Review

Hybrid Nanostructures from the Self-Assembly of Proteins and DNA

Nicholas Stephanopoulos[1,2,*]

Proteins and DNA are two commonly used molecules for self-assembling nanotechnology. In this tutorial review, we discuss the hybrid field of "protein-DNA nanotechnology," whereby proteins are integrated with DNA scaffolds for the creation of hybrid nanostructures with distinct properties of each molecular type. We first discuss bioconjugation strategies, both covalent and supramolecular, for integrating proteins with DNA nanostructures. Next, we review seminal work in four emerging areas of protein-DNA nanotechnology: (1) controlling protein orientation on DNA nanoscaffolds, (2) controlling protein function with DNA nanodevices, (3) answering biological questions with protein-DNA nanostructures, and (4) building hybrid structures that integrate both protein and DNA structural units. Finally, we close with a series of forward-looking research propositions and ideas for directions of the field. The emphasis of this work is on integrated nanostructures with precise protein orientation on DNA scaffolds, as well as hybrid assemblies that integrate the structural and functional properties of each molecule.

INTRODUCTION AND MOTIVATION FOR PROTEIN-DNA NANOTECHNOLOGY

Since the inception of nanotechnology, scientists have dreamed of the ability to create tiny structures and machines that can manipulate matter at will. For example, to this day much research in the field is driven by the concept of nanomachines or devices (or, more evocatively, "nanorobots") that can interact with biological or other systems in programmable ways. Such nanostructures could, for example, diagnose and treat disease, synthesize novel materials, harvest and shuttle energy, exert mechanical forces, store and transmit information, or arrange other molecules with atomic precision. Richard Feynman outlined this idea of nanoscience in 1959 in his groundbreaking talk "There's Plenty of Room at the Bottom," and generations of chemists, biologists, engineers, and materials scientists have since probed the limits of nanostructure synthesis and function. Not surprisingly, biology has served as one of the most fertile sources of inspiration in this endeavor. Cells are teeming with nanoscale analogs of macroscopic structures and machines, including architectural scaffolds (the cytoskeleton), programmable "robots" and assembly lines for building materials in a controlled and monodisperse manner (the ribosome, non-ribosomal peptide synthesis, and enzyme cascades), motors and other machines for exerting mechanical force (actin, myosin, and focal adhesion complexes), channels and transport mechanisms for controlling the flow of matter (ion channels and endocytosis), structural materials with exceptional strength (spider silk, bone, and nacre), adaptors that can selectively bind to a target in a crowded sea of competing molecules (antibodies and ligand receptors), and both "hardware" and "software" for information processing (DNA and RNA copying and transcription, signaling pathways, and riboswitches). Aside from cells, viruses offer another example of complex biological

The Bigger Picture

Nanotechnology as a field seeks to create structures and materials that can manipulate and influence the microscopic world in much the same way that traditional machines and devices work on the macroscopic world. For inspiration, scientists have turned to biology, which has countless examples of nanoscale structures and machines that can carry out complex functions. Biological molecules such as proteins and DNA are particularly attractive for this purpose because of their programmable nature and functional relevance. In this review, we discuss hybrid nanostructures that integrate the structural programmability of DNA nanotechnology with the chemical and functional diversity of proteins. We discuss strategies for creating complex, integrated structures with these two biomolecules, as well as four areas where they have found application. In the long term, the field of "protein-DNA nanotechnology" has the potential to create materials with capabilities that rival, or even surpass, nature.

nanodevices with their ability to enter a cell, bypass the cell's defense mechanisms, and create new copies of themselves by hijacking the host machinery.

All of these functions, and many others, are mediated either entirely or in part by proteins. Although composed primarily of the 20 canonical amino acids, proteins have a breathtakingly wide range of functions as a result of their complex folds and their ability to hierarchically self-assemble with other proteins, DNA, RNA, carbohydrates, and lipids. This complexity comes at a cost, however, because the relationship between protein sequence and function or assembly is still imperfectly understood. The past few decades have brought remarkable progress in directed protein evolution,[1] *de novo* computational design,[2] the repurposing of existing biological scaffolds such as viral capsids,[3] and the abstraction of design rules in simplified building blocks such as self-assembling peptides[4] or proteins.[5] However, there is still a need for generating nanostructures with a high degree of programmability and structural control that can capitalize on the enormous power of native protein function both for recapitulating and probing biological systems and for designing new materials that can outpace nature. Specifically, one key unmet challenge is building highly anisotropic structures *de novo* from proteins, such as a nanoscale machine or robot that can perform a function given an external stimulus. Cells possess many such multi-protein complexes, but the difficulty in predicting even monomeric protein structures means that most assemblies made to date are highly symmetric—such as polyhedral cages or extended fiber and sheet assemblies—and usually static.

By contrast, oligonucleotides have proved to be highly promising molecules for constructing anisotropic and uniquely addressable assemblies at the nanoscale, as well as imparting programmable dynamic behavior. The field of DNA nanotechnology uses these molecules as "smart" self-assembling building blocks divorced from their natural genetic role. The design rules that drive oligonucleotide hybridization (i.e., the Watson-Crick pairing rules) are well known, a vast number of orthogonal interactions (i.e., sequences) exist, and the structural and physicochemical properties of the double helix, single-stranded DNA, and Holliday junction crossovers have been determined to great precision. In the past three decades, DNA nanotechnology has reported an ever-increasing catalog of structures, including simple 1D and 2D arrays[6,7] (Figures 1A and 1B), 3D crystals (Figure 1C),[8,9] highly complex and anisotropic 2D and 3D nanostructures (commonly known as "DNA origami;" Figures 1D and 1E),[10,11,12] and dynamic[13] or logic-gated[14] machines and devices (Figure 1F). Importantly, because DNA nanotechnology relies on a small subset of key motifs for self-assembly—which are then combined independently and modularly into more complex structures—the design process can be aided by user-friendly, graphical interface software such as Cadnano.[15] This facility allows both rapid entry of non-experts into the field and the parallel design and testing of multiple structures with a high degree of precision.

The programmability and tractability of DNA, however, comes at the expense of chemical heterogeneity. With a few notable exceptions, such as aptamers or DNAzymes, DNA does not come close to mimicking the functions of proteins. Thus, in recent years there has been an explosion of interest in merging the chemical and functional diversity of proteins with the structural programmability of nucleic acid nanotechnology to forge a truly hybrid field of "protein-DNA nanotechnology." Research in this field has already shown promise in creating protein-DNA nanostructures for applications that include targeted delivery of therapeutics, biosensing, control over protein activity, elucidation of protein structure, functional biomaterials, and experiments to probe biology in new ways. In this tutorial review, we will

[1]School of Molecular Sciences, Arizona State University, Tempe, AZ 85281, USA

[2]Biodesign Center for Molecular Design and Biomimetics, Arizona State University, Tempe, AZ 85281, USA
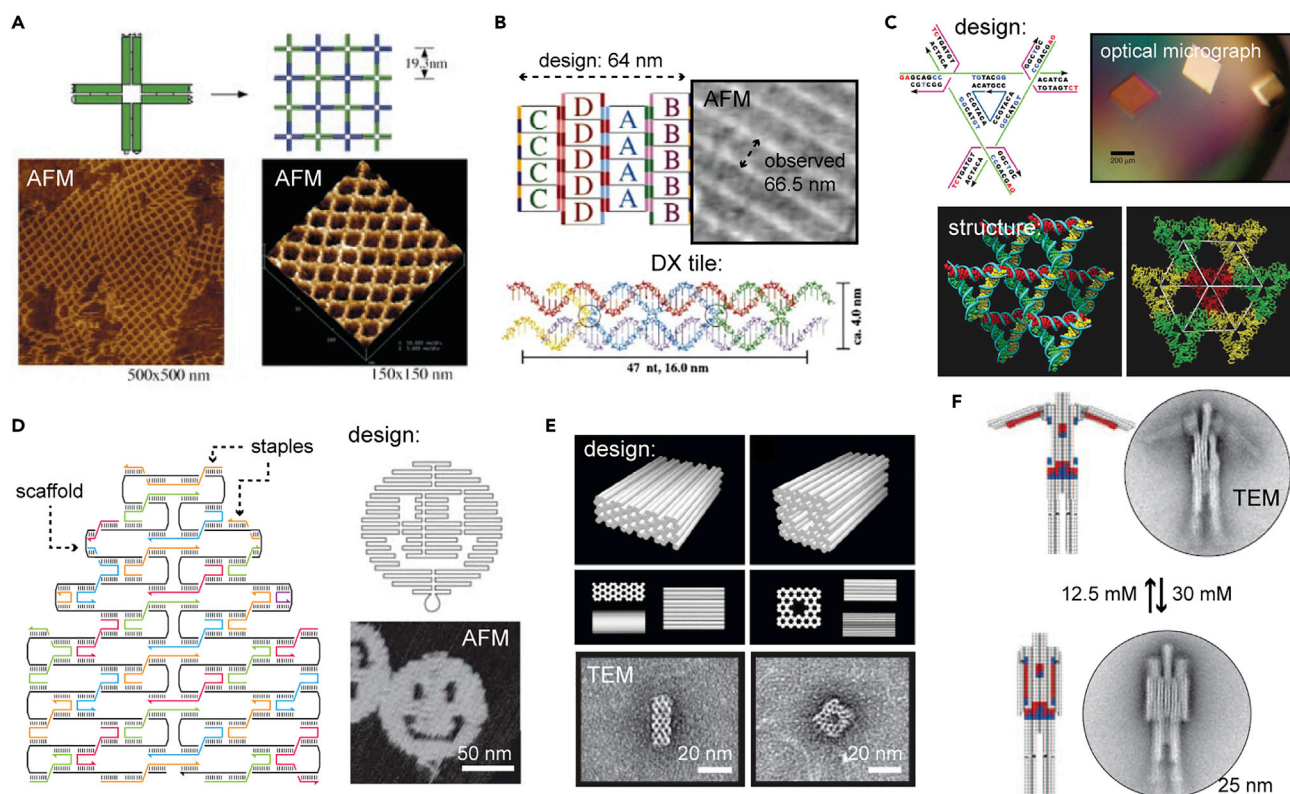
*Correspondence: nstepha1@asu.edu

Figure 1. Examples of DNA-Based Nanostructures

(A and B) 2D arrays based on cross-shaped branched junctions (A) and double-crossover ("DX") tiles (B). Reprinted with permission from Yan et al.[6] (copyright 2003 AAAS) and Winfree et al.[7] (copyright 1998 Springer Nature).

(C) 3D self-assembled crystals based on a tensegrity triangle motif. Reprinted with permission from Zheng et al.[8] Copyright 2009 Springer Nature.

(D) 2D "DNA origami" with a long scaffold strand folded by many short staple strands. Reprinted with permission from Rothemund.[10] Copyright 2006 Springer Nature.

(E) Example of 3D DNA origami shapes. Reprinted with permission from Douglas et al.[11] Copyright 2009 Springer Nature.

(F) A dynamically reconfigurable DNA origami "nanorobot" that switches between two states when the concentration of divalent magnesium is varied. Reprinted with permission from Gerling et al.[13] Copyright 2015 AAAS.

discuss recent advances in protein-DNA nanotechnology with an emphasis on truly integrated assemblies that seek a well-defined relationship between the protein and DNA scaffold. To accomplish these endeavors, researchers have had to develop a host of novel chemical methods for synthesizing hybrid nanostructures, including site-specific modification of the protein with one or more oligonucleotides, the use of binding agents (such as DNA-binding fusion proteins) that yield a defined interface between the two components, control over linker length and rigidity between the two molecules, or some combination of the aforementioned factors.

We will begin with an overview of protein-DNA bioconjugation strategies in which we emphasize chemical approaches for site-specific covalent modification and strategies for modifying proteins more than once, as well as discuss several supramolecular methods for associating the two molecules. We will then describe landmark work in four key areas that have seen an explosion of interest in recent years: (1) controlling protein orientation on a DNA scaffold, (2) dynamically controlling protein function by using DNA structures, (3) using proteins on DNA nanoscaffolds to answer biological questions, and (4) synthesizing hybrid nanoscale assemblies with both protein and oligonucleotide *structural* components. Finally, we

will close with a series of research propositions and future directions for the field, where we stress the role of novel chemical and supramolecular methods for improving the integration between proteins and DNA.

We also take this moment to mention what this review will *not* cover. First and foremost, it is not intended to provide a comprehensive overview of all protein-DNA hybrid materials; rather, it focuses on seminal examples in the four areas described. An extensive and rich literature exists on the merger of proteins with DNA nanotechnology,[12,16–21] so we have restricted the focus on the four key areas described below. Second, because of space limitations, we will not cover work on enzymatic cascades scaffolded by oligonucleotides despite the central role that these systems have played in protein-DNA nanotechnology. We will discuss some examples that pertain to the dynamic control over enzyme function in Using DNA Nanostructures to Dynamically Control Protein Function, but for a more comprehensive treatment, we refer the interested reader to several comprehensive reviews on this rich topic.[22–24] Third, although RNA will be increasingly used in the future with proteins—thanks to its richer structural diversity, its ability to be transcribed inside cells, and the existence of multiple protein-binding domains—we will restrict this discussion to DNA-based scaffolds because of their preponderance in the field. Fourth, we will discuss hybrid nanostructures containing only proteins produced by recombinant expression or from natural isolates rather than synthetic peptides conjugated to oligonucleotides.[25] Fifth, because the focus of this review is on controlled integration and positioning of proteins on DNA scaffolds—driven by chemical strategies for site-specifically modifying proteins with DNA—we will not cover examples where proteins are used to coat DNA nanostructures via electrostatic effects or non-specific DNA-binding domains.[26–29] Sixth, we will consider only systems where a protein is attached to a DNA nanostructure more complex than a duplex or a simple branched junction or forms a hybrid assembly with both structural components. We will not discuss examples where DNA is used as a barcode, as a linker to attach a protein to another material, or for proximity-based ligation or detection strategies, although these approaches have been crucial in a number of other applications reviewed elsehwere.[30,31] Finally, both the work covered and the forward-looking section at the end reflect our own research interests and excitement for future work. We apologize in advance to all the scientists whose work we are not able to discuss because of space limitations or the selection of themes. There are unquestionably many fruitful directions in this hybrid area of nanotechnology, and it is our hope that this work will both highlight emerging directions in the field and spur new ideas in previously untapped disciplines.

STRATEGIES FOR MODIFICATION OF PROTEINS WITH DNA

Given that only a small subset of proteins naturally interact with DNA or RNA, most hybrid protein-DNA nanostructures rely on one of two methods for integrating the two molecules: (1) direct covalent modification of the protein with the oligonucleotide and (2) fusion of the protein of interest (POI) with a DNA-binding protein, a protein for which an aptamer or exists, or streptavidin (which binds biotin). The first approach, which results in a covalent linkage between the two molecules, poses distinct challenges in both reactivity and site specificity of the target. Proteins and DNA are both large molecules with many potentially reactive sites, so achieving efficient coupling without compromising their function can be very difficult at the low micromolar concentrations typically used. This challenge is exacerbated when the POI is highly cationic, which can lead to non-specific aggregation with the oligonucleotide. Incomplete or over-modification is common, which in turn raises the

issue of separating the desired species from unmodified proteins or proteins with too many strands.

Below we will describe some of the key strategies for protein-oligonucleotide conjugation and outline the advantages and disadvantages of each. Although a great number of bioconjugation reactions exist,[31,32] here we will highlight only approaches already demonstrated for linking proteins to DNA. We divide these approaches into two sections: (1) methods commonly used in most protein-DNA conjugation studies, both historically and currently, and (2) less commonly used strategies that nonetheless have great potential for future applications, especially when two or more modifications with DNA are necessary. We also take a moment to stress that most monomeric proteins are quite small in relation to DNA nanostructures. For comparison, in Figure 2A, we show that green fluorescent protein (GFP), which has a molecular weight of ~27 kDa, is of comparable size to a 20-nt strand of DNA (approximately two helical turns). Thus, larger assemblies such as DNA origami (~5,000 kDa when the M13 scaffold is used) dwarf most proteins, and they often surpass even large multivalent protein assemblies such as viral capsids.[33]

Common Approaches for Protein-DNA Conjugation
Lysine Acylation
The most straightforward way to modify a protein with DNA is through lysine acylation with activated esters or iso(thio)cyanates (Figures 2B and 2C). Lysine is one of the most common surface-exposed residues on a protein surface, so most wild-type proteins can be modified without further engineering. The biggest downside of this approach is that usually multiple lysines are accessible (and potentially the N terminus as well), so a mixture of conjugates varying in the location and number of modifications is obtained. This lack of selectivity can be problematic if the DNA strand ends up attached to a critical part of the protein (such as a binding interface) or if it positions the protein on a DNA scaffold in a way that blocks its function. Lysine acylation can also be relatively slow and inefficient at the low micromolar concentrations used; with small molecules, a large excess of one coupling partner can be used to circumvent this issue, but this is often not possible with DNA and proteins. Background hydrolysis of amine-reactive reagents such as NHS esters also restricts the utility of this reaction. To link DNA to proteins via lysine chemistry, homo-bifunctional linkers such as disuccinimidyl glutarate (DSG) or tunable bis-NHS polyethylene glycol (PEG) linkers, in conjunction with amine-modified DNA, are typically used (Figure 2D). Amines are available as a common DNA modification from commercial suppliers and are introduced via solid-phase synthesis with a functionalized phosphoramidite. Modifying the DNA with a thiol and using a hetero-bifunctional crosslinker with a disulfide or maleimide moiety can avoid the crosslinking of two proteins or two oligonucleotides (see Cysteine Modification).

One key consideration with most of the conjugation strategies described herein is that they involve flexible linkers such as aliphatic carbon chains or oligoethylene glycol in the bifunctional molecule itself, between the DNA and the introduced functional group, or both. Although the flexibility of these linkers helps facilitate the reaction of the two large biomolecules, they often preclude a defined orientation between the protein and DNA scaffold, which could be problematic for some applications. Creating truly hybrid nanostructures with well-defined relationships between the protein and DNA components will most likely require a more rigid connection. Indeed, one of the key challenges in protein-oligonucleotide nanotechnology will be building complex assemblies where the two molecular scaffolds are
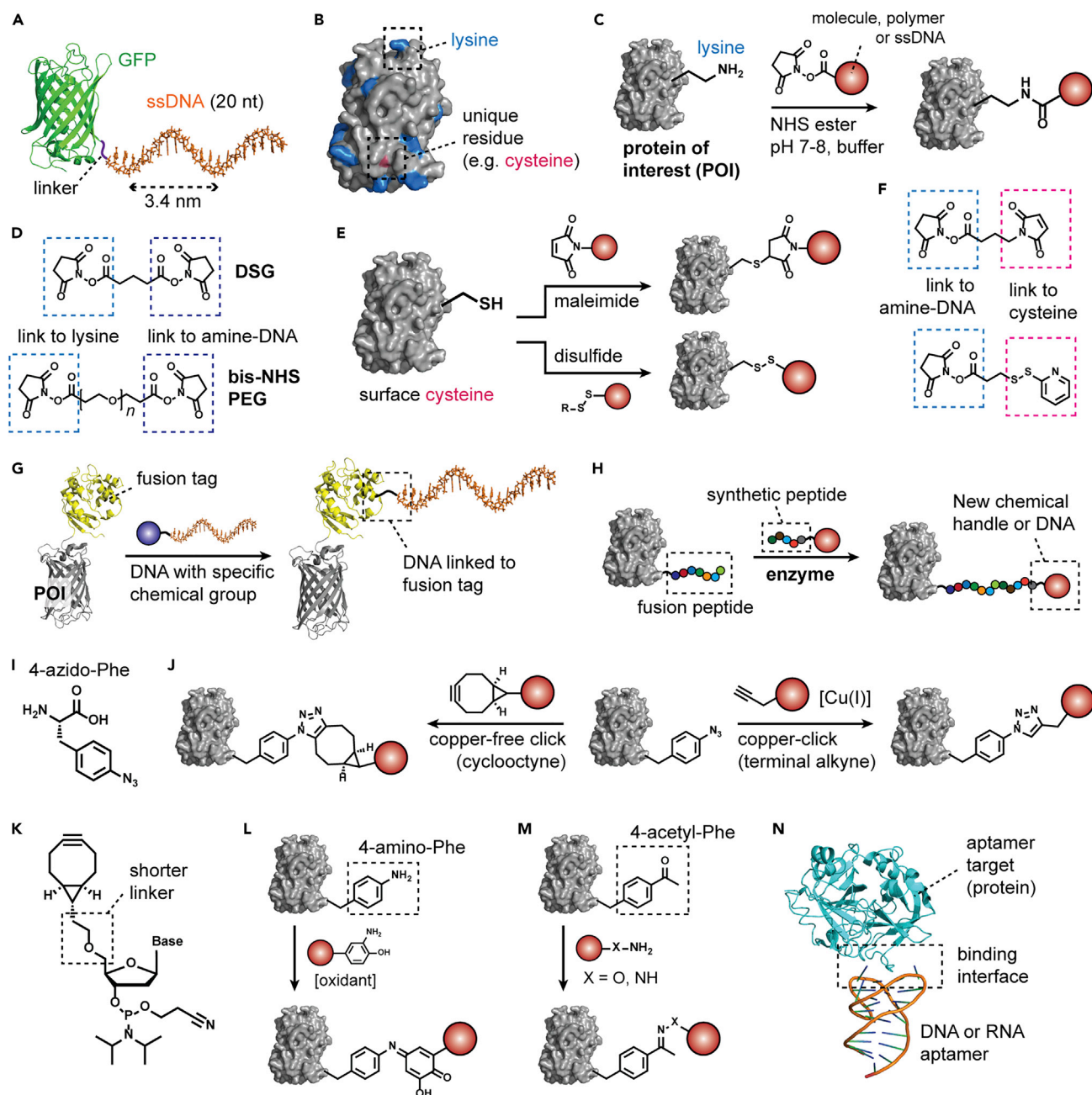
Figure 2. Bioconjugation Chemistry for Protein-DNA Hybrids

(A) Schematic of GFP-DNA conjugate shows the relative size of the two molecules.

(B) Most proteins (e.g., GFP) contain many lysine residues on their surface, whereas a unique residue such as cysteine can often be incorporated for site selectivity.

(C) Lysine acylation with NHS esters.

(D) Examples of commercially available homo-bifunctional NHS ester crosslinkers.

(E) Modification of cysteine with maleimide or disulfide reagents.

(F) Examples of hetero-bifunctional molecules for linking amines to thiols.

