ARTICLE

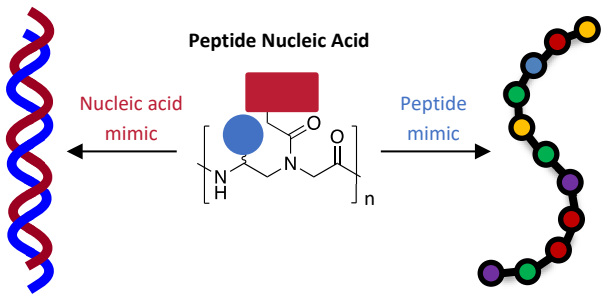
Peptide Nucleic Acids Harness Dual Information Codes in a Single Molecule

Colin S. Swenson,[a] Jennifer M. Heemstra*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Nature encodes the information required for life in two fundamental biopolymers: nucleic acids and proteins. Peptide nucleic acid (PNA), a synthetic analog comprised of nucleobases arrayed along a pseudopeptide backbone, has the ability to combine the power of nucleic acids to encode information with the versatility of amino acids to encode structure and function. Historically, PNA has been perceived as a simple nucleic acid mimic having desirable properties such as high biostability and strong affinity for complementary nucleic acids. In this feature article, we aim to adjust this perception by highlighting the ability of PNA to act as a peptide mimic and showing the largely untapped potential to encode information in the amino acid sequence. First, we provide an introduction to PNA and discuss the use of conjugation to impart tunable properties to the biopolymer. Next, we describe the integration of functional groups directly into the PNA backbone to impart specific physical properties. Lastly, we highlight the use of these integrated amino acid side chains to encode peptide-like sequences in the PNA backbone, imparting novel activity and function and demonstrating the ability of PNA to simultaneously mimic both a peptide and a nucleic acid.



Introduction

Proteins and nucleic acids are two of the essential building blocks of life, necessary for both cellular information storage and biological function. Nature has evolved these biopolymers over millions of years to create and refine their properties. The amino acid sequences in peptides and proteins encode intricate structural information that enables complex functions, and researchers have investigated and probed these codes in order to appropriate their function for a wide range of applications.[1–3] In parallel, nucleic acids are virtually unmatched in their capacity to store retrievable information necessary for life and the nucleotide code offers exceptionally predictable and precise molecular recognition properties. Scientists have extensively studied and utilized this recognition ability for naturally occurring nucleic acids,[4,5] and many different forms of unnatural nucleic acids have been developed to improve upon the physical properties of canonical DNA and RNA.[6,7]

Peptide nucleic acid (PNA) is an interesting and powerful nucleic acid analogue developed by Nielsen and coworkers in which the negatively charged phosphate backbone is replaced with a neutral pseudopeptide backbone (**Figure 1**).[8] This peptide-like backbone, combined with the ability to undergo complementary Watson-Crick-Franklin hybridization, is responsible for a number of

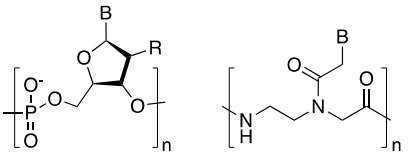


Figure 1: Chemical structures of DNA, RNA, and PNA.

desirable properties such as improved binding affinity, enhanced specificity, and strong resistance to chemical and enzymatic degradation.[8–12] Therefore, it is not surprising that PNA has garnered significant interest in a number of research areas. For example, as a tool in medicine and biology, therapeutic and antimicrobial PNAs can be designed to regulate gene expression through antisense interactions or nuclease-like activity.[13–18] PNA has been used as a delivery vehicle for nucleic acid therapeutics and PNA-based approaches to gene editing have also been described.[19–23] In molecular imaging and diagnostics, the enhanced recognition properties of PNA are promising for the development of methods for detection and *in situ* visualization of DNA and RNA.[24] Each of these uses for PNA have been extensively reviewed and rely solely upon the ability of PNA to function as a nucleic acid analogue.[25–31] Looking beyond the ability of PNA to undergo Watson-Crick-Franklin base pairing, the pseudopeptide nature of the PNA backbone also allows for direct inclusion of amino acid functionality, and researchers have

[a] *Department of Chemistry, Emory University, Atlanta, Georgia 30322, USA*

utilized this capability to introduce side chains that enhance solubility or cellular uptake.[32,33] We recognized that the amino acid sequence could also be used to encode information to direct assembly, opening the door to the use of PNA as both a nucleic acid and a peptide analogue.

In this feature article, we discuss ways in which the structure of PNA has been synthetically altered to afford increased functionality and structural information. Specifically, we first highlight examples of PNA conjugates that include terminally-appended amino acids or lipids to create peptide-like mimics or impart structural changes. We then focus on modifications that have allowed the insertion of amino acid-like side chains in the PNA backbone, including their design, synthesis, characterization, and functional properties. Finally, we discuss examples of PNA as peptide mimics and highlight a recent advance describing a PNA that can simultaneously encode nucleic acid and amino acid information in a single biopolymer to impart specific structure and function.

Conjugation of PNA

PNA suffers from inherent weaknesses such as low solubility and cell permeability, which hinder its use in many biomedical applications.[34,35] In order to overcome these limitations, researchers have developed strategies for conjugating a variety of functional groups to the termini. PNA benefits from high chemical stability and is synthesized through standard solid-phase peptide synthesis (SPPS) procedures employing Fmoc- or Boc-based strategies.[35] Therefore, it is relatively easy to functionalize the termini with modifiers that can be added using the same amino acid chemistry, including peptides and hydrophobic moieties such as lipids or alkyl chains.[36]

PNA Conjugated to Peptides

A simple method to increase the solubility of PNA is to include charged amino acid residues through the solid-phase synthesis procedure.[35] Incorporating one or more lysine or arginine residues at either terminus results in enhanced solubility, presumably due to an increase in charge and therefore hydrophilicity. The simplicity of adding additional amino acids during synthesis has led to a number of PNA-amino acid conjugates having improved solubility.

Several molecules having potential biological applications lack inherent cell permeability, and a common approach for improving the cellular uptake is to attach cell penetrating peptides (CPPs).[37] These cationic peptides facilitate movement across cell and tissue barriers that would otherwise prevent uptake. Covalent attachment of these peptides to PNA results in CPP-PNA conjugates that have improved cellular uptake, and thus increased efficacy in biological applications such as antisense gene regulation (**Figure 2**).[38] As an example, Nielsen and coworkers observed the uptake of PNA conjugated to pAnt (a 17-residue fragment of the protein antennapedia) and pTat (a 14-residue fragment of HIV Tat protein) in five cell types (SK-BR-3, HeLa, IMR-90, U937, and H9), whereas unmodified PNA showed no uptake.[39] However, in these cases the PNA primarily localized to vesicular compartments and did not escape endosomal entrapment.

