Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Nucleic Acid Mimetic Hexaplex Assembly Stabilized by the Synergistic Association of Non-Canonical Recognition Units and Peptide Backbones

Nicholas Hud

hud@chemistry.gatech.edu

Georgia Institute of Technology https://orcid.org/0000-0001-7711-6472

Sajena Kanangat Saraswathy

Georgia Institute of Technology

Suneesh karunakaran

Georgia Institute of Technology

Shibna Balakrishnan

Georgia Institute of Technology

Gary Newnam

Georgia Institute of Technology

Elizabeth Duselis

Georgia Institute of Technology

Luis Andres Casavilca Ramirez

Georgia Institute of Technology

Howard Page

Georgia Institute of Technology

Asem Alenaizan

King Fahd University of Petroleum and Minerals

Vicki Wysocki

Georgia Institute of Technology

Luke Leman

Scripps Research Institute https://orcid.org/0000-0002-1879-5900

Jonathan Chaires

Gary Schuster

Georgia Institute of Technology

Article

Keywords:

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11531588/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Abstract

Two modes of self-assembly, backbone-backbone hydrogen bonding and base-base hydrogen bonding with stacking interactions, are brought together in a class of nucleic acid mimetic peptides that form hexameric water-soluble supramolecular assemblies. These oligodipeptides contain dimeric repeats of glutamic acid (Glu) and a diaminobutyric acid (Dab) that is covalently linked to either melamine or cyanuric acid (Mel^{Dab} and Cy^{Dab}). Molecular models predict that these oligomers will form assemblies with a core of stacked, hydrogen-bonded Mel-Cy hexads that are sheathed by the six peptide backbones, creating a structure comparable to a six-stranded β-barrel. Experimental analyses confirm that such assemblies are formed upon mixing of the octapeptides $(Mel^{Dab}\text{-}Glu)_4$ and $(Cy^{Dab}\text{-}Glu)_4$. This assembly exhibits high thermal stability with pH-dependent apparent melting temperatures greater than 100°C. The chemical structure and assembly properties of these peptides support the hypothesis that compounds of similar form may have served as ancestors to RNA (i.e., proto-RNA) in the emergence of life on Earth.

Introduction

A long-standing goal of chemists and materials scientists is to mimic the self-assembly of peptides and nucleic acids in water. In multiple attempts, the canonical nucleobases were linked by a non-canonical backbone.[1–3] Other efforts have used the natural nucleic acid phosphate diester backbone but employed non-canonical bases.[4,5] We previously reported that the nucleobase melamine (Mel) associates with cyanuric acid (Cy) to form water-soluble, non-covalent supramolecular assemblies if one or both of these pairing partners is modified with a water-solubilizing moiety.[6] In the work reported here, we describe the synthesis and characterization of octapeptides formed by linking Mel and Cy nucleobases to a dipeptide backbone. These oligodipeptides self-assemble in water to form a new class of extraordinarily stable structures comprised of a hydrophobic core of hydrogen-bonded Mel and Cy nucleobases surrounded by six backbones that form a beta-barrel-like peptide sheath around the nucleobase core (see Fig. 1). The structure and the properties of these assemblies were examined by extensive computational and experimental techniques that include imaging, analytical spectrometric and X-ray diffraction methods.

Our previous studies revealed that the hydrophobic surface area of hexads formed by three units of each pairing partner (e.g. three Mel and three Cy) provides the free energy necessary for assembly in water and is sufficient to drive a phase transition between free monomers and extremely long linear assemblies of stacked hexads.[7] Earlier studies showed the formation of hexameric “rosettes” in non-polar solvents when these same bases are modified with hydrophobic appendages.[8–18] Given that the canonical nucleobases do not assemble as monomers in water, the considerable stability of supramolecular assemblies formed by hexad-forming non-canonical nucleobases prompted us to propose that such nucleobases could have facilitated the assembly of nucleic acids precursors prior to the emergence of RNA.[19] Furthermore, the discovery of a covalent backbone that links these nucleobases while assembled within a hexad stack would convert the previously reported evanescent supramolecular assemblies into a stable six-stranded structure that more closely resembles the extant

nucleic acids. The search for such backbones, with a focus on chemical structures that are also plausibly prebiotic, was the motivation for the current work.[20,21]

In 2018, Kashida, Asanuma and co-workers demonstrated the formation of parallel, six-stranded assemblies that are held together by a core of stacked hexads.[22] These investigators linked the pairing bases to a d-threoninol-phosphate backbone (d-αTNA) that is structurally similar to DNA but with an acyclic sugar. They showed that d-αTNA decamers assembled with hexad-forming nucleobases are stable in water. We decided to investigate the possibility of hexameric assembly with a backbone that may have formed more readily in a prebiotic environment than RNA or d-αTNA. We focused on the α-peptide backbone because it is utilized by proteins in life today and because amide bonds are more easily formed in prebiotic reactions than the phosphodiester linkages of nucleic acids.[23]

Model structures indicated that peptides with a dipeptide repeat (oligodipeptides) comprised of one amino acid presenting a pairing base on its side chain and a second negatively charged amino acid is compatible with the formation of a hexameric assembly with inwardly facing bases that create a central core of stacked hexads surrounded by outwardly facing charged groups that provide water solubility. Oligodipeptides with a dimeric repeat of an amino acid with a pairing base and a charged amino acid were first reported by Krishnamurthy, Eschenmoser and co-workers who observed duplex formation.[24–26] These investigators were also motivated by the possibility that such oligodipeptides could represent an ancestral form of nucleic acids. Subsequently, Bong and co-workers demonstrated that oligodipeptides of a similar form can assemble into triplexes with two strands of DNA.[27] More generally, oligodipeptides with sequence motifs defined by alternating charged and uncharged residues exhibit a propensity to self-assemble in water and may have been some of the earliest prebiotic peptides to create ordered assemblies.[28–35]

The octapeptides reported here are comprised of dimeric sequence motifs that are characterized by the first amino acid having a side chain terminated with a pairing nucleobase (Mel or Cy) and glutamic acid as the second amino acid. The resulting peptides each have four amino acids with the Cy or Mel recognition units separated by glutamic acid residues, which bears a negatively charged carboxylate group above pH 4. Significantly, the linkages between the bases and backbone of these monomers mimic the purine-pyrimidine geometry of Watson-Crick base pairs, a geometry that appears to be key for promoting both favorable base-base and backbone-backbone interactions. We found that these synergistic interactions enable the formation of hexameric assemblies with remarkable stability, compared to duplex DNA and RNA. The nucleic acid-like and protein-like features of these oligomers and their assemblies support the hypothesis that nucleic acids could have evolved from a proto-RNA that utilized a peptide backbone.[25,26]

Results

Modeling to select the backbone

Our previous X-ray fiber diffraction examination of supramolecular polymers comprised of stacked hexads formed from monomeric triaminopyrimidine (TAP) and Cy nucleobases revealed an inter-hexad distance of 3.4 Å and a helical twist angle of ca. 30° that varied with base modification.[36] We sought a backbone that covalently links the stacked bases of the hexads and thereby converts the ephemeral supramolecular polymers to more permanent oligomeric assemblies. Specifically, our objective was to identify a backbone that maintains alignment of the bases and their complementary hydrogen-bonding motif while simultaneously preserving the inter-hexad spacing and the preferred twist angle of the supramolecular polymer and includes an ionizable carboxylic acid group to assure water solubility at or above neutral pH.

Physical and computational models were constructed to probe potential structures that satisfied these requirements. This search led us to an α-oligodipeptide backbone with a dimeric repeat of a 2,4-diaminobutyric acid residue (Dab) and a glutamic acid (Glu) dipeptide unit. The covalently linked Mel and Cy bases that compose the hexads perform the same recognition function as the canonical nucleobases of RNA. They are linked to the side chains of the Dab residues (named Mel^{Dab} and Cy^{Dab}) by an ethylene bridge between these bases and the backbone α-carbon atom. Notably, our models indicate that when six of these α-DPNA (α-Dipeptide Nucleic Acids) are assembled with Mel-Cy hexads there is also the opportunity for backbone-backbone hydrogen bonding. The modelling revealed the possibility of a structure of the assembled peptides that resembles a six-stranded β-barrel wrapped around a core of stacked nucleobases (Fig. 1 and Supplementary Fig. 1).

