



ARTICLE

<https://doi.org/10.1038/s41467-020-14607-3>

OPEN

Sequence-selective dynamic covalent assembly of information-bearing oligomers

Samuel C. Leguizamón ¹ & Timothy F. Scott ^{1,2,3,4*}

Relatively robust dynamic covalent interactions have been employed extensively to mediate molecular self-assembly reactions; however, these assembly processes often do not converge to a thermodynamic equilibrium, instead yielding mixtures of kinetically-trapped species. Here, we report a dynamic covalent self-assembly process that mitigates kinetic trapping such that multiple unique oligomers bearing covalently coreactive pendant groups are able to undergo simultaneous, sequence-selective hybridization with their complementary strands to afford biomimetic, in-registry molecular ladders with covalent rungs. Analogous to the thermal cycling commonly employed for nucleic acid melting and annealing, this is achieved by raising and lowering the concentration of a multi-role reagent to effect quantitative dissociation and subsequently catalyze covalent bond rearrangement, affording selective assembly of the oligomeric sequences. The hybridization specificity afforded by this process further enabled information encoded in oligomers to be retrieved through selective hybridization with complementary, mass-labeled sequences.

¹ Department of Chemical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. ² Macromolecular Science and Engineering Program, University of Michigan, Ann Arbor, MI 48109, USA. ³ Department of Chemical Engineering, Monash University, Clayton, VIC 3800, Australia. ⁴ Department of Materials Science and Engineering, Monash University, Clayton, VIC 3800, Australia. *email: timothy.scott@monash.edu

The capacity for sequence-specific polymer strands to selectively assemble into intricate, folded structures and multimeric complexes relies upon the information borne by their residue sequences^{1,2}. Although there are several examples of the non-covalent assembly of information-bearing species^{3–6}, nucleic acids represent a particularly versatile class for producing nanostructures which, through careful consideration of their nucleobase sequence, can be designed to predictably self-assemble via the hybridization of complementary strands into arbitrary structures with nanometer-scale precision^{7–10}. Unfortunately, the versatility of nucleic acids as nanoconstruction media is tempered by the thermal and mechanical instability of the resultant structures, attributable to the weakness of the hydrogen bonds that hold the strands together¹¹. Efforts to improve the stability of double-stranded, nucleic acid-like duplexes include the utilization of abiotic, neutral backbones^{12–14}, the integration of isosteric nucleobase mimetics^{15,16}, and the incorporation of photo-induced cross-links^{17,18}; nevertheless, the assembly processes for these duplexes remain mediated by intermolecular interactions. In contrast, employing dynamic covalent interactions to mediate molecular self-assembly processes can afford inherently robust, covalently cross-linked structures^{19–26}, although impediments to connectivity rearrangement amongst the assembled components, including slow reaction rates and reaction site inaccessibility, often result in the kinetic trapping of non-equilibrium species^{27–29}.

Here, we describe a dynamic covalent assembly approach where mixtures of sequence-specific oligomers bearing covalently coreactive amine-based and aldehyde-based pendant groups selectively associate and dimerize with their oligomeric complements to afford in-registry molecular ladders with imine-based covalent rungs, mimicking the selective, information-directed hybridization of complementary nucleic acid sequences.

Results

Dynamic covalent assembly of encoded molecular ladders. Peptoids (i.e., oligo(*N*-substituted glycine)s)³⁰ were employed here as the precursor oligomers owing to their ready synthetic accessibility via Zuckerman's sub-monomer solid phase synthetic method³¹. Moreover, the peptoid backbone is inherently neutral, achiral, and lacks hydrogen bond donor sites³², ensuring that oligomer hybridization selectivity is dominated by covalent interactions between pendant groups. These oligomers were synthesized as binary sequences of amine and aldehyde pendant groups (denoted here as '1' and '0', respectively) alternating with inert spacer residues such that, owing to the zig-zag 'Σ-strand' conformation adopted by the constituent peptoid chains, the reactive groups were presented on the same side of the peptoid backbone³³. The spacer residues also increased solubility of both the precursor oligomers and hybridized molecular ladders, preventing precipitation during the self-assembly process. To enable facile identification of the hybridization product species by mass spectrometry, sequences were mass-tagged by adding extra, inert spacer residues to the peptoid ends during synthesis. Additionally, each imine bond generated by an amine/aldehyde condensation reaction reduces the molar mass of the resultant molecular ladder by 18 g/mol owing to the loss of a water molecule, enabling in- and out-of-registry species to be readily distinguished by mass. As the oligomer sequences bore multiple, covalently coreactive functional group types, ethylene acetals were used as aldehyde-protecting groups to preclude premature reaction between coreactive pendant groups during synthesis and purification, and which could then be deprotected in situ, allowing the dynamic covalent assembly reaction to proceed. Monomer residue structures are as shown in Supplementary

Table 1 and purified, acetal-protected peptoid sequences were characterized as shown in Supplementary Figs. 1–3.

