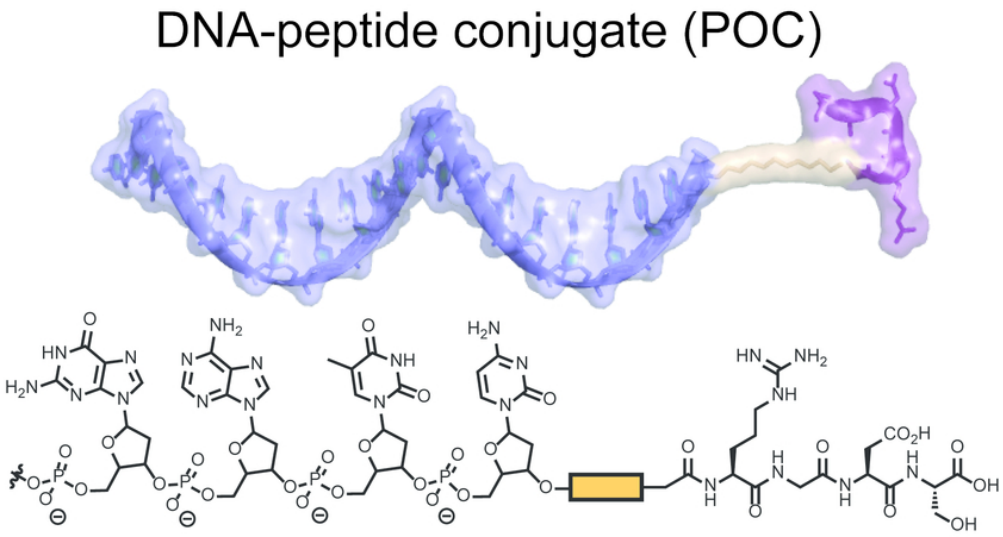


Figure 8: DNA nanostructures for scaffolding peptides. (a) DNA tile self-assembly can create programmable two-dimensional peptide arrays for detecting binding partners like antibodies.[86] (b) A self-assembled DNA tetrahedron can control the number and spatial presentation of bioactive ligands like peptides for targeted siRNA delivery.[87] (c) Example of a DNA tetrahedron modified with a tumor targeting peptide and loaded with doxorubicin for glioblastoma therapy.[88] (d,e) Design and molecular simulation model of a circular DNA scaffold that can display up to twelve membrane-spanning, pore-forming peptides.[90]



71x37mm (300 x 300 DPI)