(G) Functionalization of a genetically fused self-ligating protein (e.g., SNAP-tag) with DNA bearing a specific chemical tag (e.g., O^6-benzylguanine).

(H) Enzymatic modification of a peptide genetically fused on a protein with a novel chemical handle (or a DNA strand) with the use of a synthetic peptide.

(I) Structure of the non-canonical amino acid (NCAA) 4-azidophenylalanine (azF).

(J) Modification of azF-containing proteins with "click" chemistry.

(K) Structure of a phosphoramidite (for solid-phase DNA synthesis) with minimal linker length between the 5' end and a cyclooctyne moiety.

integrated in a seamless fashion, much like naturally occurring protein-protein interfaces. Although the simplicity of lysine chemistry makes it ubiquitous for making protein-DNA conjugates, we will for the most part not discuss this method; however, we will include a few exceptions for particularly interesting applications within the four areas described or as a secondary reaction in conjunction with a more site-specific strategy.

Cysteine Modification

A second way to modify native protein residues with DNA is through alkylation of thiols with reagents such as maleimides and iodoacetamides or via disulfide formation and exchange (Figure 2E). Because cysteine is one of the rarest amino acids exposed on a protein surface—most are either buried in an active site or tied up in disulfide bonds—this strategy is powerful for achieving site specificity. A mutagenically introduced cysteine can often be targeted selectively, allowing for DNA modification away from any potentially deleterious site on the protein, and side reactions with other nucleophiles such as amines can usually be avoided. Cysteine modification represents the best balance between selectivity and general accessibility for protein-DNA bioconjugation and is the first reaction our lab and many others working on protein-DNA nanotechnology use when site selectivity is required.[32] Furthermore, a number of commercially available linkers allow for modification of the protein with either amine- or thiol-modified DNA (Figure 2F). One potential downside of this method is that a surface-exposed cysteine can lead to protein dimerization and/or aggregation as a result of spontaneous disulfide formation.[34] These disulfides can be broken prior to modification with a reducing agent such as dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP), although one must carefully control the reaction conditions to avoid cleaving endogenous disulfides, which can in turn lead to over-modification of the protein.

Biotin-(strept)avidin and Ni-NTA

All of the above methods result in covalent modification of the protein, but strong non-covalent interactions can also be used to functionalize proteins with DNA. The most common approach uses the binding of biotin to streptavidin, which has such a high affinity ($K_d \sim 10^{-15}$ M) as to be almost irreversible. DNA strands functionalized with biotin are readily available from commercial suppliers, and tetravalent streptavidin can be used as an intermediary "glue" between DNA and a biotinylated protein; alternatively, monovalent avidin can be fused directly to the POI.[35] The interaction between a fused hexahistidine (His$_6$) tag—which is inherently site specific, though generally restricted to the protein's termini—and DNA functionalized with nickel-nitrilotriacetic acid (Ni-NTA) can also be used for modifying proteins in a reversible manner[36] and attaching them site specifically to DNA origami.[37] In one particularly elegant example, Gothelf and coworkers used a His$_6$ tag (or naturally occurring metal-binding patches) to transiently modify a protein with a DNA handle, which could in turn be used to direct a complementary strand modified with an NHS ester (Figure 3A).[38] In this way, the authors could target only a subset of lysine residues (in close proximity to the directing strand) without resorting to more complex bioconjugation strategies.
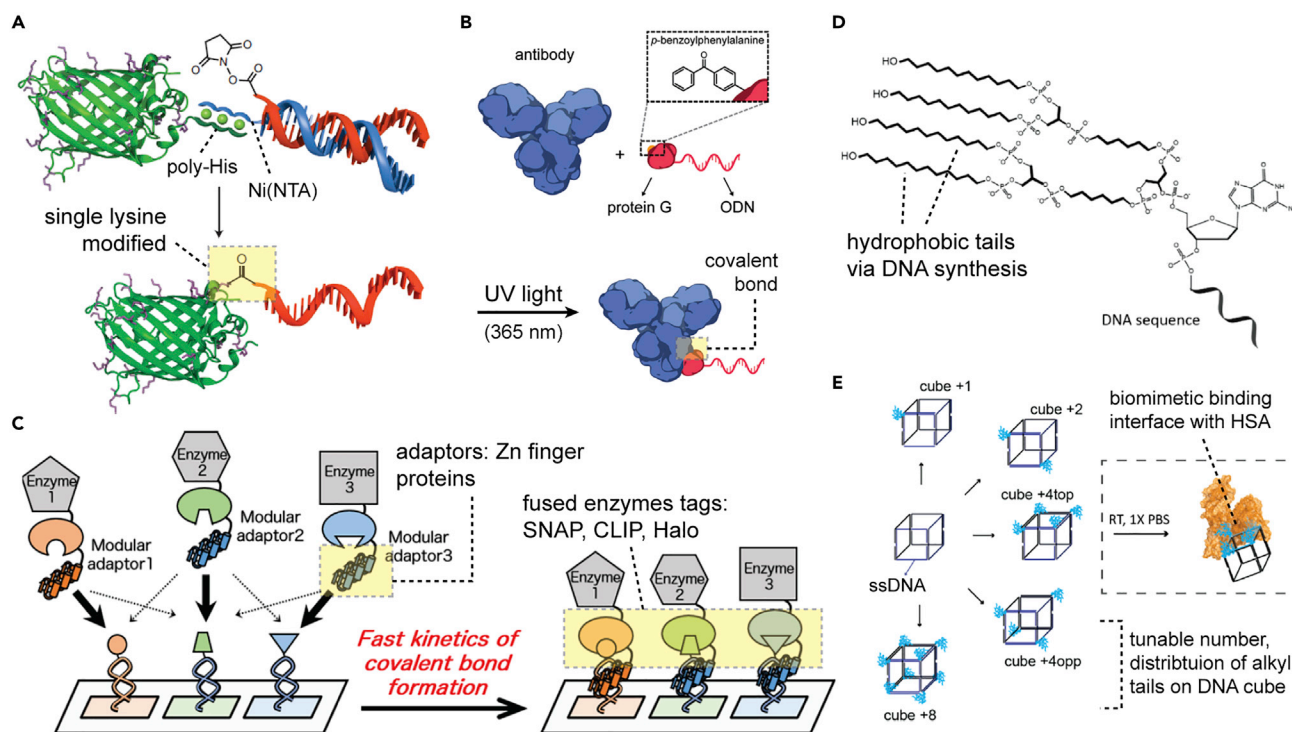
Figure 3. Additional Bioconjugation Strategies

(A) A Ni(NTA)-modified DNA strand was bound to a poly-His region on a protein surface through metal coordination; in a second step, this handle could direct an NHS-ester-functionalized complementary strand for selective modification of one lysine on the protein surface. Reprinted with permission from Rosen et al.[38] Copyright 2014 Springer Nature.

(B) Incorporation of the NCAA 4-benzoylphenylalanine into protein G bearing a DNA handle allows for covalent crosslinking to the Fc region upon UV illumination. Reprinted from Rosier et al.,[39] licensed under CC BY 3.0.

(C) Zinc-finger adaptor proteins were used to non-covalently localize three enzyme-tag fusions to orthogonal locations on a DNA origami scaffold, allowing for covalent trapping through SNAP-tag, CLIP-tag, and HaloTag reactions. Reprinted with permission from Nguyen et al.[40] Copyright 2017 American Chemical Society.

(D and E) Structure of a DNA oligonucleotide modified with a dendritic alkyl tail (D). The number and spatial distribution of alkyl dendrimers could be controlled on a DNA cube nanostructure, resulting in a controlled binding interface with human serum albumin (E). Reprinted with permission from Lacroix et al.[41] Copyright 2017 American Chemical Society.

Covalent Modification of Self-Ligating Protein Tags

An alternate common strategy for protein-DNA conjugation, which altogether avoids chemical conjugation of DNA to the POI, involves fusing the POI to protein tags that can in turn link specific substrates to their own surface. Functionalizing the target DNA with the chemical moiety accepted by these self-ligating tags results in conjugation of the oligonucleotide to the POI-tag fusion (Figure 2G). Specific examples include (1) SNAP-tags (which ligate O^6-benzylguanine groups),[42] (2) HaloTags (which ligate haloalkanes),[43] (3) CLIP-tags (which ligate O^2-benzylcytosine moieties),[44] and (4) the SpyTag/SpyCatcher system (which ligates a short peptide).[45] This approach is powerful because the fusion proteins can be readily generated through standard molecular biology techniques, and commercial suppliers offer many of the target moieties as standard DNA modifications. Because the tag ligates the DNA to a specific site on its surface, a single well-defined conjugate is formed, often efficiently and at high yield, in a single step. The tags are also orthogonal to one another, so they can be used consecutively to modify different locations on a DNA nanostructure with two different proteins (or three, if yet another orthogonal conjugation, such as biotin-streptavidin, is used).[46] Conversely, this method requires fusion with a full-length, folded protein that can be comparable to the

POI (see Figure 2G, which shows the size of a SNAP-tag relative to GFP). The self-ligating tag is also usually grafted onto the POI through a flexible amino acid linker. Thus, it is useful for applications that require single, site-specific attachment of DNA (e.g., displaying proteins on an origami scaffold) with minimal manipulation or purification of the conjugates. The method is less suitable for creating hybrid nano-structures (given that the tag itself takes up a lot of space), if rigid attachment is desired, or if the target protein must be fairly small, for example, to fit inside a DNA origami box.

Less Common Strategies for Protein-DNA Conjugation
Enzymatic Modification of Small-Molecule or Peptide Tags

A different approach to modifying proteins with DNA is the use of enzymes to ligate the two components. This strategy requires introducing a chemical handle—such as a small molecule or genetically fused peptide tag—that the enzyme can recognize to the POI. The other chemical handle is attached to the oligonucleotide, leading to selective linking of the two components (Figure 2H). A key advantage of this method is the inherent site specificity of a fusion peptide, as well as the lack of competing side reactions (like hydrolysis with NHS esters), because the target functional groups react only in the presence of the enzyme. Enzymes such as protein farnesyltransfer-ase (PFTase) can be used to graft modified isoprenoids with a new chemical handle to the short C-terminal peptide tag CVIA for subsequent secondary bioconjugation (e.g., copper click chemistry).[47] Another enzyme, sortase A, ligates an oligo-glycine peptide to the LPXTG sequence, although this strategy requires a second bio-conjugation reaction to link one of those peptides to DNA.[48] Additional enzymes that have been used to link peptide-tagged proteins to DNA include transglutami-nase[49] and methyltransferase.[50] Avoiding peptide tags altogether, Gothelf and coworkers used a terminal deoxynucleotidyl transferase (TdT) to ligate DNA strands with molecules bearing a pendant nucleotide triphosphate (NTP).[51] By first attaching the NTP to the molecule of interest (e.g., via click chemistry using an azide dCTP), the authors could efficiently attach DNA to peptides, polymers, dendrimers, or full-length proteins. Alternatively, relaxase enzymes ligate themselves to specific oligonucleotide sequences, which can be engineered into an origami scaffold;[52] the use of several different enzymes with differing sequence specificities imparts additional orthogonality to this approach. However, using this method to attach an arbitrary protein would require a fusion of the POI with the relaxase (which the authors demonstrated with fluorescent proteins). A similar approach was demon-strated with a fusion of a POI and the phi X174 Gene-A* protein, which covalently links a tyrosine residue in Gene-A* to a specific oligonucleotide sequence.[53] Expressed protein ligation—which usually links a peptide to a protein bearing an in-tein fusion—can also be used to directly ligate an oligonucleotide bearing a terminal cysteine (or a mimic thereof).[54]

Functionalization of Non-canonical Amino Acids with Bio-orthogonal Reactions

When none of the above approaches are suitable, a non-canonical amino acid (NCAA) can be introduced and functionalized with reactions that do not affect the native residues. For example, amino acids such as 4-azidophenylalanine can be installed by the Schultz amber-codon-suppression method[55] and targeted via "click" chemistry (catalyzed either by copper[56] or through a strain-promoted, "cop-per-free click;"[57] Figures 2I and 2J). DNA can be purchased with both terminal and internal azides or alkynes, usually with an intervening 6- or 12-carbon linker. Alternatively, modified phosphoramidites with a reduced linker length can be used in solid-phase DNA synthesis (Figure 2K), resulting in a more seamless transition between the oligonucleotide and the protein surface. The lack of reactivity

with native residues makes NCAA incorporation ideal for site selectivity or if two conjugation reactions are required. Furthermore, the relative inertness and resistance to hydrolysis of the reactive moieties allow for long-term storage, extended reactions times, and elevated temperatures. The primary downside to this method is that it requires additional plasmids for the non-canonical tRNA and tRNA synthetase necessary for incorporating the NCAA, and the expression yields are typically lower than with wild-type expression. Although click chemistry is by far the most common method for NCAA modification with DNA, the Francis lab has reported an attractive series of oxidative coupling reactions with residues such as 4-aminophenylalanine (Figure 2L).[58] These reactions proceed in minutes at mild aqueous conditions with oxidants such as sodium periodate or potassium ferricyanide, and their efficiency can rival that of click chemistry. Additional examples of NCAA-mediated DNA coupling include the reaction of aldehyde-containing proteins (e.g., introduction of 4-acetylphenylalanine)[59] with hydroxylamine- or hydrazine-modified DNA to form oximes or hydrazones, respectively (Figure 2M).[60]

Aptamers, Fusion Proteins, and Other Binding Agents

Aptamers—antibody-mimetic single-stranded DNA sequences that fold into a tertiary structure and bind to a target[61]—represent a particularly attractive strategy for immobilizing proteins on DNA scaffolds[62] because they can be easily introduced into a constituent strand of the structure (Figure 2N). Although aptamers bind non-covalently to proteins, they have the advantage of being inherently "site specific" as they target a unique interface on the protein. Although this strategy has not been extensively employed, fusing the POI with a protein that already has an aptamer would allow for attachment to a DNA scaffold. Compared with antibodies or other binding groups, aptamers often have modest K_d values ($\sim\mu$M). Their binding can, however, be enhanced through spatial control of two aptamers that bind to different interfaces of a protein, "clamping" proteins in a more rigid fashion.[63] It is also possible to covalently photo-crosslink an aptamer to a protein surface through reactive handles such as phenyl azides;[64] to our knowledge, this strategy has never been applied to DNA nanostructures, but it could represent an avenue for future research. As an alternative to aptamers, natural protein-protein interactions can be used to functionalize a target with DNA, as demonstrated by de Greef and coworkers with antibodies.[39] Protein G—which binds to the Fc region of antibodies and is more easily expressed, handled, and modified with DNA than a full-size IgG—was functionalized with a DNA handle and a benzophenone moiety, resulting in covalent trapping after UV irradiation (Figure 3B). This approach is particularly useful if, like for protein G, the binding interaction does not affect the function of the protein (antigen binding by the Fv region in this case). This last report was particularly interesting because it relied on site-specific incorporation of both the DNA and the benzophenone—the former through a unique cysteine residue and the latter via the Schulz method with a benzophenone NCAA—demonstrating the potential for multi-functional hybrid protein-DNA structures.

Fusing a DNA-binding protein—such as a zinc finger, leucine zipper, or transcription factor homeodomain—to the POI can also localize it on a DNA nanostructure[65] bearing the cognate DNA sequence to create, for example, enzyme cascades on a DNA origami[66] or to reconstitute functional ion channels.[67] The key drawback of this approach is the reversibility of binding, given that many DNA-binding proteins have K_d values in the nanomolar regime, comparable to the working concentrations for DNA origami. The Morii lab demonstrated an elegant workaround to this issue by using zinc-finger DNA-binding proteins to localize enzymes fused with SNAP-tags, CLIP-tags, and HaloTags to three different regions of a DNA origami modified

with the double-stranded DNA (dsDNA) targets.[40] In this fashion, the non-covalent (reversible) DNA binding enhanced the covalent (irreversible) linking of the proteins to the scaffold and enabled an over 90% yield of the modified structure to generate an enzymatic cascade (Figure 3C). Once again, we foresee that such "cooperative" approaches—using a non-covalent interaction to direct a covalent one—will be particularly useful for creating rigid interfaces with proteins and/or modifying them in more than one location. As an alternative to aptamers or DNA-binding domains, natural protein-biomolecule interactions can be used in a multivalent fashion to create a binding interface between the two components. In 2017, the Sleiman lab leveraged the affinity of human serum albumin (HSA) for lipids to bind a DNA nanocube by functionalizing it with multiple alkyl tails (Figures 3D and 3E).[41] They could tune the number of tails on the cube by modifying the constituent DNA strands, and a structure with four dendritic chains gave 5-fold stronger binding than a single chain. Instead of alkyl tails, multiple peptides that bind to a protein surface could be attached to a DNA structure to create a mimetic of a protein-protein interface. This approach was demonstrated with two different peptides that bind to different faces of the POI (templated on a linear duplex[68]) or multiple copies of the same peptide that binds to a multivalent protein (templated on a DNA origami[69]); we will discuss this more thoroughly in Controlling Protein Orientation on a DNA Scaffold below.

SEMINAL EXAMPLES OF INTEGRATED PROTEIN-DNA NANOTECHNOLOGY

In this section, we will describe four exciting areas where the integration of proteins and DNA nanoscaffolds has been used for creating novel structures. Most involve covalent protein-oligonucleotide conjugates, but several use affinity interactions instead. Many of these examples could fit into several sections, so we chose the topic where they demonstrated the greatest novelty and potential for applications.

Controlling Protein Orientation on a DNA Scaffold

Interestingly, the entire field of DNA nanotechnology was inspired by immobilizing proteins on a self-assembled DNA scaffold with a defined orientation. In 1982, Ned Seeman proposed using oligonucleotides as a structural material to create self-assembled 3D crystals through sticky-end cohesion and use these addressable frameworks to immobilize proteins in a periodic 3D array.[70,71] In this way, the structure of the guest molecule would be solved by X-ray crystallography without the laborious process of crystallizing the protein first. Although this goal has not been realized to date (and is indeed an exciting future direction for protein-DNA nanotechnology; see Structural Biology on Proteins Aided by DNA Scaffolds), the controlled positioning of proteins on DNA nanoscaffolds has continued to inspire the field ever since. Potential applications go far beyond structural biology, and indeed the other three topics covered in this review would all benefit from control between the protein orientation and the underlying nanoscaffold. For example, nanostructures that control enzyme function (e.g., latched DNA origami boxes in Using DNA Nanostructures to Dynamically Control Protein Function) will not function properly if the protein active site is occluded because it "points at" the cavity wall. Likewise, biological studies such as those described in Using Hybrid Protein-DNA Nanostructures to Answer Biological Questions must present the active portion of the protein (e.g., its binding interface) correctly to enable association with its receptor. Finally, the hybrid structures with both protein and DNA components as discussed in Building Nanostructures with Protein and DNA Structural Components must have a specific relationship between the two molecules to create

a well-defined final assembly. Future directions—such as protein-actuated DNA nanomachines—will likewise require careful integration of the two components. Even the enzyme-cascade examples not covered herein rely on the positional control of proteins with respect to one another to enhance substrate flow between them[22–24] and could benefit from orienting the respective pieces with greater precision. Finally, on a conceptual level, biology tightly controls the orientation and interaction of proteins with one another and other molecules, so investing in similarly precise synthetic protein-DNA nanostructures and devices will pay dividends in new and unexpected ways in the future.

In 2006, Turberfield and coworkers reported one of the earliest examples of a nanostructure-guided protein display by using the intrinsic helicity of the DNA duplex to control the relative placement of cytochrome c with respect to a tetrahedral cage.[72] Although the protein was conjugated to a thiolated DNA via its lysine residues in a non-specific fashion, systematically varying its attachment point on the duplex comprising the edge of the cage resulted in a smooth shift from the inside to the outside of the interior volume (Figure 4B). This change could be probed by gel electrophoresis and demonstrated an elegant mechanism for controlling the relative position of the two macromolecules. The Fan and Yan labs both recently imparted site specificity to the cytochrome c by using a mutagenically introduced cysteine and also confirmed the orientation (pointing in versus out of the cage) via cryo-electron microscopy (cryo-EM; Figure 4C).[73] By tethering the DNA cage to a gold substrate through three thiol-modified vertices, the authors could tune the orientation of the protein in relation to the surface, yielding a ~50% increase in electron-transfer rate for the cytochrome c inside the cage. This ability to regulate the spatial relationship between proteins and other functional components or interfaces will be particularly useful for creating nanosystems that control the flow of energy or matter in nanoscale factories or synthetic cells. In a related report, Fromme and coworkers demonstrated that PNA-modified proteins could be incorporated into a tetrahedral DNA cage (Figure 4D).[74] The authors did not explicitly control the protein orientation but rather found that, depending on its charge, it was either repelled from the cage (e.g., negatively charged azurin) or attracted to the anionic environment inside the cage (e.g., positively charged cytochrome c). This result raises the intriguing possibility of controlling the placement of proteins on a DNA scaffold by either manipulating their surface charge or fusing them with a highly charged partner as a "directing group" of sorts.