Sequence-selective, dynamic covalent oligomer hybridization was initially examined using a model system consisting of two complementary peptoid sequences, each bearing alternating amine-based and aldehyde-based reactive pendant groups (i.e., sequences 10101, mass-labeled with an additional methoxyethylamine residue at its *N*-terminal end, and 01010, see Fig. 1a). Our group has previously demonstrated the utilization of scandium (III) triflate at low loadings as a dual-role, Lewis acidic catalyst both to effect the in situ deprotection of oligomer-borne, acetal-protected aldehydes and to catalyze transimination and imine metathesis reactions for the rearrangement of imine bonds generated by amine/aldehyde condensation reactions, thereby successfully yielding molecular ladders with up to eight rungs via the dynamic covalent assembly of diblock, amine-bearing and aldehyde-bearing precursor oligomers. Consequently, we applied the previously determined reaction conditions (i.e., 0.2 equivalents of Sc(OTf)₃ for 2 h at 70 °C)³⁴ to the model system examined here.

For this system, the anticipated, thermodynamically favored product should emerge from the cross-reaction between the 10101 sequence and its 01010 oligomeric complement (i.e., 10101 × 01010) to selectively afford the formation of an in-registry molecular ladder with five imine-based rungs (denoted here as Im₅). As determined by MALDI mass spectrometry (Fig. 1b, bottom spectrum), the applied reaction conditions did yield complete acetal deprotection; however, these conditions resulted in the non-selective, kinetically trapped generation of multiple in- and out-of-registry molecular ladder species resulting from the incomplete coreaction of two 01010 (01010 × 01010) and two 10101 (10101 × 10101) sequences, both of which can yield molecular ladders with a maximum of four imine rungs (i.e., Im₄), in addition to those from the 10101 × 01010 cross-reaction.

Our previous work on the Sc(III)-catalyzed dynamic covalent assembly of imine-based molecular ladders revealed that the assembly reaction proceeds via a hybrid mechanism, where amine-bearing and aldehyde-bearing oligomers initially associate and bind through condensation reactions at any point along the backbone to form out of registry molecular ladders, whereupon transimination and imine metathesis reactions proceed to shuffle the imine bonds until the molecular ladders come into registry³⁵. Whereas this mechanism enabled the annealing of out-of-registry, intermediate species for simple systems composed of homo- or diblock oligomers, more complex sequences where the reactive residue type switches between amine and aldehyde pendant groups multiple times impede inter-strand bond shuffling as imine rearrangement reactions can no longer occur exclusively between adjacent residues. Rather, rearrangement reactions for such sequences would need to proceed between relatively remote functionalities to enable inter-strand shuffling. This registry mechanism is dissimilar to that commonly employed for DNA assembly where, instead of inter-strand shuffling, dissociated DNA sequences selectively hybridize with their complements upon gradual cooling from raised temperature to afford in-registry double helices at low error rates⁸. Pioneering work by the Lehn group on the catalysis of imine rearrangement reactions demonstrated that, at moderate concentrations, Lewis acidic rare-earth metal triflates do act to catalyze imine connectivity rearrangement, but at raised concentrations they shift the amine/aldehyde condensation reaction equilibrium towards the unpaired reactants³⁶. Analogous to the thermal cycling commonly employed for nucleic acid melting and annealing, raising and lowering the scandium triflate concentration was thus used to mitigate kinetic trapping by enabling