We prepared two α-DPNA octapeptides for the investigation of the ability of this system to assemble and for the examination of the properties of the assemblies. These octapeptides are comprised of four identical dimer motifs $(Mel^{Dab}\text{-}Glu)_4$ and $(Cy^{Dab}\text{-}Glu)_4$ that are designated α-$(M)_4$ and α-$(C)_4$ (Fig. 1d). The experimental investigation of these oligomers reveals that the expected β-barrel-like assembly is formed and that the cooperative actions of base pairing, hydrophobic stabilization and backbone-backbone hydrogen bonding produce a remarkably robust multi-stranded structure.

Confirmation of hexaplex formation by α-(M) and α-(C)

The association of α-$(M)_4$ and α-$(C)_4$ in solution was investigated using Circular Dichroism (CD) spectroscopy. The CD spectrum of an equilibrated 1:1 mixture of α-$(M)_4$ and α-$(C)_4$ exhibits a strong negative band with a maximum intensity at ca. 218 nm and a smaller positive band with a maximum intensity at ca. 246 nm (Fig. 2a). The appearance of these bands is indicative of strand assembly, as the CD spectra of the individual strands are of low intensity and featureless compared to that of their mixture. The CD spectrum of each individual strand exhibits a negative band near 200 nm and a weaker positive band at 218 nm. A qualitative comparison of the CD spectra with those from polypeptides and proteins of known secondary structures indicates that the individual strands exist in a random coil conformation, whereas the spectrum of the assembled strands is consistent with the adoption of the β-sheet secondary structure.[37]

A more quantitative analysis of the CD spectrum of α-$(M)_4$ and α-$(C)_4$ 1:1 mixture using the program Beta Structure Selection (BeStSel)[38] indicates 87% antiparallel β-sheet conformation (Fig. 2a, inset). This analysis supports the computational and physical models that predict an anti-parallel, β-barrel-like structure. Naturally-occurring β-barrel proteins are most often observed with eight or more strands, but six-stranded β-barrels are observed as domains within larger proteins[39] and have been formed as isolated folds by *de novo* protein design.[40] Based on CD spectroscopy alone, it is not possible to determine if the assembly formed by α-$(M)_4$ and α-$(C)_4$ is a true β-barrel with hydrogen bonding interactions between all adjacent strands. Our computational and physical models suggest that the hexad core with ethylene linkages between the bases and the backbones is of a size that the six peptide backbones in an anti-parallel arrangement are only able to accommodate hydrogen bonding between every other pair of strands (Figs. 1c, f and Supplementary Fig. 1). Therefore, this structure may be viewed as the association of three, two-stranded anti-parallel β-sheets ($[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$).

The thermal stability of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly was investigated by monitoring the temperature-dependent changes in the CD spectrum of a 1:1 sample (50 μM in each octapeptide) at pH 7.2 (Fig. 2b). A plot of the CD signal intensity at 218 nm reveals a cooperative transition (i.e., sigmoidal curve) with a midpoint temperature of 55°C. This result provides an apparent melting temperature (apparent T_m) because the assembly does not fully reform upon cooling (Supplementary Fig. 2). The change in the CD spectrum of the heat-cycled sample with cooling indicates that approximately 35% of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly is reformed when the sample is returned to 4°C. Incubation for an additional 11 hr at 4°C did not result in additional reappearance of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly signal. The partial reassembly of the α-$(M)_4$ and α-$(C)_4$ strands after heating to 84°C appears to be due to an unidentified irreversible change to one or both strands of the assembly at high temperature. This interpretation is supported by the observation that a second heating cycle of the same sample results in a further reduction of recovered assembly when the sample is returned to 4°C.

Atomic force microscopy (AFM) was also used to confirm the assembly of α-$(M)_4$ and α-$(C)_4$. An AFM image is shown in Fig. 2c of the structures formed when a 1:1 solution of these two octapeptides (initially 400 μM each strand) is dried on a mica surface. The image reveals the formation of fibers having lengths of at least 100 nm and heights of ca. 2 nm, a thickness that is consistent with a hexameric assembly. The observation of 100 nm fibers indicates that the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assemblies, which are themselves only ca. 1.4 nm long, must be associated in a co-axial manner to form much longer structures. The mechanism of association is driven primarily by the hydrophobic effect as revealed by its disruption with ethidium bromide, see below and Supplementary Fig. 4.

Polyacrylamide gel electrophoresis (PAGE), a powerful method for molecular weight analysis of charged biopolymers, was used to examine the length distribution of $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assemblies. Figure 2d presents an image of a gel loaded with α-$(M)_4$ and α-$(C)_4$ as individual strands and with $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assemblies. After electrophoresis (1 hr, 250 V, 4°C) the gel was stained with SYBR Green II, a fluorescent dye known to bind to G-quadruplexes, which was used here because G-quadruplexes have a charge

density closer to that of the expected hexaplex than to duplex nucleic acids. As expected, α-$(M)_4$ and α-$(C)_4$ as individual strands are not stained by SYBR Green II and therefore do not show detectable bands in the fluorescence image of the gel (Fig. 2d, lanes 2 and 3). In contrast, the [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies show at least seven distinct bands and slower moving unresolved species that are apparently of even higher molecular weight (Fig. 2d, lane 4). Considering the fastest moving band in lane 4 to be that of [α-$(M)_4$/α-$(C)_4$]$_3$, the slowest moving resolved band corresponds to a heptamer of [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies. Of note, the most intense band in lane 4 appears to correspond to a dimer of [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies. We suspect that the larger assemblies observed by PAGE are the result of end-to-end stacking of multiple [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies, which would be consistent with the length of fibers observed by AFM.

We have previously reported that the cationic DNA intercalator ethidium bromide (EtBr) shortens supramolecular hexad stacks formed from Mel and Cy monomers by binding to the termini of linear assemblies and thereby reducing the hydrophobic force that drives the formation of extremely long assemblies.[41] We used PAGE to assess the effect that EtBr has on the length of assemblies formed by α-$(M)_4$ and α-$(C)_4$. The gel image in Fig. 2D shows the impact of adding EtBr to α-$(M)_4$/α-$(C)_4$ mixtures. As described above, without EtBr a ladder of increasingly large assemblies is observed (Fig. 2d, lane 4). The relative amount of these larger assemblies is dramatically diminished as the concentration of EtBr in the solution increases (Fig. 2d, lanes 5–7), with essentially only the single [α-$(M)_4$/α-$(C)_4$]$_3$ assembly being observed for EtBr concentrations at 3 mM or higher. These observations provide support for our supposition that assemblies of α-$(M)_4$ and α-$(C)_4$ larger than [α-$(M)_4$/α-$(C)_4$]$_3$ are the result of hydrophobic forces driving the coaxial stacking of multiple [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies.

Native mass spectrometry (nMS) enables the direct study of biomolecules in their undisrupted “native” state where non-covalent associations are largely preserved.[42] We used nMS to analyze a 1:1 mixture of α-$(M)_4$ and α-$(C)_4$, 1 μM in each strand, with four molar equivalents of EtBr to limit the formation of multi-hexaplex assemblies (Fig. 2e and Supplementary Fig. 3). Significantly, in positive mode the most abundant peak in the mass spectrum corresponds to the [α-$(M)_4$/α-$(C)_4$]$_3$ assembly, consistent with what is observed by PAGE. This assembly appears predominantly in the + 4 charge state around m/z = 2066.75 and in the + 3 charge state around m/z = 2755.34, which corresponds to an average mass of 8262.97 Da (calculated mass is 8263.26 Da). The high mass accuracy (< 1 ppm) and distribution of the isotopic ratios of peaks for this + 3 charge state are also in excellent agreement with the assignment of a [α-$(M)_4$/α-$(C)_4$]$_3$ assembly (Fig. 2f).