A key limitation of DNA nanostructures is that their resolution for molecular placement is generally limited to the minimal distance between handles attached to staple strands, typically ~3–5 nm. Self-assembling protein scaffolds such as viral capsids, by contrast, can position chemically appended molecules with a higher resolution (~0.5–1 nm).[3] Although the first example of viral capsids immobilized on DNA origami was reported in 2010 by the Yan and Francis groups,[33] in 2018 Wang and coworkers demonstrated that the relative orientation of self-assembling protein scaffolds could be controlled with DNA origami.[75] The authors used the tobacco mosaic virus (TMV), but rather than directly modifying the individual proteins with DNA, they assembled the capsid monomers around RNA strands in a similar fashion to native capsid formation around the RNA genome (Figure 4E). The authors could precisely control the exact number and length of TMV rods by attaching one or more RNA strands to defined locations on several different DNA origami templates. Highly anisotropic assemblies could be formed by this approach (Figures 4F and 4G), demonstrating the power of DNA nanotechnology as a programmable scaffold able to retain the chemical and self-assembly properties of the protein. Moreover,
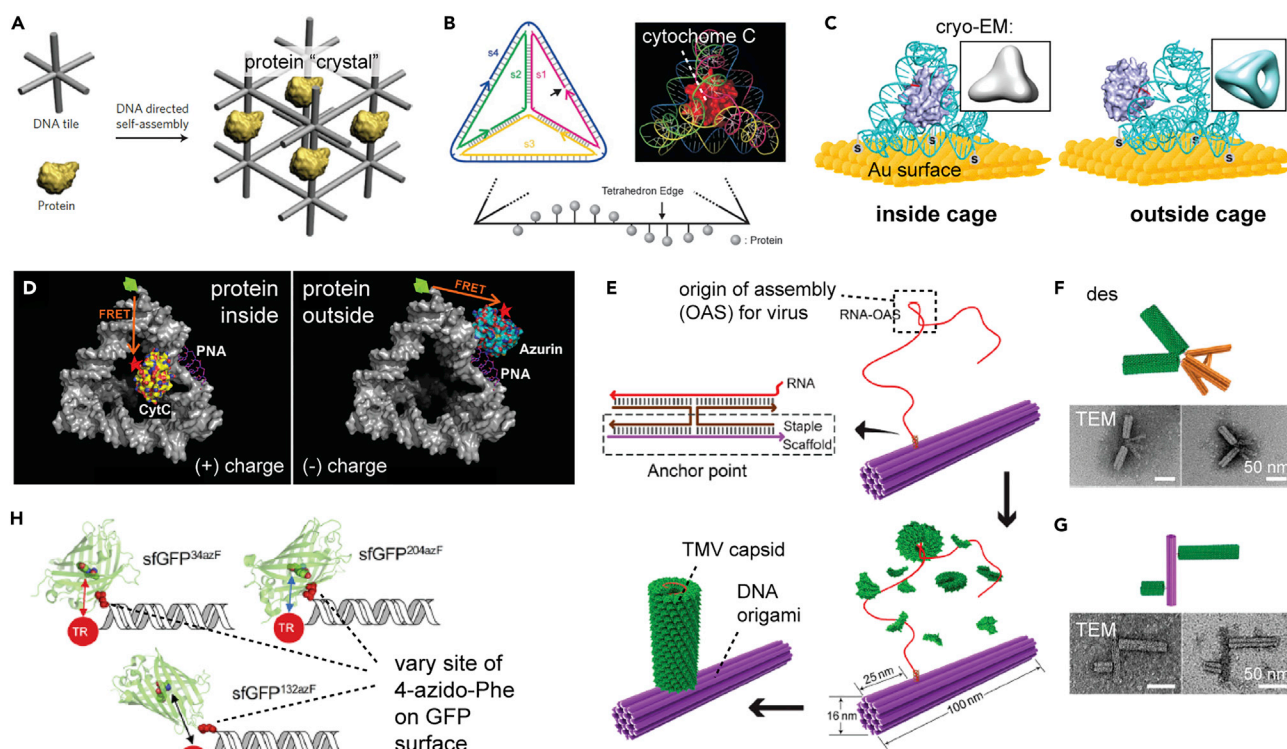
Figure 4. Controlling Protein Orientation with DNA Scaffolds

(A) The original conception of DNA nanotechnology as outlined by Ned Seeman: using a 3D self-assembled DNA array to artificially "crystallize" proteins by immobilizing them in fixed orientations. Reprinted from Pinheiro et al.[19] Copyright 2011 Springer Nature.

(B) Tuning orientation of cytochrome c relative to a DNA tetrahedron through the attachment site. Reprinted from Erben et al.[72] Copyright 2006 Wiley-VCH Verlag GmbH & Co. KGaA.

(C) Cytochrome c "inside" versus "outside" a DNA cage for tuning the electron-transfer rate to a gold surface (the insets show cryo-EM reconstructions of the protein-DNA cages). Reprinted with permission from Ge et al.[73] Copyright 2019 American Chemical Society.

(D) Cytochrome c and azurin adopt different conformations than a DNA cage as a result of differing net charges. Reprinted with permission from Flory et al.[74] Copyright 2014 American Chemical Society.

(E–G) Templating TMV capsids on a DNA origami scaffold through attachment of RNA scaffolds (E). Tuning the origami scaffold can precisely control both the spatial distribution (F) and length (G) of the capsids. Reprinted with permission from Zhou et al.[75] Copyright 2018 American Chemical Society.

(H) Varying the attachment of DNA on a protein surface through the incorporation of a site-specific NCAA. Reprinted with permission from Marth et al.[76] Copyright 2017 American Chemical Society.

the exact orientation of not just one molecule but an entire self-assembled protein scaffold was controlled with high precision. Further functionalizing the TMV monomers with bioactive molecules (e.g., peptides), polymers, or nanoparticles would yield unprecedented materials with hierarchical control over multiple length scales. DNA nanostructures can also be assembled reversibly through strand-displacement reactions, providing a temporal control not possible with many protein-based systems.

Although many researchers covered herein acknowledged the importance of site-specific protein-DNA chemistry, the Jones and Stulz groups set out to systematically probe this effect by modifying proteins with DNA in different locations and examining the relative effect on their function.[76] The authors chose superfolder GFP (sfGFP) and β-lactamase (BL) for their study, and to ensure site selectivity, they turned to copper-free click chemistry with the NCAA 4-azidophenylalanine. The NCAA was incorporated via the Schultz technique, and the cyclooctyne coupling partner was introduced into the DNA with a modified phosphoramidite. This

strategy allowed them to place the modification at any location on the protein surface without any of the restrictions imposed by fusion tags or cysteine residues. The authors first studied energy transfer (fluorescence resonance energy transfer [FRET]) between the sfGFP chromophore and a Texas Red dye attached to an oligonucleotide complementary to the protein-linked DNA handle. The efficiency of FRET could be tuned from ~90% to 75% depending on the modification site on the protein as a result of differing distances between the donor and acceptor dyes (Figure 4H). DNA was also attached to several different locations on the BL surface, and the catalytic efficiency of the protein attached to a DNA origami surface was probed. Both the site of modification and the orientation relative to the origami (pointing "up" versus "down") affected the catalytic rate, allowing for up to 30-fold enhancement relative to bulk solution. Thus, this work showed the role of protein orientation on a DNA scaffold and the importance of using a site-specific bioconjugation strategy. We also note that the cyclooctyne used (appended directly to the 5′ end of the DNA) and the use of a NCAA on the protein resulted in a short linker between the molecules and thus allowed for tight coupling between the two components.

One potentially transformational area for protein-DNA nanotechnology—which harkens back to the original motivation for the field—is the use of origami scaffolds to aid in cryo-EM characterization of proteins. Using cryo-EM for single-particle reconstruction of protein structure requires imaging tens to hundreds of thousands of individual images of the target molecule, classifying them into different orientations, averaging the electron density for each of these orientations, and then combining all the images into a 3D reconstruction of the protein. DNA origami scaffolds (which are large and easily visualized by cryo-EM) are ideal candidates for "nanoscale sample holders" to control the orientation of the protein, prevent its adsorption and denaturation at the air-water interface of the sample, and help find small particles with a low signal-to-noise ratio (see Structural Biology on Proteins Aided by DNA Scaffolds). In 2016, the Dietz and Scheres groups demonstrated this principle by using a barrel-like origami to immobilize the transcription factor p53.[77] The use of a DNA-binding protein circumvented the need for bioconjugation, and the researchers could control the orientation of the protein by changing the location of the binding site on a DNA duplex spanning the origami barrel. By relying on the helical nature of DNA, the authors could tune the exact orientation of the protein by using the origami as a "nanoscale goniometer" (Figure 5A). The DNA nanostructure protected the protein from adsorption to the air-water interface (which could denature it or bias its orientation), controlled the thickness of the ice, and provided a reference mask for selecting particles and sorting them into different classes. The authors were able to obtain a structural solution for the protein to ~15 Å resolution, which was not sufficient for atomic-scale information but did yield some new insights into the way the protein bound DNA. A number of factors prevented a higher-resolution structure, but chief among them was the lack of a sufficiently rigid and well-defined protein-DNA interface. The authors also explicitly mentioned that, although their method was amenable to DNA (or RNA) binding proteins, extending it to arbitrary proteins would require "chemical modifications of some of the DNA staples within the support with a specific tag on the target protein." In Structural Biology on Proteins Aided by DNA Scaffolds, we discuss potential approaches for addressing this exact issue for both cryo-EM and X-ray crystallography applications.

In 2018, Mao and coworkers demonstrated a different approach to protein-DNA cryo-EM studies by using a DNA nanobarrel to immobilize the membrane protein α-hemolysin (Figures 5B and 5C).[78] The authors used the DNA scaffold to attach lipids modified with complementary handles, creating a hydrophobic milieu that
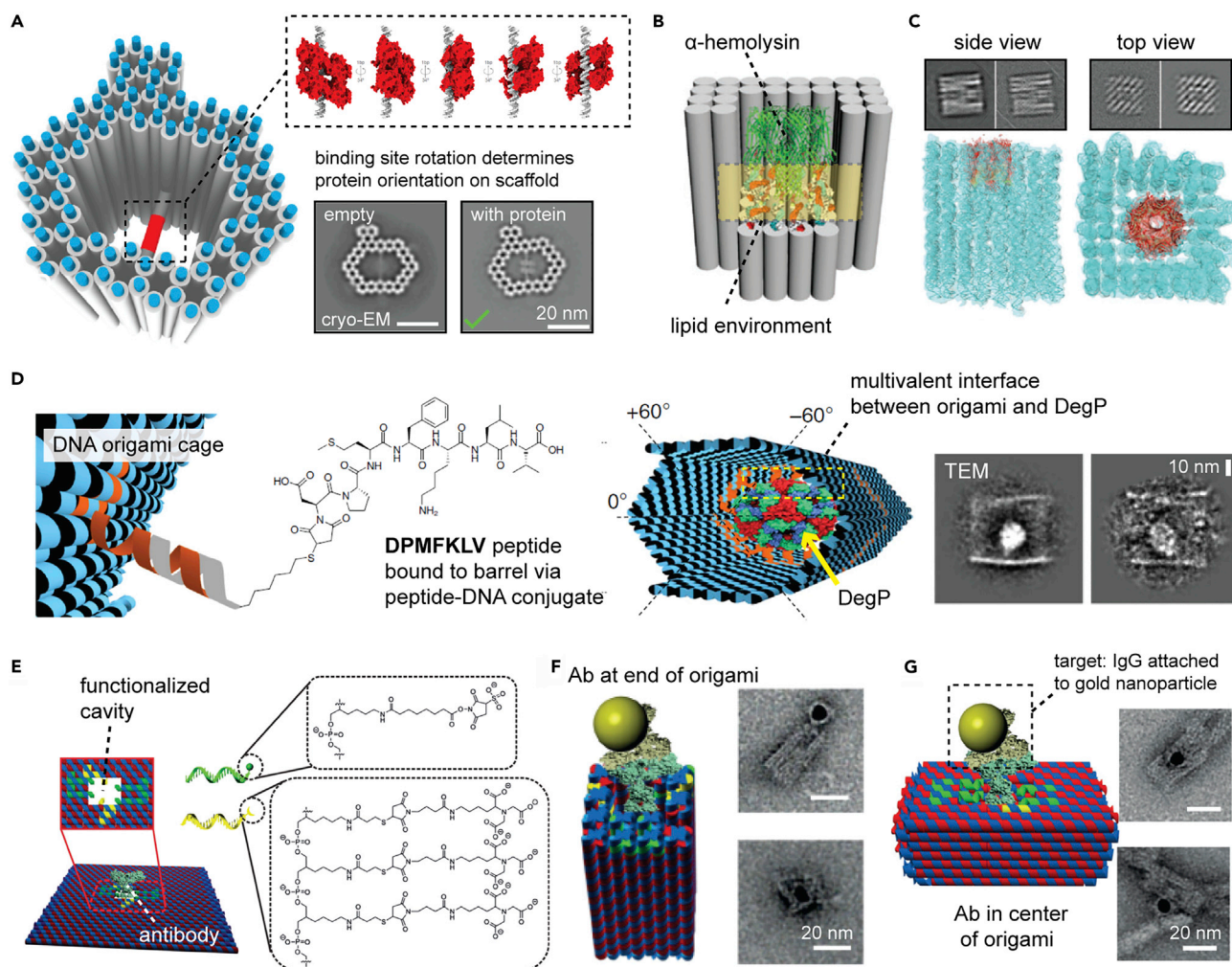
Figure 5. Controlled Attachment of Proteins to DNA Cages

(A) Structure of a DNA origami cage with a binding site for the transcription factor p53. The orientation of the protein can be tuned by the relative helical rotation of the binding site. Compared with empty cages, the protein can be clearly seen in cryo-EM images. Reprinted with permission from Martin et al.[77]

(B and C) DNA origami nanobarrel for immobilizing α-hemolysin through the attachment of lipids to the barrel interior (B) and cryo-EM images and electron density maps of the top and side views of the protein-filled nanobarrel (C). Reprinted with permission from Mao et al.[78] Copyright 2018 Wiley-VCH Verlag GmbH & Co. KGaA.

(D) Hexagonal DNA origami that "wraps" the multivalent protein DegP through multiple peptide-protein interactions scaffolded by the cage. Reprinted from Sprengel et al.,[69] licensed under CC BY 4.0.

(E–G) Attachment of antibodies into a DNA origami cavity via Ni(NTA)-mediated anchoring and NHS ester chemistry (E). The antibodies can be anchored in the end (F) or middle (G) of a 3D origami structure; the functionality of antibody binding is retained, as probed by binding to a gold-nanoparticle-modified Fc target. Reprinted with permission from Ouyang et al.[79] Copyright 2017 Wiley-VCH Verlag GmbH & Co. KGaA.

allowed anchoring of the protein in a local environment mimicking a lipid nanodisc. Key to integrating the two molecules—all while preventing aggregation due to their hydrophobic nature—was first stabilizing the protein with detergent, which could be slowly dialyzed away to promote insertion of the protein into the hydrophobic nano-structure cavity. After single-particle 2D class averaging, the structure of the DNA barrel could be visualized at 7.5 Å resolution, allowing fitting of the DNA backbone. The protein, because it lacked specific orientational control, could be resolved to only around 30 Å (and then only after application of its intrinsic C7 symmetry). How-ever, combining this method with additional bioconjugation techniques to "pin" the

protein more tightly and prevent rotation in the barrel could potentially enhance the resolution and help determine the structures of membrane proteins, for which crystallization is notoriously difficult. The DNA scaffold can also be readily tuned to match differently sized membrane proteins, an advantage not possible with systems such as nanodiscs or micelles. Indeed, using DNA origami to control liposome size and shape[80] would be a natural way to further immobilize membrane proteins.

In biological systems, the relative orientation of two interacting proteins is controlled by their binding interface, which is in turn dictated by the spatial arrangement of supramolecular interactions (e.g., hydrogen bonds, salt bridges, or hydrophobic packing). In 2017, the Sacca lab demonstrated that a similar interface could be generated between a DNA origami structure and the multivalent protease DegP—which can exist in oligomeric states ranging from 6 to 24 monomers—through the attachment of multiple short binding peptides (Figure 5D).[69] The peptides (sequence: DPMFKV) were attached to thiolated DNA handles through an N-terminal maleimide, allowing up to 18 peptides to be attached to a DNA origami structure composed of hinged sheets, which could in turn "wrap" the DegP target through multiple peptide-protein interactions. Key to this work was the controlled positioning of these peptides (driven in turn by the helicity of DNA and the exact attachment site); pointing the peptides "outward" dramatically reduced protein binding. The number of peptides on the structure could be controlled such that more peptides yielded tighter binding, effectively converting a rather weak individual peptide-protein interaction (K_d ~5 μM) into a tight, multivalent interface. Interestingly, although the size of the barrel was a better match for the 24-mer protein, the 12-mer DegP bound almost twice as well as the smaller or larger oligomers. The authors surmised that this was due to balancing the size match and energy of binding (which increases with larger assemblies) with the ease of diffusion into the cage (which increases with smaller assemblies). Extending this concept to other targets, especially monovalent proteins—by attaching several different peptides (or other binding agents) that each bind to a different face of the protein—would enable synthetic antibodies that can take on arbitrary shapes, sizes, and additional functionalization. Alternatively, generating a binding interface on a DNA scaffold could help rigidly pin down a target protein, which will be critical for structural biology studies.

One of the most impressive examples of controlling protein orientation—integrating multiple site-specific conjugation approaches with the judicious design of a DNA nanostructure—was reported by the Gothelf lab[79] and built off prior work using polyhistidine-Ni(NTA) interactions to direct an NHS ester conjugation to a specific lysine residue on an antibody (as in Figure 3A).[38] Origami scaffolds (both 2D rectangles and 3D cages) were designed with holes large enough to fit the Fc region of a mouse IgG antibody, flanked by two handles for (NTA)-functionalized DNA handles. Four additional handles for complementary strands bearing NHS esters were added to the other two sides of the cavity in order to covalently trap an antibody after binding to the (NTA) moieties via Fc histidine clusters (Figure 5E). The efficiency of IgG crosslinking to the origami (which the authors determined by adding the nickel chelator EDTA and probing for how many antibodies remained tethered to the structures) increased with the number of NHS esters until it peaked at over 90% for the constructs with four handles. Through both chemical conjugation and steric trapping, the orientation of the antibody could be controlled precisely, and the authors demonstrated two 3D origami objects with the IgG protruding from either the center or the end of the nanostructures (Figures 5F and 5G). The yield for these more complex objects was lower (~50%) but still impressive given the degree of control of a

large biomolecular complex on the nanostructure. Finally, the functionality of the antibody (antigen binding) was retained, highlighting the advantage of this approach over non-specific mechanisms that could occlude the Fv region.

Using DNA Nanostructures to Dynamically Control Protein Function

A second key area of application for protein-DNA nanotechnology is in modulating protein function with a DNA scaffold either by hiding the protein function in a cage to block its activity or by spatially restricting other co-reagents. One of the seminal examples in this field was reported in 2012 by the Church group, who developed a "nanorobot" that could control protein function in a logic-gated, stimulus-responsive manner.[14] The authors designed a hexagonal DNA origami cage mimicking a clamshell, whose top and bottom halves were held together with two DNA "locks" (Figure 6A). Placing a protein inside the DNA structure prior to closing it blocked its activity by sterically isolating it inside the origami cage. In order to render the cage stimulus responsive, the two locks holding it closed consisted of aptamers for a specific target; upon exposure to a protein "key," the aptamer-target interaction would outcompete hybridization and open the lock. Using aptamers for two different targets on the cage resulted in an AND logic gate, whereby the cage would open only when both targets were present (e.g., on a cell surface), opening the clamshell and exposing the protein cargo. Six different cells lines, expressing different combinations of three possible aptamers, were used for demonstrating the function of the system; the correct lock combination exposed an antibody fragment to human leukocyte antigen-A/B/C. The system could selectively label a target cell even in the presence of non-target cells, and various aspects of cell signaling could be modulated with antibodies that activated specific intracellular pathways. In this case, the "protein function" controlled by the origami was antibody binding, but this principle could apply to virtually any protein whose function can be blocked by the cage. Although many other targeted drug-delivery systems exist, the work by Church and colleagues was the first to demonstrate a programmable container that could be opened in a "smart" fashion.

In 2018, a collaborative team (led by the Nie, Yan, Ding, and Zhao labs) extended this "nanorobot doctor" concept to an *in vivo* application for treating tumor cells.[81] A rolled-up origami sheet locked with aptamers against nucleolin—a marker of tumor vasculature—was used to block the action of thrombin, a critical protein in blood coagulation (Figure 6B). The aptamers targeted the robot to a human breast cancer tumor, whereupon the rolled-up sheet opened to expose thrombin, resulting in localized clotting that killed the cancer cells. The aptamer served as both a targeting element and a functional "lock" for selective actuation of the structure, and the authors showed effective tumor necrosis *in vivo* in both mice and miniature pigs. The structures also did not raise an immunological response, as measured by cytokine production. One key limitation to these approaches is the requirement for aptamers that bind to a target of interest; extending this concept to protein-based locking mechanisms (e.g., antibodies, which exist for a wide range of targets) is an exciting future direction for protein-DNA nanotechnology. We also foresee great opportunities for combining this approach with protein- or polymer-based coatings to stabilize the nanobots to degradation and enhance their targeting to the desired cells.

Aside from targeted killing of diseased cells, reversible DNA cages provide a powerful way to switch protein activity on and off, allowing for dynamic control of enzyme function. In 2017, Andersen and coworkers used a DNA "nanovault" to reversibly expose and occlude a protein in order to control access to its substrate.[82] The
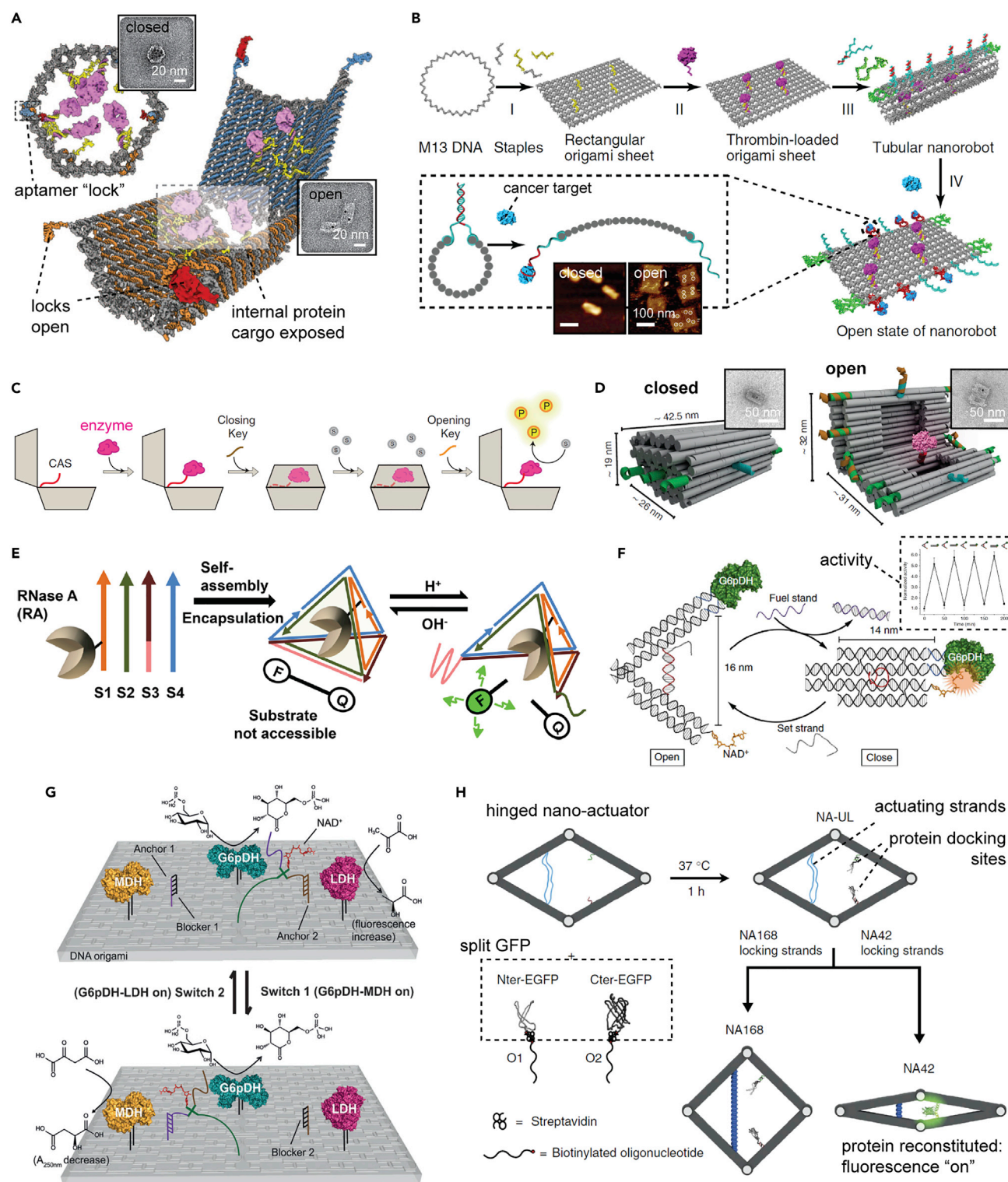
Figure 6. Dynamic Control of Protein Function with DNA Nanostructures

(A and B) DNA origami cages with clamshell (A) or rolled-up sheet (B) morphologies that expose proteins upon opening of aptamer "locks" by target binding. Reprinted with permission from Douglas et al.[14] (copyright 2012 AAAS) and Li et al.[81] (copyright 2018 Sprinter Nature).