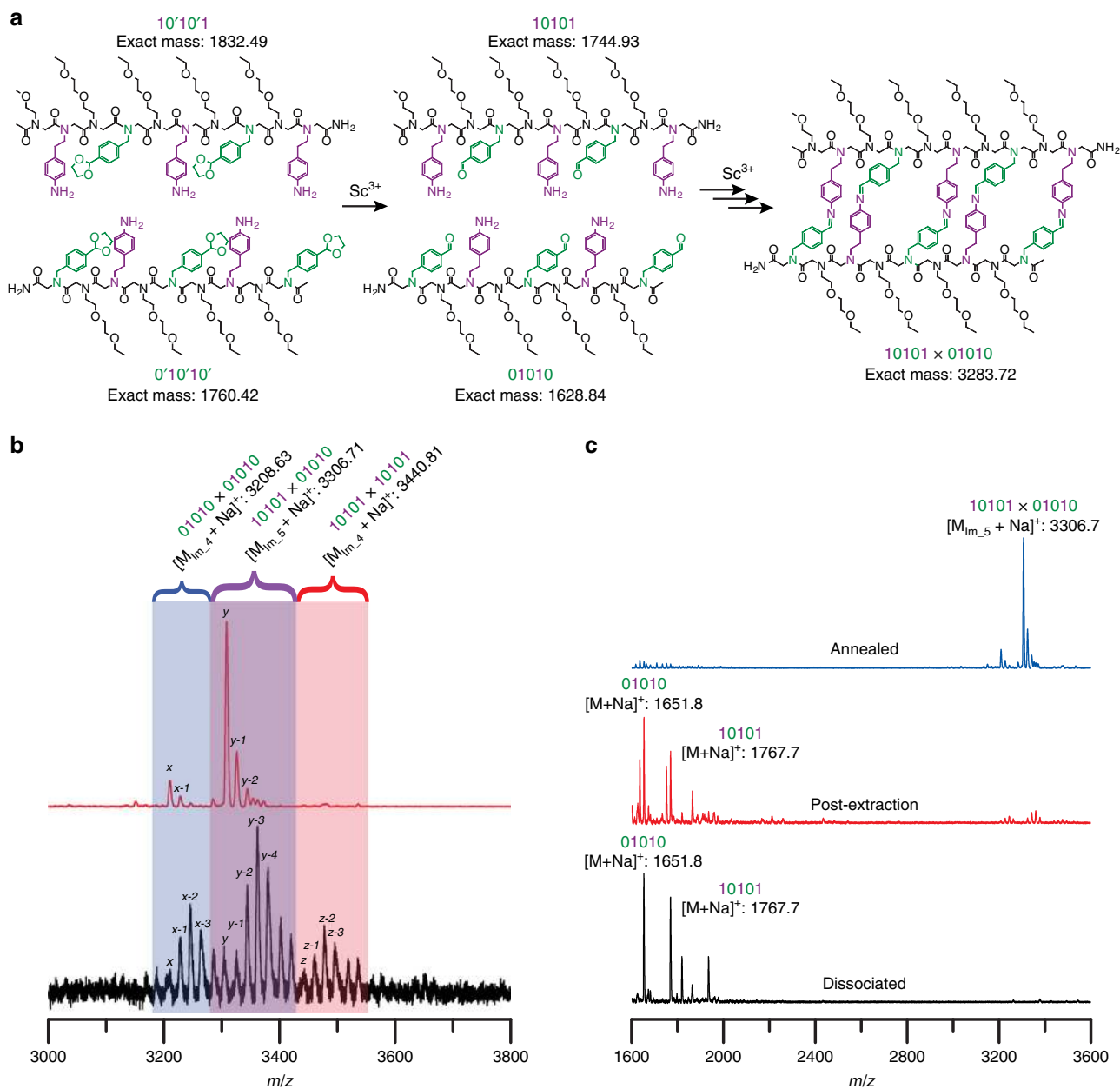


Fig. 1 Dynamic covalent assembly of encoded molecular ladders. **a** Structures of complementary, mass-labeled precursor peptoids and desired, in-registry molecular ladder product. **b** MALDI mass spectra of molecular ladder reaction mixtures generated either after a single-step, deprotection, and direct assembly process (bottom, black), or via a stepwise, dissociation/extraction/annealing assembly process (top, red). Peaks at multiples of $+18m/z$ values are attributable to ladders species with fewer rungs (e.g., y = in-registry, 5-rung 10101×01010 molecular ladder, $y-1$ = out-of-registry, 4-rung 10101×01010 molecular ladder, etc.). **c** MALDI mass spectra of the reaction mixture initially containing 10101 and 01010 sequences and a high loading of $Sc(OTf)_3$ (bottom, black), immediately after aqueous extraction (middle, red), and after annealing at $70^\circ C$ for 2 h (top, blue). Expected exact masses: $[M_{01010} + Na]^+$ = 1651.83; $[M_{10101} + Na]^+$ = 1767.92; $[M_{01010 \times 01010} + Na]^+$ (x) = 3208.63; $[M_{10101 \times 01010} + Na]^+$ (y) = 3306.71; $[M_{10101 \times 10101} + Na]^+$ (z) = 3440.81.

quantitative dissociation and subsequent, sequence-selective assembly of the 10101 and 01010 peptoid sequences to afford the desired 10101×01010 molecular ladder product (see Fig. 1b, top spectrum).

This process is detailed in Fig. 1c where MALDI mass spectra collected on reaction mixture aliquots after each process step are presented. The original reaction mixture initially contained a stoichiometric amount of ethylene acetal-protected 10101 and 01010 sequences and a high loading (i.e., 1.5 equivalents per amine/aldehyde pair) of $Sc(OTf)_3$. Under these reaction conditions, the acetals groups were quantitatively deprotected to reveal amine-reactive aldehydes; nevertheless, the peptoid strands

remained dissociated, evident from the presence of deprotected single strands at $1651.8m/z$ and $1767.7m/z$ and virtual absence of paired species from $3200-3600m/z$, owing to the influence of $Sc(OTf)_3$ on the reaction equilibrium (Fig. 1c, bottom black spectrum). This reaction mixture was subject to an aqueous extraction to selectively extract a fraction of the $Sc(OTf)_3$, reducing its relative loading. The residual scandium concentration was determined by inductively coupled plasma mass spectrometry (ICP-MS), revealing that 0.055 ± 0.002 equivalents of Sc^{3+} remained post-extraction, approaching the optimal concentration for transimination catalysis reported previously³⁶. Characterization of the reaction mixture by mass spectrometry

