Metal-Mediated DNA Nanotechnology in 3D: Structural Library by Templated Diffraction

Authors: Simon Vecchioni,[1]†* Brandon Lu,[1]† William Livernois,[4] Yoel P. Ohayon,[1] Jesse B. Yoder,[2] Chu-fan Yang,[1] Karol Woloszyn,[1] William Bernfeld,[1,3] M.P. Anantram,[4] James W. Canary,[1] Wayne A. Hendrickson,[5] Lynn J. Rothschild,[6] Chengde Mao,[7] Shalom J. Wind,[8] Nadrian C. Seeman[1]‡, Ruojie Sha[1*]

Affiliations: [1]Department of Chemistry, New York University, New York, NY 10003, USA; [2]IMCA-CAT, Argonne National Lab, Argonne, IL 60439, USA; [3]ASPIRE Program, King School, Stamford, CT, USA; [4]Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195, USA; [5]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA; [6]NASA Ames Research Center, Planetary Sciences Branch, Moffett Field, CA 94035, USA; [7] Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA; [8]Columbia University, Department of Applied Physics and Applied Math, Columbia University, New York, NY 10027, USA.

†These authors contributed equally to this work. ‡Deceased *Corresponding authors: R.S. (ruojie.sha@nyu.edu) and S.V. (sv1091@nyu.edu).

Abstract: DNA double helices containing metal-mediated DNA (mmDNA) base pairs have been constructed from Ag^{+} and Hg^{2+} ions between pyrimidine:pyrimidine pairs with the promise of nanoelectronics. Rational design of mmDNA nanomaterials has been impractical without a complete lexical and structural description. Here, we explore the programmability of structural DNA nanotechnology toward its founding mission of self-assembling a diffraction platform for biomolecular structure determination. We employed the tensegrity triangle to build a comprehensive structural library of mmDNA pairs *via* X-ray diffraction and elucidated generalized design rules for mmDNA construction. We uncovered two binding modes: N3-dominant, centrosymmetric pairs and major groove binders driven by 5-position ring modifications. Energy gap calculations showed additional levels in the lowest unoccupied molecular orbitals (LUMO) of mmDNA structures, rendering them attractive molecular electronic candidates.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/adma.202210938.

Main Text: The programmability of DNA base pairing has been investigated for nearly forty years as a platform for the self-assembly of designer nanomaterials.[1] More recently, DNA oligomers containing metal-mediated DNA (mmDNA) base pairs have been shown to possess enhanced electrical conductivity relative to their canonical analogs.[2] This suggests that incorporation of mmDNA into self-assembled DNA nanostructures by replacing Watson-Crick hydrogen bonding schemes with the precise coordination of specific metal ions at the core of the double helix[2b, 3] may offer superior electronic and magnetic properties.[4] The discovery of the mercury(II)-mediated thymine-thymine base pair ($dT:Hg^{2+}:dT$)[5] and the silver(I)-mediated cytosine-cytosine base pair ($dC:Ag^{+}:dC$)[3a] has elevated mmDNA as a means of programming bioinorganic DNA complexes using unmodified nucleobases—a critical characteristic for both biological expression and simple chemical synthesis.[6] These metal-mediated interactions exploit gaps between "mismatched" pyrimidine nucleobases to insert metal ions with high specificity, typically at the N3 position (Fig. 1C).[4a, 7] Twenty years of study elucidated the enhanced aqueous thermostability of mmDNA,[8] strong selectivity as sensors and aptamers,[9] and enhanced molecular conductivity in consecutive $dC:Ag^{+}:dC$ base pairs.[2, 10] The addition of chemical modifications to the ring structure of the pyrimidines generated a variety of binding schemes.[7c, 11] Despite several salient examples,[3b, 7a, 12] it has remained a great challenge to obtain uniform structural information of mmDNA base pairs, and this has ultimately precluded the development of complex nanotechnologies based on mmDNA.[13] Herein, we report a general method to solve this problem by directly employing structural DNA nanotechnology.

We utilized the symmetric tensegrity triangle, a DNA motif that reliably self-assembles into 3D rhombohedra via sticky end cohesion (Fig 1A).[14] At the center of each helix of the triangle, we positioned a pyrimidine:pyrimidine pair and carried out the self-assembly process with twofold excess of Ag^{+}, Hg^{2+}, or Au^{+}, resulting in macroscopic crystalline structures of diffraction quality (Fig. 1B,C). Modifications were made at four functional sites of the pyrimidine aromatic ring to assay the effects on mmDNA base pair morphology. Two families of nucleobases were chosen, those resembling cytosine with one amine moiety, one carbonyl moiety, and a deprotonated N3 heteroatom: deoxycytosine [dC], 5-methyl-deoxy*iso*cytosine [dS or IMC],[15] and 5-methyl-deoxycytosine [dC* or 5CM]; and those resembling thymine with two carbonyl moieties and a protonated N3 atom: deoxythymine [dT], deoxyuracil [dU] and 5-iodo-deoxyuracil [diU or 5IU] (Fig. 1D, Supplemental

Appendix). Methylcytosine was chosen to mimic dT, which together serve as the 5-methylated analogs to dC and dU, respectively. Methylcytosine was also chosen for its role in epigenetics and Z-DNA formation.[16] Isocytosine was employed to study the electronic effect of exchanging the cytosine functional groups at the 2 and 4 positions, as well as its recent use in 6-letter and 8-letter DNA alphabets.[15] Iodouracil was employed to study the effect of halogenation in lieu of 5-position methylation.[17] Unless otherwise noted, crystals were formed through a thermal protocol described previously[14b] in hanging drops composed of 6-12 µM DNA triangles, pH 8.0 MOPS buffer (10 mM MOPS, 125 mM MgSO4, 10 mM NaOH) and twofold excess metal ion ($AgNO_3$, $HgCl_2$) added directly before annealing. Crystals were equilibrated against a 10X reservoir while the temperature was slowly decreased from 65 ºC to 20 ºC at a rate of 0.4 ºC/ hr. Further details can be found in the SI. Structures diffracted to a maximum anisotropic[18] resolution of 2.94 Å (Table S1, Fig. S1), consistent with similar Watson-Crick designs (see PDB ID: 5W6W, 3.06 Å, PDB ID: 8D96, 2.96 Å). Metal ions were used as phasing agents to obtain unbiased phases of the structures via anomalous dispersion techniques (Fig. 2). Ion disambiguation, site assignation, and occupancy determination were carefully carried out using established anomalous f″ refinement techniques (see SI for extensive discussion, Fig. S2).[19] The resulting mmDNA structures can be found in Fig. 3 (see SI for detailed deposition information).

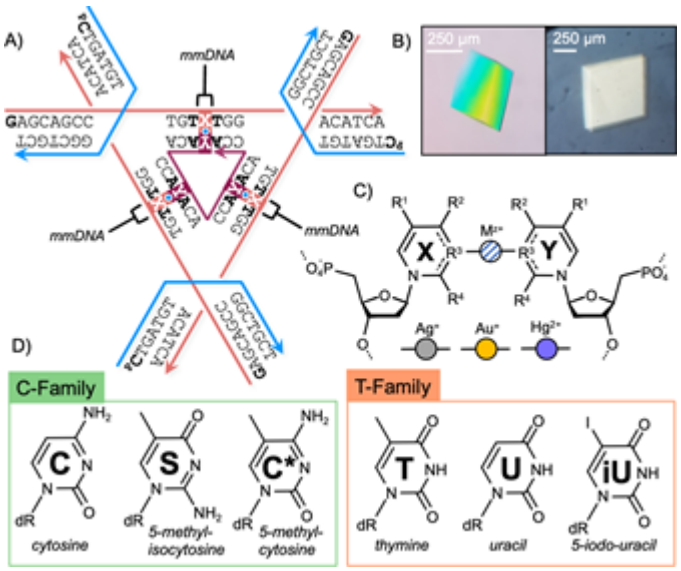


Figure 1: mmDNA design schematic used in this study. **A)** The DNA tensegrity triangle[14b] was modified to contain a metal-responsive, pyrimidine:pyrimidine (X:Y) "mismatch" at the center of each asymmetric unit. **B)** Remarkably, this motif self-assembled into crystals up to 500 μm in size (scale inset) via sticky-end cohesion. **C)** The generalized pyrimidine:pyrimidine pair is shown with the functional groups modified in this study. Atomic sites are numbered in blue. At the 5 position (blue labels inset), R^1=H, CH_3, I; at the 4 position, $R^2 = NH_2$, O; at the 3 position, unreacted T-family amines are protonated at neutral pH; at the 2 position, $R^3 = NH_2$, O. **D)** These pyrimidines fall into two general categories: C-like and T-like nucleobases.