Consistent with the PAGE results, in the absence of EtBr the major peak in the nMS spectrum of a 1:1 mixture of α-$(M)_4$ and α-$(C)_4$ corresponds to a dimer of the [α-$(M)_4$/α-$(C)_4$]$_3$ hexameric (Supplementary Fig. 5). nMS spectra of α-$(M)_4$ and α-$(C)_4$ as separate samples exhibited only the + 1 and + 2 charged states of each strand (Supplementary Figs. 6 and 7). These nMS results, the results of the gel electrophoresis experiments and the CD spectroscopy provide convincing evidence that octapeptides α-

$(M)_4$ and α-$(C)_4$ form the predicted [α-$(M)_4$/α-$(C)_4$]$_3$ hexameric assembly. Additionally, CD spectroscopy data confirm that an assembly containing a 1:1 molar ratio of strands is favored over alternative possible assemblies that contain ratios of α-$(M)_4$ and α-$(C)_4$ of 1:2 and 2:1 (Supplementary Fig. 8), such as would be expected for triplex assemblies similar to those observed by Bong and co-workers for peptides containing Mel recognitions units.[27]

X-ray Diffraction and Computational Analysis.

To evaluate more thoroughly the structure of the [α-$(M)_4$/α-$(C)_4$]$_3$ assembly, we recorded its X-ray diffraction pattern and compared it to simulated patterns calculated for the range of possible structures. Figure 3a is the experimental X-ray powder diffraction pattern of an isotropic sample of [α-$(M)_4$/α-$(C)_4$]$_3$ that was prepared by drying a 1:1 solution of α-$(M)_4$ and α-$(C)_4$ at ambient temperature and humidity. The diffraction pattern exhibits five distinct bands between 3 and 9 Å. The simulated diffraction pattern shown in Fig. 3a is the best match to the experimental data. This structure has a helical twist of 30°, peptide backbones in the β-sheet secondary structure, and an antiparallel strand orientation. The remarkable correspondence between the computed and experimental diffraction patterns supports the proposed structure and is very close to the structure shown in Fig. 1.

The best fit diffraction pattern shown in Fig. 3a is the result of a calculation based on four [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies that are stacked axially in helical register to generate a continuous sextuple helix with sixteen stacked hexads. Diffraction patterns for three or four stacked [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies more closely resemble the experimental pattern than do one or two stacked hexaplexes (Supplementary Fig. 9). The nMS, PAGE, and AFM data show that [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies associate in solution. Therefore, it is not surprising that stacking occurs for the dried sample used for the X-ray diffraction experiment.

Powder diffraction patterns are the result of the averaging of all possible orientations of the microcrystals in the sample. This rotational averaging produces a convoluted pattern for which the relationship between specific diffraction bands and distances within a molecular structure can be difficult to discern. We correlated individual diffraction bands with specific structural features by calculating diffraction patterns for subsets of the components that comprise the model structure. For example, computation of the diffraction pattern for a model structure containing all molecular components except the bases shows that the band at 3.38 Å is due to the stacked bases of the hexaplex (Supplementary Fig. 10). This distance is essentially identical to the 3.4 Å typical of the stacked nucleobases of DNA and RNA duplexes. Similarly, comparison of the experimental pattern with computed patterns from different model sub-structures reveals that the remaining bands in the powder diffraction pattern primarily depend on the backbone and Glu side chains (Supplementary Fig. 10).

A comprehensive search using the proto-Nucleic Acid Builder (pNAB) program[21] was carried out to determine if the structure shown in Fig. 1 uniquely fits the experimental diffraction data. This search explored all possible backbone conformations with helical twist angles between 0 and 60° for both parallel and antiparallel strand orientations. Four possible “families” of backbone structures were

identified (Supplementary Figs. 11 and 12). Of these, the family having a peptide β-sheet with antiparallel strand orientation provides the best match to the experimental pattern and is physically reasonable. Two of the other families also lack favorable backbone H-bonds or exhibit unfavorable electrostatic contacts.

The pNAB program reports backbone conformations based on reasonable bond angles and the absence of steric clashes without consideration of electrostatic interactions or hydrogen-bond geometries. Consequently, some of the peptide backbone conformations identified by pNAB exhibit unfavorable electrostatic interactions. For the same reason, pNAB does not preferentially identify structures with favorable H-bonds. To overcome this limitation, the most favorable structure identified by pNAB, based on X-ray powder diffraction analysis, was refined using an energy minimization calculation that improved the geometry of interstrand H-bonds (Supplementary Information). The resulting minor changes slightly improved agreement of the calculated diffraction pattern with the experimental data (Fig. 3a). The result of this comprehensive structural search supports the peptide backbone existing in a β-sheet conformation with antiparallel strand orientation and a hexad helical twist angle of between 29 and 31°.This minimally refined model structure is shown in Figs. 3b and c where it is compared with the original model presented in Fig. 1.

The stacking of four $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assemblies is the only alteration purposely made to improve the match between the theoretical and the experimental diffraction patterns. Importantly, we made no *ad hoc* adjustments to the predicted hexaplex backbone conformation to match features in the experimental diffraction pattern. A remaining structural uncertainty concerns the conformation of the flexible Glu side chains. A comparison of simulated diffraction patterns for model structures with and without Glu side chains revealed changes in the intensity of multiple diffraction bands (see Supplementary Fig. 10). The model structures used for powder diffraction simulations have fully extended Glu side chains. In the dry state, local contact between assemblies may affect Glu sidechain conformation, which could account for the differences in intensities observed between some diffraction bands in experimental and simulated diffraction patterns.

Kinetics of $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ Assembly Reveals Unique Cooperative Behavior.

The formation of the hexaplex requires the association of six strands and is expected to exhibit strongly concentration-dependent kinetic behavior. This expectation was confirmed experimentally by comparing the CD spectra of two samples both having a final total strand concentration of 140 μM. One sample was prepared by a High Concentration (HC) method where concentrated stock solutions of $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$ were mixed and then diluted. The second sample was prepared by a Low Concentration (LC) method where were $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$ were mixed to prepare directly the 140 μM solution (see Supplementary Information for details). The sample prepared by the HC assembly method shows a CD spectrum indicative of complete assembly within 13 min and remained unchanged for at least 3 days at 20°C (Figs. 4a). In contrast, the sample prepared by LC assembly method shows incomplete assembly: ca. 74% at 15 min and ca. 86% after 3 days. All studies described above used the HC assembly method to assure consistency.

The kinetic behavior of hexaplex assembly as a function of concentration was explored further by following the time dependent CD spectra (Supplementary Fig. 13) of samples prepared by the LC method having total strand concentrations ranging from 20 to 100 μM. After samples appeared to be at equilibrium, the concentration of free strands and assembled hexaplex were determined from fractional change in the CD signal relative to those of the strands alone and the fully assembled hexaplex. A plot of hexaplex concentration versus total strand concentration is linear and shows that the onset of assembly begins at about 7 μM (Fig. 4b). The assembly of the four covalently linked recognition units of the α-DPNA octapeptides can be compared with the assembly of Mel/Cy monomers (and related TAP and BA monomers) for which the onset of assembly occurs at concentrations between 3.5 and 50 mM depending on the particular solubilizing modification to the nucleobase.[6] Thus, when the monomers are linked with the peptide backbone, assembly occurs at a concentration 500–7000 lower than for the free monomers. This result reveals that the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly is driven by the synergistic cooperation between base stacking and the hydrogen-bonding of the peptide backbone. To our knowledge, this is the first example of a self-assembling structure to exhibit this cooperative behavior.

pH and Ionic Strength Dependence of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ Assembly Stability.