immediately after extraction indicated that the unpaired single strands remained dominant, although non-negligible amounts of intermediate hybridized species were observable (Fig. 1c, middle red spectrum). Owing to the chemical and molecular weight similarities between in- and out-of-registry ladders, ionization efficiencies in MALDI mass spectrometry are presumed sufficiently similar to permit relative concentrations of the individual species to be estimated^{35,37,38}. Heating for 2 h at 70 °C to anneal the system yielded the in-registry 10101 × 01010 molecular ladder as the primary reaction product (Fig. 1c, top blue spectrum). Small peaks attributable to 01010 × 01010 ladder species remained after annealing owing to a slight stoichiometric imbalance that arose during the extraction step where some of the 10101 sequence was lost in the aqueous extraction. Additionally, trace amounts of species at +18*m/z* (i.e., 4-rung ladders) were attributable to the residual Sc³⁺ affecting the reaction equilibrium rather than kinetically trapped, out-of-registry products. Notably, when the system was annealed at room temperature, fully in-registry 10101 × 01010 ladder species were observed by MALDI mass spectrometry within 1 day and emerged as the major peak within 7 days (Supplementary Fig. 4) indicating that, although raised temperatures increase reaction rates as expected, the self-assembly process does proceed under ambient conditions. Gel permeation chromatography (GPC), calibrated using low dispersity polystyrene standards, confirmed formation of the 10101 × 01010 molecular ladder as the predominant product (43.8%) by comparing elution volumes of inert, Alloc-protected single strands to hybridization mixtures (Supplementary Fig. 5a–d). Notably, this yield is lower than previously determined yields for imine rung-based molecular ladders assembled from homoblock oligomeric sequences³⁹ as these oligomers are incapable of participating in intramolecular amine/aldehyde condensation reactions, thus eliminating a potential mechanism for side product generation. Peak fractions, collected during GPC purification, were characterized by MALDI mass spectrometry to identify the fraction constituents (Supplementary Fig. 5e). The hybridization product was preceded by several minor peaks at higher molecular weights, attributable to multimeric complexes, and proceeded by single-stranded species (Supplementary Fig. 5f and g), including unreacted 10101 and 01010 strands and those with either one intramolecular imine bond to yield macrocycles, or two intramolecular imine bonds to potentially afford hairpin-like structures. The occurrence of a binding event was further verified with diffusion ordered spectroscopy (DOSY) NMR, which demonstrated a decrease in the diffusion coefficient after extraction and annealing of an initially dissociated mixture containing 10101 and 01010 strands (Supplementary Fig. 6).

To examine the effect of Sc³⁺ concentration on molecular ladder dissociation, varying amounts of scandium triflate were added to post-annealed, in-registry 10101 × 01010 reaction mixtures which were then heated at 60 °C for 6 h, left to stand at room temperature overnight to ensure system equilibration, and characterized by mass spectrometry (Supplementary Fig. 7). At 0.20 equivalents of Sc(OTf)₃ (matching the Sc³⁺ concentration employed for the single-pot, single-step assembly method), the desired in-registry molecular ladder remained the major product; however, peaks at multiples of +18*m/z*, attributable to ladders species with fewer rungs, were more apparent, suggesting a significant shift in the reaction equilibrium towards the amine and aldehyde reactants. At 0.30 equivalents of Sc³⁺ equivalents and above, the extent of ladder dissociation progressively increases and peaks attributable to dimerized species of mismatched, non-complementary strands (i.e., those consisting of 01010 × 01010 and 10101 × 10101 sequences) become more prevalent.

To examine the robustness of this dissociation/extraction/annealing assembly approach to alternative, Lewis acidic multi-role reagents, reaction mixtures of 10101 and 01010 oligomeric sequences were formulated with one of a library of rare-earth metal triflates with varying Lewis acidity, including Yb(OTf)₃, Lu(OTf)₃, Y(OTf)₃, and Sm(OTf)₃ (Supplementary Fig. 8). Whereas ytterbium triflate is more Lewis acidic than lutetium triflate, the annealing period required for in-registry 10101 × 01010 molecular ladder equilibrium was at least 3 days at 70 °C for the ytterbium triflate-containing reaction mixture while overnight at 70 °C was sufficient to yield in-registry ladders for the mixture formulated with lutetium triflate. The residual rare-earth metal triflate in the reaction mixtures post-extraction was assessed by ICP-MS and revealed that a greater concentration of lutetium triflate remained in the organic layer than that of ytterbium triflate (0.014 ± 0.0008 eq. and 0.009 ± 0.0004 eq., respectively, indicating that the higher residual concentration of Lu³⁺ more than compensated for its lower Lewis acidity than Yb³⁺). Notably, again owing to their lower Lewis acidity, three equivalents of yttrium triflate and samarium triflate (i.e., twice the required equivalents of Sc³⁺, Yb³⁺, and Lu³⁺) were necessary to fully deprotect and dissociate the oligomeric strands. After extraction, in-registry molecular ladder equilibrium was achieved after 7 days at 70 °C for yttrium triflate-containing solution, while the reaction mixture incorporating samarium triflate required an additional 2 days for complete ladder assembly. Here, ICP-MS revealed residual, post-extraction concentrations of 0.025 ± 0.001 eq. and 0.059 ± 0.003 eq. for Y³⁺ and Sm³⁺, respectively).

Competitive dynamic covalent assembly. The single reactant pair mismatch discrimination of our dynamic covalent self-assembly process was examined using combinations of peptoid strands, including the complementary pair of 10101 and 01010, and a 10001 sequence that is non-complementary with the other two. Importantly, the 10001 and 01010 sequences contain only a single mismatch located at the third residue to preclude any overhangs that might otherwise occur from influencing the 10001 × 01010 hybrid stability. Subjecting a non-complementary mixture of 10001 and 01010 strands in a 1:1 stoichiometric ratio to the dissociation/extraction/annealing reaction conditions yielded three cohorts of dimerized ladder species, 10001 × 10001, 01010 × 01010, and 01010 × 10001 (Fig. 2a, bottom red spectrum), demonstrating poor hybridization selectivity. In contrast, upon inclusion of a stoichiometric amount of the 10101 sequence to the initial reaction mixture to afford a ternary sequence system, where two of the peptoid strands are complementary (i.e., 10101 and 01010) results in the preferential generation of a complementary and in-registry 10101 × 01010 molecular ladder as the major mass spectrum peak (Fig. 2a, top blue spectrum).

