Design and analysis of nanoscale bioassemblies

Jarrod Clark, Elizabeth M. Singer, Darlynn R. Korns, and Steven S. Smith

Bionanotechnology is an emerging field in nanotechnology. In general, it uses concepts from chemistry, biochemistry, and molecular biology to identify components and processes for the construction of self-assembling materials and devices. Distant goals of the science of bionanotechnology range from developing programmable nanoscale devices that can sample or alter their environments to developing assemblies capable of Darwinian evolution. At the heart of these approaches is the concept of the production of supramolecular assemblies (SMAs; also known as supramolecular aggregates) by programmed self-assembly in an aqueous medium. Ordered arrays, planar and closed-shell tilings, dynamic machines, and switches have been designed and constructed by using DNA-DNA, protein-protein, and protein-nucleic acid biospecificities. We review the designs and the analytical techniques that have been employed in the production of SMAs that do not occur in nature.

INTRODUCTION

In a recent debate between two of the primary advocates of nanotechnology, K. Eric Drexler and Richard Smalley, Smalley pointed out (1) that the chemistry of nanoscale assemblers, often invoked as central to nanotechnology, will require enzyme-like tools that operate in a liquid medium, most likely water. He further suggested that the products of that approach “cannot be much broader than the meat and bone of biology.” His point is well taken in light of the interest in using nanoscale assemblers for the production of micro-electro-mechanical systems (MEMS) and nanoscale components for electronics. This is because the laws of chemistry require that the assemblers not only juxtapose atoms but also induce or take advantage of complementary electronic structures along a set reaction coordinate in a given medium. Thus, Smalley’s limitation suggests that mechanical nanoscale assemblers, as most often envisioned (2), are unlikely to make short-term contributions to the continuing miniaturization of MEMS and electronic components. In short, this limitation suggests that the continuing success of Moore’s Law (information on Moore’s Law is available at http://www.webopedia.com/TERM/M/Moores_Law.html) will require alternative approaches.

On the other hand, Smalley’s limitation is helpful for understanding the rapidly expanding field of bionanotechnology. Like other forms of nanotechnology, bionanotechnology seeks to define approaches to the fabrication of useful materials and devices. Modern bionanotechnologists use concepts from chemistry, biochemistry, and molecular biology to identify components and processes for the construction of self-assembling materials and devices. Distant goals of the science of bionanotechnology range from developing programmable nanoscale devices that can sample their environments to developing assemblies capable of Darwinian evolution. At the heart of these approaches is the concept of the production of supramolecular assemblies (SMAs; also known as supramolecular aggregates (3)) by programmed self-assembly in aqueous medium.

In this review, we confine our coverage to the area of bionanotechnology that involves the re-engineering of existing multirubin enzyme systems, as well as self-assembly based in the application of an external field, surface-free energy minimization, and similar forces have been thoroughly reviewed elsewhere (4–9). We discuss self-assembly of biological material into SMAs, the enabling methods in use in the analysis of these assemblies, and several of the applications that have been proposed for them. Progress in this area has been rapid in the past few years, with a variety of protein domains and oligodeoxynucleotides being adapted as components for self-assembling SMAs. In addition to the numerous oligodeoxynucleotides (10–16) that have been employed in building these constructs, a variety of proteins including streptavidin (10,17), barnase and barstar (18), influenza virus matrix protein M1 (19), ribosomal protein L9 (19), bromoperoxidase (19), carboxylesterase (19), the bacterial DNA methyltransferases M•HhaI (13), and M•EcoRII (14) have been employed.

DESIGNS

An essential component of an SMA is the linker, which is a molecule that
connects the components to create a new structure. Linkers of several types have been used in these applications. Linkers can take advantage of the specificity of protein domain interactions or nucleic acid complementarity to produce self-assembling SMAs. In general, when protein domain interactions are used, the assemblies are composed almost exclusively of protein (17). Likewise, when nucleic acid complementarity is employed, the assemblies are composed almost exclusively of nucleic acid (15). When SMAs are formed by using both protein and nucleic acid interactions, it becomes possible to use the nucleic acid as a scaffold to order proteins. In this case, the linker can be either a nucleic acid (10,20) or a protein (13).