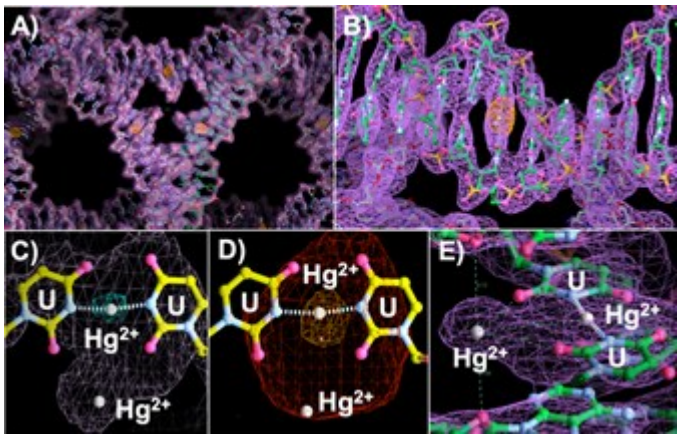


Figure 2: Electron density maps of the dU:Hg^{2+}:dU mmDNA base pair and the tensegrity triangle structure at 2.94 Å resolution (PDB ID: 7SD8). **A-B)** Whole motif and stacking planes, with the *2mFo-DFc* density map contoured at 2.5 σ, with sharpening and anomalous map contoured at 10.0 σ (Coot). **C)** Electron density shows two high-density Hg^{2+} atoms (occupancies *0.80*, *0.20*). Highest density site shown with *2mFo-DFc* at 9.0 σ (cyan map). **D)** Anomalous map contoured at 15.0 σ (orange) and 3.0 σ (red). **E)** The lower-occupancy O2-position mercury ion is strongly contained by the density, with a coordination pocket formed by minor groove oxygen atoms.

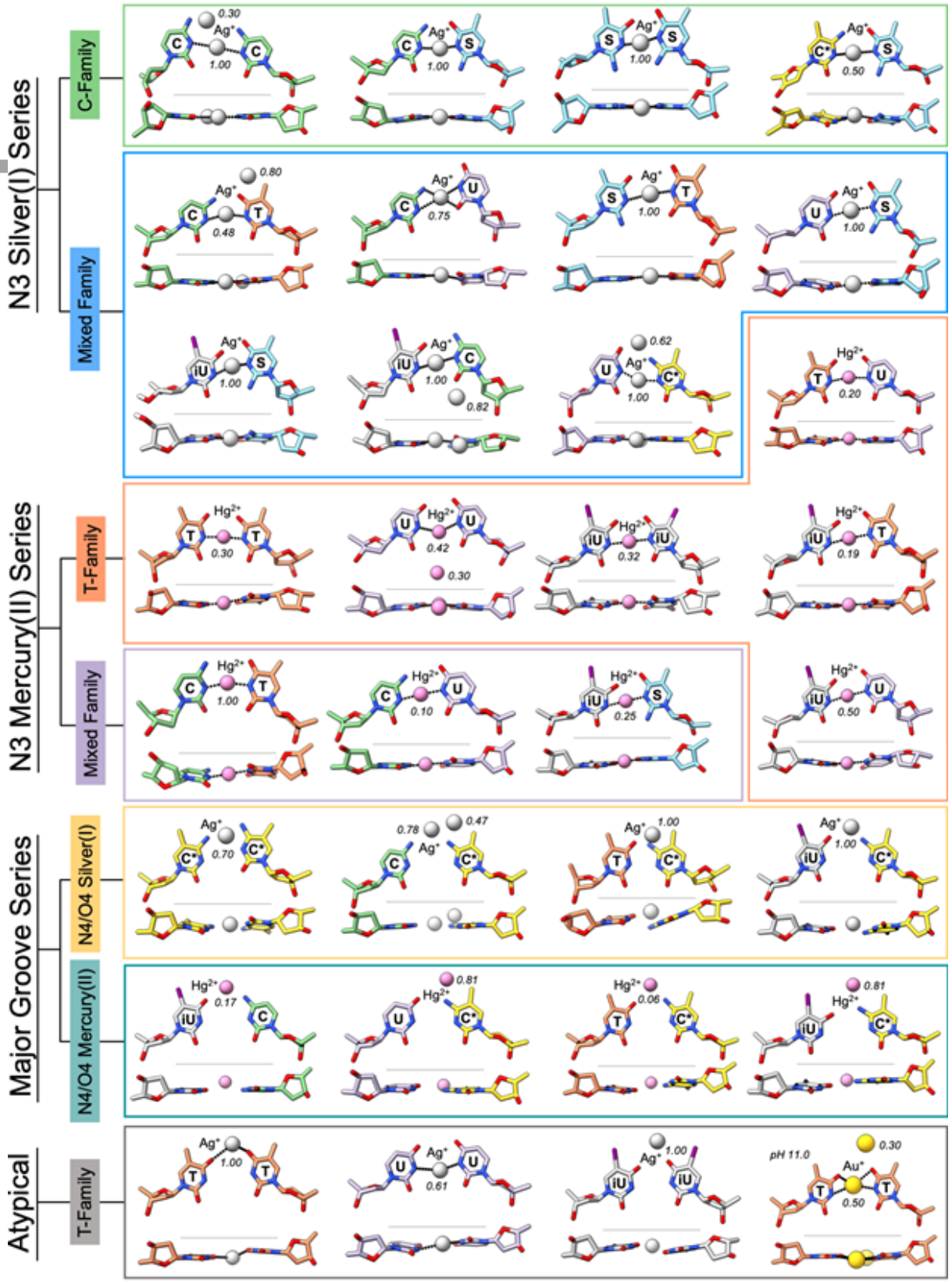


Figure 3: Structural library of all the crystallized mmDNA base pairs in this study. Nucleobase carbon atoms are color coded (cytosine (C): green; 5-methyl-isocytosine (S): cyan; 5-methyl-cytosine (C*): yellow; thymine (T): orange; uracil (U): lavender; 5-iodo-uracil (iU): white) with heteroatoms displayed (O: red, N: blue). Ag^{+} (grey), Hg^{2+} (pink) and Au^{+} ions (gold) are shown with their respective occupancies (italic numbers). Base pairs are shown down the helical axis (top image in each pair) and along the radial axis from the viewpoint of the minor groove (bottom image in each pair). Pairs are organized by binding modality (N3 or major groove), ion identity, and intra- or mixed-family pairing (hierarchy shown at left of panel). An index of base pairs by nucleotide with PDB accession codes can be found in the Supplemental Appendix.