(C and D) Turning enzyme activity "on" and "off" through the reversible opening of a DNA origami nanovault (C) and open and closed states (with TEM images) of the nanovault (D). Reprinted from Grossi et al.,[82] licensed under CC BY 4.0.

Figure 6. Continued

(E) Reversible control of RNase A with a pH-responsive DNA tetrahedron cage. Reprinted with permission from Zhou et al.[83] Copyright 2018 American Chemical Society.

(F) Controlling enzyme activity by reversibly changing the distance between the protein and its cofactor with a DNA tweezer. Reprinted from Liu et al.[84] Copyright 2013 Springer Nature.

(G) Switching between two different enzymatic cascades by spatial control of the cofactor-tethered strand. Reprinted from Yan et al.[85] Copyright 2016 Wiley-VCH Verlag GmbH & Co. KGaA.

(H) A hinged "nanoactuator" for dynamic reconstitution of the EGFP protein. Reprinted from Ke et al.,[86] licensed under CC BY 4.0.

authors designed the nanovault to fully encapsulate the enzyme α-chymotrypsin and could open and close it with DNA strands in order to control its cleavage of casein, a target protein (Figures 6C and 6D). The authors conducted a number of elegant experiments to probe the accessibility of the protein in the DNA cage and had to engineer several additional features into the structure to mitigate the inherent (and non-negligible) porosity of the DNA origami cage. In the end, they were able to enhance the activity of the enzyme by roughly 3-fold between open and closed states. Interestingly, the authors found that the best way to conjugate the enzyme to DNA was via copper click chemistry with an azide-functionalized protein (itself made through non-selective lysine chemistry with an azide-NHS ester) with an alkyne strand already incorporated into the open cage. Attempting to first modify the protein with a DNA handle and then attach it to the cage gave a lower yield of encapsulation overall. Using a more selective bioconjugation method (to ensure that the enzyme is not attached in a way that occludes its active site) could enhance its function in the future. Alternative locking mechanisms for DNA origami boxes—such as pH-switchable locking mechanisms,[87] photocleavable linkers between the structure and the POI, or photoswitchable protein latches (see Protein-Actuated DNA Nanomachines)—could further enhance dynamic protein control via this strategy. Around the same time as the Andersen nanovault work, the Kim lab reported a pH-responsive tetrahedral DNA cage that could reversibly control the activity of RNase A.[83] Their cage, composed of only four strands, was a much simpler structure than the nanovault and functioned through the reversible assembly and disassembly of one side via a pH-switchable i-motif (Figure 6E). The authors built off Turberfield's previous work[72] (Figure 4B) to ensure that the RNase A was located inside the cage, and conjugation was achieved by copper-free click chemistry after protein modification with a cyclooctyne-NHS conjugate. Encapsulation protected the enzyme from both degradation by proteases and binding to antibodies, and its activity could be switched by roughly 2-fold between the open and closed states. If the enzyme was positioned outside the cage, by contrast, no such modulation was seen, demonstrating the importance of protein orientational control for creating functional nanostructures.

Rather than expose or occlude a protein's active site, a second mechanism for switching enzymes on and off is by controlling the availability of reagents or other cofactors necessary for catalysis. In 2013, the Liu lab demonstrated the reversible regulation of an enzyme cascade by using a DNA "nanotweezer" to control the distance between the two proteins.[88] The tweezer was driven between two states with strand-displacement mechanisms, which in turn modestly affected the efficiency of intermediate transfer between the two enzymes. That same year, the Yan and Fu labs used the same tweezer structure to control the availability of an NAD^+ cofactor to the glucose-6-phosphate dehydrogenase (G6pDH) enzyme (Figure 6F).[84] Similar to Liu's work, this DNA nanomachine effectively controlled protein function in a reversible fashion, enhancing catalytic activity by ~5-fold in the "closed" versus "open" states. Shortly thereafter, Yan and coworkers reported a similar mechanism for guiding protein function: tethering $NAD^+/NADH$ on a "swinging arm" in order to

shuttle electrons between G6pDH and malic dehydrogenase (MDH) with nanoscale control.[89] The authors demonstrated both the dependence of the enzymatic rate on the distance from the swinging arm and its relative orientation to the proteins (i.e., pointing "towards" versus "away"), as controlled by the helicity of DNA. In this case, the proteins were tethered to the scaffolds by non-specific lysine chemistry, so even greater control might be possible through their careful positioning with active sites oriented toward or away from the cofactor. In two follow-up works, the Yan and Yang labs imbued these systems with greater control over protein function by using light to reversibly tether the cofactor-laden swinging arm away from the enzymes[90] or controlling its relative position between two different sets of enzymes (Figure 6G).[85] Taking a different tack, Ke, Bellot, and coworkers used a DNA origami "nanoactuator" to control the distance between two halves of a split GFP; mechanically bringing them into close proximity with DNA locking strands reconstituted the protein and turned fluorescence on (Figure 6H).[86] Using DNA nanomachines to control protein function in these ways provides a powerful way to build nanoscale "chemical plants" (or synthetic cells). Such systems can also be used to more precisely probe protein function in biological contexts, a topic to which we turn next.

Using Hybrid Protein-DNA Nanostructures to Answer Biological Questions

One of the most promising applications of DNA nanotechnology in the past decade has been using structures to answer questions of biological importance. In cells, the nanoscale distribution of proteins is critical to their function, as are their oligomerization state and the forces they apply (or that are applied upon them). DNA origami constructs are particularly good at controlling the spacing and stoichiometry of biological ligands, as well as exerting tunable biophysical forces on proteins, allowing researchers to probe systems with a precision not possible with other approaches. In conjunction with single-molecule imaging techniques, these hybrid nanostructures have opened up a new frontier in science that will undoubtedly spread to diverse subfields of biology. Early experiments with DNA tile arrays modified with peptides demonstrated that antibodies could be patterned with nanoscale accuracy by binding to their antigen,[91] highlighting the potential for single-molecule measurements via atomic force microscopy (AFM). However, it was not until the adoption of the origami approach[10] that DNA nanostructures began to be used widely for advanced biological experiments. We note that in this section we only discuss biological studies using pre-formed hybrid protein-DNA nanostructures. Examples where a DNA nanostructure alone was used for visualizing or probing the function of a protein acting upon it—such as the elegant work by the Endo and Sugiyama labs on DNA-manipulating enzymes[92–94]—will not be covered.

From its inception, one of the key applications of DNA origami has been as a "molecular breadboard" capable of controlling the positioning of other species with ~5 nm resolution. In cells, the nanoscale presentation of extracellular ligands can cluster their receptors, leading to activation of intracellular pathways. Thus, in the past 5 years there has been an explosion of interest in using DNA-scaffolded proteins to probe the distance and valence dependence of protein presentation on cellular behavior. In 2014, the Hogberg and Texiera labs used a "nanocaliper" to present two (or more) copies of ephrin A5 in order to probe the optimal distance for activation of the EphA2 receptor (Figure 7A).[95] Although other approaches for clustering the EphA2 receptor had been developed for probing its signaling (which is often disrupted in cancer), none was able to control the distance between exactly two ligands in a tunable fashion. The ephrin A5 ligand was conjugated to amine-labeled DNA with a bifunctional linker and bound to complementary handles on a rod-like origami structure. The distance between proteins was set at either 40 or
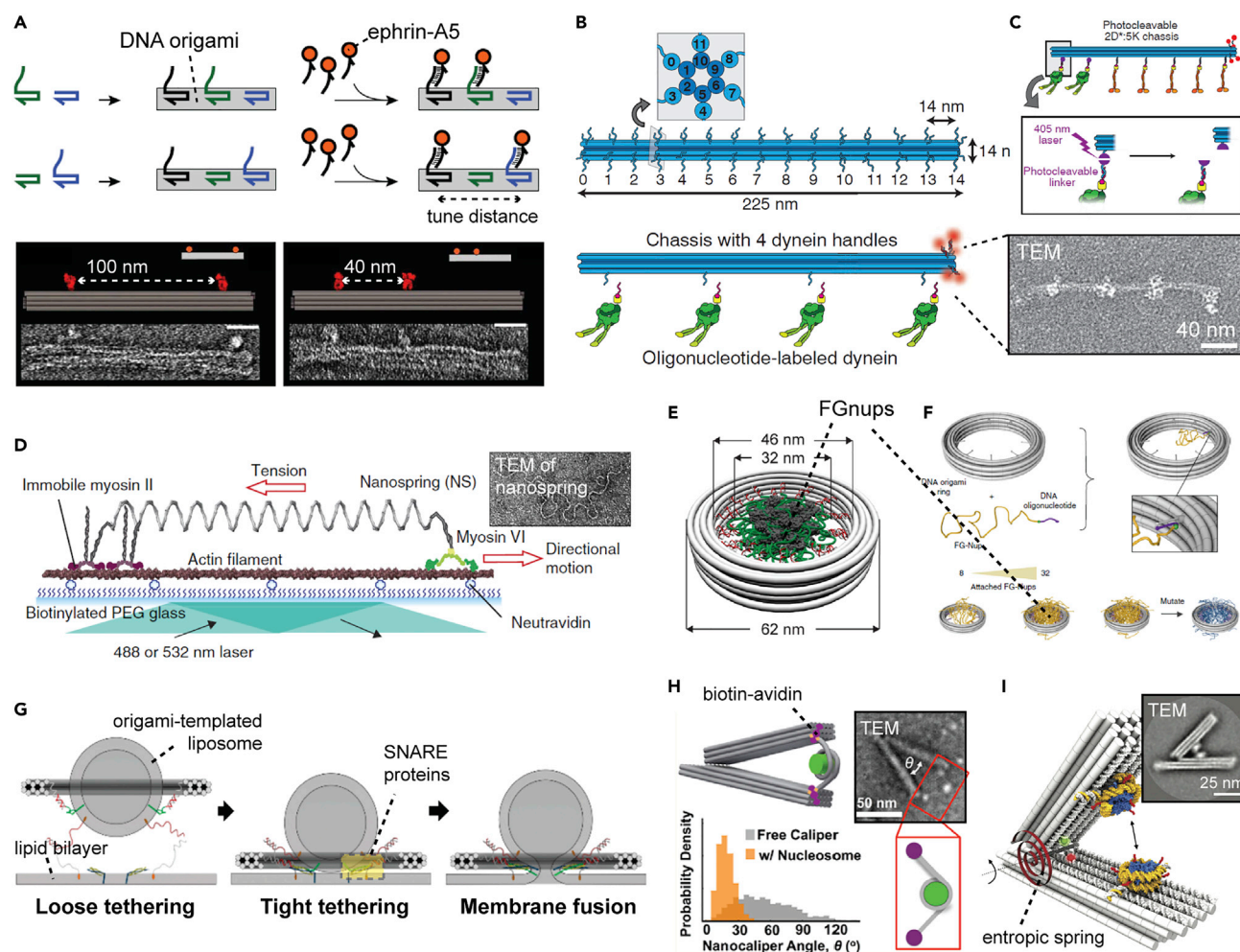
Figure 7. Probing Biological Questions with Protein-DNA Nanostructures

(A) DNA origami "nanocaliper" for controlling the distance between proteins. Reprinted with permission from Shaw et al.[95] Copyright 2014 Springer Nature.

(B and C) DNA origami rods for tethering fixed numbers of dynein and kinesin motor proteins (B). Photocleavable linkers allow for selective release of dynein motors while keeping kinesin motors attached (C). Reprinted with permission from Derr et al.[98] Copyright 2012 AAAS.

(D) DNA origami nanospring for measuring the force applied by a myosin motor. Reprinted from Iwaki et al.,[99] licensed under CC BY 4.0.

(E and F) Two similar designs of origami ring structures for assembling nuclear pore proteins (FG-nups) in a confined volume. Reprinted with permission from Fisher et al.[100] (copyright 2018 American Chemical Society) and Ketterer et al.[101] (licensed under CC BY 4.0).

(G) Probing the effect of SNARE proteins on liposome fusion with a membrane by using DNA origami rings to tether the liposomes and control the number of SNARE proteins. Reprinted with permission from Xu et al.[102] Copyright 2016 American Chemical Society.

(H and I) DNA origami nanocaplipers for measuring the interaction between nucleosomes and DNA (H) or between two nucleosomes (I). Reprinted with permission from Le et al.[103] (copyright 2016 American Chemical Society) and Funke et al.[104] (licensed under CC BY 4.0).

100 nm; structures with only a single ligand (which should not be able to dimerize the receptor) were used as controls. The multivalent structures bound more tightly to EphA2 (as measured by surface plasmon resonance), but only the origami with 40 nm spacing produced more receptor phosphorylation and downstream effects on breast cancer cells than monomeric ligands. Interestingly, the 100 nm spacing was identical to the origami bearing a single ligand, and increasing the density of ligands to eight (spaced 14 nm apart) did not yield an increase over the dimers spaced 40 nm apart. Such precise control of both the number and distance between ligands is not possible with any other system, especially when rigidity must be maintained over long (e.g., tens of nanometers) distances. In 2015, the Hogberg lab

extended this nanocaliper concept to probe the binding of antibodies to antigens immobilized with a tunable distance on origami.[96] With this system, the optimal distance between ligands was determined to be 16 nm, and the precise effect of antibody type and linker length or flexibility could be probed as well. Also in 2015, the Niemeyer lab used microarrays modified with single-stranded DNA (ssDNA) handles to immobilize several different rectangular origami scaffolds, each of which displayed a different nanoscale pattern of epidermal growth factor (EGF) ligands.[97] This hierarchical approach, which the authors termed "multiscale origami structures as interface for cells" (MOSAIC), allowed for control of EGF density and spacing at the nanometer length scale and top-down patterning of different origami scaffolds at the micron length scale. Cells were then adhered to these surfaces, enabling detailed studies of ligand distributions on bioactivity in a way not possible with other surface immobilization techniques such as supported lipid bilayers or direct surface grafting.

We next discuss three specific fields where protein-DNA nanoassemblies have been used for probing biology: (1) the collective action of motor proteins, (2) density-dependent function of confined proteins, and (3) nucleosome assembly. In 2012, Reck-Peterson and coworkers demonstrated that a rigid DNA origami bundle could be used for attaching the molecular motors dynein and kinesin-1, which walk along microtubules but with opposite polarities.[98] By using a SNAP-tag fusion, the authors could precisely pattern both the number and distribution of these two proteins on the origami "chassis," allowing for unprecedented analysis of their movement on myosin at the single-molecule level (Figures 7B and 7C). In particular, the authors could probe the "tug of war" between these two oppositely oriented motors and see which one dominated as a result of differences in affinity and stall force; attaching one protein type with a photocleavable linker enabled their dynamic release and dominance of the other motor type. Without an addressable scaffold such as DNA origami, it would not be possible to create such controlled protein assemblies. We also note that the size of DNA origami (approximately tens to hundreds of nanometers) is ideally suited for positioning multiple proteins (approximately one to tens of nanometers in size) without interference; smaller nanoscaffolds would be hard pressed to retain such precision. In 2015, Sivaramakrishnan and coworkers used DNA origami scaffolds to attach two different motor proteins, myosin V and myosin VI, with controlled spacing and number.[105] In this way, they could probe the collective action of multiple motors but on a defined system that avoided the complexities of natural actin-myosin ensembles (for example, in muscle fibers). The spacing between motors could be tuned (14, 28, or 42 nm) to match the spacing of various natural filaments. The authors found that neither myosin density nor the number affected the gliding speed of the origami on surfaces coated with actin filaments, which they attributed to the ensemble of motors acting as an "energy reservoir" that allowed them to function more consistently over a larger range of loads. One year later, the Iwaki lab reported a DNA origami as not merely a scaffold but rather a "nanospring" force sensor for myosin V and VI.[99] A coiled nanostructure—with a precise force-extension curve determined with optical tweezers—was attached to motor proteins bound to actin filaments, and the extension of the spring was used as an output for the force exerted on it (Figure 7D). The nanostructure matched the stall force of the motor proteins but over a much shorter distance than with dsDNA, demonstrating the utility of origami. The authors were able to not only probe the tug of war between myosin V and VI but also demonstrate that myosin VI switched between hand-over-hand- and inchworm-type motions depending on the force exerted on it. The origami-based experiment was also more tractable

than traditional approaches using optical tweezers and allowed for precise protein patterning not possible with other methods.

The second key area where protein-DNA nanostructures have played a key role is in determining the function of proteins under nanoscale confinement. DNA origami scaffolds are ideal for creating a defined volume, and the exact number of molecules entrapped therein can be tuned through protein-DNA conjugates that orient the proteins into that volume. In 2018, two reports—one by the Lin and Lusk labs[100] and the other by the Dietz and Dekker labs[101]—used DNA origami nanorings to assemble the proteins that make up the nuclear pore complex (Figures 7E and 7F). Both studies probed the function of FG-nups, disordered repeat proteins rich in phenylalanine-glycine ("FG") residues, and which control the flow of molecules such as transcription factors or ribosome components into the cell nucleus. However, the complexity of the nuclear pore, combined with the unstructured nature of the FG-nup proteins, has made studying their properties (e.g., how they selectively control the flow of different molecules into the nucleus) difficult. The two studies discussed were able to incorporate 32–48 copies of the protein in an inward-directed fashion (through site-specific cysteine chemistry) and probed their assembly by using AFM, transmission electron microscopy (TEM), cryo-EM, and molecular modeling approaches. The dynamics of protein occlusion of the ring could be visualized, and control experiments with outwardly oriented proteins or mutants bearing hydrophilic residues demonstrated that both the geometry and chemical composition of the proteins are important for blocking the nuclear pore. Critically, the DNA origami rings provided a scaffold highly similar to the dimensions and shape of the nuclear pore, allowing the FG-nups to assemble in a biomimetic fashion not possible with other templates. These systems also open up the possibility of probing the function of multiple different FG-nups or creating assemblies with multiple types of proteins (as is seen in the native nuclear pore) in order to probe the effect of protein architecture on function and selectivity of molecular transport.

Although the above works investigated nuclear pore proteins, DNA origami provides an attractive platform for nucleating other protein assemblies as well. In 2016, the Shih and Rothman labs used ring-like origami to template SNARE proteins,[102] which drive membrane fusion in processes such as vesicle and neurotransmitter transport. Building off the work by Lin and Shih to template liposomes with DNA origami,[80] the nanoring scaffolds could simultaneously encapsulate a spherical liposome and a defined number of SNARE proteins through programmable DNA handles (Figure 7G). Additional DNA handles could be used to dock the origami-encircled liposomes with a lipid bilayer, allowing for a detailed study of membrane fusion as a function of SNARE number. By decoupling the docking and fusion steps, the authors showed that only one or two SNARE proteins were necessary for the process, resolving an outstanding debate in the field. It is particularly important to highlight that this system combined several key aspects of protein-DNA nanotechnology: (1) controlled orientation of SNARE proteins, (2) assembly of a defined number of proteins with controlled spacing, and (3) integration of proteins with other molecules, such as lipids, in a highly biomimetic fashion. Aside from SNARE proteins, in 2019 the Fan and Zhong labs demonstrated that DNA-templated CsgB proteins could be used to nucleate bacterial curli nanofibers by polymerizing CsgA proteins.[106] This approach was highly mimetic of natural fiber nucleation and growth from the surface of *E. coli* cells, and using the origami nucleator as an easily visualized "molecular landmark" allowed the growth kinetics of the fibers to be determined by high-speed AFM.

A third recent subfield in biology where DNA nanostructures have found applications for single-molecule biophysics is in probing nucleosome assembly and forces. Three reports in 2016—one by the Poirier and Castro groups[103] and two by the Dietz lab[104,107]—used DNA origami hinge-like structures to investigate either the wrapping of dsDNA by nucleosomes (Figure 7H) or the interaction force of two nucleosomes as they were brought into contact (Figure 7I). In all three cases, the origami served as an easily visualized output for these supramolecular interactions—a lever-like "amplifier" of a much smaller-scale molecular interaction—through direct imaging (via TEM) or indirect methods (such as FRET between two dyes attached to the devices). The effect of parameters such as DNA length, salt concentration, nucleosome number, or transcription factor binding on nucleosome wrapping (or the effect of histone acetylation on the interaction forces between nucleosomes) could be probed with unprecedented precision. These approaches all relied on a thorough understanding of the biophysical properties of the hinged origami structures, which are governed by electrostatic and entropic spring effects, and how these properties relate to the force applied to their ends. However, the researchers demonstrated that DNA origami constructs are uniquely suited for single-molecule experiments in that they combine rigidity, ease of modification at precise locations, and multiple possible output modes. Future experiments that directly apply tunable forces at multiple points on a protein's surface (e.g., the "Bohr-radius" resolution tweezer developed by the Dietz lab)[108] would allow for precise unfolding experiments akin to optical tweezers but with a much simpler setup. We hasten to add that for monomeric proteins, accomplishing this goal will require modification of the protein in two distinct locations with high site specificity, short linkers, and defined rigidity (as discussed in Controlled and Rigid Orientation of Proteins on DNA Nanoscaffolds).