**Protein-Protein SMAs**

The streptavidin-biotin system has been used to link proteins together (21,22) or to join duplex DNA into networks (23). Streptavidin is a tetrameric protein that displays biotin binding sites in a roughly tetrahedral arrangement (24). It has been used successfully to display SMAs with four engineered antibody Fv fragments as protein fusions (Figure 1) linked by the four streptavidin subunits (17). The system can yield a nearly 35-fold enhancement in antigen binding avidity. The RNase barnase and its proteinaceous inhibitor barstar form a tightly interacting system that has been used to link barnase-antibody fusions into dimeric and trimeric SMAs (18), much in the same way that streptavidin has been used in that application (17). Here again, an enhancement in antigen binding avidity was observed (18). In the barnase-barstar system, two or three antibody fragments are displayed in a flexible dimer or trimer, while the streptavidin arrangement is fixed and roughly tetrahedral. Both of these multivalent complexes have potential in medical applications (17,18).

Additional symmetries can be obtained by constructing fusion proteins (18,19) that display protein-protein interaction domains. Unlike the streptavidin system, which displays fused functionalities as tetramers (17) or can provide a node of degree 4 in a biotin-linked network (23), fused protein-protein interaction domains can be used to form linked arrays (18) and filaments (18,19). Moreover, fusion proteins designed for dimer and trimer bonding (Figure 2) have been used to construct a spherical shell comprising 12 subunits in tetrahedral symmetry (19). The principles of spherical shell assembly (25) suggest that these systems should also be capable of forming spherical shells comprising 24 identical subunits in octahedral symmetry and 60 identical subunits in icosahedral symmetry if appropriate angles between the fusions can be obtained. However, the stability of these systems is also highly dependent on the protein-protein interaction affinity of the subunit domains (19).

L-rhamnulose-1-phosphate aldolase (rhuA) can serve as a protein network node of degree 4 (21). In this case, three mutations were introduced in each subunit of tetrameric aldolase to create two biotin binding sites for a total of eight biotins per rhuA molecule. Each subunit was fused to a His6-tag at the C terminus so the aldolase tetramer could be bound to a planar surface. Streptavidin tetramers could interact with each of two biotins on the rhuA surface. Because the tetramer was rigid enough to prevent streptavidin interaction around a corner, the system tiled the plane in a roughly Cartesian pattern as seen in negatively stained electron micrographs (21). More complex patterns could be obtained by mixing bis-

---

**Figure 1.** A multivalent supramolecular assemblies for Fv antibody display. Fusion constructs of streptavidin (STV) and Fv antibody fragments display antibody binding sites in a roughly tetrahedral arrangement.

**Figure 2.** Construction of tiled shells using fusion proteins that display protein-protein interaction domains. Tiles can be formed from fused protein domains (19). (A) When the angle between the fusion domains is 180°, the tiles assemble along monovalent or divalent interfaces into extended structures as shown (arrows). (B) Restricting the angle between the fusion proteins to approximate that required for tetrahedral symmetry allowed tiling of a closed shell of 12 subunits.
biotin-labeled streptavidin molecules as building blocks.

An extended network was created using a fusion between streptavidin and S-layer protein (22), which self-assembles into a monomolecular lattice. The streptavidin-S-layer complex assembled into an array and displayed the streptavidin, which was capable of binding biotinylated proteins as seen in transmission electron microscopy (TEM). This is a unique approach to immobilize compounds in an ordered spacing on a surface. However, the networks created by the streptavidin-biotin system are not homogenous because neither the level of biotinylation in proteins nor the binding of biotin to all four sites in streptavidin is uniform. In addition, it is difficult to analyze products and remove impurities. To create uniform arrays, a different method of assembly will be required.