Cytosine-family N3 silver(I) pairs

Structures of mmDNA base pairs reported here were classified in two categories based on the highest density ion site: N3-binders and major-groove binders, with minor overlap between the groups. Within these groups, we differentiated them by ion identity, intra-family or inter-family base pairs (C-like and T-like together). To our surprise, we found that the dC:Ag^{+}:dC base pair possesses both binding modes, the well-studied N3 structure,[6, 8a, 12b] as well as a degenerate N4-Ag^{+}-N4 bond that draws the ion toward one dominant N4 amine. In our study, we found that any small perturbation to the buffer chemistry or sample preparation (such as pH elevation, Cl^{-} contamination, etc.) drives the occupancy to strongly favor the N4 position over the commonly cited N3 site. This supports early findings that the dC:Ag^{+}:dC bond is somewhat ill-behaved.[8b, 20] By contrast, the isocytosine performed as the ideal silver(I)-pairing nucleobase. The exchange of a soft N4 with a harder O4 moiety drastically reduced the preponderance of major groove ion pairing, overriding even the nearly universal tendency of methylcytosine to disfavor N3 (see below). The dC:Ag^{+}:dS and dS:Ag^{+}:dS base pairs were found to have ideal geometries and occupancies as they appear to be both robust and isomorphic with an average Ag^{+}-N3 distances of 1.8 Å (Table S16), in good agreement with previously reported thermodynamic studies.[21] Finally, the isocytosine nucleobase strongly disfavors mercury(II) base pairs, eluding crystallization altogether (in contrast with non-metalated mismatches that can still crystallize—see Fig. S5). In our experimental work, only one isocytosine-mercury(II) crystal was ever obtained, which yielded the structure of diU:Hg^{2+}:dS and exhibited N3 geometry. The N3 site dominance, together with major groove site reduction and orthogonality to Hg^{2+}, makes isocytosine the ideal Ag^{+}-binding nucleotide.

Thymine-family N3 mercury(II) pairs

In keeping with existing structures, dT:Hg^{2+}:dT exhibits a single, N3-bound ion. Together with the other mmDNA base pairs in this series, we found generally low occupancy Hg^{2+} ions with symmetric N3 geometry at a bond distance of ~2.1 Å (Table S3). Lower occupancies in Hg^{2+} mmDNA agrees with previous studies.[12a] Overall, the diffraction resolution of these pairs was higher than Ag^{+} mmDNA (Fig. 2, Table S2), and at sub-3 Å resolution, we were able to observe two sites in the dU:Hg^{2+}:dU base pair: a standard N3 ion (occupancy *0.80*) with a bond distance of 2.1 Å, and a minor groove, O2 binding mercury ion (occupancy *0.20*), which formed a good coordination pocket with both the O2 atoms (at distances of 3.4 Å and 3.6 Å), the ribosyl O4′ atom (3.3 Å), and probable waters in the solvent shell. Similar geometry in the dC:Hg^{2+}:dT base pair was observed in the F_o-F_c density maps, but diffraction resolution (3.63 Å) was insufficient to assign the site. These results suggest that Hg^{2+} has an affinity for other minor groove sites throughout the helix, which may be elucidated at higher resolution. The ability of uracil to bind more than one ion is underscored in its interactions with other nucleobases shown below.

Mixed-family N3 mmDNA pairs

N3 base pairs composed of one member each of the C- and T-family nucleotides, "mixed" family pairs, showed surprising structural diversity. These pairs have N3-M distances of 1.8-2.4 Å. Of particular note were the secondary binding sites in the major and minor grooves, which have been predicted in similar structures of A-form dC:Ag^{+}/Hg^{2+}:dT and more exotic pyrimidines.[12a, 22] From these structures, we note that Ag^{+} binding appeared more labile than Hg^{2+}—the structural diversity of Ag^{+} pairs is higher, supporting multiple coplanar ions per base pair in several cases (dC:Ag^{+}:dT, diU:Ag^{+}:dC, dU:Ag^{+}:dC*). Additional Ag^{+} may manifest both at the O4/N4 and O2/N2 sites, while secondary Hg^{2+} predominated at the minor groove O2/N2 positions. These secondary groove-binding ions appeared in many instances within covalent-bonding distances (see Table S16). Isocytosine supports high-occupancy Ag^{+} mmDNA with all the T-family nucleobases and a low-occupancy Hg^{2+} bond with iodouracil. Of particular note is the behavior of uracil: with the exception of isocytosine (which does not form mercury(II) pairs) and methylcytosine (major groove N4-Hg^{2+}), uracil can form a symmetric N3 mmDNA pair with every other tested nucleobase. Uracil even overrides the tendency of methylcytosine to form an N4-only Ag^{+} bond, forming a co-stabilized N3-Ag^{+}-N3 and O4-Ag^{+}-N4

mmDNA pair. We further note that uracil forms an N3 homopair with Ag^+. With these data, it can be concluded that uracil is a versatile, near-universal mmDNA-forming nucleobase.

Major Groove Series

To our surprise, eight major-groove-only mmDNA base pairs were observed. Seven of these N4 pairs involved methylcytosine, and the final an Hg^{2+} pair between cytosine and iodouracil. This leads to the unexpected conclusion that modifications to pyrimidines away from the base pairing face, at the 5 position, can reconfigure the electronic structure of the base pairing functional groups to pull the metal ions out of the center of the helix. More specifically, the electron donating 5-methyl modification consistently activated the 4-position amine to render it more electron rich. The resulting metal bonds tend to favor the methylcytosine N4 amino group over that of its pairing partner, with bond distances of 1.9-2.6 Å (Table S16). The increased metal proclivity of the 4-position functional group is consistent with other studies observing thermostability changes in the presence of 5-position thymine modifications,[17] and poor metal ion capture for methylcytosine.[23] The limitations of earlier studies are surely related to the unexpected geometry and general asymmetry of these base pairs—the metal ion tightly bound to a primary amine may not be forming a true base pair, but rather an intermediary, electrostatically-stabilized interaction. It is thus unlikely that these N4 pairs strongly affect the thermodynamic stability of the duplex, leading to their relative invisibility to other analyses.

Non-Paradigmatic, Atypical Base Pairs

We further scanned homopyrimidine pairs against their non-paradigmatic metal partners (C-family and Hg^{2+}, T-family and Ag^+). No metal sites were observed for C-family pairs annealed with Hg^{2+} in the conditions tested here, but the T-family pairs were reliably able to coordinate Ag^+ ions. This phenomenon has been previously observed by diffraction under select conditions, such as the non-B-form nanowire reported by Kondo and colleagues, as well as in 4-thio-thymine homopairs [3b, 22]. However, this binding mode for Ag^+ T-family bases that has not been previously detected via solution-based studies using NMR and thermal denaturation.[24] Here we show that in a B-form geometry, U:U forms a stable N3 Ag^+ base pair, consistent with other uracil base pairs in our library (Fig. 3). By contrast, T:T and iU:iU pairs form O4, major groove Ag^+ pairs, underscoring the effect of 5-position functional groups on ion coordination in DNA.

Finally, we synthesized a soluble Au^{+} salt (AuCl(TPPTS), see SI) to test whether base pairs in our library could bind a non-standard transition metal with N3 coordination. We performed a scan across pH and buffer composition, and identified a stable dT:$(Au^{+})_2$:dT di-metal base pair at pH 11.0, in which two Au^{+} ions are stabilized at the N3 and O4 positions (Fig. 3). A previous study identified a stable rC:Au^{3+}:rG RNA base pair with square planar geometry. In the present case, the ligand exchange between AuCl(TPPTS) and the dT-dT base pair in this work likely preserves the charge state on Au^{+}, as crystals were not obtained with Au^{3+} ions. The ensuing base pair exhibits N3-Au^{+}-N3 geometry, with possible square planar coordination with the O4 groups. This result suggests that future work may be carried out to elucidate the effects of pH and non-paradigmatic metal coordination in DNA structures.