A key feature of DNA and RNA is their ability to reversibly form duplexes and higher order structures in response to a change in temperature, ionic strength and pH. We expected similar behavior for the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly. Mel has a pK_a of 5.0 (conjugate acid) and Cy has a pK_a of 6.9. Consequently, it was anticipated that the stability of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ hexaplex would exhibit a maximum at ca. pH 6.0 (the pH at which ionization of the bases is at a minimum) [43–45]. We examined the pH dependence of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly by CD spectroscopy (Fig. 4c and Supplementary Fig. 14), which reveals an increase in apparent T_m as the pH is reduced from 8.0 to 5.2. The assembly is so stable at pH between 6.5 and 5.2 that a melting transition cannot be observed within the monitored temperature range. However, a qualitative comparison of melting curves with those acquired at higher pH indicates that the melting transition of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly for pH below 6.5 is at least 100°C at the scan rate used. The increased stability below pH 6.0 is attributed to the partial protonation of the carboxylate groups of the Glu sidechains. Protonation reduces the overall charge density of the assembly, which is expected to increase stability. In addition, partial protonation enables hydrogen bonding between carboxylate groups on adjacent strands. The distance between Glu carboxylate groups in the models of $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assemblies indicates the potential for such hydrogen bonds between a carboxylic acid and a carboxylate group (a monoanion pair), which can be particularly strong[46] and would stabilize the assembly.

We also investigated the effect of ionic strength on the stability $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assemblies by CD spectroscopy. The results are shown in Fig. 4d and Supplementary Fig. 15. In a 200 mM phosphate buffer solution (pH 7.2), which contains 46 mM Na^+ the assembly exhibits an apparent T_m of ca. 55°C. This value increases when the Na^+ concentration is increased to 250 mM with NaCl, and this increase

continues as the Na^+ concentration is further increased to 1 M. This observation supports the expectation that the high charge density of hexameric assemblies from the closely spaced Glu side chains is a factor in determining the thermal stability.

Conclusion

We have demonstrated that α-DPNA octapeptides comprised of four diaminobutyric acids that are linked to either a melamine or cyanuric acid moiety (Mel^{Dab} or Cy^{Dab}) with interposing glutamic acids residues (Glu) readily assemble in water to form extraordinarily stable hexameric nucleic acid mimetics. The assembly formed by the sequences α-$(M)_4$ and α-$(C)_4$ exhibit a CD spectrum and an x-ray powder diffraction pattern that is consistent with the prediction from computational and physical models of a structure comprised of a stacked nucleobase hexad core surrounded by a hydrogen-bonded anti-parallel hydrophilic peptide backbone. The assembled structure is especially stable due to the additive effects of the complementary hydrogen bonding of the core and the peptide β-barrel-like interactions of the surrounding hydrophilic peptide backbones. It is particularly noteworthy that while these hexameric assemblies are more stable than comparable RNA or DNA duplexes, they are similarly reversibly disassembled by changes in pH and ionic strength. Near neutral pH the assembly of only four stacked nucleobase hexads exhibits an apparent T_m of 55°C. The T_m decreases at higher pH but strongly increases as the pH is reduced to 6 or lower. The strong pH and ionic strength dependence appears to result from the necessity of maintaining neutrality of the pairing bases, the pH-dependence of the backbone hydrogen bonding interactions and the high negative charge density of the assembly due to the carboxylate anions of the Glu. The imaging, electrophoresis and native mass spectroscopy results show that the [α-$(M)_4$/α-$(C)_4$]$_3$ assemblies associate further through coaxial stacking to form long linear polymers.

These findings demonstrate that the hexaplex-forming oligomers reported here are exceptionally valuable compounds for the generation of rationally designed self-assembling structures in aqueous solution. The hexaplex model structure revealed by the X-ray diffraction data can be extended indefinitely without incurring steric clash or misregistration (See SI), which indicates that much longer hexaplex structures are possible.

Of further interest is their relevance to the prebiotic origin of informational polymers. While there have been substantial advances towards finding a prebiotic route to RNA polymers,[47] some challenges remain. For example, a proposed prebiotic synthesis for the canonical pyrimidine nucleotides has been demonstrated,[48] but an analogous synthesis for the canonical purine nucleotides remains to be demonstrated.[49] Additionally, finding abiotic reactions for the polymerization of the canonical nucleotides, even with pre-activation, remains a challenge. Recent efforts to facilitate the polymerization of activated nucleotides include the use of nucleotides with modified sugars (i.e., 3'-amino-3'-deoxythreose nucleosides)[50] and solution conditions expected to promote RNA degradation (i.e., pH 10).[51] Thus, informational polymers with greater chemical stability that could have been synthesized

more easily on the prebiotic Earth are of interest as possible forerunners of RNA (or pre-RNA).[52] The oligomers presented in this study utilize a peptide backbone, which is attractive given that peptides have been generated in various model prebiotic reactions,[23] and the recognition units attached to this backbone, melamine and cyanuric acid, where also likely present on the early Earth based on their appearance in meteorites and model prebiotic reactions.[53,54] Thus, the spontaneous assembly of these oligomers in water into structures that broadly resemble extant nucleic acids further supports their viability as candidates for polymers that could have participated in the origin of life before the emergence of RNA.

Methods

Materials

Chemicals, reagents and anhydrous solvents used for the synthesis of Fmoc-protected monomers and solid phase peptide synthesis (SPPS) were purchased from Sigma-Aldrich/MilliporeSigma, Fisher Scientific, AK Scientific or VWR, and used without further purification. The Kaiser test kit used in the SPPS was purchased from AAPPTEC. Thin layer chromatography (TLC) was performed using Silica gel 60 F_{254} TLC Aluminum sheets purchased from Supelco/MilliporeSigma. Sodium phosphate monobasic monohydrate ($NaH_2PO_4{\cdot}H_2O$) and sodium phosphate dibasic heptahydrate ($Na_2HPO_4{\cdot}7H_2O$) used for preparing sodium phosphate buffer were purchased from BDH. Cacodylic acid sodium salt trihydrate purchased from Alfa-Aesar was used to prepare its buffer for pH dependent studies.

Synthesis, purification and characterization of α-DPNA oligomers

Fmoc-protected Mel^{Dab} was prepared from Boc-(l)-2,4-diaminobutyric acid and 6-chloro-1,3,5-triazine-2,4-diamine followed by replacement of the Boc protecting group with Fmoc. Fmoc-protected Cy^{Dab} was prepared in five steps from N-tert-butoxycarbonyl-(l)-homoserine. Both monomers were purified by column chromatography and characterized by NMR and High-Resolution Mass Spectrometry. The syntheses of α-DPNA octapeptides were performed manually using Wang resin following the modified Fmoc protection-deprotection protocol with repetition of the coupling steps to maximize yields. Both α-DPNA octapeptides were purified by chromatography and characterized spectroscopically. Details of the syntheses and characterization are provided in the Supplementary Information (Supplementary Schemes. 1 and 2, Supplementary Figs. 16–29).

Declarations

Competing interests

The authors declare no competing interests.

Funding

This work was supported by The Alfred P. Sloan Foundation (grant no. G-2021-16757) and by NASA Exobiology (80NSSC23K1522).

Author contributions

S.K.S., S.C.K., and N.V.H. designed research; S.K.S. performed research; S.B. and L.J.L. contributed to synthesis; G.N. performed gel electrophoresis; E.D. and L.A.C.R. performed and analyzed nMS data; V.W. supervised nMS; H.P., A.A., and N.V.H. performed computational analysis; S.K.S., J.B.S., G.B.S., and N.V.H. interpreted the data and wrote the paper. All authors have given approval to the final version of this manuscript.

Acknowledgements

We thank Dr. John Basca of the Emory University X-ray Diffraction Center.

Data availability

A detailed description of methods, CD spectra, NMR spectra, HPLC chromatograms, Native MS analysis, and LC-MS analysis are provided in the supplementary information and in the accompanying data files.