In addition to assembling SMAs of proteins, protein-protein interactions have been used to assemble colloidal gold particles into SMAs (26). In this case, the leucine zipper-forming capacity of synthetic peptides arrayed on the surface of the gold particles has been used to associate the gold particles into SMAs in a fashion analogous to the use of complementary DNA sequences in a similar application (27).

### DNA-DNA SMAs

Although the conformation space occupied by stable and metastable single-stranded DNAs of a given length is demonstrably smaller than that occupied by stable and metastable proteins or single-stranded RNAs of the same length (28), it is still quite large, as shown by studies with DNA aptamers (29). However, single-stranded DNA conformation has not been extensively exploited (30) as a linker system in the formation of SMAs. By and large, the capacity of duplex DNAs to form stable or immobile junctions has served in this role (12,16). While the Holliday, or four-arm junction, is the best-known of these in biology, junctions with as many as six arms have been studied (31,32). In principle, these structures permit the formation of linear or branched three-dimensional (3-D) SMAs.

Perhaps the most celebrated example of an SMA formed exclusively of nucleic acid is the DNA cube (15). The remarkable plasticity of DNA has been further explored in the construction of molecules with the connectivity of certain platonic solids (33), and the connectivity of the three interlaced circles that form a Borromean ring knotted topology (34,35). Extended 2-D planar arrays of DNA have been produced with a form of tiling with synthetic DNA double-crossover molecules. These building blocks are created by using sticky ends that associate by Watson-Crick complementarity. They are spatially constrained by the crossover interconnections into planar units with complementary edges suitable for association during tiling (12). Specific periodic (12) subperiodic (36), and aperiodic (37) patterns have been generated. In one study, a DNA hairpin was introduced into one of the double-crossover molecules to help visualize the pattern (12) by atomic force microscopy (AFM). Specific periodic patterns were observed at the nanometer scale, but the existence of other structures suggested that side products were also formed. In principle, if a pure preparation of tiles attached to a molecule of interest could be obtained, that molecule could be arranged in the observed tiling.

DNA-DNA interaction has also been used to link inorganic moieties into SMAs. Extended aggregates of nanoparticles have been formed based on mixing modified particles linked to spherically arrayed oligodeoxynucleotides (27,38). In a manner similar to that envisioned for the assembly of the earliest prebiotic SMAs (39), ordered
patterns have also been obtained with DNA alone by simply attaching observable moieties (e.g., gold nanocrystals) to single strands of DNA and allowing them to assemble at set positions on a complementary strand (20).

DNA has also been used in the construction of several devices. One system (Figure 3) used a modification of the double-crossover molecules to capitalize on the B-Z transition in DNA to generate a salt-dependent switch (40). Telemetry on the position of the arms of the switch was obtained by using resonance energy transfer between dye molecules at the ends of each double crossover. Another DNA-DNA nanodevice used the concept of strand displacement to alter the position of dye molecules at the ends of a DNA strand in a cyclic fashion. A molecular tweezer structure was formed by the hybridization of three oligodeoxynucleotide strands with overhanging sticky strands. A fourth (displacing) oligodeoxynucleotide hybridized with the sticky ends caused the tweezers to close the system. A fifth strand was then added to reopen the system, and the process was repeated. Cycling was visualized by polyacrylamide gel electrophoresis (PAGE), where the open and closed forms run at different positions. The strands were labeled with two different fluorescent dyes at the 3′ and 5′ ends so the motion of the device could be monitored by resonance energy transfer in real time (11). Similar devices based on DNA crossover (41) or DNA duplex-quadruplex interconversions have also been described (42).