Additional quantification of the hybridization selectivity was performed using peptoid strands functionalized with 5,6-carboxyfluorescein (FAM) and 4-((4-(dimethylamino)phenyl)azo)benzoyl (DABCYL) groups as a Förster resonance energy transfer (FRET) fluorophore and fluorescence quencher pair, respectively. These groups were selected for FRET studies owing to their ready incorporation as oligomer end groups and their stability to resin cleavage conditions. Here, the oligomer sequence 01010 was labeled with FAM (denoted as 01010-FAM) and the sequences 10001 and 10101 were labeled with DABCYL (10001-DABCYL and 10101-DABCYL, respectively). All assays were normalized to a non-quenched fluorescence intensity ceiling, determined using a solution of 01010-FAM as a single-strand negative control. A control solution incorporating 01010-FAM and 10101-DABCYL, fully dissociated with 1.5 eq. of scandium triflate, was assessed to confirm the absence of fluorescence

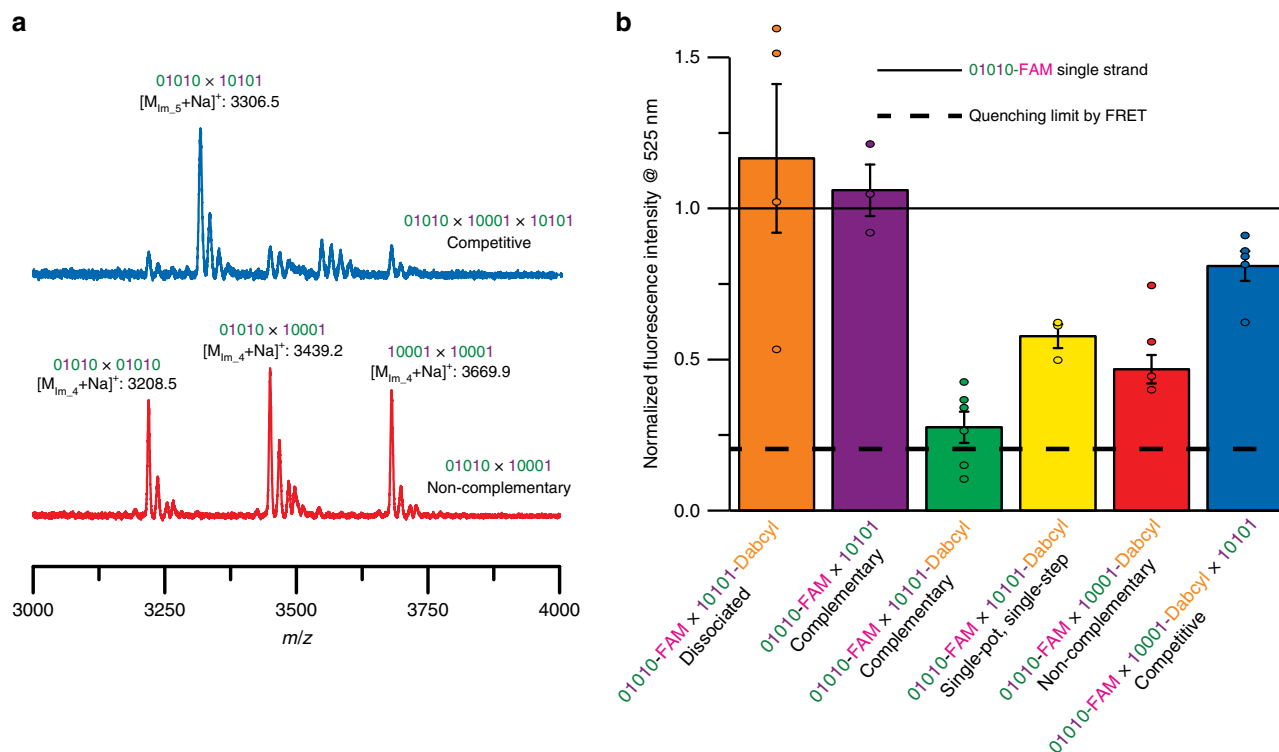


Fig. 2 Competitive dynamic covalent assembly. **a** MALDI mass spectra of reaction mixtures generated after applying the dissociation/extraction/annealing process to non-complementary 01010 and 10001 sequences (bottom, red) and to a ternary mixture of a competitive environment consisting of a complementary sequence and a non-complementary sequence via the dissociation/extraction/annealing process (top, blue). Peaks at multiples of $+18m/z$ values are attributable to ladders species with fewer rungs. **b** Sequence-selective hybridization assessed by FRET and normalized to the fluorescence intensity of a 01010-FAM sequence single strand. Assembly solutions were excited at 495 nm and fluorescence emission was measured at 425 nm. Error bars represent standard error. The horizontal solid and dashed lines 1 and 0.2, respectively, represent the unquenched 01010-FAM control and the reported FAM/DABCYL fluorescence quenching efficiency of 80%⁴⁰, respectively.