In order to ascertain the cellular internalization of PNA, Bendifallah *et. al.* employed a luciferase-based activity assay for CPP-

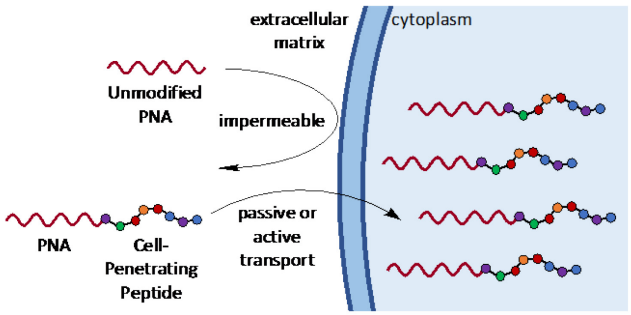


Figure 2: Modification of PNA with a cell-penetrating peptide increases cell permeability via passive or active transport.

PNAs.[40] Using a series of unique conjugation techniques, they examined seven distinctive CPPs including a nuclear localization sequence (NLS) and a simple stretch of positively charged arginine residues ($R_{7\text{-}9}$). Transportan-conjugated PNA as well as the arginine-based conjugates displayed increased activity. Interestingly, delivery efficacy was dependent on the chemical nature and position of the covalent linker. An NLS sequence had previously been shown to increase uptake, likely due to the increased positive charge.[41] However, in the aforementioned study, the NLS-PNA conjugates showed minimal or no activity. This could be explained by the difference in cell types and targets used for each study. An expanded number of CPP-PNAs and their delivery efficiency in different cell types have been studied and reported but further investigation is required to fully elucidate the relationship between identity and efficacy.[42–44]

The attachment of CPPs to PNA resulted in an abundance of knowledge regarding methods to fuse peptides to PNA to enhance the properties of this nucleic acid analogue for biomedical applications. However, in these examples the peptide and PNA portions are designed to act autonomously, each imparting their own chemical and biological properties independent of the other. Looking beyond this, researchers envisioned PNA conjugates in which the properties of each segment modulate and enhance the properties of the architecture as a whole, and to this end developed amphiphiles in which the PNA and appended functionality serve to direct assembly.

Amphiphilic PNA Conjugates

Peptide amphiphiles are an attractive class of molecules that are capable of assembling into supramolecular architectures at the nanoscale.[45,46] These typically consist of a short hydrophilic peptide sequence ligated to a hydrophobic domain such as a lipid or alkyl chain.[47] PNA comprised of a neutral polyamide backbone and canonical nucleobases is inherently hydrophobic, but strategic placement of amino acids and hydrophobic lipid or alkyl domains can induce amphiphilic behavior and promote controlled assembly analogous to that of peptide amphiphiles (**Figure 3A**).[48] Stupp and coworkers demonstrated this phenomenon through the synthesis of an amphiphilic PNA-peptide conjugate.[49] A β-sheet promoting peptide sequence (KGGGAAAK) followed by a palmitoyl alkyl chain was conjugated to the C-terminus of a polythymine PNA heptamer using SPPS methods. Transmission electron microscopy (TEM) and circular dichroism (CD) revealed the formation of uniform nanofibers under aqueous conditions (pH > 7), consistent with similar peptide amphiphiles. Thermal denaturation experiments revealed a significant increase in affinity and specificity for the fully matched

bly

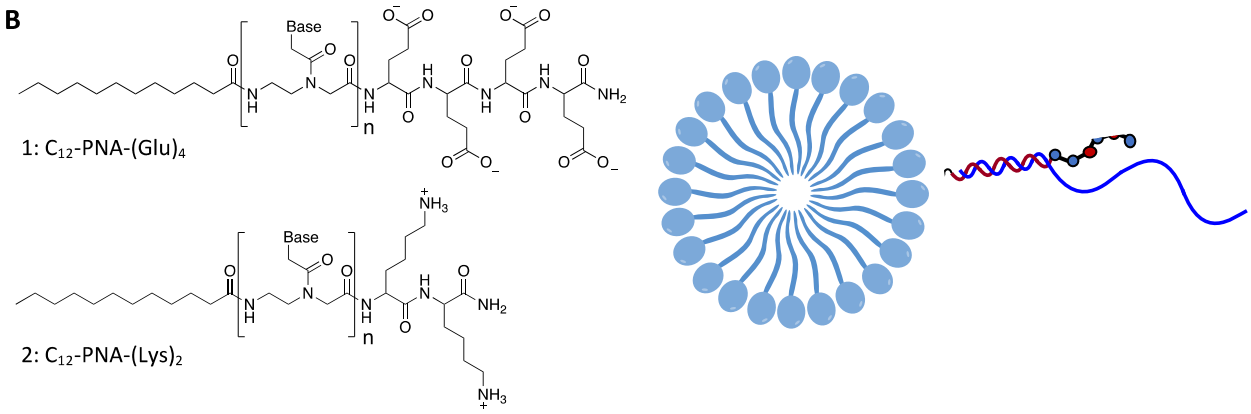


Figure 3: (A) Structure and self-assembly of amphiphilic PNA conjugates having a PNA segment between a hydrophobic domain and a hydrophilic peptide. (B) Chemical structures of (1) negatively charged and (2) positively charged PNA amphiphiles. (C) TX-100/PNA micelles can be used for purification of DNA PCR products. Adapted with permission from Grosser, S.T. et al. *ACS Macro Lett.* **2014**, *3* (5), 467–471. Copyright 2014 American Chemical Society.

polyadenine DNA when self-assembled into nanostructures as opposed to acting as free PNA in solution. Taken together, these results are promising for applications such as biosensing and purification, which require high specificity and strong binding to oligonucleotide targets.[49]

Amphiphiles wherein the PNA oligomer is positioned between a hydrophilic peptide segment and a hydrophobic domain have also been described (**Figure 3A**). Schneider and coworkers produced a series of PNA amphiphiles having varied PNA sequences, hydrophilic peptides, and alkyl chains.[50,51] Specifically, they explored the use of positively charged lysine and negatively charged glutamic acid as the hydrophilic head groups (**Figure 3B**). As expected, the longer the PNA sequence and alkyl chain, the greater the number of charged amino acids required for solubility. Inclusion of negatively charged glutamic acids produced PNA amphiphiles having a lower critical micelle concentration (CMC) compared to the positively charged lysines. Additionally, the position of PNA in the center of the constructs did not affect the binding stability or selectivity for a complementary DNA sequence.[50] These studies guided the design of DNA-binding liposomes through the use of a di-alkyl PNA amphiphile.[52] Incorporating the di-alkyl hydrophobic domain allowed co-extrusion of amphiphiles with phospholipids, enabling the PNA sequence to hybridize to target DNA for the development of highly sensitive bioanalytical devices. Later work showed that through co-assembly with a common surfactant, Triton X-100, micelles capable of binding to specific nucleic acid oligomers are formed (**Figure 3C**).[53] Utilizing this design, purification of single-stranded PCR products of varying lengths was accomplished by micellar electrokinetic chromatography in which the PNA-surfactant micelle acts as a "drag-tag" to shift the mobility, providing a rapid and gel-free DNA separation method.

In more recent work, Zhang and coworkers explored the assembly properties of semi-self-complementary PNA amphiphiles equipped with a hydrophobic alkyl chain and a negatively charged glutamic acid head group.[54] Their results indicated the ability of PNA:PNA duplexes to exist within a micellar architecture, presumably the result of a stabilizing or "cross-linking" effect through hybridization and stronger base stacking interactions. Although applications for this structural phenomenon are less immediately apparent, the authors suggest potential uses for this technology in oligonucleotide purification, gene-related therapeutics, and biomedical detection devices.

