Bilingual Peptide Nucleic Acids: Encoding the Languages of Nucleic Acids and Proteins in a Single Self-Assembling Biopolymer

Nature encodes information, structure, and function in two basic forms of biopolymers: nucleic acids and proteins. Nucleotide sequences in DNA or RNA encode both genetic information and complementary molecular recognition properties, while amino acid sequences in peptides and proteins convey complex information for structure and function. The robust and adaptable performance of both of these biopolymer structures has been a fundamental driving force for evolution on Earth and explains their continued ubiquitous presence in Nature. Moreover, the straightforward design principles and privileged physicochemical properties of both nucleic acids and proteins position them as attractive materials to leverage and integrate in applications beyond their canonical roles. Although Nature has evolved machinery to read the nucleic acid code and translate it into amino acid code, the extant biopolymers are restricted to encoding amino acid or nucleotide sequences separately, limiting their potential applications in medicine and biotechnology. Here we describe the design, synthesis, and stimuli-responsive assembly behavior of a bilingual biopolymer that integrates both amino acid and nucleobase sequences into a single peptide nucleic acid (PNA) scaffold to enable tunable storage and retrieval of tertiary structural behavior and programmable molecular recognition capabilities. Incorporation of a defined sequence of amino acid side-chains along the PNA backbone yields amphiphiles having a “protein code” that directs self-assembly into micellar architectures in aqueous conditions. However, these amphiphiles also carry a “nucleotide code” such that subsequent introduction of a complementary RNA strand induces a sequence-specific disruption of assemblies through hybridization. Together, these properties establish bilingual PNA as a powerful biopolymer that combines two information systems to harness structural responsiveness and sequence recognition. The PNA scaffold and our synthetic system are highly generalizable, enabling fabrication of a wide array of user-defined peptide and nucleotide sequence combinations for diverse future biomedical and nanotechnology applications.