References

1. Egholm M, Buchardt O, Nielsen PE, Berg RH (1992) Pepide nucleic acids (PNA). Oligonucleotide analogues with an achiral peptide backbone. J Am Chem Soc 114:1895–1897
2. Eschenmoser A (1999) Chemical etiology of nucleic acid structure. Science 284:2118–2124
3. Anosova I et al (2015) The structural diversity of artificial genetic polymers. Nucleic Acids Res 44:1007–1021. https://doi.org/10.1093/nar/gkv1472
4. Hoshika S et al (2019) Hachimoji DNA and RNA: A genetic system with eight building blocks. Science 363:884–887. https://doi.org/10.1126/science.aat0971
5. Krueger AT, Lu H, Lee AH, Kool ET (2007) Synthesis and properties of size-expanded DNAs: toward designed, functional genetic systems. Acc Chem Res 40:141–150. https://doi.org/10.1021/ar068200o
6. Schuster GB, Cafferty BJ, Karunakaran SC, Hud NV (2021) Water-soluble supramolecular polymers of paired and stacked heterocycles: Assembly, structure, properties, and a possible path to pre-RNA. J Am Chem Soc 143:9279–9296. https://doi.org/10.1021/jacs.0c13081

7. Cafferty BJ et al (2013) Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. J Am Chem Soc 135:2447–2450. https://doi.org/10.1021/ja312155v
8. Seto CT, Whitesides GM (1990) Self-assembly based on the cyanuric acid melamine lattice. J Am Chem Soc 112:6409–6411. https://doi.org/10.1021/ja00173a046
9. Lehn J-M, Mascal M, Decian A, Fischer J (1990) Molecular recognition directed self-assembly of ordered supramolecular strands by cocrystallization of complementary molecular components. J Chem Soc Chem Commun 479–481. https://doi.org/10.1039/c39900000479
10. Lehn JM, Mascal M, Decian A, Fischer J (1992) Molecular ribbons from molecular recognition directed self-assembly of self-complementary molecular-components. J Chem Soc Perkin Trans 2:461–467. https://doi.org/10.1039/p29920000461
11. Zerkowski JA, Seto CT, Whitesides GM (1992) Solid-state structures of rosette and crinkled tape motifs derived from the cyanuric acid melamine lattice. J Am Chem Soc 114:5473–5475. https://doi.org/10.1021/ja00039a096
12. Whitesides GM et al (1995) Noncovalent synthesis - using physical-organic chemistry to make aggregates. Acc Chem Res 28:37–44. https://doi.org/10.1021/ar00049a006
13. Bohanon TM et al (1995) Barbituric-acid 2,4,6-triaminopyrimidine aggregates in water and their competitive interaction with a monolayer of barbituric-acid lipids at the gas-water interface. Angew Chem Int Ed Engl 34:58–60. https://doi.org/10.1002/anie.199500581
14. Koyano H, Bissel P, Yoshihara K, Ariga K, Kunitake T (1997) Effect of melamine-amphiphile structure on the extent of two-dimensional hydrogen-bonded networks incorporating barbituric acid. Chem Eur J 3:1077–1082. https://doi.org/10.1002/chem.19970030715
15. Marchi-Artzner V, Jullien L, Gulik-Krzywicki T, Lehn JM (1997) Molecular recognition induced aggregation and fusion between vesicles containing lipids bearing complementary hydrogen bonding head-groups. Chem Comm 117–118. https://doi.org/10.1039/a606340c
16. Huck WTS, Hulst R, Timmerman P, vanVeggel F, Reinhoudt DN (1997) Noncovalent synthesis of nanostructures: Combining coordination chemistry and hydrogen bonding. Angew Chem Int Ed Engl 36:1006–1008. https://doi.org/10.1002/anie.199710061
17. Ren YZ et al (1997) Structural study on the alternating Langmuir-Blodgett films deposited with two complementary barbituric acid and melamine derivatives. Thin Solid Films 293:170–174. https://doi.org/10.1016/s0040-6090(96)09066-9
18. Hanabusa K, Miki T, Taguchi Y, Koyama T, Shirai H (1993) 2-Component, small-molecule gelling agents. J Chem Soc Chem Commun 1382–1384. https://doi.org/10.1039/c39930001382
19. Cafferty BJ, Hud NV (2015) Was a pyrimidine-pyrimidine base pair the ancestor of Watson-Crick base pairs? Insights from a systematic approach to the origin of RNA. Isr J Chem 55:891–905. https://doi.org/10.1002/ijch.201400206
20. Fialho DM et al (2021) Depsipeptide nucleic acids: Prebiotic formation, oligomerization, and self-assembly of a new candidate proto-nucleic acid. J Am Chem Soc 143:13525–13537. https://doi.org/10.1021/jacs.1c02287

21. Alenaizan A, Barnett JL, Hud NV, Sherrill CD, Petrov AS (2021) The proto-Nucleic Acid Builder: a software tool for constructing nucleic acid analogs. Nucleic Acids Res 49:79–89. https://doi.org/10.1093/nar/gkaa1159
22. Kashida H et al (2018) Bifacial nucleobases for hexaplex formation in aqueous solution. J Am Chem Soc 140:8456–8462. https://doi.org/10.1021/jacs.8b02807
23. Frenkel-Pinter M, Samanta M, Ashkenasy G, Leman LJ (2020) Prebiotic peptides: Molecular hubs in the origin of life. Chem Rev 120:4707–4765. https://doi.org/10.1021/acs.chemrev.9b00664
24. Eschenmoser A (2005) Searching for nucleic acid alternatives. Chimia 59:836–850
25. Mittapalli GK et al (2007) Mapping the landscape of potentially primordial informational oligomers: oligodipeptides tagged with 2,4-disubstituted 5-aminopyrimidines as recognition elements. Angew Chem Int Ed Engl 46:2478–2484. https://doi.org/10.1002/anie.200603209|ISSN
26. Mittapalli GK et al (2007) Mapping the landscape of potentially primordial informational oligomers: oligodipeptides and oligodipeptoids tagged with triazines as recognition elements. Angew Chem Int Ed Engl 46:2470–2477
27. Zeng Y, Pratumyot Y, Piao X, Bong D (2012) Discrete assembly of synthetic peptide-DNA triplex structures from polyvalent melamine-thymine bifacial recognition. J Am Chem Soc 134:832–835. https://doi.org/10.1021/ja2099326
28. Brack A, Orgel LE (1975) Beta structures of alternating polypeptides and their possible prebiotic significance. Nature 256:383–387
29. Brack A (1977) Beta-structures of alternating polypeptides and their possible role in chemical evolution. BioSystems 9:99–103. https://doi.org/10.1016/0303-2647(77)90018-1
30. Mandel-Gutfreund Y, Gregoret LM (2002) On the significance of alternating patterns of polar and non-polar residues in beta-strands. J Mol Biol 323:453–461. https://doi.org/10.1016/s0022-2836(02)00973-7
31. Rubinov B, Wagner N, Rapaport H, Ashkenasy G (2009) Self-Replicating Amphiphilic β-Sheet Peptides. Angew Chem Int Ed Engl 48:6683–6686. https://doi.org/10.1002/anie.200902790
32. Bowerman CJ, Nilsson BL (2012) Review self-assembly of amphipathic β-sheet peptides: Insights and applications. Biopolymers 98:169–184. https://doi.org/10.1002/bip.22058
33. Xiong Q, Stupp SI, Schatz GC (2021) Molecular Insight into the β-Sheet Twist and Related Morphology of Self-Assembled Peptide Amphiphile Ribbons. J Phys Chem Lett 12:11238–11244. https://doi.org/10.1021/acs.jpclett.1c03243
34. Rout SK, Friedmann MP, Riek R, Greenwald J (2018) A prebiotic template-directed peptide synthesis based on amyloids. Nat Commun 9:8. https://doi.org/10.1038/s41467-017-02742-3
35. Kuroiwa K, Tanimura Y, Wang MF, Morisue M, Hasegawa Y (2025) Diblock Copolypeptide Amphiphiles-Mediated Assembly and Structural Modulation of Porphyrin Dimers in Aqueous Media. Langmuir 12. https://doi.org/10.1021/acs.langmuir.5c04712