Comparison with Other mmDNA Structures

We present a detailed comparison of those mmDNA pairs which have structural analogs in previous studies, and we find that the mmDNA pairs reported here generally agree with prior work (Fig. 4). Earlier reports of dC:Ag^{+}:dC mmDNA pairs include pseudo-inifinite nanowires composed of Ag^{+} ions, both with and without full continuity between ions (PDB ID's: 7XKM, 5IX7).[3b, 3c] These arrays, as well as a the dC:Ag^{+}:dC pair in a parallel helical context (5XJZ),[25] all contain N3-dominated, nearly-planar bonds. The B-form, antiparallel dC:Ag^{+}:dC base pair found here in tensegrity triangles (7SDG) exhibits the standard N3-Ag^{+}-N3 geometry as well as an N4-Ag^{+} binding mode.

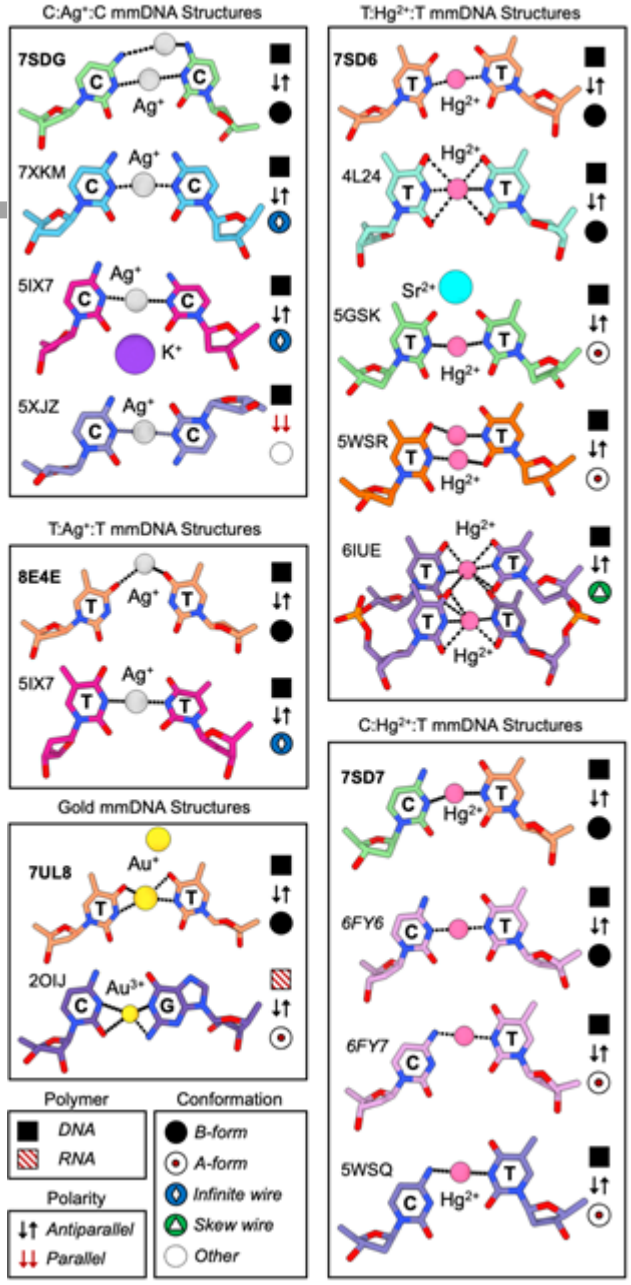


Figure 4: Comparison of mmDNA base pairs in this study with NMR and x-ray structures from previous studies. PDB ID's are indicated, with bold text indicating those reported here (7SDG, 8E4E, 7UL8, 7SD6, 7SD7) italic text delineating NMR structures (6FY6, 6FY7), and plain text indicating x-ray structures by other groups (7XKM, 5IX7, 5XJZ, 2OIJ, 4L24, 5GSK, 5WSR, 6IUE, 5WSQ). The legend symbols describe the polymer (DNA, RNA), the polarity (parallel, antiparallel) and the conformation (B-form, A-form, infinite nanowire, skew nanowire, and other).

The structure of dT:Ag^{+}:dT in the nanowire reported by Kondo and colleagues[3b] demonstrated N3-Ag^{+}-N3 geometry, likely mediated by metallophilic stacking of ions within the wire (5IX7). In an isolated B-form context, we observe exclusively N4-Ag^{+}-N4 coordination (Fig. 4, 8E4E). A variety of dT:Hg^{2+}:dT structures have been reported, and the tensegrity triangle structure (7SD6) shows good alignment of binding geometry and distances with a prior B-form structure (4L24).[26] A subsequent study demonstrated A-form N3-Hg^{2+}-N3 geometries stabilized by divalent cations (Sr^{2+}—PDB ID: 5GSK, Ba^{2+}—PDB ID: 5WSS)[12a] The absence of heavy divalent ions produced "slip" geometries, with N3-Hg^{2+}-O2 and O4-Hg^{2+}-N3 bonds (Fig. 4). Owing to the low observed Hg^{2+} occupancies, it is likely that these slip modes contained either—but not both—Hg^{2+} ions in each asymmetric unit. Owing to the necessity of excess Mg^{2+} in the current study to drive self-assembly, we anticipate that Mg^{2+} may coordinate in a similar way in the major grooves of tensegrity triangle mmDNA. Finally, a complex N3 binding mode was observed in the skew Hg^{2+} nanowire reported by Ono and colleagues (6IUE).[3d]

Schmidt and colleagues reported two distinct binding modes for dC:Hg^{2+}:dT using NMR: a B-form N3-Hg^{2+}-N3 confirmation (Fig. 4: PDB ID 6FY6), and an A-form N4-Hg^{2+}-N3 geometry (6FY7) that aligns with the earlier crystallographic A-form study by Liu and colleagues (5WSQ).[12a, 27] In the present study we enforced B-form geometry using the tensegrity triangle frame and consequently obtained N3-Hg^{2+}-N3 geometry (7SD7) that agrees with the NMR B-form structure. However, several structures in our library show some shearing, such as dC:Ag^{+}:dU, where the O4 or N4 group participates in coordination.

The first gold base pair was identified by Ennifar and colleagues in 2003, where an RNA duplex was screened against metal ion effects. This base pair coordinated trivalent gold ions in a rC:Au^{3+}:rG

square planar arrangement (Fig. 4: PDB ID 2OIJ). Owing to the charge state and smaller ion size, the Au^{3+} base pair was observed to fit between a dC-dG base pair through N1 deprotonation. The monovalent gold ion in dT:Au^{+}:dT (7UL8) exhibits some evidence of square planar arrangement between the N3 and O4 sites, though the low working resolution makes it difficult to assign the geometry. We note that the ion is only briefly stable in aqueous conditions before ligand decomposition and the oxidation of Au^{+} to Au^{3+}, at which point no mmDNA pair is formed. The specific pH conditions (11.0) and ambient temperature of the reaction reported here enable the ligand exchange reaction to charge the dT-dT pair. We highlight the similarity between the dT:Au^{+}:dT, dC:Ag^{+}:dT and dC:Ag^{+}:dC structures reported in this library.