The chemical stability and synthetic versatility of PNA make it a promising candidate for the design of myriad conjugates having unique physical and structural characteristics. To this extent, researchers have taken advantage of the ease of covalent modification to install amino acids and lipids within PNA constructs. However, these designs require the addition of information-poor hydrophobic domains and long strings of charged amino acids that act independently of the inherently information-rich PNA segment. Furthermore, current strategies focus on utilizing the nucleic acid information of PNA while ignoring its peptide-like properties. A design that incorporates functional groups similar to amino acid side chains to alter the properties of the PNA oligomer itself would

benefit from the dual identity of PNA as both a nucleic acid and a peptide.

Functional Modifications to the PNA Backbone

The PNA scaffold is highly versatile, allowing the addition of many different functional groups through post-synthetic conjugation. Alternatively, functional groups may also be incorporated into the backbone through the use of modified PNA monomers in oligomer synthesis, and this approach offers the benefit of incorporating these modifications along the full length of PNA rather than just at the termini.

Synthesis and Properties of Backbone-Modified PNA

PNA monomers can be synthesized in a few steps and on a large scale using relatively inexpensive starting materials.[55] Typically, the amine-protected backbone portion of the monomer is synthesized and then coupled to a nucleobase acetic acid to produce monomers compatible with SPPS. The PNA backbone presents a desirable location for modifications, as it can be synthesized using amino acid precursors, allowing a number of different functional groups to be included.[56] For this reason, the PNA backbone has intrigued researchers and inspired the design of many modified PNA monomers. Specifically, chiral PNA monomers have been developed to include modifications at the α-, β-, and γ-positions, each with unique synthetic routes and resulting oligomer properties (**Figure 4**).[57]

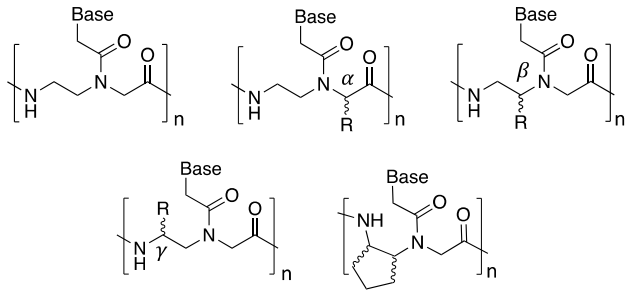


Figure 4: Chemical structures of unmodified PNA, α-PNA, β-PNA, γ-PNA, and cyclic PNA.

α-Modified PNA (α-PNA) was the first backbone-modified PNA analogue to be designed and was reported by Nielsen *et. al.* In this initial example, glycine was replaced with L- or D-alanine during monomer synthesis.[58] The D-form was discovered to hybridize to complementary DNA with slightly higher affinity than the L-form, displaying a preference for a specific stereochemistry. Later functional groups incorporated at the α-position included lysine, serine, and arginine, the latter of which showed enhanced cellular uptake compared to unmodified PNA, similar to that of the TAT transduction domain peptide.[33,59,60] Bulky hydrophobic groups such as phenylalanine, tryptophan, valine, and leucine have also been incorporated at this position, displaying the versatility of α-PNA for functionalization.[61]

The development of β-modified PNA (β-PNA) is a relatively new field, likely due to the significant synthetic barriers compared to that of α-PNA. However, monomers containing methyl groups at the β-position have been produced and their hybridization properties explored.[62] Cyclic PNA monomers containing chiral centers at both the β- and γ-positions have also been synthesized.[63–68] Cyclopentane-derived PNAs have been shown to be promising diagnostic agents because of their improved binding affinity and selectivity for DNA or RNA.[63,64,67,68]

γ-Modified PNA (γ-PNA) is perhaps the most widely explored and promising candidate for including functional groups in the backbone, and modifications have been explored to improve a range of physicochemical properties such as solubility, cell permeability, and hybridization stability. The first γ-PNA monomer was synthesized by reacting an alanine amino acid aldehyde with glycine through reductive amination.[69] Similar to α-PNA, the identity of the stereocenter was found to be important for hybridization affinity to complementary DNA. γ-PNA synthesized from L-amino acids exhibited a significantly higher affinity than those produced from D-amino acids, indicating that stereochemical purity is important for binding affinity.[70] Understanding this significance, Ly and coworkers developed a synthetic route based on a Mitsunobu coupling reaction to produce optically pure γ-PNA for improvements towards antigene applications.[71] Regardless of the route, γ-PNA has been produced containing a variety of hydrophilic and hydrophobic side chains, including lysine, guanidine, mercaptomethyl, hydroxymethyl, sulfate, aspartic acid, miniPEG, valine, leucine, and phenylalanine.[32,33,72–77] This structurally diverse range of functional groups able to be incorporated at this position enables researchers to fine-tune specific physical characteristics for many applications.

PNA hybridizes to single- or double-stranded DNA in different arrangements depending on the sequence and backbone modifications (**Figure 5**).[8,10,78–83] Cytosine-rich homopyrimidine PNAs have been shown to bind double-stranded DNA in a triple helix, or triplex, configuration through Hoogsteen base pairing.[78] Alternatively, thymine-rich homopyrimidine PNA was shown to bind in a triplex wherein two PNAs hybridize to a single homopurine DNA strand by Watson-Crick-Franklin and Hoogsteen base pairing.[79,80] This occurs through a strand invasion process in which one DNA strand in the duplex is displaced by the competing PNA strands.[8,80] It was later discovered that single mixed purine and pyrimidine PNA strands are also capable of forming stable Watson-Crick-Franklin duplexes with complementary DNA in a 1:1 ratio.[10] However, a single unmodified PNA of mixed purine and pyrimidines is incapable of invading double stranded linear DNA.[81] It is possible to form double duplexes through the use of two PNA strands targeting separate DNA strands in the duplex, but this approach requires the use of

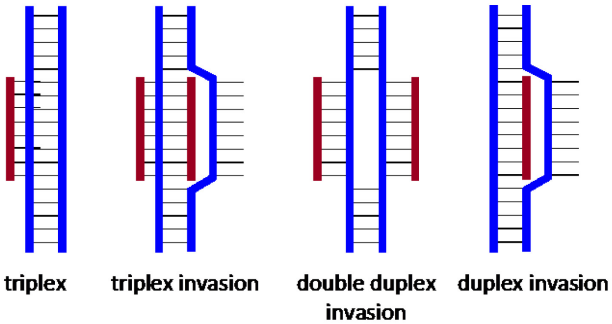


Figure 5: Hybridization modes of PNA to complementary nucleic acids. Adapted from Lohse, J. et al. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (21), 11804–11808. Copyright 1999 National Academy of Sciences.

pseudocomplementary nucleobases and multiple PNA strands, complicating synthesis and design.[82] Including modifications in the backbone significantly increases binding affinity and allows a single PNA to target double stranded DNA or RNA in a sequence-specific manner through a strand invasion mechanism. As an example of this, incorporating a methyl group at the γ-position in the PNA scaffold revealed the formation of PNA:DNA duplexes by strand invasion.[83] Interestingly, the backbone modified PNA hybridized to the DNA solely in a double helix arrangement, and no evidence of triplex or double duplexes was observed.