36. Alenaizan A et al (2021) X-ray Fiber Diffraction and Computational Analyses of Stacked Hexads in Supramolecular Polymers: Insight into Self-Assembly in Water by Prospective Prebiotic Nucleobases. J Am Chem Soc 143:6079–6094. https://doi.org/10.1021/jacs.0c12010
37. Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc 1:2876–2890. https://doi.org/10.1038/nprot.2006.202
38. Micsonai A et al (2022) BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. Nucleic Acids Res 50:W90–W98. https://doi.org/10.1093/nar/gkac345
39. Castillo RM et al (1999) A six-stranded double-psi beta barrel is shared by several protein superfamilies. Structure 7:227–236. https://doi.org/10.1016/s0969-2126(99)80028-8
40. Kim DE et al (2023) De novo design of small beta barrel proteins. *Proc. Natl. Acad. Sci. USA* 120, e2207974120 https://doi.org/10.1073/pnas.2207974120
41. Karunakaran SC et al (2018) Exquisite regulation of supramolecular equilibrium polymers in water: chain stoppers control length, polydispersity and viscoelasticity. Polym Chem 9:5268–5277. https://doi.org/10.1039/c8py01146j
42. Karch KR, Snyder DT, Harvey SR, Wysocki VH (2022) Native Mass Spectrometry: Recent Progress and Remaining Challenges. Annu Rev Biophys 51:157–179. https://doi.org/10.1146/annurev-biophys-092721-085421
43. Cafferty BJ, Avirah RR, Schuster GB, Hud NV (2014) Ultra-sensitive pH control of supramolecular polymers and hydrogels: pK_a matching of biomimetic monomers. Chem Sci 5:4681–4686. https://doi.org/10.1039/c4sc02182g
44. Krishnamurthy R (2012) Role of pKa of nucleobases in the origins of chemical evolution. Acc Chem Res 45:2035–2044. https://doi.org/10.1021/ar200262x
45. Ma M, Bong D (2011) Determinants of cyanuric acid and melamine assembly in water. Langmuir 27:8841–8853. https://doi.org/10.1021/la201415d
46. Cassidy CS, Lin J, Tobin JB, Frey PA (1998) Low-Barrier Hydrogen Bonding in Aqueous and Aprotic Solutions of Dicarboxylic Acids: Spectroscopic Characterization. Bioorg Chem 26:213–219. https://doi.org/https://doi.org/10.1006/bioo.1998.1102
47. Hirakawa Y et al (2025) Interstep compatibility of a model for the prebiotic synthesis of RNA consistent with Hadean natural history. *Proceedings of the National Academy of Sciences* 122, e2516418122 https://doi.org/10.1073/pnas.2516418122
48. Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242
49. Yadav M, Kumar R, Krishnamurthy R (2020) The chemistry of abiotic nucleotide synthesis. Chem Rev 120:4766–4805
50. Sammons SR, Giurgiu C, Szostak JW, O'Flaherty DK (2026) Potentially Prebiotic Synthesis of a 3'-Amino-3'-deoxyribonucleoside. J Am Chem Soc 148:206–216. https://doi.org/10.1021/jacs.5c08614

51. Caimi F et al (2025) High-Yield Prebiotic Polymerization of 2',3'-Cyclic Nucleotides under Wet–Dry Cycling. ACS Cent Sci 11:1546–1557. https://doi.org/10.1021/acscentsci.5c00488
52. Hud NV, Cafferty BJ, Krishnamurthy R, Williams LD (2013) The origin of RNA and 'My Grandfather's Axe'. Chem Biol 20:466–474
53. Hayatsu R, Studier MH, Oda A, Fuse K, Anders E (1968) Origin of organic matter in early solar system 2. Nitrogen compounds. Geochim Cosmochim Acta 32:175–190. https://doi.org/10.1016/s0016-7037(68)80003-1
54. Menor-Salvan C, Ruiz-Bermejo DM, Guzman MI, Osuna-Esteban S, Veintemillas-Verdaguer S (2009) Synthesis of pyrimidines and triazines in ice: Implications for the prebiotic chemistry of nucleobases. Chem Eur J 15:4411–4418. https://doi.org/10.1002/chem.200802656

Figures

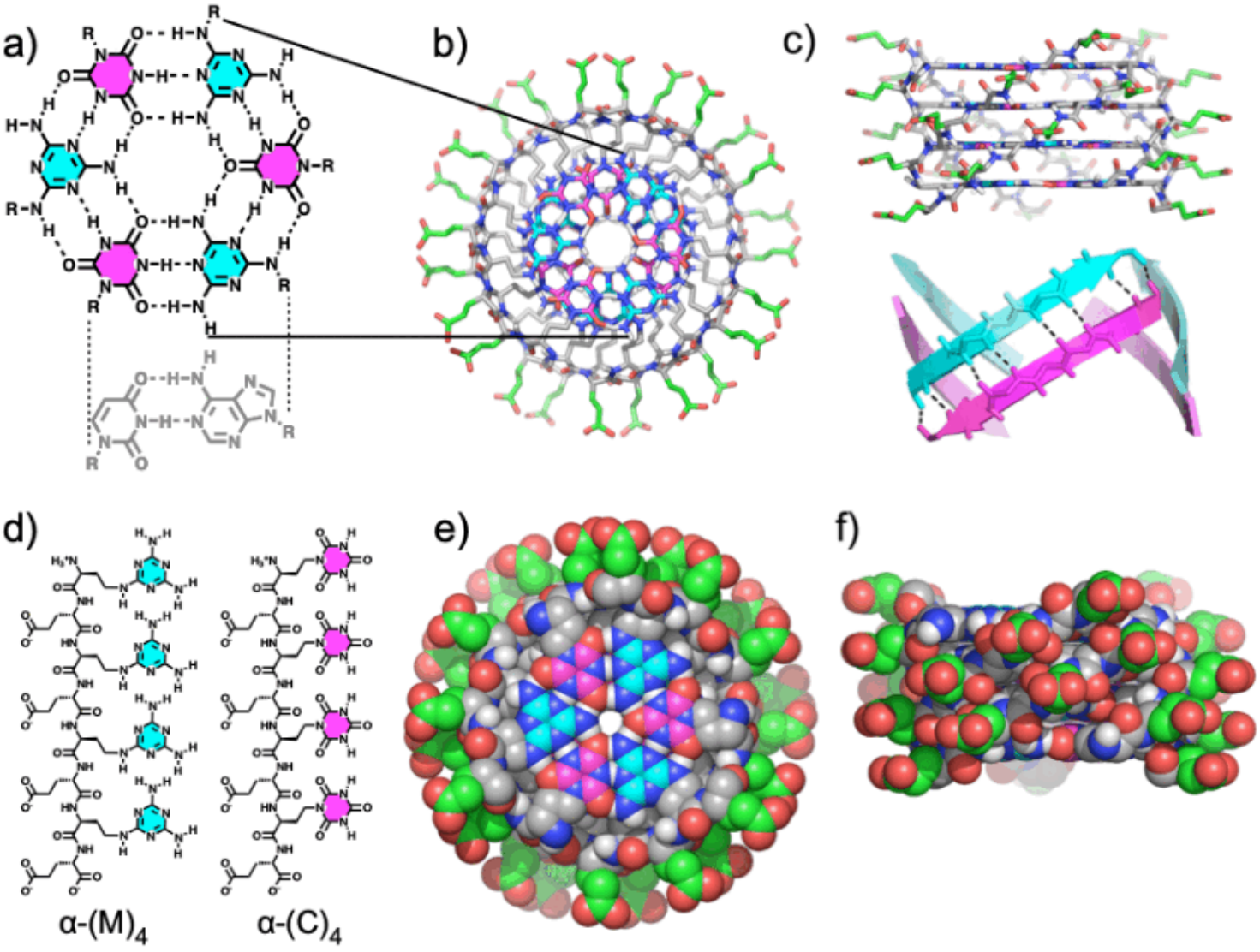


Figure 1

a H-bonded pairing within a Mel-Cy hexad. The Watson-Crick base pair presented below the hexad illustrates how the connection of a base nitrogen atom of Cy and an exocyclic nitrogen atom of Mel provide a similar pairing geometry. **b** Stick representation of a computational model created for the anti-parallel hexaplex $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ (axial view). Straight lines to (a) connect to the hexad formed by the pairing bases at the center of the hexaplex. **c** Side view of $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ as stick model and with backbones as β-sheet arrows with backbone atoms and interstrand H-bonds shown for the two strands in front. **d** Chemical structures of $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$. **e**, **f** Axial and side views of space-filling representation of $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$. Color coding: Mel rings are cyan in (a) and (d); Cy rings are magenta in (a) and (d). In 3D computational models, carbons of Mel bases are cyan, and carbons of Cy bases are magenta.