Building Nanostructures with Protein and DNA Structural Components

All of the examples presented in the previous three sections employed a pre-formed DNA scaffold upon which proteins were arrayed for either probing or controlling their function. In this section, we turn to a conceptually distinct area of protein-DNA nanotechnology: integrating proteins and DNA into nanostructures that contain both molecules as *structural* components. We focus specifically on assemblies where each plays a unique role in the final assembly and could serve as a scaffold for other materials of molecules. Although such structures present new challenges—namely balancing the self-assembly of two different macromolecules, each with their individual requirements and physicochemical behavior—they also offer several distinct advantages. First and foremost, proteins possess a wide range of unique structural motifs that DNA does not—such as α helices, coiled coils, β sheets, and collagen triple helices—with varying mechanical properties and nanoscale display of chemical functionality. Second, proteins have the potential for functionality ranging from catalysis to protein binding to stimulus-responsive behavior. Third, the chemical diversity of the 20 canonical residues (and dozens of reported non-canonical amino acids)[55] opens up novel chemical attributes beyond the uniformly anionic phosphate backbone of DNA. Fourth, proteins are a potentially "higher-resolution" scaffold than DNA, allowing functional groups to be positioned in closer proximity; for example, a typical α helix has a pitch of 0.54 nm, compared with 3.4 nm for the B-form DNA helix. Fifth, most proteins do not require the elevated (and supraphysiological) concentrations of divalent cations such as magnesium that many complex DNA nanostructures do, and protein complexes can form highly specific structures in cellular environments without high-temperature annealing. Although the reports below focus on full-length folded proteins, we point out that synthetic peptides—such as collagen-mimetic peptides,[109] coiled coils,[110,111]
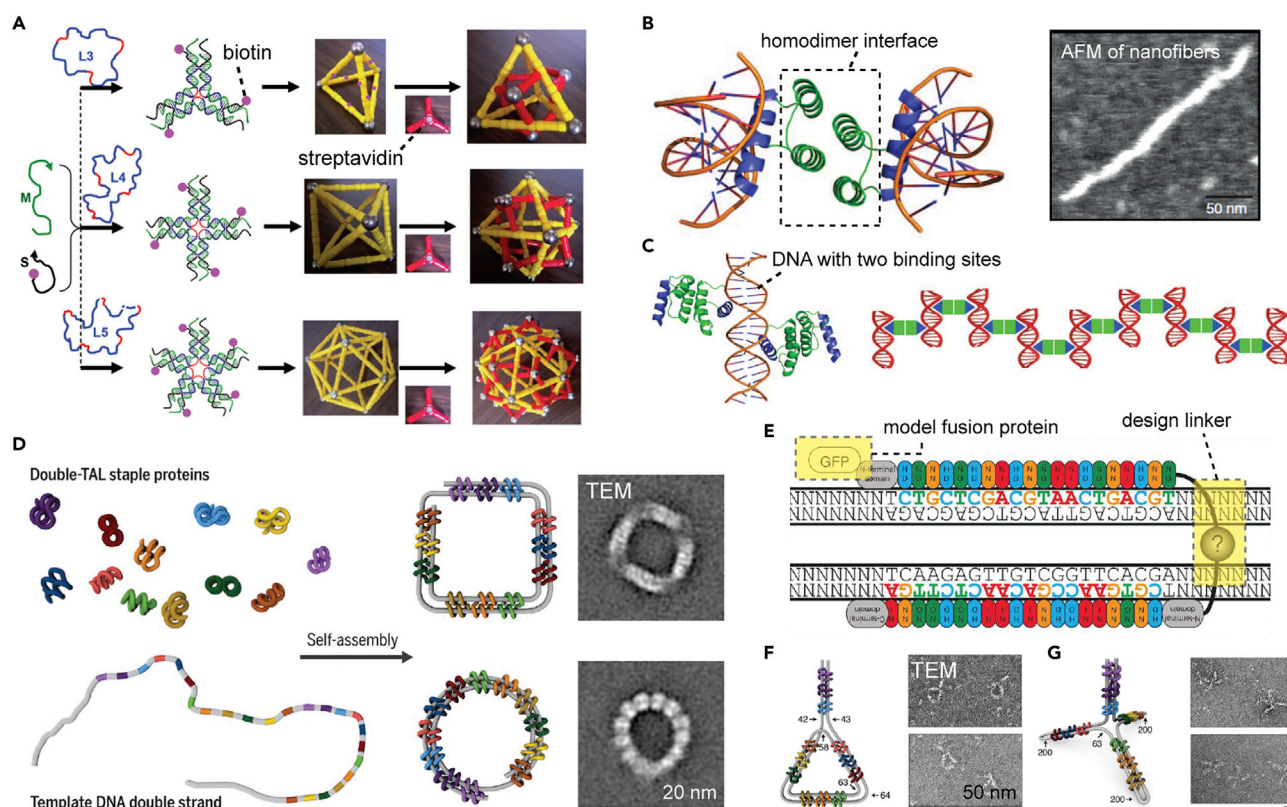
Figure 8. Hybrid Nanostructure from Non-covalent Protein-DNA Interactions

(A) Wireframe polyhedral cages (with cryo-EM reconstructions) bearing biotin; the pores are "capped" with tetravalent streptavidin. Reprinted with permission from Mao et al.[113] Copyright 2012 Wiley-VCH Verlag GmbH & Co. KGaA.

(B and C) Structure of an engineered C2-symmetric homodimer of a DNA-binding domain (B). Self-assembly of the homodimer with dsDNA bearing two binding sites results in an extended 1D nanofiber (C). Reprinted with permission from Mou et al.[115] Copyright 2015 Springer Nature.

(D–G) "Folding" a long dsDNA template with site-specific TAL fusion protein "staples" (D). Schematic of TAL staple design and specificity of binding to a dsDNA template (E). Examples of a 2D Drigalski spatula (F) and a 3D tetrahedral shape (G) via the TAL-staple approach. Reprinted with permission from Praetorius and Dietz.[116] Copyright 2017 AAAS.

and peptide amphiphiles[112]—have also been integrated with DNA for the creation of nanoassemblies with both structural motifs and represent a promising direction that takes advantage of polypeptide properties while skirting some of the challenges of recombinant expression.

One of the first examples of a nanostructure comprising both DNA and protein structural elements was reported in 2012 by Mao and coworkers, who used polyhedral cages constructed from the symmetry-guided self-assembly of branched DNA tiles.[113] The authors attached biotin to one of the strands, and as a result of the symmetry of assembly, this approach yielded three of these molecules projecting into the triangular cavity. Adding streptavidin (which can bind up to four biotin molecules) in a second step effectively "plugged" each cavity in a highly multivalent fashion (Figure 8A). Several protein-DNA cages—including tetrahedral, octahedral, and icosahedral geometries—were designed, and their 3D structure was confirmed by cryo-EM to 29 Å resolution. However, these results were reported prior to the "resolution revolution" in cryo-EM arising from the direct electron detector,[114] so it is likely that today much higher-resolution structures of protein-DNA cages can be obtained. The ability to extend beyond streptavidin to more structurally and functionally complex proteins would yield structures that mimic the dense protein shell of

viral capsids but with programmable sizes and novel symmetries. However, to rival protein assemblies such as viral capsids (e.g., to minimize undesired porosity), a rigid and preferably tight interface between the DNA frame and the protein "walls" will be necessary. Nonetheless, this example demonstrates that hybrid nanocages could be constructed from a small set of building blocks in two steps with a combination of symmetry and multivalency through well-defined protein-DNA interfaces.

In 2015, the Mayo lab achieved a simple yet elegant approach to protein-DNA hybrid structures by using the engrailed homeodomain (ENH) and dsDNA.[115] The ENH protein binds dsDNA (sequence: TAATNN) with nanomolar affinity, and the authors computationally reengineered its surface with Rosetta so that the protein formed a C2-symmetric homodimer. Co-assembling this dimeric protein with dsDNA bearing two identical binding sites rotated by 180° yielded 1D nanofibers (Figures 8B and 8C); importantly, no structure would be possible without the protein (or with only monomeric protein) or without the DNA, so the fibers were truly hybrid nanostructures. The relatively small size of the two components resulted in narrow fibers (~15 nm, although the length could extend up to 300 nm), but extending this approach to larger proteins capped with DNA-binding domains, as well as DNA nanostructure tiles or origami, could give significantly larger assemblies. As we will discuss in Protein-Actuated DNA Nanomachines, such structures are prime candidates for integrating computational protein design—for example, to control the angles and rigidity between DNA-binding domains and the structural domains of a protein building block—with DNA nanotechnology to create components that assemble without chemical modification of the protein.

In 2017, Dietz and Praetorius reported an alternative, and far more complex, method for hybrid protein-DNA structures by using sequence-specific DNA-binding proteins.[116] In DNA origami, short single-stranded "staples" are used to fold the long single-stranded bacteriophage M13 "scaffold" strand into arbitrary shapes.[10,11,12] The authors realized that the staple strands could be replaced with sequence-specific DNA-binding proteins called transcription-activator-like (TAL) effector proteins, which bound to two distal parts of a *double*-stranded scaffold, to fold it into distinct shapes through an analogous mechanism (Figure 8D). TAL proteins consist of 34 amino acid repeats, each of which can bind to a specific DNA base pair (Figure 8E), so concatenating 21 distinct repeats enabled binding of two turns of B-form DNA. By linking two such binding domains with a flexible linker, the authors could bring distal parts of the scaffold into close proximity. In an experimental tour de force, the authors carried out extensive characterization of this system to probe the exact design of the staples and the design strategies to give well-formed structures, generating shapes such as circles, squares, a Drigalski spatula, and a four-armed tetrahedral structure (Figures 8F and 8G). The approach could be extended to multi-layered assemblies (e.g., a four-helix bundle) reminiscent of 3D DNA origami,[11] and structures could even be generated with a cell-free transcription and translation system from plasmids encoding the staple proteins. This last result was particularly important because it strongly suggests that these nanostructures could be formed isothermally (i.e., without annealing) in a cellular environment, paving the way for protein-DNA nanotechnology *in vivo*. The authors also fused staples with GFP as a model cargo to demonstrate that the final structures could, in principle, display functional proteins. Overall, this landmark work demonstrates several key elements of protein-DNA nanotechnology: orientational control (given that TAL proteins bind rigidly and in a defined manner), hybrid structures with both protein and DNA components, and the possibility for highly functional assemblies for probing biology. Future work

integrating this system with other proteins has the potential to generate user-defined, intracellular nanomachines that can carry out complex functions.

In contrast with the non-covalent protein-DNA interactions used in the reports above, recent years have seen additional interest in building hybrid structures from covalent protein-DNA conjugates. In two reports in 2018, the Aida and Mirkin labs independently modified multivalent proteins with multiple DNA strands and then linked them together either directly or with complementary linkers (Figures 9A and 9B).[117,118] Although the proteins used were different (GroEL by the Aida lab and β-galactosidase by the Mirkin lab), both approaches relied on site-specific chemistry—namely, alkylation of mutagenically introduced cysteine residues at defined locations—to generate a well-defined protein-DNA hybrid molecule. But rather than attach these proteins to a preformed DNA scaffold, their DNA-mediated polymerization generated 1D nanofibers alternating between protein and DNA structural units. In both cases, the multivalent association of DNA strands drove the formation of fairly rigid linear structures, and the site specificity gave the assemblies a defined directionality that would not have been possible with non-specific approaches such as lysine acylation. The fibers could also be de-polymerized through the addition of displacement strands to break the DNA hybridization. Protein surfaces are inherently asymmetric, so mutagenesis can create anisotropic functionalization in a way not possible with more isotropic surfaces, such as those of inorganic nanoparticles. Merging these approaches with bioactive proteins, or multivalent DNA nanostructures that can control the diameter and stiffness of the protein fibers, would be particularly useful in functional biomaterials (see Protein-DNA Bionanomaterials).

Aside from linear arrays of proteins, 3D assemblies and crystals with both protein and DNA components can be created from oligonucleotide-functionalized virus capsids[123,124] or smaller multivalent proteins.[119] Although crystals are not nanostructures per se, the systems described herein are composed of nanostructured repeating units, so we feel it is appropriate to include them in this section. In 2015, the Mirkin lab reported that catalase proteins (tetrameric, heme-containing enzymes) could be densely modified with DNA through a two-step strategy: functionalization of surface lysines with an azide-NHS ester and subsequent copper-free click chemistry with cyclooctyne-labeled DNA.[119] This two-step protocol was presumably necessary to get a higher yield (up to 15 strands per tetramer) than with a DNA-NHS ester conjugate directly. Combining two sets of enzymes with complementary sticky ends and thermally annealing them resulted in 3D crystal lattices with body-centered cubic (BCC) symmetry (Figures 9C and 9D). The enzymes were still functional in the crystals and could be co-assembled with DNA-modified nanoparticles for the creation of a hybrid lattice bearing both components. In an elegant follow-up work, the authors could switch the exact lattice symmetry of the hybrid protein-nanoparticle crystals from a BCC to an AB_2 packing by moving the modification of DNA from lysines (which were evenly and spherically distributed) to cysteines (which yielded fewer handles that were more asymmetrically distributed), demonstrating the power of site-specific bioconjugation to yield a tunable protein-DNA building block.[125] Creating "Janus" protein-DNA nanoparticles by merging cysteine-specific DNA-based dimerization with a dense lysine-specific DNA coating yielded a hexagonal symmetry (Figure 9E).[120] All together, these reports show the great potential of geometrically defined assemblies—where the DNA display and valence are controlled by the protein surface—for creating hybrid protein-DNA nanostructures.
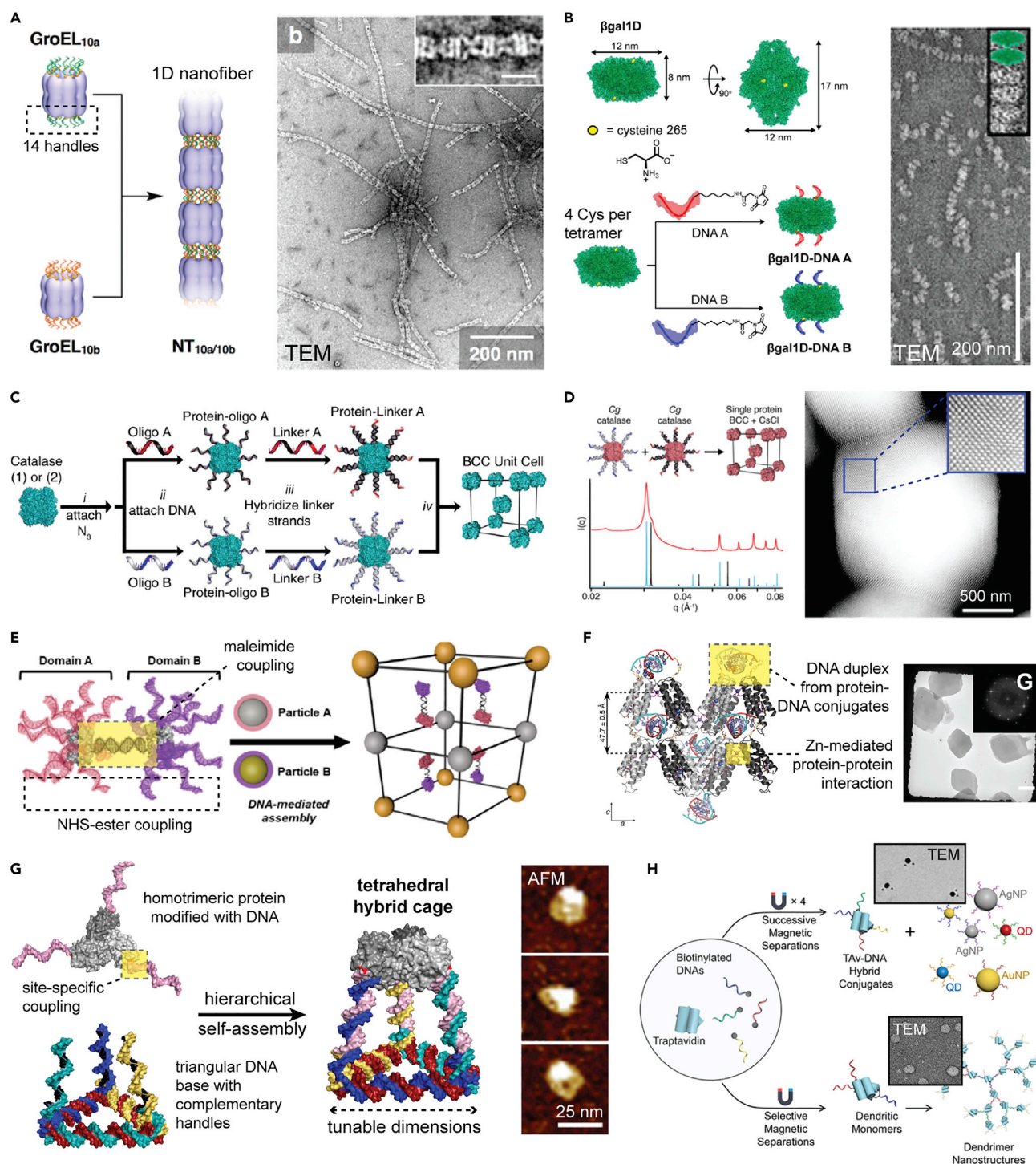
Figure 9. Hybrid Nanostructure from Covalent Conjugation or Biotin-Traptavidin Interactions

(A and B) 1D nanofibers from multivalent proteins, either GroED (A) or β-galactosidase (B), modified with complementary DNA handles. Reprinted with permission from Kashiwagi et al.[117] and McMillan et al.[118] Copyright 2018 American Chemical Society.

(C and D) Strategy for creating a 3D crystal from proteins modified with multiple DNA handles bearing sticky ends (C) and SAXS profile and TEM micrograph of protein-DNA crystals (D). Reprinted with permission from Brodin et al.[119]

(E) Creating a "Janus particle" by using two proteins for assembly into complex lattice geometries. A unique cysteine residue on the proteins results in an anisotropic homodimer, whereas lysine modification gives a dense DNA shell for crystal formation. Reprinted from Hayes et al.[120] Copyright 2018 American Chemical Society.

Figure 9. Continued

(F) Crystal assembly driven by multiple interactions between proteins (DNA hybridization and zinc coordination), as well as a TEM micrograph of the crystals formed. Reprinted with permission from Subramanian et al.[121] Copyright 2018 American Chemical Society.

(G) Self-assembly of tunable nanoscale cages composed of both protein and DNA building blocks. A homotrimeric protein modified with DNA handles assembled with a triangular DNA base bears complementary ssDNA arms. Reprinted with permission from Xu et al.[34] Copyright 2019 American Chemical Society.

(H) Protein-DNA building blocks with four orthogonal DNA handles via biotin-traptavidin association; both nanoparticle clusters and dendrimers can be assembled. Reprinted with permission from Kim et al.[122] Copyright 2019 American Chemical Society.

Also in 2018, Tezcan and coworkers described a different approach for assembling 3D protein-DNA crystals. Rather than relying solely on DNA hybridization interactions to drive the assembly of proteins, the authors used proteins that included engineered intermolecular interactions.[121] For this purpose they selected the protein RIDC3—which the Tezcan lab had redesigned to self-assemble into crystalline structures upon the addition of zinc ions—and modified it with DNA at a unique, mutagenically introduced cysteine. Mixing proteins with complementary oligonucleotide handles resulted in the rapid assembly of crystalline materials driven by both DNA hybridization and zinc-mediated protein-protein interactions (Figure 9F). This approach differed from many others described in this section in that the authors did not specifically design the system to form one particular assembly; indeed, many highly divergent arrangements of the protein and DNA building blocks were possible. The authors combined a suite of structural characterization techniques—such as AFM, scanning electron microscopy, small-angle X-ray scattering (SAXS), and cryo-EM—with computational simulation of over 50,000 protein orientations to yield four possible models for the hybrid assemblies. By systematically knocking out putative metal-binding interactions via alanine scanning, they determined that only one model fit all the experimental data perfectly. Interestingly, the final model contained several pH-dependent protein-DNA interactions that had not been explicitly designed, including both hydrogen bonds and salt bridges between protein side chains and the phosphate backbone. This result demonstrated both the complexity and difficulty of creating hybrid protein-DNA assemblies with multiple interaction "modes," as well as the opportunities for creating complex protein-DNA interfaces that more directly mimic those between proteins. Improving simulation tools for protein-DNA hybrid nanostructures to explicitly engineer these non-obvious interactions would enable a number of the complex applications described in Future Research Directions in Protein-DNA Nanotechnology and endow protein-DNA nanotechnology with many of the impressive capabilities already possible with Rosetta-based protein design.[2]

In 2019, our lab reported a novel design approach for hybrid protein-DNA nanostructures by constructing a tetrahedral cage self-assembled from a triangular DNA structure with three identical ssDNA handles and a homotrimeric protein modified with complementary oligonucleotides (Figure 9G).[34] Both the protein and the triangular DNA "base" are integral components of the nanostructure: in the absence of either, no cage will form. In this case, site-specific chemistry is critical to give a well-defined structure, and the location must be chosen carefully to avoid deleterious strain or steric hindrance (see below). For the protein, we chose the C3-symmetric KDPG aldolase, which is stable to over 80°C and readily amenable to mutagenesis and recombinant production in *E. coli*. We attached 21-nt DNA handles to mutagenic cysteines with a heterobifunctional crosslinker and purified the triply modified trimer away from incompletely modified proteins via anion-exchange chromatography. This protein-DNA building block was then co-assembled with the triangular DNA base bearing

complementary handles, resulting in a wireframe DNA cage "capped" by the protein trimer. The DNA base could be readily tuned in size (three and four turns yielded structures 10 and 14 nm on an edge, respectively), and a number of indirect experiments demonstrated the cage structure as designed. To prove the versatility of our approach (and avoid disulfide-induced aggregation of the proteins), we also attached the DNA by using copper-free click coupling with trimers bearing a 4-azidophenylanine residue, introduced by the Schultz method. Interestingly, the site of modification affected the yield of cage formation: if the DNA handles were placed too close together, only one arm of the base could bind to the trimer because of electrostatic repulsion. These hybrid protein-DNA cages are, to our knowledge, the first example of a discrete and monodisperse hybrid nanostructure (i.e., not an extended nanofiber or crystal) made from chemical conjugation of oligonucleotides to a protein surface. Future studies fusing targeting peptides or proteins to the trimer, creating nanostructures with multiple copies of the protein, and incorporating stimulus-responsive proteins will yield highly functional structures for multiple applications.