Rotational Geometry

In addition to the local base pair geometry, we investigated the overall effect of mmDNA on the helical twist of B-form DNA. To do this, we calculated the rotational period of the full triangle edge (see Fig. 5A,B) for each structure, benchmarked against the Watson-Crick tensegrity triangle standards (PDB IDs: 5W6W, 3GBI). We additionally compared our mmDNA structures to those with unfilled mispairs (dC^{+}-dC, dC*$^{+}$-dC*, dC-dS, dT-dT; Fig S5). We further computed the twist values, which include the contribution of each branched junction to determine how the twist differences propagate across the motif. The tensegrity triangle standards are rotationally strained inside the triangle, twisting ~10 bp/ helical turn (Fig. S6). This value adjusts to about 10.47 bp/turn when the junction is accounted for, in agreement with both the imposed period of 10.50 bp/turn and with earlier studies showing that helical twisting and bending can be absorbed by Holliday junctions (Fig. 5C).[28] Unfilled mismatches twist a mean of about 9.75 bp/turn and adjust to about 10.30 bp/turn when accounting for the junctions. The major-groove binding series agreed with this trend, suggesting that the poly-pyrimidine sequence (TpYpT—see Fig. 1A, bold) cause stacking interactions that contribute to an altered rotational period. Many N3-Ag^{+} structures twist similarly, leading to the conclusion that major groove and some N3 metal pairs have little long-range effect on DNA rotation. By contrast, many members of the mixed-family N3 series, the T-family mercury(II) series, and the atypical pairs have larger twist perturbations, dropping to as little as 9.38 bp/turn (diU:Ag^{+}:dS), or 9.42 bp/turn (dC:Hg^{2+}:dT) between crossovers, and 10.10 and 10.22 bp/turn, respectively, with the branched junction contribution (Figs. 5C, S6). When the twist was computed across the two dinucleotide steps comprising the mmDNA pair (3 bp), the helicity was generally closer to that of the standards (Fig. 5D). These "local" twist values showed significant perturbations for the C-T family Ag^{+} pairs, but the

overall twist effects of single mmDNA pairs were more prominent as long-range propagations. A full table of helicity calculations can be found in Tables S14-S15. It is clear that many mmDNA base pairs cause the double helix to rotate more than an average canonical base pair along its axis. As such, the benchmark values identified here for single mmDNA pairs can assist in the design of future motifs to drive controlled topological self-assembly.

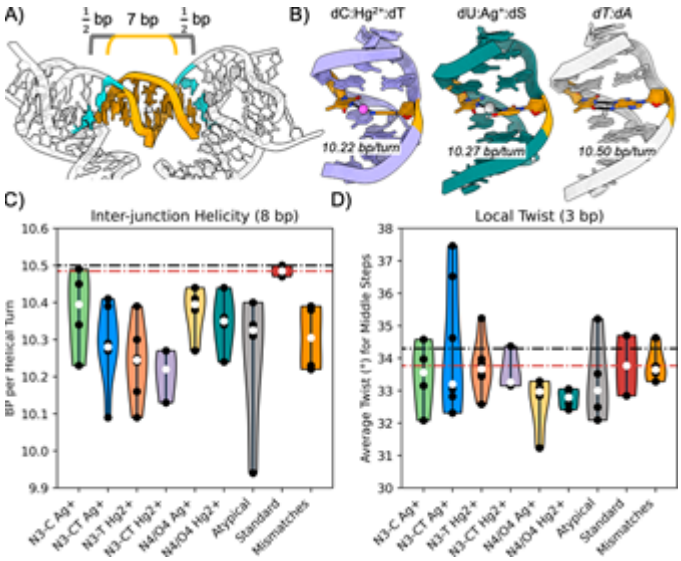


Figure 5: Helicity measurements of mmDNA base pairs. The helicity values were generated from the "simple" non-Watson-Crick values using the Web X3 DNA webserver.[29] **A)** The twist was averaged over the 7 bp edge of the triangle (orange) (Fig. S6). To account for the junctions, half of each base pair step across the Holliday junction (cyan) was also included to generate an 8-bp twist value. **B)** Structural polymorphism was observed between different mmDNA pairs, with examples dC:Hg^{2+}:dT, dU:Ag^{+}:dS shown for 8-bp helical comparison against the Watson-Crick triangle (PDB ID: 5W6W). **C)** Divided and colored by groups (see Fig. 3), the junction-inclusive helicity was calculated and compared to Watson-Crick tensegrity triangle standards (PDB ID: 3GBI, 5W6W; red). The empty mismatches (dC^{+}-dC, dC-dS, dT-dT, dC^{*+}-dC*; orange) were included to elucidate the effect of the local sequence environment on helicity. The median of each group is shown as a white circle, and the benchmark lines are drawn for the standard group (red line) and the imposed B-form helicity (10.50 bp/turn, black line). **D)** The twist calculations for the two dinucleotide steps comprising the mmDNA

base pair showed salient changes for several base pairs, but the twist alteration appeared to be distributed across the whole tensegrity triangle edge.

Molecular Electronics and Further Manipulation of mmDNA Crystals

The use of DNA as a scaffold to organize nanoelectronic components was proposed in 1987.[30] This idea of a DNA scaffold has carried through the intervening years, and DNA has been used as a structural scaffold for a variety of functional nanostructures and nanodevices, including metallic,[31] semiconducting[32] and magnetic nanoparticles, carbon nanotubes[33] and carbon nanotube transistors.[10, 33a, 34] Here we describe a system in which the active elements are embedded within the scaffold itself, creating intrinsically functional structures with *sub-nanometer* precision. Evidence of electronic functionality in DNA oligomers containing mmDNA base pairs has been observed for dC:Ag^+:dC.[2, 10] Using the 3D self-assembly process described here, it should be possible to design materials that exploit this embedded molecular electronics strategy. To demonstrate the feasibility of this approach, we designed a tensegrity triangle which could accommodate arm extensions via sticky-ended linkers, which can be tailored for helicity and metalation sequence using a recently established approach (Fig. 6A).[35] We employed a one-turn (either 10 or 11 bp) duplex linker to extend each asymmetric unit with one dT:dT and one dC:dC pair separated by canonical base pairs (Fig. 6B) to accommodate both Hg^{2+} and Ag^+ in the same motif, for total of 6 ions in each triangular assembly. Macroscopic self-assembly was not observed in the absence of metals, while the designed rhombohedral crystal morphology was obtained when both ions were added (Fig. 6C). This demonstrates the requirement, as well as the success, of intercalation of both Hg^{2+} and Ag^+ in a single motif. We further note that the 10mer linker accommodated Ag^+ only, in a slower yet larger, self-assembly product, showcasing that T:T pairs can be coaxed to form Ag^+-stabilized interactions. This linker-based system suggests a means for increasing the complexity within a single motif through helicity tuning and careful sequence design.