**Protein-DNA SMAs**

Streptavidin has the capacity to bind four biotins, and it should therefore be able to act as a linker between four biotinylated DNA fragments. In principle, the streptavidin molecule could serve as a node of degree 4 in forming an SMA by connecting an extended network of DNA oligodeoxynucleotides (Figure 4). In practice, however, gel electrophoresis and AFM detected only two or three streptavidin connections (43) in circularized or networked oligodeoxynucleotides, suggesting that steric hindrance prevents the nodes from accommodating four DNA termini.

DNA linkers can also be biotinylated, and streptavidin can be used to couple biotinylated proteins to linkers with different DNA sequences. These DNA streptavidin-protein complexes can then be ordered along a guide DNA strand (10,44), forming ordered linear assemblies (Figure 5) that mimic those thought to have been components of prebiotic metabolosomes (39). Here the DNA serves as a scaffold for the assembly of the linked proteins.

DNA (cytosine-5) methyltransferases can also be used as linkers in constructing ordered nucleoprotein-based SMAs (13). This system takes advantage of the sequence specifici-

![Figure 5. Ordered linear supramolecular assemblies using a DNA linker.](image)

The DNA-streptavidin protein complexes can be ordered along a DNA guide strand by Watson-Crick interactions with single-stranded linker DNAs. Biotinylated proteins couple to the DNA streptavidin (STV) linker most often with a 1:1 stoichiometry of protein to streptavidin, although the 2:1 and 3:1 stoichiometries shown can occur, suggesting that obtaining a homogenous 1:1 stoichiometry would require further purification (10,44).

![Figure 6. Ordered supramolecular assemblies (SMAs).](image)

(A) An ordered linear SMA using DNA methyltransferases (MTase) as linkers. DNA methyltransferase fusion proteins (FP) are ordered via a covalent bond formation at methyltransferase recognition sites carrying 5-fluorocytosine at the target base in the sequence (13,45). (B) Ordered multi-arm SMAs using DNA methyltransferase fusion proteins as a linker. DNA methyltransferase fusion proteins are covalently linked at methyltransferase recognition sites, containing 5-fluorocytosine along the arms of a Y-junction DNA, creating an ordered multi-arm SMA (14).
ties of the known DNA (cytosine-5)-methyltransferases to target functional proteins in a specified way along a DNA scaffold. The mechanism of action of these enzymes makes it possible to covalently trap them at multiple sites on duplex DNA structures to create ordered assemblies. Multiple methyltransferases have been targeted to a single DNA molecule by placing multiple recognition sites in linear (Figure 6A) or branched (Figure 6B) DNA structures (13,14,45). In addition, three different methyltransferases (HhaI, EcoRII, and MspI) have been targeted to their specific recognition sites on a single DNA structure (13,14). Methyltransferase peptide fusion proteins have been successfully placed on a DNA scaffold and detected by standard gel electrophoretic mobility shift assay (EMSA) of 32P-labeled DNA (13,14), microfluidics chip-based EMSA (14), and immunologically with Western blot analysis using an antibody to the fusion peptide (13). Microfluidics analysis of the gel shift associated with a given SMA makes it especially easy to determine whether the selected target sites are completely occupied.

This methyltransferase technique has several advantages over other DNA-protein assembly technologies. The linkage between the DNA and methyltransferase is covalent, so the assembly is very stable; the linking reaction takes place at 37°C, where most proteins are stable and the fusion proteins can be placed at preselected sites on a DNA scaffold of almost arbitrary structure. The most significant advantage is the possibility of creating methyltransferase fusion proteins containing peptides that bind to receptors for directed targeting of the nanostructure (46). Immobilization of linking proteins on two different Y-junctions could also lead to the formation of a carcerand (46,47). For this system to be used successfully, the protein fused to the methyltransferase must be expressible as a fusion, and it is also necessary to produce significant amounts of the building blocks. DNA methyltransferases are not unique in their capacity to form covalent complexes with DNA. However, their sequence specificity in this application does set them apart. Of the several non-methyltransferase DNA binding proteins now known, only the Cre recombinase from bacteriophage P1 has well-characterized sequence specificity and the capacity to produce a covalent link to DNA. It recognizes the 34-bp loxP sites without the requirement of accessory proteins shown by other members of the tyrosine recombinase family (48). It can be trapped on DNA as a covalent complex by presenting a symmetrical (loxA) recognition site containing 3' DNA nicks that the enzyme recognizes. In this case, the enzyme forms a stable covalent intermediate that does not undergo further reaction (49).