Also in 2019, the Song lab used protein-DNA hybrids to create dendrimeric nanoparticles composed of both DNA and protein structural units.[122] To avoid chemical conjugation of DNA, the authors used tetrameric traptavidin (a more stable mutant of streptavidin) and modified it with four distinct DNA handles through the corresponding biotin conjugates. Key to their approach was the purification of tetramers bearing exactly four distinct handles from conjugates stemming from the statistical mixture of all possible combinations, which they accomplished by using a sequential magnetic-bead purification method (Figure 9H). Although low yielding for the final tetra-functionalized protein (~8%), this method avoids the myriad challenges of site-specific bioconjugation, and the traptavidin-biotin interaction is virtually irreversible. The authors used the multivalent protein-DNA conjugates to assemble dimers, trimers, and tetramers of DNA-functionalized gold particles, as well as hierarchical, size-controlled dendrimers. The dendrimers in particular represent a new paradigm in DNA-mediated protein assembly, which is to create nanostructures with a defined size and both DNA and protein building blocks; extending this approach to multivalent proteins with greater functionality (e.g., through genetic fusions to an oligomeric protein) will greatly expand the applications of these structures. One next logical step for both these traptavidin-DNA conjugates and our protein-DNA cages will be to use addressable DNA scaffolds to direct the conjugation of additional strands and thereby break the intrinsic symmetry of the protein assemblies without laborious purification.

FUTURE RESEARCH DIRECTIONS IN PROTEIN-DNA NANOTECHNOLOGY

The four areas listed above demonstrate the great diversity in both fundamental and applied nanostructures that can be achieved through the merging of proteins and DNA scaffolds. We next turn to several key areas of future investigation and application in this field that are of particular interest to our lab and many others. We place a special emphasis on creating nanostructures where the protein and DNA scaffold make up a continuous, hybrid "macromolecule" (through either covalent or supramolecular interfaces). This goal, in turn, will require advancements in bioconjugation methods or the integration of these methods with approaches such as affinity interactions (e.g., binding peptides and aptamers) for further control. The net outcome will be to create a set of protein-DNA building blocks that can be combined in a fully

modular fashion (akin to the Lego-like DNA "bricks" developed by the Yin group)[126] in a highly predictable and computationally designable manner and integrate protein functionality and diversity.

Controlled and Rigid Orientation of Proteins on DNA Nanoscaffolds

Although the examples in Controlling Protein Orientation on a DNA Scaffold demonstrated the potential for controlling protein orientation on DNA scaffolds, many opportunities remain in this field both for chemical approaches and for novel applications. To further enhance the defined relationship between proteins and DNA scaffolds, two key criteria will be crucial (Figure 10A): (1) the reduction of linker length between the DNA strand and the protein surface and (2) the attachment of the protein at two or more points on the resulting nanostructure. We can accomplish the first goal by avoiding commercial linkers (which generally include 6–12 bonds between the two components) and synthesizing custom phosphoramidites with bioconjugation handles directly off the 5′ or 3′ ends or attached to the backbone (Figure 10B). To compensate for decreased efficiency due to shortened linkers, powerful reactions such as inverse-electron-demand Diels-Alder reactions between tetrazines and *trans*-cyclooctenes[127] or oxidative couplings[58] in conjunction with non-canonical amino acids might be necessary. It will be particularly difficult to attach a second (or third) unique DNA strand to a protein surface, especially if longer strands are used, because of electrostatic and steric effects. In these cases, single nucleotides or short strands can be attached, and then enzymatic ligations can be applied to extend them (Figure 10C), as in the terminal deoxynucleotidyl transferase strategy[51] or a splint-based ligation strategy. Alternatively, the first DNA modification can be used to tether a protein on a DNA structure and enhance the local concentration of the second strand (Figure 10D); strand displacement to remove the nanostructure can then follow. This strategy is similar to Gothelf's polyhistidine-Ni(NTA) strategy[38] or the photo-affinity labeling of proteins with DNA with photo-crosslinkable moieties such as diazirines[128] and creates the potential for the size and shape of the nanostructure to be used for directing the second modification.

Although a protein can be attached to a DNA nanostructure at multiple points through two sequential and site-specific bioconjugation steps, another option is to use a binding moiety—such as a peptide or aptamer—to help "anchor" the protein on the structure. Protein-protein interactions generally rely on a binding interface composed of multiple weak interactions between the two molecules; DNA structures are ideal nanoscale scaffolds for positioning multiple weak binders to tightly immobilize a protein. Such affinity interactions could be used in concert with a covalent linkage and could bind directly to the protein surface or to an engineered domain such as a fusion (for which binding peptides already exist), a coiled-coil peptide (Figure 10E), or a peptide that binds to certain DNA sequences, such as an "A-T hook."[129] Novel affinity molecules can in turn be discovered through methods such as phage display (for peptides) or SELEX (for aptamers), and introducing photo-crosslinking moieties can convert a non-covalent interaction into a covalent bond in order to lock it in place. Small, protein-based binding agents such as nanobodies or scFv antibody fragments—which can bind to the protein or a fusion thereof and are more amenable to recombinant expression and DNA modification—can also be used (Figure 10F). Alternatively, multiple short binding peptides (which each target a different part of the protein surface)[68] could be spatially arranged on a DNA scaffold to best "fit" a protein and bind it with high affinity and rigidity (Figure 10G). The work by Sacca and coworkers[69] mentioned in Controlling Protein Orientation on a DNA Scaffold (Figure 5D) is an example of such DNA-enabled mimics of
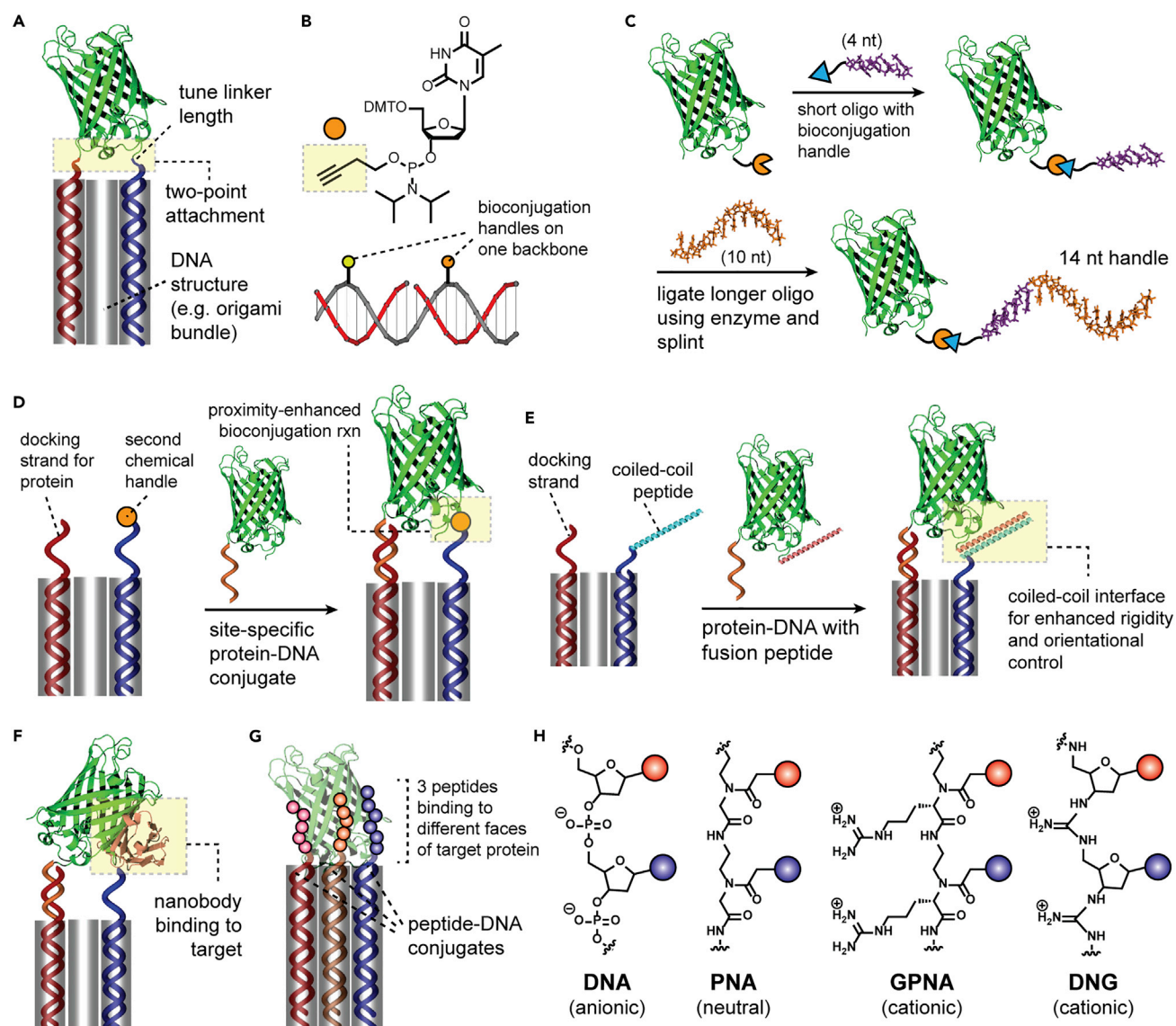
Figure 10. Future Chemical Strategies for Protein-DNA Hybrid Materials

(A) Pinning a protein on a DNA scaffold in two locations—with control over the linker length and rigidity—will enhance the orientational control of the protein.

(B) Example of a phosphoramidite for introducing an alkyne bioconjugation handle into the DNA backbone. Attaching two such residues 10–11 nt apart will allow for "pinning" a protein on a DNA backbone.

(C) Attaching a short (4-nt) DNA strand to a protein and then elongating it with an enzyme (e.g., via splint ligation) could circumvent the challenges with attaching full-length DNA strands to a protein.

(D) Rather than modifying a protein with two DNA strands, the first strand can be used to position it close to the second bioconjugation handle, resulting in a proximity-enhanced second reaction.

(E) A coiled-coil association between a protein fusion and a peptide linked to a DNA structure can help enhance a rigid interface without requiring a second bioconjugation reaction.

(F) A small binding agent such as a nanobody can be used to help anchor a protein on a DNA scaffold.

(G) DNA-scaffold-templated peptides bind to different faces of a protein in order to "pin" it, akin to antibody-antigen binding.

(H) Comparison of the structure of natural DNA (which is anionic) with those of PNA and DNG, which are neutral and cationic, respectively. The colored spheres represent the natural DNA bases (A, T, C, and G).

traditional protein-protein interfaces, albeit with multiple copies of a *single* peptide sequence for binding to a multivalent assembly. Extending this approach to monovalent proteins would greatly expand the palette of applications possible.

In parallel with these chemical methods, new strategies will need to be developed for purifying multiply modified proteins or non-covalent protein-DNA assemblies, including chromatography—for example, anion-exchange chromatography, which can be particularly effective given the large number of negative charges introduced by appending DNA—or gradient-ultracentrifugation methods. The DNA handles can also be used as affinity tags to pull the desired conjugate out of a mixture through either modification of a solid support with the complementary handle or temporary attachment of the protein to a larger DNA nanostructure to aid purification (as demonstrated by Fromme and coworkers).[74] Given that many proteins are cationic (or have cationic domains, such as heparin- or DNA-binding modules), additional issues could arise as a result of non-specific aggregation with the DNA during the conjugation reaction. In this case, uncharged oligonucleotides, such as peptide nucleic acid (PNA), or oligonucleotides with cationic backbones, such as guanidine-PNA[130] (GPNA) or deoxynucleic guanidine[131] (DNG), can be used instead (Figure 10H).

Structural Biology on Proteins Aided by DNA Scaffolds

As mentioned in Controlling Protein Orientation on a DNA Scaffold, Ned Seeman conceived of DNA nanotechnology as a way to solve protein structures by positioning them in 3D on a repeating oligonucleotide scaffold. Several groups have in fact immobilized proteins in the cavities of 2D and 3D DNA scaffolds, though not with sufficient rigidity to yield a structural solution.[132–134] The Yan lab, in collaboration with our own, is actively pursuing novel designs to increase the cavity size and improve the crystal resolution[9,135,136] in order to immobilize proteins or peptides to solve their structure. As more designs with larger cavities and channels are reported, ever larger guest molecules can be immobilized in the self-assembled lattices. The key to solving protein structure on such assemblies will be rigid attachment in a defined manner on the DNA that composes the crystal, as well as high occupancy of the available cavities. The techniques and advances outlined in Controlled and Rigid Orientation of Proteins on DNA Nanoscaffolds will be critical to accomplishing a rigid and identical linkage between the protein and the DNA scaffold; binding interfaces (or molecules such as aptamers or nanobodies) could be particularly helpful in avoiding modification of the target protein with multiple DNA handles first. In addition to traditional X-ray crystallography (which requires crystals tens to hundreds of micrometers in size), new methods such as X-ray free-electron lasers[137] and cryo-EM electron diffraction[138] can be used with much smaller DNA crystals bearing proteins (tens to hundreds of nanometers). These approaches—which dramatically reduce the distance the protein must diffuse from the outside to the interior—will be useful if the proteins are not stable to the temperatures used for DNA self-assembly and thus must be soaked into the crystals after assembly.

A second field where DNA nanoscaffolds can aid in protein structural determination is cryo-EM. As described in Controlling Protein Orientation on a DNA Scaffold, the Dietz lab demonstrated the first use of a DNA origami structure as a fiducial marker in cryo-EM to select particles, protect the protein from adsorption to the air-water interface, and tune the orientation of the target. In order to solve the structure of proteins that do not intrinsically bind DNA, additional methods will be necessary for rigidly attaching them to the origami scaffold. Binding interfaces, or the ability to chemically (and seamlessly) transition from the DNA origami platform into the protein and back out again, will be critical to enforcing a defined relationship between the protein and the scaffold. DNA nanostructures have several key advantages for this purpose. First, the Shih and Lin labs have demonstrated the templating

of liposomes[80] or lipid nanodiscs[139] on DNA nanostructures, paving the way for cryo-EM characterization of membrane proteins. Second, asymmetric DNA scaffolds can be used for determining the absolute orientation of the appended protein (as long as it is rigidly attached), thereby aiding particle averaging. Third, nanostructures with repeating sites for proteins attachment (in a linear,[6] 2D,[6,132] or helical fashion) can be designed, which will in turn allow many orientations of the target to be visualized from a single particle.

One reasonable criticism of the above proposals is that structural biology methods are constantly improving, rendering the need for a DNA scaffold such as a crystal or a cryo-EM nanogoniometer moot. Furthermore, rigid attachment of a protein to a DNA scaffold assumes some level of knowledge of the structure to begin with, so using those scaffolds to *solve* the structure does not add any additional value. This second point can be addressed with the use of protein-binding agents whose structures *are* known (e.g., nanobodies and aptamers) for binding a target with unknown structure or probing protein-protein (or protein-DNA or protein-RNA) interactions that are not known even if the individual partners are. However, even if emerging structural biology methods obviate the need for a DNA scaffold altogether, controlled attachment of proteins to crystals or large nanostructures has a range of additional advantages. For example, a 3D crystal can be used to immobilize enzymatic cascades of proteins or as a material to protect them from degradation, neither of which requires solving the structure of the protein on the crystal. Cryo-EM characterization of protein-DNA nanoassemblies could also be useful for characterizing the spacing and orientation of DNA-scaffolded proteins or the exact structure of hybrid protein-DNA nanomachines or nanostructures (as described in the next two sections). Neither of these applications requires atomic resolution, making them useful with current capabilities. Currently, researchers routinely characterize DNA nanostructures by cryo-EM to resolutions ~10 Å in order to probe their structure in solution (e.g., to prove that an origami cage is actually a 3D object with an interior cavity), so there is great value in applying these methods to the hybrid structures described herein.

Protein-Actuated DNA Nanomachines

One of the most enduring inspirations for nanotechnology is the idea of a "nanorobot" that can manipulate molecular targets in programmable ways. DNA nanotechnology is arguably the most powerful method for building complex, highly anisotropic nanostructures (including those that even resemble macroscopic robots; Figure 1F),[13] and a number of actuation mechanisms have been designed to switch them between two or more states. Most of these rely on modulation of DNA hybridization or stacking, but stimulus-responsive proteins represent a powerful and unexplored alternative mechanism. Such nanostructures would be particularly useful in triggered cargo delivery or in switching protein activity "on" and "off" by opening a DNA box containing the protein. If photoswitchable proteins are used to open and close the box—e.g., proteins that reversibly dimerize under two different wavelengths of light[140]—such structures would in effect control the protein activity by light, an approach that could be termed "nano-optogenetics" (Figure 11A). Combining multiple triggers in one cage would allow it to activate the protein only upon binding an intracellular target (e.g., a protein or mRNA, as demonstrated in the two key nanorobot papers covered in Using Hybrid Protein-DNA Nanostructures to Answer Biological Questions)[14,81] and in the presence of light. It is hard to achieve this complexity with other supramolecular systems, especially if reversibility of protein function is desired. In fact, in 2018 Famulok and coworkers reported a protein-driven DNA machine composed of a T7 RNA polymerase attached to a
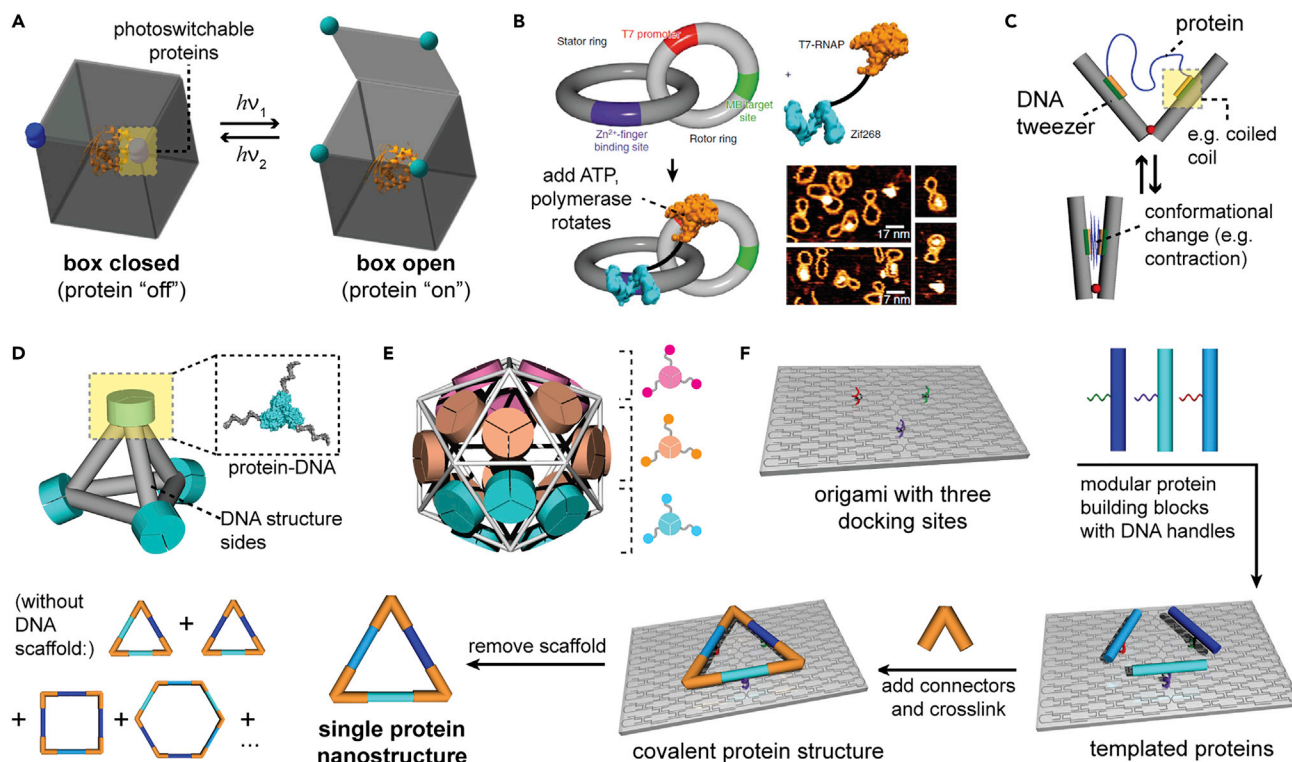
Figure 11. Future Applications for Protein-DNA Nanotechnology

(A) Opening and closing a DNA origami box with reversibly photoswitchable protein "latches" allows for activation of the protein within.

(B) A "catenane nanoengine" composed of a zinc finger and RNA polymerase undergoes constant rotary motion upon the addition of ATP. Reprinted with permission from Valero et al.[141] Copyright 2018 Springer Nature.

(C) Reversible actuation of a DNA tweezer nanostructure with a protein that undergoes a stimulus-responsive conformational change.

(D) Example of a protein-DNA tetrahedral cage with protein structural units at all vertices.

(E) Plugging the holes of an icosahedral wireframe DNA origami cage with three different trimeric proteins bearing DNA handles with different sequences.

(F) Using a DNA origami scaffold to position, link, and then release protein-DNA building blocks in order to create a unique protein nanostructure. Without the DNA scaffold, many assemblies and oligomeric states would be possible.

catenated DNA nanoring that functioned as a "nanoengine" by consuming fuel (e.g., ATP) to drive the protein motion in a continuous circular fashion (Figure 11B).[141] Other stimulus-responsive proteins that undergo a conformational change—such as calmodulin or elastin-like polypeptides (ELPs)—could be used for reconfiguring DNA nanostructures in a reversible fashion, akin to "muscles" acting on a DNA "skeleton" (Figure 11C). Fuel-dependent enzymes such as molecular motors would allow for mechanical motion to be stimulus responsive and out of equilibrium. Any proteins actuating a DNA structure would have to be tethered at two (or more) site-specific locations, through relatively rigid or short linkers, in order to exert force on a DNA nanostructure in an efficient and directionally defined manner. The exact relationship between the protein and DNA could be probed with a combination of moderate-resolution cryo-EM (as described in Controlled and Rigid Orientation of Proteins on DNA Nanoscaffolds) and computational simulations.