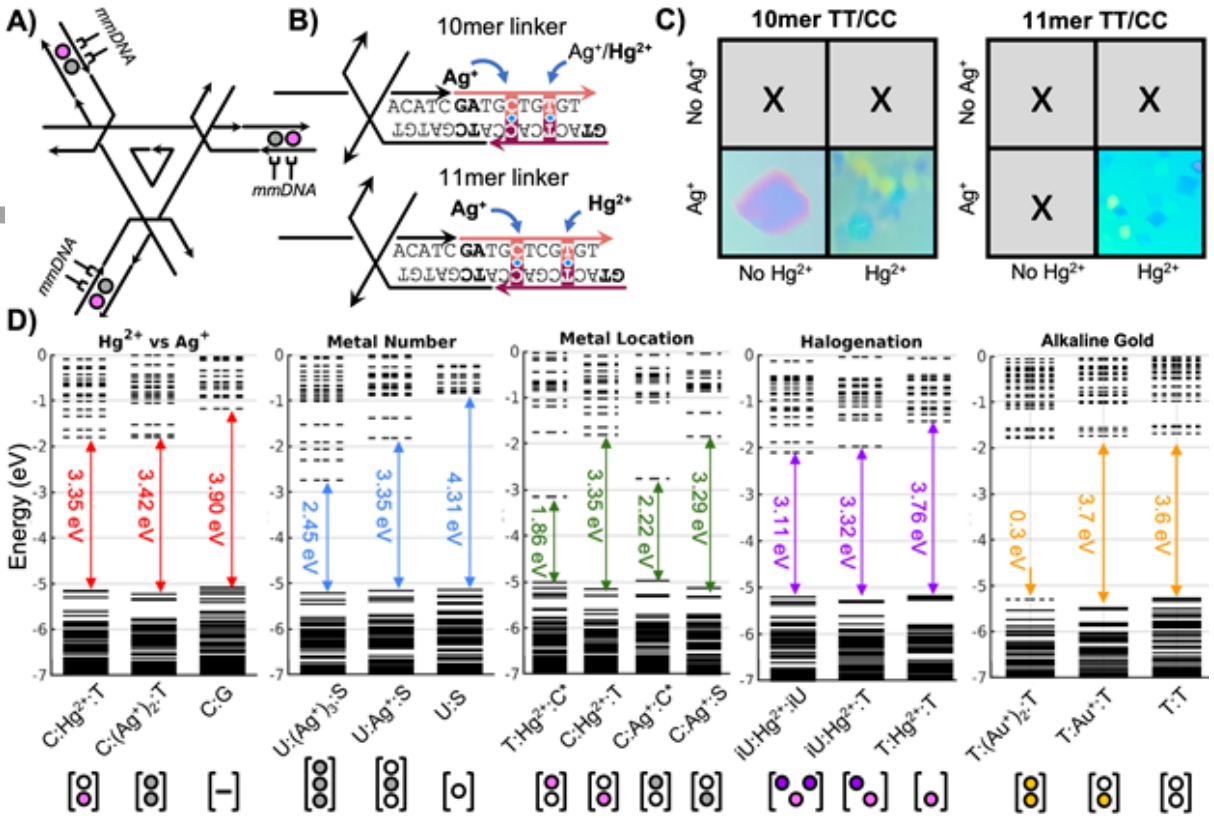


Figure 6: Applications of mmDNA crystals. **A)** Sticky-end-mediated extensions of each tensegrity triangle edge are designed to accommodate heterogenous, higher-density metals in a one-turn DNA linker with non-consecutive dC:dC and dT:dT mismatches. **B)** Owing to measured differences in helical twist, both a 10-mer and an 11-mer linker were employed. **C)** Self-assembly of DNA crystals was performed in four conditions: in the absence of metals, with Ag^+ only, Hg^{2+} only, and in the presence both Ag^+ and Hg^{2+}. Metal-mediated self-assembly of six metal sites per triangle assembly is observed, half Ag^+ and half Hg^{2+}. Affinity of T-family bases to Ag^+ under select helical conditions is noted. Each image is 200 μm wide. **D)** Energy gap calculations of the middle seven base pairs of energy-minimized mmDNA crystal structures from our library are carried out using DFT: Hg^{2+} vs. Ag^+ vs. WC (red); number of metals per base pair from 0-3, using extra sites from the dU:Ag^+:dS refinement (blue); N3 vs. major groove N4/O4 metalation (green); iterative replacement of methyl groups with iodine atoms (purple); and Au^+ site count at alkaline pH. Pictorial representations of each

base pair are drawn below, with Hg^{2+} (pink), Ag^{+} (silver), Au^{+} (gold), iodine (purple), unfilled sites (empty circles) and WC pairs (—) indicated. Metal-mediated energy levels are observed in LUMO.

Finally, we carried out energy level calculations of the triangle core using density functional theory (DFT) to elucidate an electronic fingerprint of the different classes of structures in our mmDNA library (Fig. 6D). We found that all mmDNA pairing modes (N3, N4/O4, Hg^{2+}, Ag^{+}, methylated, halogenated) add levels to the lowest unoccupied molecular orbital (LUMO)—akin to the conduction band in a bulk material. Even the addition of a single mmDNA pair causes a distinct lowering of the energy gap in these structures. A comparison of dC:Hg^{2+}:dT, dC:$(Ag^{+})_2$:dT and dC:dG shows similar addition of LUMO states and reduction of the energy gap from 3.90 eV to 3.35-3.42 eV for both metals. Using the extra, unassigned Ag^{+} sites found during dU:Ag^{+}:dS structure refinement, we observed a dramatic decrease in the gap to 2.45 eV for the three-metal base pair case, versus 3.35 and 4.31 eV for the one and zero metal cases, suggesting that harnessing and optimizing this multi-ion effect may lead to further electronic enhancement. A comparison of N3- and N4-bound Ag^{+} and Hg^{2+} carried out by analyzing the methylcytosine analogs to dC:Ag^{+}:dS and dC:Hg^{2+}:dT. Surprisingly, we observed a large reduction in energy gap through a single LUMO state in the major-groove-bound mmDNA structures relative to N3 ones (3.35 eV to 1.86 eV for Hg^{2+}; 3.29 eV to 2.22 eV for Ag^{+}), likely a result of tilting the nucleotides together at the top toward the metal. This offers the possibility of a precessing metal chain, rather than the traditional straight, 1D chain of N3 metals that might conduct efficiently along the double helix.[3b] Finally, we analyzed dT:Hg^{2+}:dT against similar base pairs with one and two 5-position iodines replacing the thymidine methyl groups. Each iodine added more states to LUMO, dropping the gap from 3.76 to 3.32 to 3.11 eV for zero, one and two iodines, respectively. Further investigation of the molecular orbitals showed that the lowest LUMO levels in these structures remain centered around the N3 mercury ion, and that addition of the iodine atoms serves to delocalize—but not draw away—this orbital to the halogens (Fig S8). The presence of a single Au^{+} ion relative to the unfilled T:T mismatch had little effect on the band gap (3.7 eV vs 3.6 eV), which was determined to be an effect of the highly-alkaline environment. Reverting the effective pH via protonation of thymines showed a population of LUMO states in the single Au^{+} base pair (3.4 eV vs 4.3 eV at neutral pH, Fig. S9). By contrast, the presence of both identified Au^{+} sites closed the band gap to less than 0.8 eV, irrespective of pH. This was the result of a single LUMO state which was localized to the major groove Au^{+} ion (Fig. S9). This single low-energy state might facilitate electron hopping across the molecule. Finally, a calculation of the density of states indicated an increase of several orders of magnitude for metalated dC:dT pairs relative to WC controls, increasing

the expected transmission across the molecule (Fig. S10). We anticipate that increasing the number of mmDNA pairs and the tuning and exploitation of these electronic effects will allow for the design of electrically-active DNA nanomaterials.

Conclusions

In this study, we utilized self-assembling DNA crystals to elucidate the following mmDNA base pairing rules:

1) Isocytosine drives N3-exclusive Ag^+-base pairing with near orthogonality to Hg^{2+};

2) Uracil is a near-universal and highly-versatile N3 Ag^+- and Hg^{2+}-pairing nucleobase;

3) Methylcytosine strongly disfavors N3 ion pairs and forms major groove mmDNA interactions as a result of its modification at the 5-position;

4) Ag^+ pairs form a larger number of structures with greater secondary site diversity, while secondary Hg^{2+} ions tend to be found, if at all, in the minor groove;

5) Contrary to previous studies, cytosine has a tendency to form N4-metal bonds, supporting both N3 and N4 in its homopair, and perhaps driving a major-groove-only Hg^{2+} pair with iodo-uracil;

6) T-family nucleobases can form N3 or O4 homopairs with Ag^+, but C-family bases cannot form homopairs with Hg^{2+} in the conditions studied here;

7) Alterations to buffer identity and pH may drive the formation of unexpected mmDNA pairs, such as dT:Au^+:dT, and may be further exploited;

8) No nucleobase in this study paired exclusively with Hg^{2+}, leaving open the search for Ag^+ orthogonality.