quenching for un-hybridized strands, while another solution of 01010-FAM \times 10101 (i.e., omitting a quencher) was examined to confirm that the hybridization of the FAM-bearing strand with its DABCYL-free complementary sequence does not negatively affect fluorescence (see Fig. 2b, orange and purple bars, respectively). Upon treatment with the dissociation/extraction/annealing process, the 01010-FAM \times 10101-DABCYL reaction mixture afforded a normalized fluorescence intensity of 0.275 ± 0.050 (i.e., a $72.5 \pm 5.0\%$ reduction in fluorescence, see Fig. 2b, green bar), a value approaching the reported FAM/DABCYL fluorescence quenching efficiency of 80%⁴⁰. Conversely, solutions of both the same 01010-FAM and 10101-DABCYL oligomer pair after treatment with the single-pot, single-step assembly method, and the non-complementary 01010-FAM and 10001-DABCYL sequences after the dissociation/extraction/annealing process yielded moderate reductions in fluorescence intensity (Fig. 2b, yellow and red bars, respectively); however, the extent of fluorescence quenching for these systems was curtailed by the generation not only of hybridized species bearing both FAM and DABCYL but also of hybrids exclusively bearing either FAM or DABCYL groups (i.e., equivalent to the hybridized species between non-complementary strands apparent in Fig. 1b, bottom black, and Fig. 2a, bottom red spectra, respectively). Finally, hybridization selectivity was examined using a ternary system, equivalent to that described above, employing stoichiometric amounts of the sequences 01010-FAM, 10001-DABCYL, and quencher-free 10101. After the dissociation/extraction/annealing process, the normalized fluorescence intensity was determined to be 0.810 ± 0.055 (Fig. 2b, blue bar), demonstrating the

preferential hybridization of complementary sequences enabled by this process to selectively afford the 10101 \times 01010-FAM molecular ladder bearing 5 imine-based rungs (Fig. 2b, green bar) over a non-complementary, 4 imine-rung 10001-DABCYL \times 01010-FAM molecular ladder (Fig. 2b, red bar).

Sequence-selective hybridization. The assembly of sophisticated, multimeric macromolecular complexes requires the selective and concurrent hybridization of multiple information-encoded components^{7–10}. To examine the potential for this dynamic covalent assembly system to achieve concurrent hybridization selectivity, three unique pairs of mass-labeled, complementary peptoids sequences (shown in Fig. 3a) were initially each subjected to the stepwise, dissociation/extraction/annealing assembly process described above. The mass spectra of each reaction mixture demonstrate the successful assembly of the target 10101 \times 01010, 11100 \times 00011, and 11111 \times 00000 molecular ladders from their respective precursor strands (see Fig. 3b, bottom three black, red, and blue spectra, respectively); for comparison, mass spectra of the 00111 \times 11000 and 11111 \times 00000 molecular ladder reaction mixtures using the previously determined, single-step reaction conditions are shown in Supplementary Fig. 9. Notably, assembly of the precursor strands in either a parallel or anti-parallel orientation can afford in-registry molecular ladders for both 10101 \times 01010 and 11111 \times 00000 systems; however, the in-registry 00111 \times 11000 ladder product can only be achieved through parallel-oriented strands. Whereas hybridization of 11111 and 00000 homo-oligomers achieves registry under either reaction conditions (Supplementary Fig. 9a), the 00111 \times 11000 system