### ANALYTICAL METHODS

The nano-realm (less than 100 nm) is difficult to visualize. Bionanotechnology has therefore employed several of the established techniques of molecular science and is fueling the development of several others (Table 1).

### Computer-Aided Design

X-ray crystallography has contributed numerous 3-D structures to molecular science. Reference to the constraints...
imposed by the 3-D structure of the components in a contemplated geometric system is an invaluable approach to determine what is possible and what can be expected when biomolecules or their domains are employed (13,17–19,21). More detailed molecular modeling of predicted SMAs has also proven useful (13,14,18,19,45), as has electronic structure calculation (13).

Centrifugation

The use of equilibrium sedimentation has been employed to determine the molecular mass of SMAs (19). Equilibrium density gradient centrifugation (50) and velocity gradient sedimentation (51) have not yet been employed in studies of nanoassemblies. Given the success of these methods in the separation of multisubunit proteins, topoisomers of DNA, phage particles, and subcellular organelles, it is reasonable to expect that they will find useful application in both the characterization and purification of SMAs.

Gel Electrophoresis

Gel electrophoresis has been one of the most widely applied methods in the analysis of SMAs. The ability to size and identify structures in a quick and efficient manner, at minimal cost for reagents and equipment, has made this a key technique in bionanotechnology. Polyacrylamide gels have been used in the identification of the various complexes formed by directed assembly. In the characterization of a nanoscale camshaft (45), PAGE played a pivotal role in the detection of the various forms of decorated duplex DNA. In the same manner, Niemeyer et al. (10) used electrophoresis to interpret the architecture of their streptavidin-biotin complexes. Likewise, Yan and others (41) were able to resolve the components of their complex B-Z DNA switching device and to define the morphology created in their lattice designs (12). Although gel electrophoresis can identify the intermediates in assembling an SMA, because of the equivalency of intermediates in symmetrical systems, it does not provide a clear view of correct component placement or how closely the SMA resembles a given molecular model (14,45). Moreover, the topological edge connectivity of a platonic solid can be demonstrated with gel electrophoresis, but the geometric shape of the edges cannot (15).

Microfluidics

With the advent of microfluidics chip-based capillary electrophoresis, the speed and resolution with which one can identify structures has been taken a step further. This system furnishes all of the information provided by gel electrophoresis with significant savings in time and sample volume. Data are digital and therefore easily analyzed. Clark et al. (14) have successfully used this technique to quickly screen for the accurate formation of nucleoprotein assemblies by using microfluidics-based EMSA. The microfluidics system (14) also provided improved resolution over previously employed EMSA methods (10,13).

Resonance Energy Transfer

Resonance energy transfer has been put to good use in the analysis of dynamic SMAs (11,40,42). In this technique, energy is transferred from a donor dye molecule to an acceptor dye molecule. It is often detected by the appearance or the quenching of fluorescence in the two-dye system, often referred to as Förster or fluorescence resonance energy transfer (FRET). Resonance energy transfer is dependent on the inverse of the sixth power of the intermolecular separation and is very efficient within the Förster radius (2–6 nm), which is on a scale that can be achieved with most biological macromolecules (52). If the donor wavelength is quenched, then the observed structures are within a Förster radius. Resonance energy transfer is attractive because it is inexpensive and somewhat impervious to contaminating molecules. Moreover, once a protein or DNA molecule is labeled, detecting the signal can be achieved swiftly. Instead of merely viewing the stages of assembly, a signal indicates the expected function. However, identification of the moving component is difficult to determine solely by this method.
Surface Plasmon Resonance