The founding objective of DNA nanotechnology is the self-assembly of designer crystal scaffolds for the structure determination of guest biomolecules.[36] In this study, we validated this approach, demonstrating that programmable DNA crystals can be used to rapidly screen a library for structural

information at the molecular level. We showed that we can generate systematized information from composable parts in a rapid, uniform crystalline system. This is a critical demonstration for structural DNA nanotechnology, and through this platform we were able to investigate and realize an expanded material language for self-assembly. We anticipate that this crystallographic approach will not be limited to mmDNA, and that we will be able to screen other inorganic and non-canonical nucleic acid modifications.

We further anticipate the integration of this expanded semantomorphic alphabet into DNA-based architectures for the exploitation of their electronic, magnetic, and catalytic properties.[30] It is clear that the integration of electronically-active atoms into metal organic frameworks is an attractive technological foundation, but synthetic difficulties and high-energy pattern transfer or waveguide techniques have hampered the implementation of these technologies thus far. To address this problem, the present study identifies a means of programming—and arranging—individual metal ions of interest within a 3D lattice via self-assembly. Future modifications can be made within this framework by altering the pyrimidine reaction center to diversify the metal chemistry;[11] exposing them to external fields, oxidation states and hydride concentrations;[4a, 37] and fundamentally, to develop molecular electronics based on the topological self-assembly of DNA.[20, 30, 38]

References and Notes

[1] a) X.-L. Yang, H. Sugiyama, S. Ikeda, I. Saito, A. H.-J. Wang, *Biophys J* **1998**, 75, 1163; b) C. D. Mao, W. Q. Sun, N. C. Seeman, *J Am Chem Soc* **1999**, 121, 5437.

[2] a) E. Toomey, J. Xu, S. Vecchioni, L. Rothschild, S. Wind, G. E. Fernandes, *J Phys Chem C* **2016**, 120, 7804; b) N. Fardian-Melamed, L. Katrivas, G. Eidelshtein, D. Rotem, A. Kotlyar, D. Porath, *Nano Letters* **2020**, 20, 4505.

[3] a) A. Ono, S. Cao, H. Togashi, M. Tashiro, T. Fujimoto, T. Machinami, S. Oda, Y. Miyake, I. Okamoto, Y. Tanaka, *Chem Commun (Camb)* **2008**, 4825; b) J. Kondo, Y. Tada, T. Dairaku, Y. Hattori, H. Saneyoshi, A. Ono, Y. Tanaka, *Nature Chem* **2017**, 9, 956; c) T. Atsugi, A. Ono, M. Tasaka, N. Eguchi, S. Fujiwara, J. Kondo, *Angew Chem Int Ed* **2022**, 61, e202204798; d) A.

Ono, H. Kanazawa, H. Ito, M. Goto, K. Nakamura, H. Saneyoshi, J. Kondo, *Angew Chem Int Ed* **2019**, 131, 16991.

[4] a) G. H. Clever, S. J. Reitmeier, T. Carell, O. Schiemann, *Angew Chem Int Ed* **2010**, 49, 4927; b) G. H. Clever, M. Shionoya, *Coord Chem Rev* **2010**, 254, 2391.

[5] a) Y. Miyake, H. Togashi, M. Tashiro, H. Yamaguchi, S. Oda, M. Kudo, Y. Tanaka, Y. Kondo, R. Sawa, T. Fujimoto, T. Machinami, A. Ono, *J Am Chem Soc* **2006**, 128, 2172; b) S. Katz, *Nature* **1962**, 195, 997.

[6] S. Vecchioni, M. C. Capece, E. Toomey, L. Nguyen, A. Ray, A. Greenberg, K. Fujishima, J. Urbina, I. G. Paulino-Lima, V. Pinheiro, *Sci Rep* **2019**, 9, 6942.

[7] a) H. Yamaguchi, J. Sebera, J. Kondo, S. Oda, T. Komuro, T. Kawamura, T. Dairaku, Y. Kondo, I. Okamoto, A. Ono, J. V. Burda, C. Kojima, V. Sychrovsky, Y. Tanaka, *Nucleic Acids Res* **2014**, 42, 4094; b) A. Ono, H. Torigoe, Y. Tanaka, I. Okamoto, *Chem Soc Rev* **2011**, 40, 5855; c) Y. Takezawa, K. Nishiyama, T. Mashima, M. Katahira, M. Shionoya, *Chem Eur J* **2015**, 21, 14713; d) Y. Takezawa, A. Suzuki, M. Nakaya, K. Nishiyama, M. Shionoya, *J Am Chem Soc* **2020**, 142, 21640; e) H. Torigoe, J. Kondo, F. Arakawa, *Journal of Inorganic Biochem* **2023**, 112125.

[8] a) H. Torigoe, I. Okamoto, T. Dairaku, Y. Tanaka, A. Ono, T. Kozasa, *Biochimie* **2012**, 94, 2431; b) S. M. Swasey, L. E. Leal, O. Lopez-Acevedo, J. Pavlovich, E. G. Gwinn, *Sci Rep* **2015**, 5, 10163.

[9] Z. Wang, J. H. Lee, Y. Lu, *Chem Commun (Camb)* **2008**, 6005.

[10] N. Fardian-Melamed, G. Eidelshtein, D. Rotem, A. Kotlyar, D. Porath, *Adv Mat* **2019**, 31, 1902816.

[11] Y. Takezawa, J. Müller, M. Shionoya, *Chem Lett* **2017**, 46, 622.

[12] a) H. Liu, C. Cai, P. Haruehanroengra, Q. Yao, Y. Chen, C. Yang, Q. Luo, B. Wu, J. Li, J. Ma, J. Sheng, J. Gan, *Nucl Acids Res* **2017**, 45, 2910; b) J. Kondo, Y. Tada, T. Dairaku, H. Saneyoshi, I. Okamoto, Y. Tanaka, A. Ono, *Angew Chem Int Ed* **2015**, 54, 13323; c) T. Dairaku, K. Furuita, H. Sato, J. Šebera, K. Nakashima, J. Kondo, D. Yamanaka, Y. Kondo, I. Okamoto, A. Ono, *Chem Eur J* **2016**, 22, 13028.

[13] S. Vecchioni, R. Sha, N. C. Seeman, L. J. Rothschild, S. J. Wind, *J Self-Assembly Mol Elect* **2022**, 17.

[14] a) J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginell, C. Mao, N. C. Seeman, *Nature* **2009**, 461, 74; b) Y. P. Ohayon, C. Hernandez, A. R. Chandrasekaran, X. Wang, H. O. Abdallah, M. A. Jong, M. G. Mohsen, R. Sha, J. J. Birktoft, P. S. Lukeman, *ACS Nano* **2019**, 13, 7957; c) S. Vecchioni, B. Lu, J. Janowski, K. Woloszyn, N. Jonoska, N. C. Seeman, C. Mao, Y. P. Ohayon, R. Sha, *Small* **2022**, In press, 2206511; d) B. Lu, S. Vecchioni, Y. P. Ohayon, J. W. Canary, R. Sha, *Biophys J* **2022**, 121, 4759.

[15] S. Hoshika, N. A. Leal, M.-J. Kim, M.-S. Kim, N. B. Karalkar, H.-J. Kim, A. M. Bates, N. E. Watkins, H. A. SantaLucia, A. J. Meyer, *Science* **2019**, 363, 884.

[16] a) C. Mao, W. Sun, Z. Shen, N. C. Seeman, *Nature* **1999**, 397, 144; b) R. Holliday, *Cell Biophys* **1989**, 15, 15.

[17] I. Okamoto, K. Iwamoto, Y. Watanabe, Y. Miyake, A. Ono, *Angew Chem Int Ed* **2009**, 48, 1648.