In contrast to unmodified PNA, which is flexible and can bind complementary nucleic acids in a parallel or anti-parallel fashion (with the N terminus of PNA being defined as equivalent to the 5' terminus of DNA or RNA),[10] chiral modified PNA pre-organizes into a left- or right-handed helix depending on the position and stereochemistry of the side chain (**Figure 6**).[84,85] This effect can stabilize or destabilize hybridization, presumably due to the entropic cost associated with rearrangement, and in cases of stabilization is likely responsible for the increased strand invasion capability.[83]

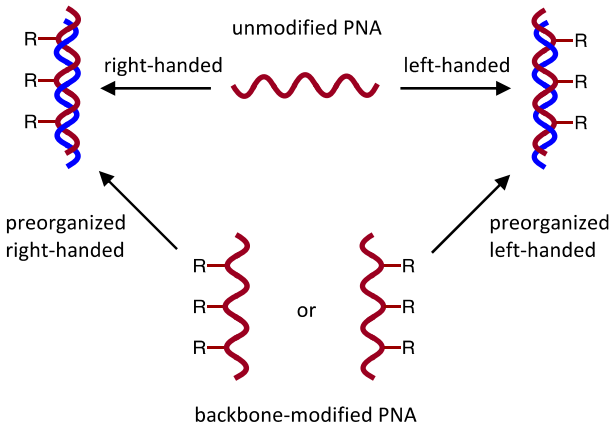


Figure 6: Schematic representation of the helical properties of unmodified PNA compared to backbone-modified PNA.

Interestingly, a longstanding hypothesis in the field had been that the increased hybridization stability of PNA with DNA or RNA could be attributed to the lack of charge repulsion with the neutral backbone. However, our group explored the incorporation of negatively charged side chains into the PNA backbone and found that while affinity to DNA and RNA was dependent upon ionic strength, PNAs having negatively or positively charged side chains had similar affinities at physiological ionic strength. From these data, we hypothesize that it is pre-organization of the backbone through hydrogen bonding interactions rather than the lack of charge that is responsible for the strong binding affinity of PNA to complementary nucleic acids.[76] This finding is important for applications such as charged-based delivery methods and the development of PNA therapeutics or materials, as it suggests that PNA can be synthesized to electrostatically mimic DNA and RNA and thus take advantage of the cellular delivery methods that are rapidly being developed for these nucleic acids.

Applications of Backbone-Modified PNA

The exploration of structurally diverse backbone-modified PNAs has allowed researchers to gain a better understanding of structure-function relationships for this artificial biopolymer and expand the repertoire of applications that are possible. For example, the enhanced affinity and specificity afforded to modified PNAs increases their capacity to function as biosensors and capture probes. Marchelli *et al.* described the use of α-PNA containing a chiral D-lysine box to develop a method using capillary electrophoresis to detect a single point mutation in DNA related to cystic fibrosis.[86] Including the lysine functional groups increased the specificity and allowed researchers to distinguish between healthy, mutated homozygous, and heterozygous individuals. In a different example, rational design of a cyclopentane-derived cyclic β-PNA to adopt a pre-organized structure led to higher affinity and sequence specificity,[63,64] which was then used to detect *pag* anthrax DNA through a sandwich-based assay using two β-PNA probes.[68]

Backbone-modified PNAs have also shown significant potential for use in antisense and imaging applications. Incorporating arginine side chains in α-PNA or γ-PNA increases the solubility, hybridization, and cell permeability of PNA probes.[33,87] Manicardi *et al.* synthesized a series of 18-mer PNA probes containing eight arginine side chains at varying positions (α or γ) along the backbone (consecutive or alternating) that target miRNA-210 (**Figure 7A**).[88] Fluorescence-activated cell sorting revealed that each modified probe was internalized by cells, with α-modified probes exhibiting the highest level of uptake. Real-time quantitative PCR demonstrated the ability to inhibit the target miRNA-210 in leukemic K562 cells, presenting the first example of anti-miR activity using backbone-modified PNA. Importantly, resistance to enzymatic degradation was increased in the backbone-modified PNAs over an arginine-rich CPP-PNA conjugate, highlighting the advantage of internal side chain inclusion for *in vivo* applications. Similarly, Ly and coworkers decorated an α-PNA with arginine side chains to generate what they call guanidinium peptide nucleic acid (GPNA) and

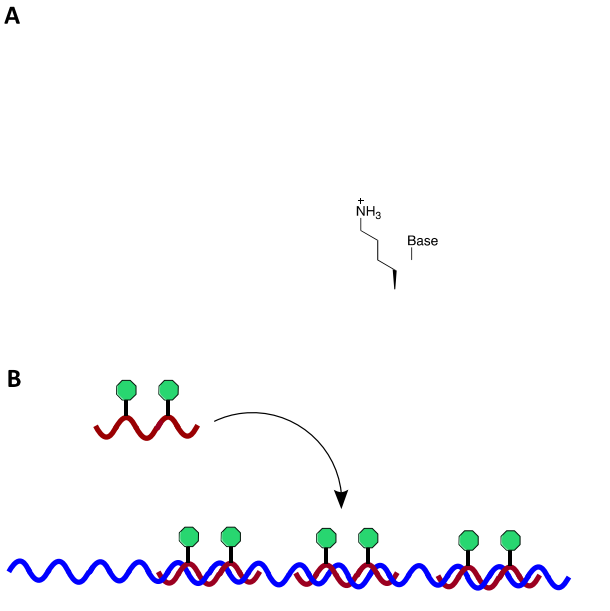


Figure 7: (A) Chemical representations of α-PNA containing a guanidinium side chain and γ-PNA containing a guanidinium, miniPEG, or lysine side chains. (B) Schematic representation of multivalent ligand display.

demonstrated the ability to sequence-specifically inhibit transcription of the E-cadherin gene in live cells (**Figure 7A**).[89] Notably, they found that including the arginine side chains within the PNA backbone resulted in less cytotoxicity than conjugation of unfunctionalized PNA to an arginine-rich CPP. This group has also pioneered the invention of miniPEG PNA, the first example of a charge-neutral yet strongly hydrophilic modification in the PNA backbone (**Figure 7A**).[32] This modification was shown to allow the direct Watson-Crick-Franklin recognition of double helical B-DNA through a strand invasion mechanism, desirable for *in vivo* applications.[90] Gupta *et al.* used a γ-PNA containing miniPEG side chains to silence the activity of miR-210 and observed anti-tumor activity in mouse models.[91] Additionally, miniPEG γ-PNAs showed promise as imaging agents for intracellular staining of telomeric DNA.[92] The improved physicochemical properties of backbone-modified PNA greatly augments its use for each of these *in vivo* applications.