A range of applications exist for the types of nanomachines proposed above. For example, functionalizing such a nanomachine with additional proteins that bind to biological receptors (as in Figure 7A) would allow for single-molecule studies of the forces required for biological activation. A protein-DNA nanomachine could also serve as a "nanoinjector" for a cell by embedding itself into the membrane

and then poking it with a protein-actuated motion, similarly to viruses such as bacteriophage T4. Machines that bind to two different faces of a protein (or protein complex) inside cells could also reversible (de)activate them by applying precise forces. Photoswitchable protein-DNA nanostructures could control the assembly (e.g., of extracellular matrix [ECM] fibers) upon exposure to light to create photo-reversible hydrogels. Proteins that are not intrinsically light responsive, such as ELPs, could be rendered so by the attachment of gold nanorods (which locally generate heat upon illumination with infrared light) to the DNA scaffolds, creating truly hybrid assemblies that integrate multiple molecular functionalities. Finally, hierarchical organization of protein-DNA nanomachines into bundles or hydrogels that span multiple length scales could result in stimulus-responsive mechanical materials such as artificial muscles.

More Complex Protein-DNA Nanostructures

We also foresee the development of more complex hybrid nanostructures with both protein and DNA components. For example, our lab's work with a trimeric protein building block "capping" a DNA structure[34] (Figure 9G) could be extended to cages with proteins at *all* vertices (Figure 11D). Tuning the rigidity of the protein-DNA interface could lead to cages of varying sizes and symmetries, akin to reports using all-DNA tiles.[142] Protein building blocks could also be used to "plug" symmetry-matched holes in a wireframe DNA origami structure (similar to Mao's work with streptavidin-capped cages),[113] creating a semi-closed protein shell akin to a virus capsid. Unlike virus capsids, however, each cavity of these structures could be modified with a different protein-DNA building block, enabling highly anisotropic protein shells and Janus-like particles (Figure 11E). The protein building blocks can also contain fusion peptides or proteins for biological activity (e.g., drug delivery, artificial vaccines, and catalytic cascades) and can be removed selectively through toehold-mediated strand displacement. In addition to covalent attachment of DNA strands, alternative strategies for integrating self-assembling proteins with DNA could yield a tighter and more rigid interface. All of these applications would especially benefit from the introduction of design software that can accurately model both building blocks. Such software—for example, extensions of Rosetta (for proteins)[143] or oxDNA (for oligonucleotides)[144] that incorporate representations of the other macromolecular type—will enable rapid *in silico* testing of designs to avoid laborious synthetic trial and error.

DNA-Scaffold-Templated Synthesis of Protein Nanostructures

One of the key advantages of DNA nanotechnology is that it allows for scaffolds with a high degree of addressability because of the unique nature of the strands that compose them. Most engineered protein nanostructures, by contrast, are highly symmetric[2] because engineering multiple specific and orthogonal interactions is difficult. Thus, a unique opportunity for protein-DNA nanotechnology is to use oligonucleotide scaffolds to position proteins in an asymmetric fashion with complete stoichiometric control, to covalently or non-covalently link them, and to then remove them from the DNA scaffold by using cleavable linkers (Figure 11F). The DNA scaffold in effect serves as a "supramolecular mold" for building protein structures that could not otherwise be created in solution, all without having to re-engineer (multiple) protein self-assembly interfaces. This approach will require additional bioconjugation strategies beyond those necessary to attach the proteins to DNA in the first place in order to link the pieces together into higher-order structures, in much the same way that traditional organic synthesis requires multiple reactions to link various functional groups together in a selective and site-specific fashion. However, the final outcome will be an all-protein nanostructure with the complexity

of DNA origami and the diversity of proteins. Beyond the biological and catalytic function of proteins, such nanostructures could also integrate *structural* proteins (e.g., actin fibers and collagen fibrils), thereby moving beyond the properties of the DNA double helix while retaining the modularity that makes DNA nanotechnology so powerful. It is still an open question what the best protein building blocks would be for this purpose, but the *de novo* designed, tunable, and highly stable assemblies employed by protein engineers (such as repeat proteins[145] or helical bundles[146]) are possible starting points.

Protein-DNA Bionanomaterials

In addition to the myriad areas discussed in Using Hybrid Protein-DNA Nanostructures to Answer Biological Questions, two areas of biology that will benefit greatly from nanomaterials that merge the structural tunability of DNA with the bioactivity of proteins are (1) targeted delivery of therapeutic cargo to cells and (2) biomaterials for regenerative medicine. In this regard, protein-DNA nanostructures will create tunable analogs of (1) viruses and (2) the ECM. For both of these applications, a DNA "skeleton" can be coated with a protein or polymer "skin" to allow independent control of size, shape, and rigidity (via the DNA) and bioactivity (via the polypeptide or polymer coating). Functionalizing these assemblies with proteins to enhance stability, modulate surface charge, and facilitate targeting, uptake, endosomal escape, and subcellular localization will allow for targeted cargo delivery into cells. For biomaterials, a DNA-based nanofiber could control the diameter, stiffness, and nanoscale morphology of a fiber to mimic the ECM (akin to collagen or laminin), whereas the proteins can interface with cell-surface proteins, such as integrins or growth factor receptors, to influence migration, differentiation, or regeneration. For example, fibronectin is a protein composed of multiple individually folded domains, much like beads on a string, each with a different biological role.[147] Creating a DNA nanostructure coated with these domains would allow for the precise control of bioactivity, such as cell adhesion or growth factor signaling, with simultaneous control over the mechanical and morphological properties of these assemblies. Such a hybrid nanostructure would strike a balance between using much simpler structures that often fail to recapitulate biological complexity and using full-length proteins.

For both targeted delivery and biomaterials, the ability of DNA to spatially control the presentation of multiple signals will be key, e.g., by "matching" the spacing of receptors or co-localizing several proteins to enhance bioactivity.[95] DNA also allows for the dynamic presentation of proteins through strand-displacement reactions or reversible crosslinking of DNA fibers, opening up a wealth of possibilities for adaptive biomaterials with spatiotemporal control.[112,148] A combination of non-specific and specific (via direct tethering or binding to defined locations) will most likely be necessary for effectively coating DNA nanoscaffolds with proteins. It should also be noted that recent breakthroughs in biotechnological production of DNA origami (~$100/g)[149] have opened up the possibility of scalable production of DNA nanostructures that match those of recombinant proteins.

Selective Modification of Proteins with a Supramolecular Scaffold

One especially powerful application of the hybrid field described herein is to extend the ideas outlined by Gothelf and coworkers (Figures 3A and 5E)[38,79] to a general platform for DNA-scaffold-enabled, site-specific protein modification. By positioning a protein on a DNA nanostructure, it should be possible to selectively modify it on one face, even on a single residue, akin to the way that enzymes such as kinases can phosphorylate a specific site by binding to a unique location on their target. This

modification could be something completely synthetic (e.g., a drug, polymer, or fluorophore) or a natural moiety such as a post-translational modification (e.g., phosphorylation, glycosylation, or lipidation). As with many of the propositions described herein, a somewhat rigid and controlled orientation of the protein on the DNA scaffold will most likely be necessary to prevent the modification of an incorrect site. Chaining together several DNA-based modules with unique site-specific chemistries, and passing the target protein sequentially between them, would in effect create a "molecular assembly line" similar to non-ribosomal peptide synthesis. The assembly of these modules could in turn be controlled dynamically with DNA-based circuits and computational elements such as logic gates, creating a truly cell-mimetic molecular factory (see below) that goes far beyond current DNA-templated enzyme cascade capabilities. We highlight that obtaining suitable quantities of proteins with these materials will require the use of simple tiles or scaffolds or the scaling up of DNA origami to gram quantities, as recently reported.[149]

Synthetic Cells and Nanoscale "Chemical Plants"

Creating an artificial cell with complexity rivaling biology but with completely synthetic components is a holy grail of nanotechnology. Such a nanoscale "chemical plant,"[150] with the ability to produce novel molecules, sense and respond to the environment, and even self-replicate and evolve, would be the ultimate realization of Feynman's iconic dream. This vision is still far away, but we believe that protein DNA nanostructures will play a key role in its realization. Applications include switching molecular assembly lines (e.g., enzymes attached to DNA scaffolds) on and off; integrating with DNA molecular computing networks for signaling, feedback, and control; using proteins on DNA "nanorobotic" arms to functionalize other molecular species; localizing proteins on DNA cages as nanoreactors; using motor proteins on dynamically controllable DNA tracks to transport cargo; and spatially controlling multiple proteins on addressable DNA "cytoskeletons." Key to these endeavors will be the integrated, hybrid protein-DNA nanostructures we have discussed herein, enabled by novel chemical tools and self-assembly methods.

CONCLUSIONS

We hope that this review has demonstrated the rich potential that lies at the interface of DNA nanotechnology and protein chemistry and engineering. A truly hybrid field of protein-DNA nanotechnology will build on the ever-increasing advances in DNA and RNA nanotechnology, *de novo* protein design, bioconjugation chemistry, supramolecular self-assembly, and computational simulation of biomolecular systems. The long-term potential for this field is to create a self-assembled, biomolecule-based analog to synthetic organic chemistry: using a palette of building blocks and reactions to build complex final structures with complete control down to the atomic level. The key differences are that protein-DNA nanotechnology will extensively use supramolecular interactions (as opposed to purely covalent bonds), and the final "molecules" will be complex assemblies and devices with functions that rival those of biology. This goal will require an intimate interplay between chemists, biologists, engineers, physicists, and materials scientists, but the potential is limitless; the ultimate goal is to create nanostructures and nanosystems that one day rival natural enzymes, cells, and perhaps entire organisms. There truly is plenty of room left at the bottom, and protein-DNA nanotechnology is ready to fill it.

ACKNOWLEDGMENTS

REFERENCES

1. Packer, M.S., and Liu, D.R. (2015). Methods for the directed evolution of proteins. Nat. Rev. Genet. 16, 379–394.

2. Huang, P.S., Boyken, S.E., and Baker, D. (2016). The coming of age of de novo protein design. Nature 537, 320–327.

3. Dedeo, M.T., Finley, D.T., and Francis, M.B. (2011). Viral capsids as self-assembling templates for new materials. Prog. Mol. Biol. Transl. Sci. 103, 353–392.

4. Stephanopoulos, N., Ortony, J.H., and Stupp, S.I. (2013). Self-assembly for the synthesis of functional biomaterials. Acta Mater. 61, 912–930.

5. Luo, Q., Hou, C., Bai, Y., Wang, R., and Liu, J. (2016). Protein assembly: versatile approaches to construct highly ordered nanostructures. Chem. Rev. 116, 13571–13632.

6. Yan, H., Park, S.H., Finkelstein, G., Reif, J.H., and LaBean, T.H. (2003). DNA-templated self-assembly of protein arrays and highly conductive nanowires. Science 301, 1882–1884.

7. Winfree, E., Liu, F., Wenzler, L.A., and Seeman, N.C. (1998). Design and self-assembly of two-dimensional DNA crystals. Nature 394, 539–544.

8. Zheng, J., Birktoft, J.J., Chen, Y., Wang, T., Sha, R., Constantinou, P.E., Ginell, S.L., Mao, C., and Seeman, N.C. (2009). From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. Nature 461, 74–77.

9. Simmons, C.R., Zhang, F., Birktoft, J.J., Qi, X., Han, D., Liu, Y., Sha, R., Abdallah, H.O., Hernandez, C., Ohayon, Y.P., et al. (2016). Construction and structure determination of a three-dimensional DNA crystal. J. Am. Chem. Soc. 138, 10047–10054.

10. Rothemund, P.W.K. (2006). Folding DNA to create nanoscale shapes and patterns. Nature 440, 297–302.

11. Douglas, S.M., Dietz, H., Liedl, T., Högberg, B., Graf, F., and Shih, W.M. (2009). Self-assembly of DNA into nanoscale three-dimensional shapes. Nature 459, 414–418.

12. Hong, F., Zhang, F., Liu, Y., and Yan, H. (2017). DNA origami: scaffolds for creating higher order structures. Chem. Rev. 117, 12584–12640.

13. Gerling, T., Wagenbauer, K.F., Neuner, A.M., and Dietz, H. (2015). Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components. Science 347, 1446–1452.

14. Douglas, S.M., Bachelet, I., and Church, G.M. (2012). A logic-gated nanorobot for targeted transport of molecular payloads. Science 335, 831–834.

15. Douglas, S.M., Marblestone, A.H., Teerapittayanon, S., Vazquez, A., Church, G.M., and Shih, W.M. (2009). Rapid prototyping of 3D DNA-origami shapes with caDNAno. Nucleic Acids Res. 37, 5001–5006.

16. Saccà, B., and Niemeyer, C.M. (2011). Functionalization of DNA nanostructures with proteins. Chem. Soc. Rev. 40, 5910–5921.

17. Wilner, O.I., and Willner, I. (2012). Functionalized DNA nanostructures. Chem. Rev. 112, 2528–2556.

18. Madsen, M., and Gothelf, K.V. (2019). Chemistries for DNA nanotechnology. Chem. Rev. 119, 6384–6458.

19. Pinheiro, A.V., Han, D., Shih, W.M., and Yan, H. (2011). Challenges and opportunities for structural DNA nanotechnology. Nat. Nanotechnol. 6, 763–772.

20. Hu, Y., and Niemeyer, C.M. (2019). From DNA nanotechnology to material systems engineering. Adv. Mater. 31, e1806294.

21. Zhou, K., Dong, J., Zhou, Y., Dong, J., Wang, M., and Wang, Q. (2019). Toward precise manipulation of DNA-protein hybrid nanoarchitectures. Small 15, e1804044.

22. Teller, C., and Willner, I. (2010). Organizing protein-DNA hybrids as nanostructures with programmed functionalities. Trends Biotechnol. 28, 619–628.

23. Rajendran, A., Nakata, E., Nakano, S., and Morii, T. (2017). Nucleic-acid-templated enzyme cascades. ChemBioChem 18, 696–716.

24. Linko, V., Nummelin, S., Aarnos, L., Tapio, K., Toppari, J.J., and Kostiainen, M.A. (2016). DNA-based enzyme reactors and systems. Nanomaterials (Basel) 6, 139.

25. MacCulloch, T., Buchberger, A., and Stephanopoulos, N. (2019). Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials. Org. Biomol. Chem. 17, 1668–1682.

26. Mikkilä, J., Eskelinen, A.-P., Niemelä, E.H., Linko, V., Frilander, M.J., Törmä, P., and Kostiainen, M.A. (2014). Virus-encapsulated DNA origami nanostructures for cellular delivery. Nano Lett. 14, 2196–2200.

27. Auvinen, H., Zhang, H., Nonappa, Kopilow, A., Niemela, E.H., Nummelin, S., Correia, A., Santos, H.A., Linko, V., and Kostiainen, M.A. (2017). Protein coating of DNA nanostructures for enhanced stability and immunocompatibility. Adv. Healthc. Mater. 6, 1700692.

28. Hernandez-Garcia, A., Kraft, D.J., Janssen, A.F.J., Bomans, P.H.H., Sommerdijk, N.A., Thies-Weesie, D.M.E., Favretto, M.E., Brock, R., de Wolf, F.A., Werten, M.W.T., et al. (2014). Design and self-assembly of simple coat proteins for artificial viruses. Nat. Nanotechnol. 9, 698–702.

29. Hernandez-Garcia, A., Estrich, N.A., Werten, M.W.T., Van Der Maarel, J.R.C., LaBean, T.H., de Wolf, F.A., Cohen Stuart, M.A., and de Vries, R. (2017). Precise coating of a wide range of DNA templates by a protein polymer with a DNA binding domain. ACS Nano 11, 144–152.

30. Niemeyer, C.M., Sano, T., Smith, C.L., and Cantor, C.R. (1994). Oligonucleotide-directed self-assembly of proteins: semisynthetic DNA–streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates. Nucleic Acids Res. 22, 5530–5539.

31. Niemeyer, C.M. (2010). Semisynthetic DNA-protein conjugates for biosensing and nanofabrication. Angew. Chem. Int. Ed. 49, 1200–1216.

32. Stephanopoulos, N., and Francis, M.B. (2011). Choosing an effective protein bioconjugation strategy. Nat. Chem. Biol. 7, 876–884.

33. Stephanopoulos, N., Liu, M., Tong, G.J., Li, Z., Liu, Y., Yan, H., and Francis, M.B. (2010). Immobilization and one-dimensional arrangement of virus capsids with nanoscale precision using DNA origami. Nano Lett. 10, 2714–2720.

34. Xu, Y., Jiang, S., Simmons, C.R., Narayanan, R.P., Zhang, F., Aziz, A.M., Yan, H., and Stephanopoulos, N. (2019). Tunable nanoscale cages from self-assembling DNA and protein building blocks. ACS Nano 13, 3545–3554.

35. Aldaye, F.A., Senapedis, W.T., Silver, P.A., and Way, J.C. (2010). A structurally tunable DNA-based extracellular matrix. J. Am. Chem. Soc. 132, 14727–14729.

36. Goodman, R.P., Erben, C.M., Malo, J., Ho, W.M., McKee, M.L., Kapanidis, A.N., and Turberfield, A.J. (2009). A facile method for reversibly linking a recombinant protein to DNA. ChemBioChem 10, 1551–1557.

37. Shen, W., Zhong, H., Neff, D., and Norton, M.L. (2009). NTA directed protein nanopatterning on DNA Origami nanoconstructs. J. Am. Chem. Soc. 131, 6660–6661.

38. Rosen, C.B., Kodal, A.L.B., Nielsen, J.S., Schaffert, D.H., Scavenius, C., Okholm, A.H., Voigt, N.V., Enghild, J.J., Kjems, J., Tørring, T., and Gothelf, K.V. (2014). Template-directed covalent conjugation of DNA to native antibodies, transferrin and other metal-binding proteins. Nat. Chem. 6, 804–809.

39. Rosier, B.J.H.M., Cremers, G.A.O., Engelen, W., Merkx, M., Brunsveld, L., and de Greef, T.F.A. (2017). Incorporation of native antibodies and Fc-fusion proteins on DNA

nanostructures via a modular conjugation strategy. Chem. Commun. (Camb.) *53*, 7393–7396.

40. Nguyen, T.M., Nakata, E., Saimura, M., Dinh, H., and Morii, T. (2017). Design of modular protein tags for orthogonal covalent bond formation at specific DNA sequences. J. Am. Chem. Soc. *139*, 8487–8496.

41. Lacroix, A., Edwardson, T.G.W., Hancock, M.A., Dore, M.D., and Sleiman, H.F. (2017). Development of DNA nanostructures for high-affinity binding to human serum albumin. J. Am. Chem. Soc. *139*, 7355–7362.

42. Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat. Biotechnol. *21*, 86–89.

43. Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., et al. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem. Biol. *3*, 373–382.

44. Gautier, A., Juillerat, A., Heinis, C., Corrêa, I.R., Jr., Kindermann, M., Beaufils, F., and Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. Chem. Biol. *15*, 128–136.

45. Min, D., Arbing, M.A., Jefferson, R.E., and Bowie, J.U. (2016). A simple DNA handle attachment method for single molecule mechanical manipulation experiments. Protein Sci. *25*, 1535–1544.

46. Saccà, B., Meyer, R., Erkelenz, M., Kiko, K., Arndt, A., Schroeder, H., Rabe, K.S., and Niemeyer, C.M. (2010). Orthogonal protein decoration of DNA origami. Angew. Chem. Int. Ed. *49*, 9378–9383.

47. Duckworth, B.P., Chen, Y., Wollack, J.W., Sham, Y., Mueller, J.D., Taton, T.A., and Distefano, M.D. (2007). A universal method for the preparation of covalent protein-DNA conjugates for use in creating protein nanostructures. Angew. Chem. Int. Ed. *46*, 8819–8822.

48. Hess, G.T., Guimaraes, C.P., Spooner, E., Ploegh, H.L., and Belcher, A.M. (2013). Orthogonal labeling of M13 minor capsid proteins with DNA to self-assemble end-to-end multiphage structures. ACS Synth. Biol. *2*, 490–496.

49. Tominaga, J., Kemori, Y., Tanaka, Y., Maruyama, T., Kamiya, N., and Goto, M. (2007). An enzymatic method for site-specific labeling of recombinant proteins with oligonucleotides. Chem. Commun. (Camb.) *2007*, 401–403.

50. Metelev, V.G., Kubareva, E.A., Vorob'eva, O.V., Romanenkov, A.S., and Oretskaya, T.S. (2003). Specific conjugation of DNA binding proteins to DNA templates through thiol-disulfide exchange. FEBS Lett. *538*, 48–52.

51. Sørensen, R.S., Okholm, A.H., Schaffert, D., Kodal, A.L.B., Gothelf, K.V., and Kjems, J. (2013). Enzymatic ligation of large biomolecules to DNA. ACS Nano *7*, 8098–8104.

52. Sagredo, S., Pirzer, T., Rafat, A.A., Goetzfried, M.A., Moncalian, G., Simmel, F.C., and de la Cruz, F. (2016). Orthogonal protein assembly on DNA nanostructures using relaxases. Angew. Chem. Int. Ed. *55*, 4348–4352.

53. Mashimo, Y., Maeda, H., Mie, M., and Kobatake, E. (2012). Construction of semisynthetic DNA-protein conjugates with Phi X174 Gene-A* protein. Bioconjug. Chem. *23*, 1349–1355.

54. Lovrinovic, M., Seidel, R., Wacker, R., Schroeder, H., Seitz, O., Engelhard, M., Goody, R.S., and Niemeyer, C.M. (2003). Synthesis of protein-nucleic acid conjugates by expressed protein ligation. Chem. Commun. (Camb.) *2003*, 822–823.

55. Liu, C.C., and Schultz, P.G. (2010). Adding new chemistries to the genetic code. Annu. Rev. Biochem. *79*, 413–444.

56. Meldal, M., and Tornøe, C.W. (2008). Cu-catalyzed azide-alkyne cycloaddition. Chem. Rev. *108*, 2952–3015.

57. Jewett, J.C., and Bertozzi, C.R. (2010). Cu-free click cycloaddition reactions in chemical biology. Chem. Soc. Rev. *39*, 1272–1279.

58. ElSohly, A.M., and Francis, M.B. (2015). Development of oxidative coupling strategies for site-selective protein modification. Acc. Chem. Res. *48*, 1971–1978.

59. Kazane, S.A., Sok, D., Cho, E.H., Uson, M.L., Kuhn, P., Schultz, P.G., and Smider, V.V. (2012). Site-specific DNA-antibody conjugates for specific and sensitive immuno-PCR. Proc. Natl. Acad. Sci. USA *109*, 3731–3736.