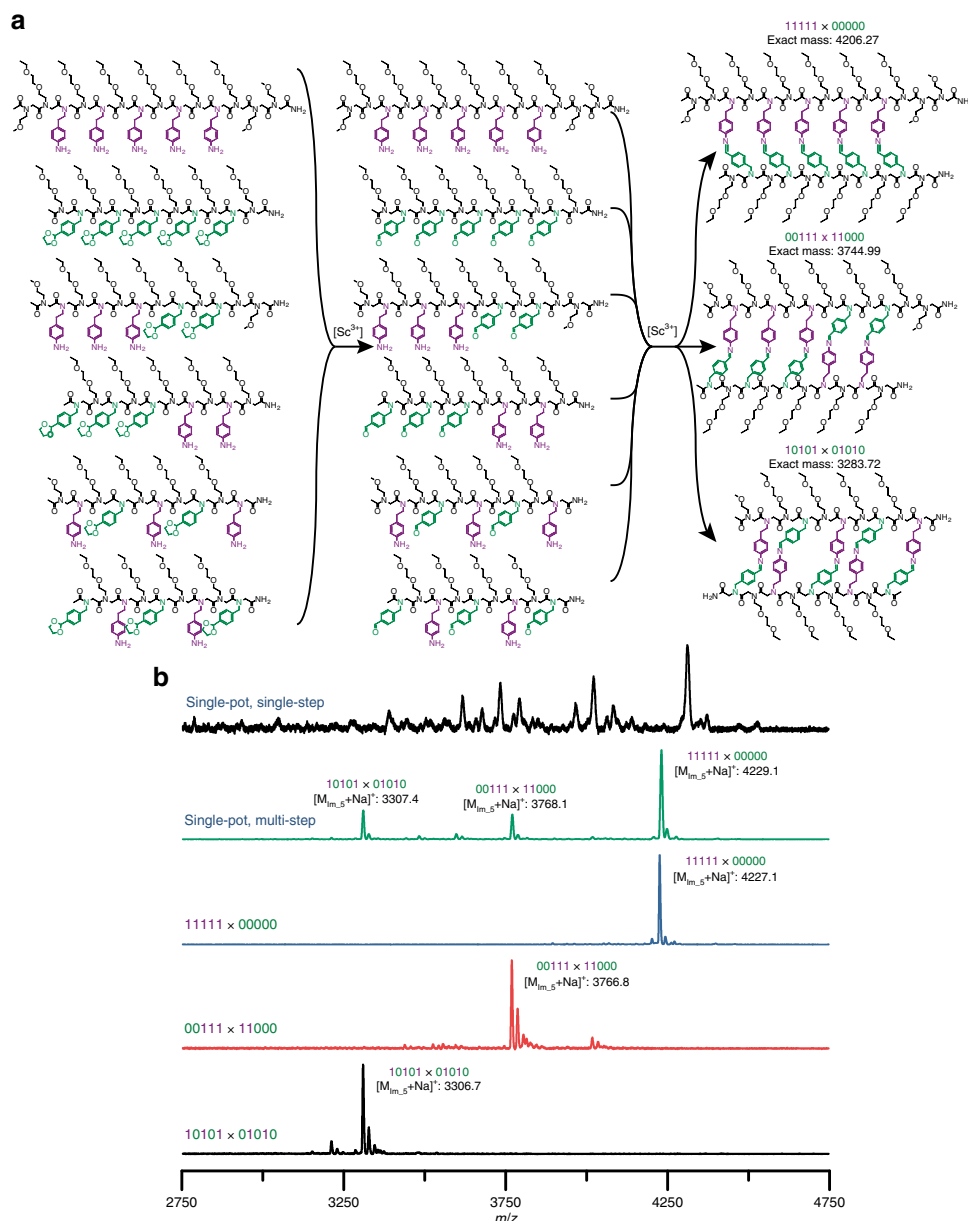


Fig. 3 Sequence-selective hybridization. **a** Structures of mass-labeled, ethylene acetal-protected precursor peptoid sequences, deprotected, dissociated sequences upon treatment with $\text{Sc}(\text{OTf})_3$, and in-registry molecular ladders upon $[\text{Sc}^{3+}]$ extraction and annealing. **b** MALDI mass spectra of individual encoded molecular ladders assembled via the dissociation/extraction/annealing process, including 10101 × 01010 (bottom, black), 00111 × 11000 (second from bottom, red), 11111 × 00000 (middle, blue), and a single-pot solution of all six oligomers to yield three in-registry molecular ladders (second from top, green). A single-pot solution of the six oligomers after the single-step, deprotection and direct assembly process (top, black) is shown for comparison. Expected exact masses: $[\text{M}_{10101 \times 01010} + \text{Na}]^+ = 3306.71$; $[\text{M}_{00111 \times 11000} + \text{Na}]^+ = 3767.93$; $[\text{M}_{11111 \times 00000} + \text{Na}]^+ = 4229.25$. Peaks at multiples of $+18m/z$ values are attributable to ladders species with fewer rungs.

results in poor hybridization selectivity under the direct deprotection and assembly reaction conditions (Supplementary Fig. 9b, bottom black spectrum), while the dissociation/extraction/annealing process selectively affords the in-registry, five-rung molecular ladder resulting from the hybridization of the 00111 strand with its 11000 complement in a parallel orientation (Supplementary Fig. 9b, top red spectrum).

Concurrent hybridization selectivity for these three sets of complementary peptoid pairs was subsequently examined by allowing equimolar amounts of the six oligomeric sequences (i.e., Fig. 3a) to react simultaneously in a single pot reaction mixture. Here, excluding out-of-registry and multimeric reaction products, this system of six individual strands could potentially yield 19

unique duplexes (as the 11111 and 00000 sequences cannot self-hybridize); nevertheless, application of the dissociation/extraction/annealing assembly process enabled each strand to preferentially hybridize with its residue sequence complement, resulting in the selective assembly of three in-registry molecular ladder products, including 10101 × 01010, parallel-oriented 00111 × 11000, and 11111 × 00000 (Fig. 3b, second from top green spectrum). In contrast, the single-step, deprotection and direct assembly reaction conditions again impeded hybridization selectivity of this system, yielding a non-equilibrium mixture of multiple, unidentified dimeric species (Fig. 3b, top black spectrum), highlighting the extraordinary capacity for the dissociation/extraction/annealing process to alleviate the kinetic

trapping that commonly accompanies dynamic covalent assembly (Supplementary Fig. 10). The full range of this spectrum (m/z of 1000–10,000, see Supplementary Fig. 11) again revealed small peaks attributable to single-stranded species, including both unreacted strands and those with either one or two intramolecular imine bonds, while multimeric complexes at higher molecular weights were not observed. Moreover, the concurrent assembly of 10101, 01010, 00111, and 11000 sequences exclusively affords 10101 \times 01010 and parallel-oriented 00111 \times 11000 ladders rather than the non-complementary cross-species 10101 \times 11000 and 01010 \times 00111, both of which could potentially form five imine bonds, confirming that pendant group sequence dictates the hybridization selectivity rather than simply maximizing the number of interstrand imine bonds, and strongly suggesting an unconvoluted, ladder-like conformation.