Beyond typical hybridization-based applications, backbone-modified PNAs have been used as well-defined scaffolds for multivalent ligand display (**Figure 7B**).[93–96] For example, Appella and coworkers first synthesized a γ-PNA containing lysine side chains and found that it was a convenient position to attach bulky modifications without disrupting DNA and RNA binding (**Figure 7A**).[72,93] Using this technology, they demonstrated programmable multivalent display of a cyclic RGD analogue to block metastatic melanoma cell attachment, and modulated dopamine D_2 receptor activity by displaying a known agonist, (±)-PPHT.[94,95] These examples highlight the power of modifying the side chain sequence of PNA to display specific amino acid functional groups at precise spatial locations on the biopolymer.

In each of the examples presented, the inclusion of amino acid-like side chains into the PNA backbone resulted in a change of internal properties and the ability to adopt a pre-organized structure or spatially present functional groups in unique ways. Thus, addition of amino acid side chains to the backbone has benefitted PNA in numerous applications. However, the design and application of these modifications has largely been to enhance the nucleic acid recognition capability by increasing solubility, hybridization efficiency, or cellular uptake. In contrast, the abundant diversity of functional groups that can be included at multiple defined positions in the PNA scaffold represents a yet unrealized potential and underexplored avenue to encode information for activity, assembly, or structure analogous to peptides.

Encoding Amino Acid Information in the PNA Backbone

The development of backbone-modified PNA revealed the underlying potential of PNA to act as both a peptide and nucleic acid mimic. By combining the information derived from PNA-peptide conjugates with the invention of functionally customized monomers, new information-rich constructs can be designed and investigated for a wide range of biological applications.

The use of PNA as a peptide mimic is still in its infancy and there are very few examples where the embedded amino acid sequence information is the main driver for structure or function. Marchelli and coworkers pioneered the consideration of PNA as a "peptide" through the first known example of inserting an amino acid sequence into the PNA backbone to promote activity.[97] Using a submonomeric synthetic strategy, a model PNA trimer conjugated to a rhodamine dye was synthesized containing an NLS sequence (PKKKRKV) at alternating γ- and α-positions in the backbone (**Figure 8**). Fluorescence microscopy images revealed high intracellular and nuclear localized levels of the modified PNA in Rhabdomyosarcoma cells, similar to that of the canonical NLS peptide. Unmodified PNA was not detected within the nuclei, providing evidence that the embedded amino acid information was responsible for the uptake and nuclear localization. The authors speculated that the amino acid sequence is able to interact with the receptor protein importin, resulting in importin-mediated nuclear transport. While further experimentation is required to ascertain the effects of both protein and nucleic acid binding to the PNA peptide mimic, this work demonstrated the feasibility of utilizing a complex peptide sequence for function, opening the doors to numerous applications in bioorganic chemistry and molecular biology.

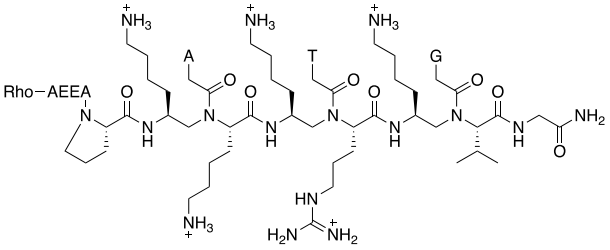


Figure 8: Chemical structure of PNA trimer with embedded NLS peptide sequence PKKKRKV. Side chains alternate between γ- and α-positions. "Rho" indicates a rhodamine dye.

Building upon the concept that PNA can be interpreted as a peptide mimic, our lab imagined the possibility of deriving controlled structural behavior through complex amino acid sequences. Inspired by peptide amphiphiles, we hypothesized that integrating a defined sequence of hydrophobic and hydrophilic amino acids along the PNA backbone would impart amphiphilic character. This "protein code" would then promote self-assembly under aqueous conditions. To demonstrate this assembly functionality, we synthesized γ-PNA monomers containing alanine and lysine side chains and incorporated them as a complex amino acid sequence in the backbone of a PNA oligomer complementary to the commonly utilized oncogenic miRNA-21 (**Figure 9A**).[98] We also synthesized and incorporated a monomer having the solvatochromic dye 4-dimethylamino-naphthalimide (4-DMN) at the hydrophobic terminus of the PNA amphiphile, speculating that such placement would lead to an increase in fluorescence upon assembly due to the change in the solvation state of the dye, enabling us to monitor the assembly process in solution. Fluorescence spectroscopy confirmed the assembly of PNA amphiphiles under aqueous conditions and TEM and DLS validated the formation of discrete micellar architectures (**Figure 9C**). Experimental observations also indicated that unmodified PNA forms a small number of amorphous aggregates while diblock PNA-amino acid conjugates form larger, indeterminate aggregates. Together, these data support our hypothesis that internal side chains are uniquely able to impart controlled assembly analogous to peptides. We then envisioned accessing the "nucleotide code" of the PNA amphiphiles through target hybridization to drive stimuli-responsive disassembly or morphological changes (**Figure 9B**). CD and thermal denaturation experiments confirmed the ability of the target miRNA-21 to bind the PNA amphiphile assemblies and form PNA:RNA duplexes.

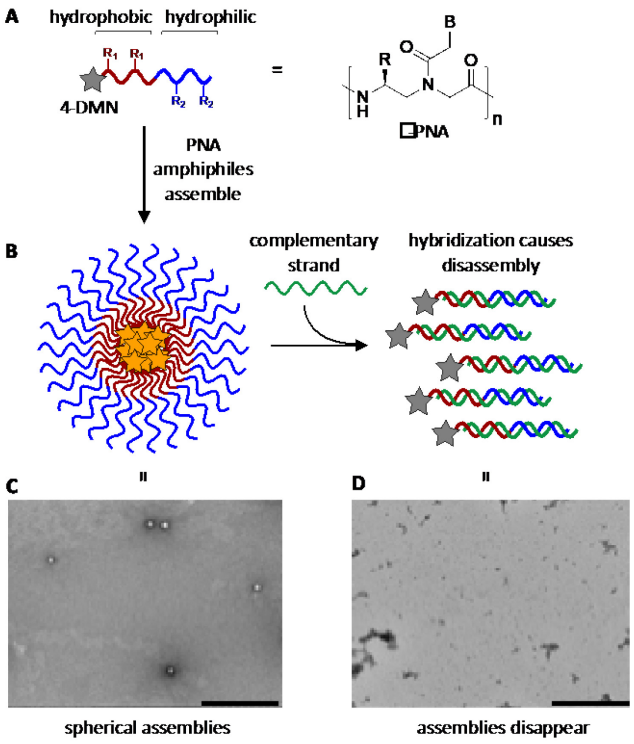


Figure 9: (A) Cartoon representation of bilingual PNA. (B) Self-assembly and disassembly by target hybridization. (C) TEM image of assembled PNA amphiphiles. (D) TEM image showing lack of assemblies after target hybridization. Adapted with permission from Swenson, C.S. et al. *J. Am. Chem. Soc.* **2019**, *141* (48), 19038–19047. Copyright 2019 American Chemical Society.

Fluorescence spectroscopy and TEM revealed the disappearance of assemblies with addition of a target sequence, while assemblies were retained in the presence of a scrambled sequence (**Figure 9D**). Thus, we designed a "bilingual biopolymer" using the PNA scaffold to encode complex structural information through a backbone-embedded amino acid code that simultaneously contains a nucleic acid code capable of directing stimuli-responsive supramolecular conformational changes through specific molecular recognition. This demonstrates the versatile and information-rich dual nature of PNA as a peptide and nucleic acid mimic.