Structural models of α-DPNA oligomers and their expected assemblies.

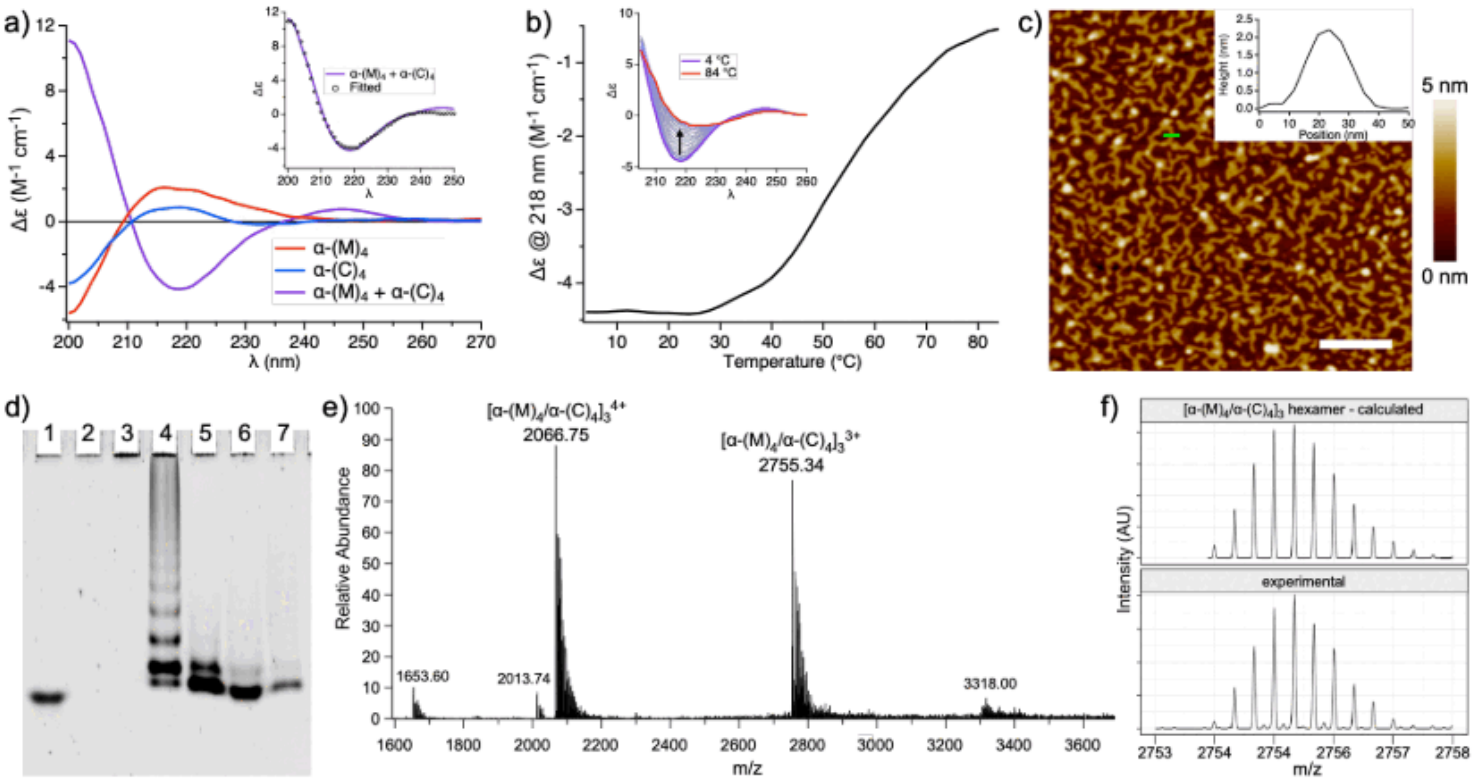


Figure 2

Analysis of the assembly formed by a 1:1 mixture of $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$.

a CD spectra of $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$ separately (50 μM in strand) and as a 1:1 mixture (50 μM in each strand). The distinctive CD spectrum of the mixture compared to the individual strands indicates their association in solution. Inset of (a) CD spectrum of the 1:1 mixture of $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$ (solid line) is fit well by the program BeStSel[38] (circles), which indicates a β-sheet secondary structure. **b** Thermal denaturation of $\alpha\text{-}(M)_4/\alpha\text{-}(C)_4$ assembly (50 μM in each strand) as monitored by the intensity of CD signal at 218 nm. Inset of (b) CD spectra of the 1:1 mixture of $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$ heated from 4 to 84 °C with a

temperature interval of 2 °C. **c** AFM image of assemblies formed by the 1:1 mixture of α-$(M)_4$ and α-$(C)_4$ when dried on a freshly cleaved mica substrate. Color coding of height measurements show that all assemblies are 2 to 3 nm thick, as expected. Inset of (c) Height measurement of an individual assembly at the location indicated by the green horizontal line in AFM image. Scale bar is 200 nm. **d** Image of gel from PAGE analysis of the assembly formation by α-$(M)_4$ and α-$(C)_4$. Lane1: d(TGGGGT) as mobility standard; Lane 2: α-$(M)_4$; lane 3: α-$(C)_4$; lane 4: 1:1 mixture of α-$(M)_4$ and α-$(C)_4$; lanes 5, 6, and 7: 1:1 mixture of α-$(M)_4$ and α-$(C)_4$ with 1 mM, 3 mM and 10 mM Ethidium Bromide (EtBr), respectively. Lanes 2-7 were loaded with 2 μL of samples that is 50 μM in each strand indicated. **e** Positive mode native Mass Spectrometry (nMS) of a 1:1 mixture of α-$(M)_4$ and α-$(C)_4$, 1 μM in each strand, with 4 μM EtBr added after strand mixing. Major peaks correspond to triply and quadruply charged species of the hexaplex composed of three strands each of α-$(M)_4$ and α-$(C)_4$ (i.e., [α-$(M)_4$/α-$(C)_4$]$_3$), along with its Na^+ and K^+ adducts (shown in Supplementary Fig. 3). **f** Isotopic analysis of the triply charged [α-$(M)_4$/α-$(C)_4$]$_3$ ion. Top: calculated isotopic mass distribution. Bottom: experimentally observed isotopic mass distribution. Additional details for each experiment described here are provided in Supporting Information.