60. Kölmel, D.K., and Kool, E.T. (2017). Oximes and hydrazones in bioconjugation: mechanism and catalysis. Chem. Rev. *117*, 10358–10376.

61. Darmostuk, M., Rimpelova, S., Gbelcova, H., and Ruml, T. (2015). Current approaches in SELEX: an update to aptamer selection technology. Biotechnol. Adv. *33*, 1141–1161.

62. Liu, Y., Lin, C., Li, H., and Yan, H. (2005). Aptamer-directed self-assembly of protein arrays on a DNA nanostructure. Angew. Chem. Int. Ed. *44*, 4333–4338.

63. Rinker, S., Ke, Y., Liu, Y., Chhabra, R., and Yan, H. (2008). Self-assembled DNA nanostructures for distance-dependent multivalent ligand-protein binding. Nat. Nanotechnol. *3*, 418–422.

64. Vinkenborg, J.L., Mayer, G., and Famulok, M. (2012). Aptamer-based affinity labeling of proteins. Angew. Chem. Int. Ed. *51*, 9176–9180.

65. Nakata, E., Liew, F.F., Uwatoko, C., Kiyonaka, S., Mori, Y., Katsuda, Y., Endo, M., Sugiyama, H., and Morii, T. (2012). Zinc-finger proteins for site-specific protein positioning on DNA-origami structures. Angew. Chem. Int. Ed. *51*, 2421–2424.

66. Ngo, T.A., Nakata, E., Saimura, M., and Morii, T. (2016). Spatially organized enzymes drive cofactor-coupled cascade reactions. J. Am. Chem. Soc. *138*, 3012–3021.

67. Kurokawa, T., Kiyonaka, S., Nakata, E., Endo, M., Koyama, S., Mori, E., Tran, N.H., Dinh, H., Suzuki, Y., Hidaka, K., et al. (2018). DNA origami scaffolds as templates for functional tetrameric Kir3 K$^+$ channels. Angew. Chem. Int. Ed. *57*, 2586–2591.

68. Williams, B.A.R., Diehnelt, C.W., Belcher, P., Greving, M., Woodbury, N.W., Johnston, S.A., and Chaput, J.C. (2009). Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds. J. Am. Chem. Soc. *131*, 17233–17241.

69. Sprengel, A., Lill, P., Stegemann, P., Bravo-Rodriguez, K., Schöneweiß, E.C., Merdanovic, M., Gudnason, D., Aznauryan, M., Gamrad, L., Barcikowski, S., et al. (2017). Tailored protein encapsulation into a DNA host using geometrically organized supramolecular interactions. Nat. Commun. *8*, 14472.

70. Seeman, N.C. (1982). Nucleic acid junctions and lattices. J. Theor. Biol. *99*, 237–247.

71. Seeman, N.C. (2003). DNA in a material world. Nature *421*, 427–431.

72. Erben, C.M., Goodman, R.P., and Turberfield, A.J. (2006). Single-molecule protein encapsulation in a rigid DNA cage. Angew. Chem. Int. Ed. *45*, 7414–7417.

73. Ge, Z., Su, Z., Simmons, C.R., Li, J., Jiang, S., Li, W., Yang, Y., Liu, Y., Chiu, W., Fan, C., and Yan, H. (2019). Redox engineering of cytochrome c using DNA nanostructure-based charged encapsulation and spatial control. ACS Appl. Mater. Interfaces *11*, 13874–13880.

74. Flory, J.D., Simmons, C.R., Lin, S., Johnson, T., Andreoni, A., Zook, J., Ghirlanda, G., Liu, Y., Yan, H., and Fromme, P. (2014). Low temperature assembly of functional 3D DNA-PNA-protein complexes. J. Am. Chem. Soc. *136*, 8283–8295.

75. Zhou, K., Ke, Y., and Wang, Q. (2018). Selective in situ assembly of viral protein onto DNA origami. J. Am. Chem. Soc. *140*, 8074–8077.

76. Marth, G., Hartley, A.M., Reddington, S.C., Sargisson, L.L., Parcollet, M., Dunn, K.E., Jones, D.D., and Stulz, E. (2017). Precision templated bottom-up multiprotein nanoassembly through defined click chemistry linkage to DNA. ACS Nano *11*, 5003–5010.

77. Martin, T.G., Bharat, T.A.M., Joerger, A.C., Bai, X.C., Praetorius, F., Fersht, A.R., Dietz, H., and Scheres, S.H.W. (2016). Design of a molecular support for cryo-EM structure determination. Proc. Natl. Acad. Sci. USA *113*, E7456–E7463.

78. Dong, Y., Chen, S., Zhang, S., Sodroski, J., Yang, Z., Liu, D., and Mao, Y. (2018). Folding DNA into a lipid-conjugated nanobarrel for controlled reconstitution of membrane proteins. Angew. Chem. Int. Ed. *57*, 2072–2076.

79. Ouyang, X., De Stefano, M., Krissanaprasit, A., Bank Kodal, A.L., Bech Rosen, C., Liu, T., Helmig, S., Fan, C., and Gothelf, K.V. (2017). Docking of antibodies into the cavities of DNA origami structures. Angew. Chem. Int. Ed. *56*, 14423–14427.

80. Yang, Y., Wang, J., Shigematsu, H., Xu, W., Shih, W.M., Rothman, J.E., and Lin, C. (2016). Self-assembly of size-controlled liposomes on DNA nanotemplates. Nat. Chem. *8*, 476–483.

81. Li, S., Jiang, Q., Liu, S., Zhang, Y., Tian, Y., Song, C., Wang, J., Zou, Y., Anderson, G.J., Han, J.Y., et al. (2018). A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. Nat. Biotechnol. *36*, 258–264.

82. Grossi, G., Dalgaard Ebbesen Jepsen, M., Kjems, J., and Andersen, E.S. (2017). Control of enzyme reactions by a reconfigurable DNA nanovault. Nat. Commun. *8*, 992.

83. Kim, S.H., Kim, K.R., Ahn, D.R., Lee, J.E., Yang, E.G., and Kim, S.Y. (2017). Reversible regulation of enzyme activity by pH-responsive encapsulation in DNA nanocages. ACS Nano *11*, 9352–9359.

84. Liu, M., Fu, J., Hejesen, C., Yang, Y., Woodbury, N.W., Gothelf, K., Liu, Y., and Yan, H. (2013). A DNA tweezer-actuated enzyme nanoreactor. Nat. Commun. *4*, 2127.

85. Ke, G., Liu, M., Jiang, S., Qi, X., Yang, Y.R., Wootten, S., Zhang, F., Zhu, Z., Liu, Y., Yang, C.J., and Yan, H. (2016). Directional regulation of enzyme pathways through the control of substrate channeling on a DNA origami scaffold. Angew. Chem. Int. Ed. *55*, 7483–7486.

86. Ke, Y., Meyer, T., Shih, W.M., and Bellot, G. (2016). Regulation at a distance of biomolecular interactions using a DNA origami nanoactuator. Nat. Commun. *7*, 10935.

87. Ijäs, H., Hakaste, I., Shen, B., Kostiainen, M.A., and Linko, V. (2019). Reconfigurable DNA origami nanocapsule for pH-controlled encapsulation and display of cargo. ACS Nano *13*, 5959–5967.

88. Xin, L., Zhou, C., Yang, Z., and Liu, D. (2013). Regulation of an enzyme cascade reaction by a DNA machine. Small *9*, 3088–3091.

89. Fu, J., Yang, Y.R., Johnson-Buck, A., Liu, M., Liu, Y., Walter, N.G., Woodbury, N.W., and Yan, H. (2014). Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. Nat. Nanotechnol. *9*, 531–536.

90. Chen, Y., Ke, G., Ma, Y., Zhu, Z., Liu, M., Liu, Y., Yan, H., and Yang, C.J. (2018). A synthetic light-driven substrate channeling system for precise regulation of enzyme cascade activity based on DNA origami. J. Am. Chem. Soc. *140*, 8990–8996.

91. Williams, B.A.R., Lund, K., Liu, Y., Yan, H., and Chaput, J.C. (2007). Self-assembled peptide nanoarrays: an approach to studying protein-protein interactions. Angew. Chem. Int. Ed. *46*, 3051–3054.

92. Endo, M., Katsuda, Y., Hidaka, K., and Sugiyama, H. (2010). Regulation of DNA methylation using different tensions of double strands constructed in a defined DNA nanostructure. J. Am. Chem. Soc. *132*, 1592–1597.

93. Endo, M., Katsuda, Y., Hidaka, K., and Sugiyama, H. (2010). A versatile DNA nanochip for direct analysis of DNA base-excision repair. Angew. Chem. Int. Ed. *49*, 9412–9416.

94. Suzuki, Y., Endo, M., Katsuda, Y., Ou, K., Hidaka, K., and Sugiyama, H. (2014). DNA origami based visualization system for studying site-specific recombination events. J. Am. Chem. Soc. *136*, 211–218.

95. Shaw, A., Lundin, V., Petrova, E., Fördős, F., Benson, E., Al-Amin, A., Herland, A., Blokzijl, A., Högberg, B., and Teixeira, A.I. (2014). Spatial control of membrane receptor function using ligand nanocalipers. Nat. Methods *11*, 841–846.

96. Shaw, A., Hoffecker, I.T., Smyrlaki, I., Rosa, J., Grevys, A., Bratlie, D., Sandlie, I., Michaelsen, T.E., Andersen, J.T., and Högberg, B. (2019). Binding to nanopatterned antigens is dominated by the spatial tolerance of antibodies. Nat. Nanotechnol. *14*, 184–190.

97. Angelin, A., Weigel, S., Garrecht, R., Meyer, R., Bauer, J., Kumar, R.K., Hirtz, M., and Niemeyer, C.M. (2015). Multiscale origami structures as interface for cells. Angew. Chem. Int. Ed. *54*, 15813–15817.

98. Derr, N.D., Goodman, B.S., Jungmann, R., Leschziner, A.E., Shih, W.M., and Reck-Peterson, S.L. (2012). Tug-of-war in motor protein ensembles revealed with a programmable DNA origami scaffold. Science *338*, 662–665.

99. Iwaki, M., Wickham, S.F., Ikezaki, K., Yanagida, T., and Shih, W.M. (2016). A programmable DNA origami nanospring that reveals force-induced adjacent binding of myosin VI heads. Nat. Commun. *7*, 13715.

100. Fisher, P.D.E., Shen, Q., Akpinar, B., Davis, L.K., Chung, K.K.H., Baddeley, D., Šarić, A., Melia, T.J., Hoogenboom, B.W., Lin, C., and Lusk, C.P. (2018). A programmable DNA origami platform for organizing intrinsically disordered nucleoporins within nanopore confinement. ACS Nano *12*, 1508–1518.

101. Ketterer, P., Ananth, A.N., Laman Trip, D.S., Mishra, A., Bertosin, E., Ganji, M., van der Torre, J., Onck, P., Dietz, H., and Dekker, C. (2018). DNA origami scaffold for studying intrinsically disordered proteins of the nuclear pore complex. Nat. Commun. *9*, 902.

102. Xu, W., Nathwani, B., Lin, C., Wang, J., Karatekin, E., Pincet, F., Shih, W., and Rothman, J.E. (2016). A programmable DNA origami platform to organize SNAREs for membrane fusion. J. Am. Chem. Soc. *138*, 4439–4447.

103. Le, J.V., Luo, Y., Darcy, M.A., Lucas, C.R., Goodwin, M.F., Poirier, M.G., and Castro, C.E. (2016). Probing nucleosome stability with a DNA origami nanocaliper. ACS Nano *10*, 7073–7084.

104. Funke, J.J., Ketterer, P., Lieleg, C., Schunter, S., Korber, P., and Dietz, H. (2016). Uncovering the forces between nucleosomes using DNA origami. Sci. Adv. *2*, e1600974.

105. Hariadi, R.F., Sommese, R.F., Adhikari, A.S., Taylor, R.E., Sutton, S., Spudich, J.A., and Sivaramakrishnan, S. (2015). Mechanical coordination in motor ensembles revealed using engineered artificial myosin filaments. Nat. Nanotechnol. *10*, 696–700.

106. Mao, X., Li, K., Liu, M., Wang, X., Zhao, T., An, B., Cui, M., Li, Y., Pu, J., Li, J., et al. (2019). Directing curli polymerization with DNA origami nucleators. Nat. Commun. *10*, 1395.

107. Funke, J.J., Ketterer, P., Lieleg, C., Korber, P., and Dietz, H. (2016). Exploring nucleosome unwrapping using DNA origami. Nano Lett. *16*, 7891–7898.

108. Funke, J.J., and Dietz, H. (2016). Placing molecules with Bohr radius resolution using DNA origami. Nat. Nanotechnol. *11*, 47–52.

109. Jiang, T., Meyer, T.A., Modlin, C., Zuo, X., Conticello, V.P., and Ke, Y. (2017). Structurally ordered nanowire formation from co-assembly of DNA origami and collagen-mimetic peptides. J. Am. Chem. Soc. *139*, 14025–14028.

110. Jin, J., Baker, E.G., Wood, C.W., Bath, J., Woolfson, D.N., and Turberfield, A.J. (2019). Peptide assembly directed and quantified using megadalton dna nanostructures. ACS Nano *13*, 9927–9935.

111. Buchberger, A., Simmons, C.R., Fahmi, N.E., Freeman, R., and Stephanopoulos, N. (2020). Hierarchical assembly of nucleic acid/coiled-coil peptide nanostructures. J. Am. Chem. Soc. *142*, 1406–1416.

112. Freeman, R., Han, M., Álvarez, Z., Lewis, J.A., Wester, J.R., Stephanopoulos, N., McClendon, M.T., Lynsky, C., Godbe, J.M., Sangji, H., et al. (2018). Reversible self-assembly of superstructured networks. Science *362*, 808–813.

113. Zhang, C., Tian, C., Guo, F., Liu, Z., Jiang, W., and Mao, C.D. (2012). DNA-directed three-dimensional protein organization. Angew. Chem. Int. Ed. *51*, 3382–3385.

114. Nogales, E. (2016). The development of cryo-EM into a mainstream structural biology technique. Nat. Methods *13*, 24–27.

115. Mou, Y., Yu, J.Y., Wannier, T.M., Guo, C.L., and Mayo, S.L. (2015). Computational design of co-assembling protein-DNA nanowires. Nature *525*, 230–233.

116. Praetorius, F., and Dietz, H. (2017). Self-assembly of genetically encoded DNA-protein hybrid nanoscale shapes. Science *355*, eaam5488.

117. Kashiwagi, D., Sim, S., Niwa, T., Taguchi, H., and Aida, T. (2018). Protein nanotube selectively cleavable with DNA: supramolecular polymerization of "DNA-appended molecular chaperones.". J. Am. Chem. Soc. *140*, 26–29.

118. McMillan, J.R., and Mirkin, C.A. (2018). DNA-functionalized, bivalent proteins. J. Am. Chem. Soc. *140*, 6776–6779.

119. Brodin, J.D., Auyeung, E., and Mirkin, C.A. (2015). DNA-mediated engineering of multicomponent enzyme crystals. Proc. Natl. Acad. Sci. USA *112*, 4564–4569.

120. Hayes, O.G., McMillan, J.R., Lee, B., and Mirkin, C.A. (2018). DNA-encoded protein Janus nanoparticles. J. Am. Chem. Soc. *140*, 9269–9274.

121. Subramanian, R.H., Smith, S.J., Alberstein, R.G., Bailey, J.B., Zhang, L., Cardone, G.,

Suominen, L., Chami, M., Stahlberg, H., Baker, T.S., and Tezcan, F.A. (2018). Self-assembly of a designed nucleoprotein architecture through multimodal interactions. ACS Cent. Sci. *4*, 1578–1586.

122. Kim, Y.Y., Bang, Y., Lee, A.H., and Song, Y.K. (2019). Multivalent traptavidin-DNA conjugates for the programmable assembly of nanostructures. ACS Nano *13*, 1183–1194.

123. Strable, E., Johnson, J.E., and Finn, M.G. (2004). Natural nanochemical building blocks: icosahedral virus particles organized by attached oligonucleotides. Nano Lett. *4*, 1385–1389.

124. Cigler, P., Lytton-Jean, A.K.R., Anderson, D.G., Finn, M.G., and Park, S.Y. (2010). DNA-controlled assembly of a NaTl lattice structure from gold nanoparticles and protein nanoparticles. Nat. Mater. *9*, 918–922.

125. McMillan, J.R., Brodin, J.D., Millan, J.A., Lee, B., Olvera de la Cruz, M., and Mirkin, C.A. (2017). Modulating nanoparticle superlattice structure using proteins with tunable bond distributions. J. Am. Chem. Soc. *139*, 1754–1757.

126. Ke, Y., Ong, L.L., Shih, W.M., and Yin, P. (2012). Three-dimensional structures self-assembled from DNA bricks. Science *338*, 1177–1183.

127. Blackman, M.L., Royzen, M., and Fox, J.M. (2008). Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. J. Am. Chem. Soc. *130*, 13518–13519.

128. Li, G., Liu, Y., Liu, Y., Chen, L., Wu, S., Liu, Y., and Li, X. (2013). Photoaffinity labeling of small-molecule-binding proteins by DNA-templated chemistry. Angew. Chem. Int. Ed. *52*, 9544–9549.

129. Huth, J.R., Bewley, C.A., Nissen, M.S., Evans, J.N.S., Reeves, R., Gronenborn, A.M., and Clore, G.M. (1997). The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. Nat. Struct. Biol. *4*, 657–665.

130. Zhou, P., Wang, M., Du, L., Fisher, G.W., Waggoner, A., and Ly, D.H. (2003). Novel binding and efficient cellular uptake of guanidine-based peptide nucleic acids (GPNA). J. Am. Chem. Soc. *125*, 6878–6879.

131. Linkletter, B.A., Szabo, I.E., and Bruice, T.C. (1999). Solid-phase synthesis of deoxynucleic guanidine (DNG) oligomers and melting point and circular dichroism analysis of binding fidelity of octameric thymidyl oligomers with DNA oligomers. J. Am. Chem. Soc. *121*, 3888–3896.

132. Selmi, D.N., Adamson, R.J., Attrill, H., Goddard, A.D., Gilbert, R.J.C., Watts, A., and Turberfield, A.J. (2011). DNA-templated protein arrays for single-molecule imaging. Nano Lett. *11*, 657–660.

133. Geng, C., and Paukstelis, P.J. (2014). DNA crystals as vehicles for biocatalysis. J. Am. Chem. Soc. *136*, 7817–7820.

134. Brady, R.A., Brooks, N.J., Foderà, V., Cicuta, P., and Di Michele, L. (2018). Amphiphilic-DNA platform for the design of crystalline frameworks with programmable structure and functionality. J. Am. Chem. Soc. *140*, 15384–15392.

135. Zhang, F., Simmons, C.R., Gates, J., Liu, Y., and Yan, H. (2018). Self-assembly of a 3D DNA crystal structure with rationally designed six-fold symmetry. Angew. Chem. Int. Ed. *57*, 12504–12507.

136. Simmons, C.R., Zhang, F., MacCulloch, T., Fahmi, N., Stephanopoulos, N., Liu, Y., Seeman, N.C., and Yan, H. (2017). Tuning the cavity size and chirality of self-assembling 3D DNA crystals. J. Am. Chem. Soc. *139*, 11254–11260.

137. Spence, J.C.H. (2017). XFELs for structure and dynamics in biology. IUCrJ *4*, 322–339.

138. Nannenga, B.L., and Gonen, T. (2019). The cryo-EM method microcrystal electron diffraction (MicroED). Nat. Methods *16*, 369–379.

139. Zhao, Z., Zhang, M., Hogle, J.M., Shih, W.M., Wagner, G., and Nasr, M.L. (2018). DNA-corralled nanodiscs for the structural and functional characterization of membrane proteins and viral entry. J. Am. Chem. Soc. *140*, 10639–10643.

140. Zhou, X.X., Fan, L.Z., Li, P., Shen, K., and Lin, M.Z. (2017). Optical control of cell signaling by single-chain photoswitchable kinases. Science *355*, 836–842.

141. Valero, J., Pal, N., Dhakal, S., Walter, N.G., and Famulok, M. (2018). A bio-hybrid DNA rotor-stator nanoengine that moves along predefined tracks. Nat. Nanotechnol. *13*, 496–503.

142. He, Y., Ye, T., Su, M., Zhang, C., Ribbe, A.E., Jiang, W., and Mao, C. (2008). Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. Nature *452*, 198–201.

143. Das, R., and Baker, D. (2008). Macromolecular modeling with rosetta. Annu. Rev. Biochem. *77*, 363–382.

144. Doye, J.P.K., Ouldridge, T.E., Louis, A.A., Romano, F., Šulc, P., Matek, C., Snodin, B.E.K., Rovigatti, L., Schreck, J.S., Harrison, R.M., and Smith, W.P.J. (2013). Coarse-graining DNA for simulations of DNA nanotechnology. Phys. Chem. Chem. Phys. *15*, 20395–20414.

145. Brunette, T.J., Parmeggiani, F., Huang, P.S., Bhabha, G., Ekiert, D.C., Tsutakawa, S.E., Hura, G.L., Tainer, J.A., and Baker, D. (2015). Exploring the repeat protein universe through computational protein design. Nature *528*, 580–584.

146. Huang, P.S., Oberdorfer, G., Xu, C., Pei, X.Y., Nannenga, B.L., Rogers, J.M., DiMaio, F., Gonen, T., Luisi, B., and Baker, D. (2014). High thermodynamic stability of parametrically designed helical bundles. Science *346*, 481–485.

147. Singh, P., Carraher, C., and Schwarzbauer, J.E. (2010). Assembly of fibronectin extracellular matrix. Annu. Rev. Cell Dev. Biol. *26*, 397–419.

148. Freeman, R., Stephanopoulos, N., Álvarez, Z., Lewis, J.A., Sur, S., Serrano, C.M., Boekhoven, J., Lee, S.S., and Stupp, S.I. (2017). Instructing cells with programmable peptide DNA hybrids. Nat. Commun. *8*, 15982.

149. Praetorius, F., Kick, B., Behler, K.L., Honemann, M.N., Weuster-Botz, D., and Dietz, H. (2017). Biotechnological mass production of DNA origami. Nature *552*, 84–87.

150. Stephanopoulos, N., Solis, E.O.P., and Stephanopoulos, G. (2005). Nanoscale process systems engineering: toward molecular factories, synthetic cells, and adaptive devices. AIChE J. *51*, 1858–1869.