As this example shows, the PNA scaffold is highly programmable, granting access to both nucleic acid and amino acid sequences for information storing and processing. Our bilingual biopolymer motif utilizes both information codes to undergo self-assembly and to conformationally respond to environmental stimuli in the form of target nucleic acids. The capacity to respond to a variety of specified nucleic acid targets offers significant potential for applications in therapeutics, diagnostics, and biotechnology. Though initially focusing on micellar architectures, we anticipate this approach to be viable for organization into many other proteinogenic structural motifs such as sheets or coils, which can in turn be of use for the creation of bioinspired materials. Moreover, considering the possibility that proteins can interact with peptide-like PNAs, we envision the construction of tunable, user-defined adapters between specific protein and nucleic acid targets.

Conclusions and Future Perspectives

Nearly three decades have passed since Nielsen and coworkers first invented PNA and described it as the "molecule with two identities."[99] Since then, researchers have leveraged organic, biological, and analytical chemistry to develop and implement technologies using the unique properties of PNA. Specifically, PNA-peptide conjugates and amphiphiles capable of self-assembly have garnered substantial interest owing to their strong chemical tolerance and ease of synthesis. PNA conjugates displaying improved cell permeability, self-assembly, and nucleic acid hybridization have also been demonstrated. Taking a different approach, backbone-modified PNAs have been developed by integrating specific side chains having unique chemical properties into the sequence of the PNA oligomer itself. Myriad functional groups can be included at precise positions to allow the tuning of physicochemical properties such as cell permeability, solubility, and affinity for complementary nucleic acids.

Each of the major approaches to alter the properties of PNA, either through conjugation or backbone modifications, have primarily focused on enhancing its function as a nucleic acid. The perception of PNA as a peptide mimic, the other "identity" expressed by Nielsen and coworkers, represents a relatively new adventure in the world of PNA design. Our lab and others have demonstrated sophisticated examples wherein a complex amino acid code embedded into the PNA backbone is capable of driving activity or intermolecular assembly for structure and function. However, additional investigation is required to fully understand the underlying properties and further advance this technology. Specifically, design rules will need to be established that allow for predictable structural or functional outcomes dependent on the PNA sequence, length, and identity of the embedded amino acid modifications. For example, does increasing the hydrophobicity of the side chains alter the size, critical assembly concentrations, or morphology of amphiphilic assemblies? Are total charge and amino acid identity important for assembly, stability, and subsequent activity? Can we access higher order structural motifs such as coils and sheets through incorporation of different amino acid sequences? With advancing knowledge about the synthesis and properties of PNA, it will become possible to answer these questions in the near future. With the multitude of natural and unnatural amino acid side chains available for precise placement along the PNA scaffold, the capabilities of this dual information coding biopolymer are limited only by our imagination.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors gratefully acknowledge support from the National Science Foundation (DMR 1822262 J.M.H.).

Notes and references

1 A. C. L. Lee, J. L. Harris, K. K. Khanna and J. H. Hong, *Int. J. Mol. Sci.*, 2019, **20**, 1–21.

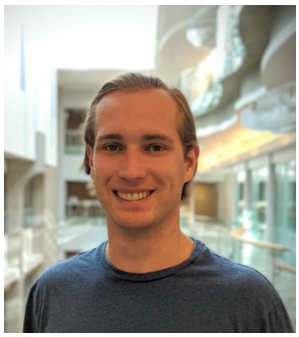
2 P. S. Huang, S. E. Boyken and D. Baker, *Nature*, 2016, **537**,