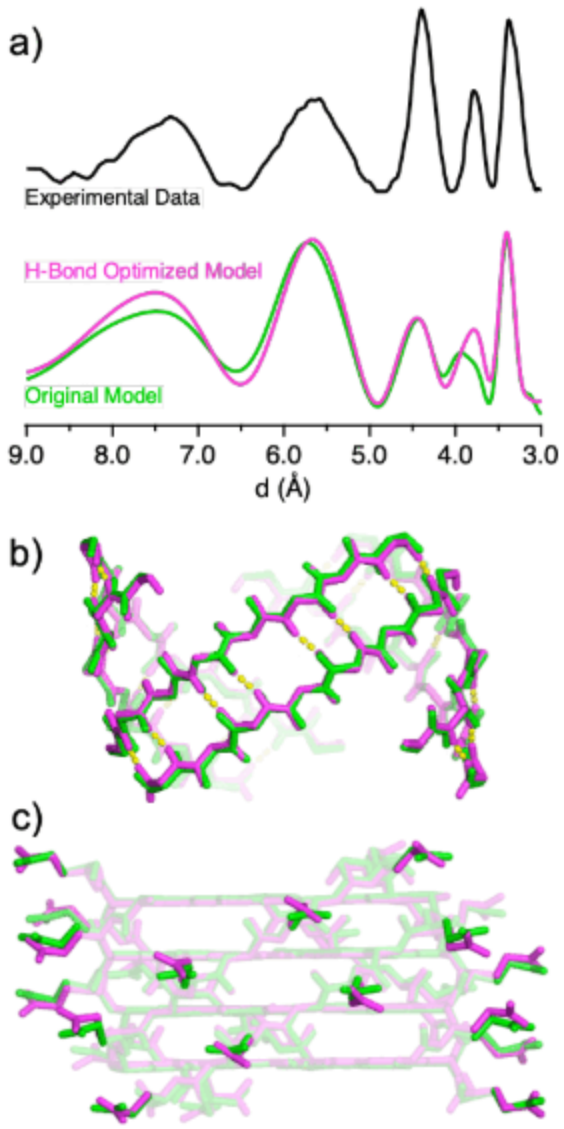


Figure 3

X-ray powder diffraction analysis of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly.

a Black curve: experimental powder diffraction data of a sample containing equal molar amounts of α-$(M)_4$ and α-$(C)_4$ that was dried at ambient temperature and humidity prior to collection of diffraction data; Magenta curve: simulated diffraction pattern of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ hexaplex model identified by the pNAB program as the most favorable structure after energy minimization with restraints to improve inter-strand H-bond geometry; Green curve: simulated diffraction pattern of the original $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ hexaplex model presented in Fig. 1. **b** Backbones in stick representation of (magenta) the model structure of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly identified as the most favorable structure by the pNAB program after optimization of H-bond geometry and (green) the original model structure presented in Fig. 1. **c** Same structures as in (b), but with all atoms shown in low contrast and with the side chains of the Glu residues in the foreground in high contrast are emphasized to illustrate the greatest difference between the two structures.

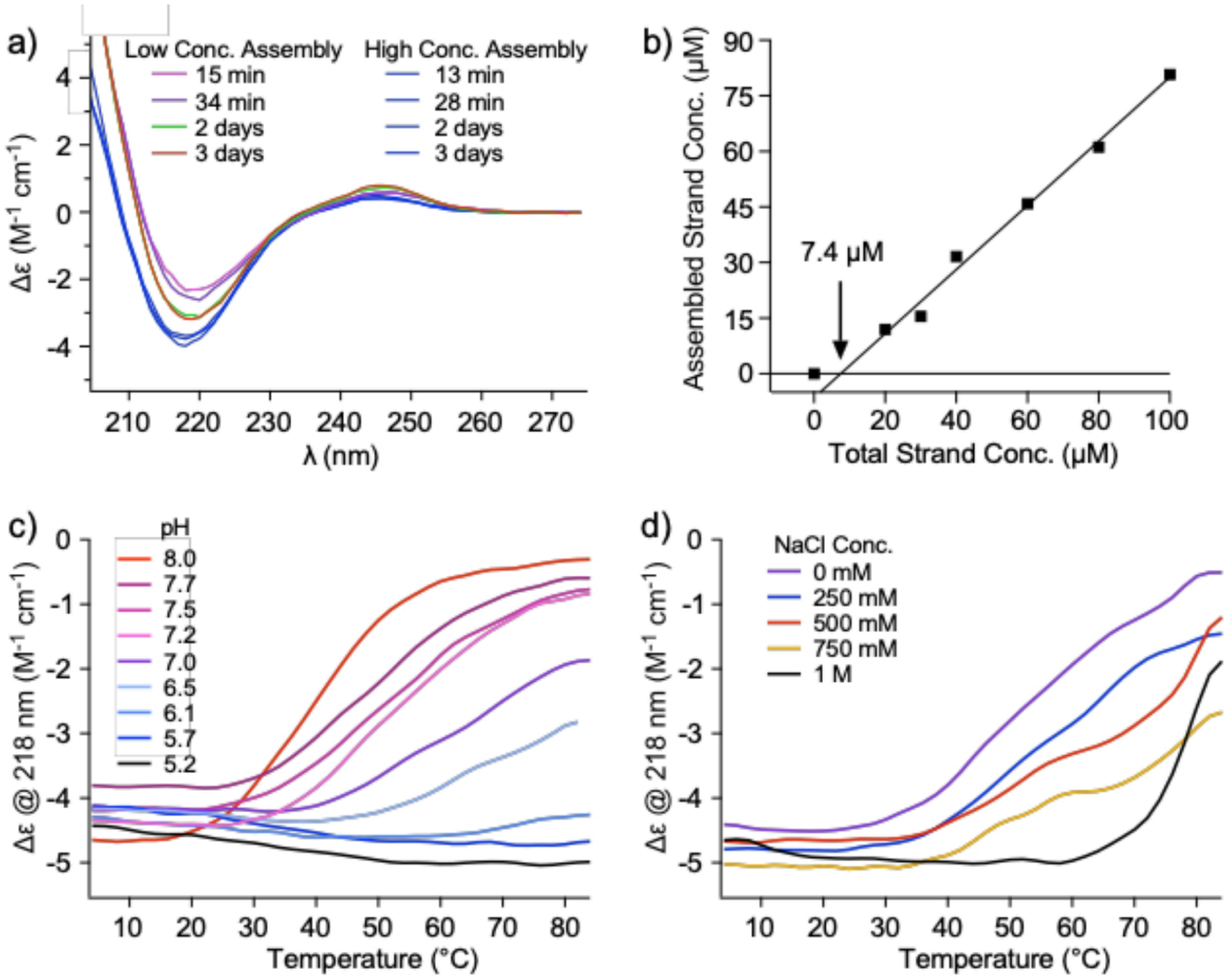


Figure 4

Analysis of $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly kinetics, equilibrium, and thermal stability as a function of pH and ionic strength.

a CD spectra illustrating the impact of strand concentration on the kinetics of maximal hexaplex formation. All samples contained 70 μM of each $\alpha\text{-}(M)_4$ and $\alpha\text{-}(C)_4$ strand, for a 140 μM total strand concentration. Spectra labeled “Low Conc. Assembly” were generated by mixing stock strand solutions that were initially 140 μM in each strand, diluted to 70 μM for each strand upon mixing. Spectra labeled “High Conc. Assembly” were generated by mixing strand stock solutions of 10 mM to create a sample that was 327 μM in each strand before dilution to 70 μM. The magnitude of negative band at 218 nm in the Low Conc. Assembly sample grew over time at 20 °C, showing that it required more than 3 days to equilibrate to maximal assembly. The greater magnitude of the negative band at 218 nm in spectra of the High Conc. Assembly sample and the constancy of the magnitude of this band over time indicates that a sample prepared by the high concentration protocol equilibrated and reached a maximal assembly

within 13 minutes. From these data, a value of -30.4 ± 2 $M^{-1}cm^{-1}$ was determined for molar circular dichroism, Δε, of the assembled particle (expressed in strand concentration). **b** Equilibrium of strand assembly. The concentrations of free strand and hexaplex concentrations were determined from fully equilibrated CD spectra as described in the text. The plot of hexaplex concentration versus total strand concentration show a “hockey stick” shape characteristic of concerted strand assembly, analogous to micelle assembly, with the onset of assembly at 7.4 μM. **c** Thermal stability measurements of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly as monitored by CD signal at 218 nm for various values of solution pH. **d** Thermal stability measurements of the $[\alpha\text{-}(M)_4/\alpha\text{-}(C)_4]_3$ assembly as monitored by CD signal at 218 nm for various concentrations of NaCl added to an initial buffer of sodium phosphate, pH 7.2. Additional details on experiments shown are provided in the Supporting Information.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TetraMelTetraCyOctamers.txt
- HomoMelHomoCy32mers.txt
- SajenaSaraswathySIOct26.pdf