Resonance energy transfer can also occur between the evanescent wave produced by the electromagnetic component of reflected light and surface plasmons (collective excitations of electrons at the interface between a conductor and an insulator), thus reducing the intensity of reflected polarized light at a specific angle of incidence. Surface plasmon resonance (SPR) can be used to measure the interaction between a protein and its ligand. SPR is induced when light is reflected from a chip containing a thin layer of gold overlaid with glass. A specific ligand for the molecule of interest is bound to the gold layer. When molecules from the sample bind to the immobilized ligand, the local concentration changes and the alteration in the SPR response is measured. Plotting the SPR changes over the course of the experiment can give a quantitative measure of the interaction between the ligand and the protein being analyzed. Binding constants can be obtained from either rate constant measurements or the steady state level of binding as a function of sample concentration. SPR has been used to measure the increase in antigen binding avidity due to the formation of SMAs with multiple valences. Increased binding avidity was observed in the SMA containing a streptavidin with four linked antibodies (17) and in the barnase-barstar system, which displayed three antibodies (18). However, for multivalent SMAs, only the avidity constants could be measured accurately.

Electron Microscopy

Electron microscopy has been used to view protein filaments, closed shells, and 2-D assemblies (19,21), and has also been used to display large networks of streptavidin and their fused macromolecular components (22). Electron microscopy requires that the samples be exposed to nonphysiological conditions. Electrons operate efficiently in a vacuum, so the sample must be exposed to a very high vacuum to be visualized. In TEM, the samples must be very thin (50–100 nm) for the electrons to pass through them, unlike scanning electron microscopy (SEM), which measures the reflection of electrons. In TEM, contrast depends on atomic number. Because most biological samples are composed of low atomic numbers, they are often stained with the salts of heavy metals and subsequently dried. Staining can often be avoided, especially with cryo-TEM, where low doses of electrons are used to build up an image without staining. However, when staining is used, distortion of the original structure can limit the analysis of the image.

Atomic Force Microscopy

The most widely used method for visualizing and probing nanostructures is AFM, which is also known as scanning force microscopy (SFM). This method was developed in the mid 1980s (3) and has been used in bionanotechnology as an alternative to electron microscopy. AFM does not require a current that can destroy certain samples and therefore can image insulators, which can be difficult in electron microscopy. In general, AFM is used to obtain a topographical map of a given sample. By adjusting the force between the tip and sample, non-destructive imaging can be achieved. This method has been widely applied in the exploration of SMAs. Individual nanoscale devices have been visualized (23,41) as well as large 2-D planes (12,21). Moreover, AFM can be used to explore the shape of an assembly. For example, analysis of curvature in putatively linear assemblies such as the nanoscale camshaft (45) or other assemblies ordered on linear DNA (10) by gel electrophoresis would require concatenation of complex assemblies or the use of Ferguson plots involving multiple electrophoretic procedures. AFM is an alternate and more visual approach, while microscopy often provides a more accessible approach (23).

Each of these techniques is valuable in the analysis of SMAs, but when combined, they offer a relatively circumstantial picture of the structure at hand. Practical considerations associated with the nature of each SMA and the desired structural information will dictate which of these tools can be best applied in structural analysis.

ACKNOWLEDGMENTS

This work was supported by grant no. 0395 from the Smokeless Tobacco Research Council; by National Institutes of Health (NIH) grant no. 1G08LM06722-01A1; from the National Library of Medicine, by NIH grant no. CA 91234 from the National Cancer Institute; and by a grant from the Ella Fitzgerald Foundation (to S.S.S.).

REFERENCES

14. Clark, J., T. Shevchuk, P.M. Swiderski, R.


Address correspondence to:
Jarrod Clark
Kaplan Clinical Research Laboratory
City of Hope National Medical Center
1500 E. Duarte Road
Duarte, CA, USA
e-mail: jclark@coh.org