320–327.
3 K. Sato, M. P. Hendricks, L. C. Palmer and S. I. Stupp, *Chem. Soc. Rev.*, 2018, **47**, 7539–7551.
4 N. C. Seeman and H. F. Sleiman, *Nat. Rev. Mater.*, 2018, **3**, 17068.
5 S. Agrawal and M. J. Gait, in *Advances in Nucleic Acid Therapeutics*, The Royal Society of Chemistry, 2019, pp. 1–21.
6 K. Morihiro, Y. Kasahara and S. Obika, *Mol. Biosyst.*, 2017, **13**, 235–245.
7 Q. Ma, D. Lee, Y. Q. Tan, G. Wong and Z. Gao, *Polym. Chem.*, 2016, **7**, 5199–5216.
8 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science (80-.).*, 1991, **254**, 1497–1500.
9 V. V Demidov, V. N. Potaman, M. D. Frank-Kamenetskii, M. Egholm, O. Buchardt, S. H. Sonnichsen and P. E. Nlelsen, *Biochem. Pharmacol.*, 1994, **48**, 1310–1313.
10 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566–568.
11 E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chemie Int. Ed.*, 1998, **37**, 2796–2823.
12 P. E. Nielsen, M. Egholm and O. Buchardt, *Bioconjug. Chem.*, 1994, **5**, 3–7.
13 A. Whitney, G. Gavory and S. Balasubramanian, *Chem. Commun.*, 2003, **9**, 36–37.
14 M. Murtola, M. Wenska and R. Strömberg, *J. Am. Chem. Soc.*, 2010, **132**, 8984–8990.
15 M. Gaglione, G. Milano, A. Chambery, L. Moggio, A. Romanelli and A. Messere, *Mol. Biosyst.*, 2011, **7**, 2490–2499.
16 L. Petersen, M. C. de Koning, P. van Kuik-Romeijn, J. Weterings, C. J. Pol, G. Platenburg, M. Overhand, G. A. van der Marel and J. H. van Boom, *Bioconjug. Chem.*, 2004, **15**, 576–582.
17 A. Ghosal and P. E. Nielsen, *Nucleic Acid Ther.*, 2012, **22**, 323–334.
18 H. Narenji, O. Teymournejad, M. A. Rezaee, S. Taghizadeh, B. Mehramuz, M. Aghazadeh, M. Asgharzadeh, M. Madhi, P. Gholizadeh, K. Ganbarov, M. Yousefi, A. Pakravan, T. Dal, R. Ahmadi and H. Samadi Kafil, *Microb. Pathog.*, 2020, **139**, 103907.
19 P. E. Nielsen and T. Shiraishi, *Artif. DNA PNA XNA*, 2011, **2**, 90–99.
20 M. G. Svahn, K. E. Lundin, R. Ge, E. Törnquist, E. O. Simonson, S. Oscarsson, M. Leijon, L. J. Brandén and C. I. E. Smith, *J. Gene Med.*, 2004, **6**, S36–S44.
21 L. J. Brandén, A. J. Mohamed and C. I. E. Smith, *Nat. Biotechnol.*, 1999, **17**, 784–787.
22 A. F. Faruqi, M. Egholm and P. M. Glazer, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 1398–1403.
23 R. Bahal, N. Ali McNeer, E. Quijano, Y. Liu, P. Sulkowski, A. Turchick, Y. C. Lu, D. C. Bhunia, A. Manna, D. L. Greiner, M. A. Brehm, C. J. Cheng, F. López-Giráldez, A. Ricciardi, J. Beloor, D. S. Krause, P. Kumar, P. G. Gallagher, D. T. Braddock, W. Mark Saltzman, D. H. Ly and P. M. Glazer, *Nat. Commun.*, 2016, **7**, 13304.
24 P. E. Nielsen and M. Egholm, *Curr. Issues Molec. Biol.*, 1999, **1**, 89–104.
25 F. Pellestor and P. Paulasova, *Eur. J. Hum. Genet.*, 2004, 12, 694–700.
26 R. D'Agata, M. C. Giuffrida and G. Spoto, *Molecules*, 2017, 22, 1951.
27 R. Gambari, *Expert Opin. Ther. Pat.*, 2014, **24**, 267–294.
28 A. Gupta, A. Mishra and N. Puri, *J. Biotechnol.*, 2017, **259**, 148–159.
29 J.-C. Wu, Q.-C. Meng, H.-M. Ren, H.-T. Wang, J. Wu and Q. Wang, *Acta Pharmacol. Sin.*, 2017, **38**, 798–805.
30 H. Narenji, P. Gholizadeh, M. Aghazadeh, M. A. Rezaee, M. Asgharzadeh and H. S. Kafil, *Biomed. Pharmacother.*, 2017, **93**, 580–588.
31 J. Saarbach, P. M. Sabale and N. Winssinger, *Curr. Opin. Chem. Biol.*, 2019, **52**, 112–124.
32 B. Sahu, I. Sacui, S. Rapireddy, K. J. Zanotti, R. Bahal, B. A. Armitage and D. H. Ly, *J. Org. Chem.*, 2011, **76**, 5614–5627.
33 P. Zhou, M. Wang, L. Du, G. W. Fisher, A. Waggoner and D. H. Ly, *J. Am. Chem. Soc*, 2003, **125**, 6878–6879.
34 P. Wittung, J. Kajanus, K. Edwards, P. Nielsen, B. Nordén and B. G. Malmström, *FEBS Lett.*, 1995, **365**, 27–29.
35 D. A. Braasch and D. R. Corey, *Methods*, 2001, **23**, 97–107.
36 T. Ljungstrøm, H. Knudsen and P. E. Nielsen, *Bioconjug. Chem.*, 1999, **10**, 965–972.
37 G. Guidotti, L. Brambilla and D. Rossi, *Trends Pharmacol. Sci.*, 2017, 38, 406–424.
38 J. J. Díaz-Mochón, L. Bialy, J. Watson, R. M. Sánchez-Martín and M. Bradley, *Chem. Commun.*, 2005, 3316–3318.
39 U. Koppelhus, S. K. Awasthi, V. Zachar, H. U. Holst, P. Ebbesen and P. E. Nielsen, *Antisense Nucleic Acid Drug Dev.*, 2002, **12**, 51–63.
40 N. Bendifallah, F. W. Rasmussen, V. Zachar, P. Ebbesen, P. E. Nielsen and U. Koppelhus, *Bioconjug. Chem.*, 2006, **17**, 750–758.
41 G. Cutrona, E. M. Carpaneto, M. Ulivi, S. Roncella, O. Landt, M. Ferrarini and L. C. Boffa, *Nat. Biotechnol.*, 2000, **18**, 300–303.
42 B. Ndeboko, N. Ramamurthy, G. J. Lemamy, C. Jamard, P. E. Nielsen and L. Cova, *Mol. Ther. - Nucleic Acids*, 2017, **9**, 162–169.
43 G. Barkowsky, A. L. Lemster, R. Pappesch, A. Jacob, S. Krüger, A. Schröder, B. Kreikemeyer and N. Patenge, *Mol. Ther. - Nucleic Acids*, 2019, **18**, 444–454.
44 Y. Turner, G. Wallukat, P. Säälik, B. Wiesner, S. Pritza and J. Oehlke, *J. Pept. Sci.*, 2010, **16**, 71–80.
45 H. Cui, M. J. Webber and S. I. Stupp, *Biopolymers*, 2010, **94**, 1–18.
46 M. P. Hendricks, K. Sato, L. C. Palmer and S. I. Stupp, *Acc. Chem. Res.*, 2017, **50**, 2440–2448.
47 A. Dasgupta and D. Das, *Langmuir*, 2019, **35**, 10704–10724.
48 O. Berger and E. Gazit, *Biopolymers*, 2017, **108**, e22930.
49 M. O. Guler, J. K. Pokorski, D. H. Appella and S. I. Stupp, *Bioconjug. Chem.*, 2005, **16**, 501–503.
50 J. P. Vernille, L. C. Kovell and J. W. Schneider, *Bioconjug. Chem.*, 2004, **15**, 1314–1321.
51 C. Lau, R. Bitton, H. Bianco-Peled, D. G. Schultz, D. J.