[18] I. Tickle, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, C. Vonrhein, G. Bricogne, *Cambridge, United Kingdom: Global Phasing Ltd* **2018**.

[19] Q. Liu, W. Hendrickson, *Acta Crystall Sect D: Biol Cryst* **2013**, 69, 1314.

[20] S. Vecchioni, M. C. Capece, E. Toomey, L. Rothschild, S. J. Wind, *J Self-Assembly Mol Elect* **2018**, 6, 61.

[21] H. Urata, E. Yamaguchi, Y. Nakamura, S. Wada, *Chem Commun (Camb)* **2011**, 47, 941.

[22] J. Kondo, T. Sugawara, H. Saneyoshi, A. Ono, *Chem Commun (Camb)* **2017**, 53, 11747.

[23] a) I. Kang, Y. Wang, C. Reagan, Y. Fu, M. X. Wang, L.-Q. Gu, *Sci Rep* **2013**, 3, 2381; b) Y. Wang, B. Ritzo, L.-Q. Gu, *RSC Adv* **2015**, 5, 2655.

[24] I. Okamoto, K. Iwamoto, Y. Watanabe, Y. Miyake, A. Ono, *Angew Chem Int Ed* **2009**, 48, 1648.

[25] H. Liu, C. Cai, P. Haruehanroengra, Q. Yao, Y. Chen, C. Yang, Q. Luo, B. Wu, J. Li, J. Ma, J. Sheng, J. Gan, *Angew Chem Int Ed* **2017**, 56, 9430.

[26] J. Kondo, T. Yamada, C. Hirose, I. Okamoto, Y. Tanaka, A. Ono, *Angew Chem Int Ed* **2014**, 53, 2385.

[27] O. P. Schmidt, S. Jurt, S. Johannsen, A. Karimi, R. K. Sigel, N. W. Luedtke, *Nature Commun* **2019**, 10, 1.

[28] a) C. Hernandez, J. J. Birktoft, Y. P. Ohayon, A. R. Chandrasekaran, H. Abdallah, R. Sha, V. Stojanoff, C. Mao, N. C. Seeman, *Cell Chem Biol* **2017**, 24, 1401; b) B. Lu, S. Vecchioni, Y. P. Ohayon, R. Sha, K. Woloszyn, B. Yang, C. Mao, N. C. Seeman, *ACS Nano* **2021**, 15, 16788; c) B. Lu, S. Vecchioni, Y. P. Ohayon, K. Woloszyn, T. Markus, C. Mao, N. C. Seeman, J. W. Canary, R. Sha, *Small* **2022**, 19, 2205830.

[29] S. Li, W. K. Olson, X.-J. Lu, *Nucl Acids Res* **2019**, 47, W26.

[30] B. H. Robinson, N. C. Seeman, *Protein Engin Design Sel* **1987**, 1, 295.

[31] a) B. Ding, Z. Deng, H. Yan, S. Cabrini, R. N. Zuckermann, J. Bokor, *J Am Chem Soc* **2010**, 132, 3248; b) V. V. Thacker, L. O. Herrmann, D. O. Sigle, T. Zhang, T. Liedl, J. J. Baumberg, U. F. Keyser, *Nature Commun* **2014**, 5, 3448.

[32] a) D. Huang, M. Freeley, M. Palma, *Sci Rep* **2017**, 7, 45591; b) R. Wang, C. Nuckolls, S. J. Wind, *Angew Chem Int Ed Engl* **2012**, 51, 11325.

[33] a) H. T. Maune, S.-p. Han, R. D. Barish, M. Bockrath, W. A. Goddard III, P. W. Rothemund, E. Winfree, *Nature Nanotechnol* **2010**, 5, 61; b) Z. Zhao, Y. Liu, H. Yan, *Org Biomol Chem* **2013**, 11, 596.

[34] a) A. Gopinath, P. W. Rothemund, *ACS Nano* **2014**, 8, 12030; b) P. W. Majewski, A. Michelson, M. A. Cordeiro, C. Tian, C. Ma, K. Kisslinger, Y. Tian, W. Liu, E. A. Stach, K. G. Yager, O. Gang, *Sci Adv* **2021**, 7, eabf0617.

[35] a) K. Woloszyn, S. Vecchioni, Y. P. Ohayon, B. Lu, Y. Ma, Q. Huang, E. Zhu, D. Chernovolenko, T. Markus, N. Jonoska, *Adv Mat* **2022**, 2206876; b) B. Lu, K. Woloszyn, Y. P. Ohayon, B. Yang, C. Zhang, C. Mao, N. C. Seeman, S. Vecchioni, R. Sha, *Angew Chem Int Ed* **2022**.

[36] N. C. Seeman, *J Theor Biol* **1982**, 99, 237.

[37] K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shionoya, *Science* **2003**, 299, 1212.

[38] G. I. Livshits, A. Stern, D. Rotem, N. Borovok, G. Eidelshtein, A. Migliore, E. Penzo, S. J. Wind, R. Di Felice, S. S. Skourtis, *Nature Nanotechnol* **2014**, 9, 1040.

Acknowledgements

ASPIRE Program students at King School, CT, USA are mentored by Dr. Victoria K. Schulman.

Dedication

We are deeply grateful to mentorship and friendship of Professor Ned Seeman, who passed away during the writing of this manuscript, and whose foundational work and guidance to DNA Nanotechnology was a light to all.

Funding:

Office of Naval Research grant N000141912596 (NCS, JWC, RS)

Department of Energy grant DE-SC0007991 (NCS, RS)

National Science Foundation grant 2106790 (NCS, RS)

Human Frontiers Science Program grant RPG0010/2017 (NCS, RS)

The MRSEC Program of the National Science Foundation, grant DMR-1420073 (NCS, SV)

The 2020 NASA Ames Center Innovation Fund Program (LJR, SV)

The King School Advanced Mathematics and Science Study Program Fund (WB)

Author Contributions: All authors analyzed data and wrote the paper. S.V. devised the project, modified the motif, grew crystals, collected data; B.L. and K.J.W. grew crystals and collected data; R.S. devised and guided the project and synthesized DNA oligomers; W.L. and M.P.A performed DFT calculations; Y.P.O. modified the motif; C.Y. synthesized the gold ligand. J. Y. and W.A.H. devised the collection strategy; C.M. devised the motif; S.J.W. designed the project; N.C.S. initiated and guided the project.

Data Availability: Atomic coordinates and experimental structure factors have been deposited within the Protein Data Bank and are accessible under codes found in Tables S4-S13 and the Supplementary Appendix.

Competing Interests: Authors declare that they have no competing interests.

Supplementary Materials:

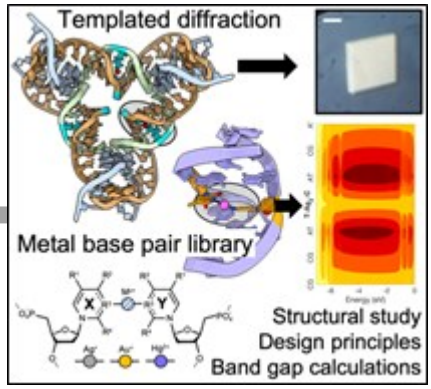
Materials and Methods

Supplementary Text

Supplementary Figures 1-10

Supplementary Tables 1-18

Supplementary Appendix of Base Pairs



The authors demonstrate the founding objective of structural DNA nanoscience by using self-assembled DNA scaffolds for the structure determination of biomolecules by x-ray diffraction. The authors assemble a structural library of 32 metal-mediated DNA base pairs (mmDNA) with Ag^{+}, Hg^{2+} and Au^{+} and further elucidate fundamental design rules for an expanded mmDNA coding system toward the development of molecular electronics and designer nanomaterials.