Cookson, S. T. Grosser and J. W. Schneider, *J. Phys. Chem. B*, 2006, **110**, 9027–9033.
52 B. F. Marques and J. W. Schneider, *Langmuir*, 2005, **21**, 2488–2494.
53 S. T. Grosser, J. M. Savard and J. W. Schneider, *Anal. Chem.*, 2007, **79**, 9513–9519.
54 L.-H. Liu, Z.-Y. Li, L. Rong, S.-Y. Qin, Q. Lei, H. Cheng, X. Zhou, R.-X. Zhuo and X.-Z. Zhang, *ACS Macro Lett.*, 2014, **3**, 467–471.
55 T. A. Feagin, N. I. Shah and J. M. Heemstra, *J. Nucleic Acids*, 2012, **2012**, 354549.
56 C. Sharma and S. K. Awasthi, *Chem. Biol. Drug Des.*, 2017, **89**, 16–37.
57 M. Moccia, M. F. A. Adamo and M. Saviano, *Artif. DNA PNA XNA*, 2016, **5**, e1107176.
58 K. L. Dueholm, K. H. Peterson, D. K. Jensen, M. Egholm, P. E. Nielsen and O. Buchardt, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 1077–1080.
59 P. E. Nielsen, G. Haaima, A. Lohse and O. Buchardt, *Angew. Chemie Int. Ed. English*, 1996, **35**, 1939–1942.
60 A. R. Katritzky and T. Narindoshvili, *Org. Biomol. Chem.*, 2008, **6**, 3171–3176.
61 A. Püschl, S. Sforza, G. Haaima, O. Dahl and P. E. Nielsen, *Tetrahedron Lett.*, 1998, **39**, 4707–4710.
62 T. Sugiyama, Y. Imamura, Y. Demizu, M. Kurihara, M. Takano and A. Kittaka, *Bioorganic Med. Chem. Lett.*, 2011, **21**, 7317–7320.
63 M. C. Myers, M. A. Witschi, N. V. Larionova, J. M. Franck, R. D. Haynes, T. Hara, A. Grajkowski and D. H. Appella, *Org. Lett.*, 2003, **5**, 2695–2698.
64 J. K. Pokorski, M. A. Witschi, B. L. Purnell and D. H. Appella, *J. Am. Chem. Soc.*, 2004, **126**, 15067–15073.
65 J. K. Pokorski, M. C. Myers and D. H. Appella, *Tetrahedron Lett.*, 2005, **46**, 915–917.
66 P. Lagriffoule, M. Eriksson, K. K. Jensen, P. E. Nielsen, P. Wittung, B. Nordén and O. Buchardt, *Chem. - A Eur. J.*, 1997, **3**, 912–919.
67 T. Govindaraju, V. A. Kumar and K. N. Ganesh, *Chem. Commun.*, 2004, **10**, 860–861.
68 C. M. Micklitsch, B. Y. Oquare, C. Zhao and D. H. Appella, *Anal. Chem.*, 2013, **85**, 251–257.
69 L. Kosynkina, W. Wang and T. C. Liang, *Tetrahedron Lett.*, 1994, **35**, 5173–5176.
70 T. Tedeschi, S. Sforza, R. Corradini and R. Marchelli, , DOI:10.1016/j.tetlet.2005.09.157.
71 A. Manna, S. Rapireddy, G. Sureshkumar and D. H. Ly, *Tetrahedron*, 2015, **71**, 3507–3514.
72 E. A. Englund and D. H. Appella, *Org. Lett.*, 2005, **7**, 3465–3467.
73 M. C. de Koning, L. Petersen, J. J. Weterings, M. Overhand, G. A. van der Marel and D. V. Filippov, *Tetrahedron*, 2006, **62**, 3248–3258.
74 S. Pensato, M. Saviano, N. Bianchi, M. Borgatti, E. Fabbri, R. Gambari and A. Romanelli, *Bioorg. Chem.*, 2010, **38**, 196–201.
75 C. Avitabile, L. Moggio, G. Malgieri, D. Capasso, S. Di Gaetano, M. Saviano, C. Pedone and A. Romanelli, *PLoS One*, 2012, **7**, e35774.
76 N. T. S. De Costa and J. M. Heemstra, *PLoS One*, 2013, **8**, e58670.
77 M. J. Crawford, S. Rapireddy, R. Bahal, I. Sacui and D. H. Ly, *J. Nucleic Acids*, 2011, **2011**, 652702.
78 P. Wittung, P. Nielsen and B. Nordén, *Biochemistry*, 1997, **36**, 7973–7979.
79 M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895–1897.
80 P. E. Nielsen, M. Egholm and O. Buchardt, *J. Mol. Recognit.*, 1994, **7**, 165–170.
81 S. Rapireddy, G. He, S. Roy, B. A. Armitage and D. H. Ly, *J. Am. Chem. Soc.*, 2007, **129**, 15596–15600.
82 J. Lohse, O. Dahl and P. E. Nielsen, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11804–11808.
83 G. He, S. Rapireddy, R. Bahal, B. Sahu and D. H. Ly, *J. Am. Chem. Soc.*, 2009, **131**, 12088–12090.
84 A. Dragulescu-Andrasi, S. Rapireddy, B. M. Frezza, C. Gayathri, R. R. Gil and D. H. Ly, *J. Am. Chem. Soc*, 2006, **128**, 10258–10267.
85 M. D. Verona, V. Verdolino, F. Palazzesi and R. Corradini, *Sci. Rep.*, 2017, **7**, 42799.
86 T. Tedeschi, M. Chiari, G. Galaverna, S. Sforza, M. Cretich, R. Corradini and R. Marchelli, *Electrophoresis*, 2005, **26**, 4310–4316.
87 P. Zhou, A. Dragulescu-Andrasi, B. Bhattacharya, H. O'Keefe, P. Vatta, J. J. Hyldig-Nielsen and D. H. Ly, *Bioorganic Med. Chem. Lett.*, 2006, **16**, 4931–4935.
88 A. Manicardi, E. Fabbri, T. Tedeschi, S. Sforza, N. Bianchi, E. Brognara, R. Gambari, R. Marchelli and R. Corradini, *ChemBioChem*, 2012, **13**, 1327–1337.
89 A. Dragulescu-Andrasi, S. Rapireddy, G. He, B. Bhattacharya, J. J. Hyldig-Nielsen, G. Zon and D. H. Ly, *J. Am. Chem. Soc.*, 2006, **128**, 16104–16112.
90 R. Bahal, B. Sahu, S. Rapireddy, C.-M. Lee and D. H. Ly, *ChemBioChem*, 2012, **13**, 56–60.
91 A. Gupta, E. Quijano, Y. Liu, R. Bahal, S. E. Scanlon, E. Song, W. C. Hsieh, D. E. Braddock, D. H. Ly, W. M. Saltzman and P. M. Glazer, *Mol. Ther. Nucleic Acids*, 2017, **9**, 111–119.
92 A. Orenstein, A. Berlyoung, E. Rastede, H. Pham, E. Fouquerel, C. Murphy, B. Leibowitz, J. Yu, T. Srivastava, B. Armitage and P. Opresko, *Molecules*, 2017, **22**, 2117.
93 E. A. Englund and D. H. Appella, *Angew. Chemie Int. Ed.*, 2007, **46**, 1414–1418.
94 E. A. Englund, D. Wang, H. Fujigaki, H. Sakai, C. M. Micklitsch, R. Ghirlando, G. Martin-Manso, M. L. Pendrak, D. D. Roberts, S. R. Durell and D. H. Appella, *Nat. Commun.*, 2012, **3**, 614.
95 A. V Dix, J. L. Conroy, K. M. G. Rosenker, D. R. Sibley and D. H. Appella, *ACS Med. Chem. Lett.*, 2015, **6**, 425–429.
96 P. Gupta, E. E. Rastede and D. H. Appella, *Bioorganic Med. Chem. Lett.*, 2015, **25**, 4757–4760.
97 S. Sforza, T. Tedeschi, A. Calabretta, R. Corradini, C. Camerin, R. Tonelli, A. Pession and R. Marchelli, *European J. Org. Chem.*, 2010, **2010**, 2441–2444.
98 C. S. Swenson, A. Velusamy, H. S. Argueta-Gonzalez and J. M. Heemstra, *J. Am. Chem. Soc.*, 2019, **141**, 19038–19047.

99 P. E. Nielsen, *Acc. Chem. Res.*, 1999, **32**, 624–630.

Author Biographies



Colin S. Swenson was born in Missoula, Montana, USA in 1992. He received his B.S. in Chemistry at the University of Utah in 2015. In that same year, he joined the lab of Dr. Jennifer Heemstra and began his Ph.D. studies synthesizing γ-modified PNA amphiphiles for exploring information-rich dual-recognition assemblies. In 2017, he transferred with Dr. Heemstra to continue his graduate research at Emory University. His current research interests lie in the broad chemical biology of proteins and nucleic acids, seeking to further understand the molecular recognition properties to utilize these influential biopolymers.

Jennifer (Jen) Heemstra received her B.S. in Chemistry from UC Irvine in 2000, where she performed research with Prof. James Nowick, which instilled in her a love of supramolecular chemistry. She then moved to the University of Illinois, Urbana-Champaign, where she completed her Ph.D. with Prof. Jeffrey Moore in 2005. At Harvard University, she was a postdoctoral researcher with Prof. David Liu. Dr. Heemstra is currently an Associate Professor in the Department of Chemistry at Emory University. Research in the Heemstra lab is focused on harnessing the molecular recognition and self-assembly properties of biomolecules for applications in biosensing and bioimaging.