Information storage and retrieval. Inspired by the facile nucleic acid sequence identification provided by contemporary DNA microarray technology^{41,42} and the hybridization selectivity of the encoded sequences demonstrated above, we were motivated to explore the versatility of this model system as a potential information retrieval mechanism. Here, messages in the form of binary-encoded peptoid sequences are challenged against a known library of unpaired, mass-labeled oligomers. Selective hybridization of a message peptoid with its complementary sequence should afford a molecular ladder with a pre-determined mass such that the message can be identified through mass spectrometry on the reaction mixture. To validate this concept, the mass-labeled, acetal-protected peptoid oligomers 01010, 11000, and 11111 were used in equimolar amounts as a known library of potential complementary sequences (Fig. 4a). Initially, treatment of this library to the dissociation/extraction/annealing assembly process in the absence of a message strand resulted in non-specific binding among the library strands, yielding five sets of dimeric ladder species with varying numbers of rungs (see Fig. 4b, bottom black spectrum). Upon addition of a sub-stoichiometric amount of a 00111 message strand to a fresh sequence library and application of the dissociation/extraction/annealing process, characterization of the reaction mixture by mass spectrometry revealed a dominant new peak attributable to a selectively hybridized, in-registry ladder composed of the message oligomer and its complementary strand from the known library (Fig. 4b, second from top blue spectrum), successfully demonstrating the capacity of this system to convey information despite the presence of competitive binding species in excess. Similar results were obtained for 01010 and 00000 message strands when challenged against the oligomer library (see Fig. 4b, second from bottom red and top green spectra, respectively). To determine the detection limit for this message identification approach, varying sub-stoichiometric amounts of a 10101 message strand were added to the known sequence library described above (see Supplementary Fig. 12a) and, after treatment with the dissociation/extraction/annealing process, the mixtures were characterized again by mass spectrometry. Whereas peaks attributable to the 10101 \times 01010 ladder product were readily apparent in the mass spectrum at 0.2 equivalents of the message strand, addition of 0.75 equivalents of the message were necessary for a peak attributable the in-registry ladder product generated from the message strand and its complementary sequence to become dominant in the mass spectrum (Supplementary Fig. 12b and Supplementary Table 2).

The ability to direct the self-assembly of oligomeric strands based on their residue sequence and mediated by dynamic covalent interactions as demonstrated here is a crucial step towards the fabrication of complex, unimolecular constructs from modest, synthetically accessible precursors. By employing a step-wise process technique involving Sc(OTf)₃ as a multi-role reagent to

deprotect acetal-protected aldehyde groups, affect the equilibrium of the amine/aldehyde condensation reaction, and catalyze the rearrangement of inter-strand imine bonds, the kinetic trapping of non-equilibrium species that often prevails in dynamic covalent assembly systems is mitigated, allowing the system to converge towards thermodynamic equilibrium through sequence-selective assembly of molecular ladders from oligo(peptoid)s. By employing a dissociation/annealing assembly process, other dynamic covalent interactions could be utilized to mediate selective sequence assembly (e.g., boronic acid/diol condensation, Diels–Alder cycloaddition, etc.), similarly mimicking that of nucleic acids. Although this study involved molecular ladders bearing covalent rungs, this multi-step approach to dynamic covalent assembly process may also be useful for other applications in which the alleviation or elimination of kinetic trapping is critical. Specifically, this process will provide significantly improved synthetic access to robust, complex covalent nanostructures, such as molecular cages and crystalline, porous polymer networks.

Methods

General experimental procedures. Unless otherwise stated, chemicals and reagents were obtained from commercial sources and used as received. ¹H NMR spectra were collected using Varian MR400 and Varian Inova 500 spectrometers. DOSY NMR spectra were collected using a Bruker 400 MHz NMR spectrometer. Chemical shifts were measured in δ (ppm) relative to residual solvent signals as internal standards (CDCl₃—7.24 for ¹H; CD₃CN—1.94 for ¹H). Matrix-assisted laser desorption/ionization (MALDI) mass spectra were recorded using a Bruker Autoflex mass spectrometer, whereas electrospray ionization (ESI) mass spectra were recorded using an Agilent Q-TOF 1200 series spectrometer. MALDI analyses were performed in reflectron positive ion mode using 2-(4-hydroxyphenylazo) benzoic acid (HABA) as the matrix, where 2 μ L of a solution of the sample (1 mM) was mixed with 6 μ L of a mixture of 10 mg matrix in 200 μ L acetonitrile, spotted on a MALDI sample plate (Bruker), and allowed to air dry. Reverse phase high performance liquid chromatography (RP-HPLC) was performed using a Shimadzu LC-6AD HPLC pump, equipped with a Shimadzu FRC 70A fraction collector, using analytical and preparative reversed phase Phenomenex Luna C18(2) columns with a linear gradient of water and acetonitrile as the eluent at 30 °C, and monitored with a Shimadzu Prominence UV/vis detector at 214 nm. GPC was performed using a Shimadzu LC-20AD HPLC pump equipped with three Phenogel GPC/SEC columns (length 300 mm \times diameter 7.8 mm, pore sizes of 500, 100, and 50 Å) in series, 94:4:2 (v/v/v) CHCl₃:MeOH:Et₃N as the eluent at room temperature, and monitored with a Shimadzu Prominence UV/vis detector at 313 nm. GPC calibration was performed using low dispersity polystyrene standards (low molecular weight ReadySet, Fluka).

Monomer synthesis. Primary amine monomers were purchased or synthesized according to published approaches^{34,35} and can be seen in Supplementary Table 1.

Npal: Npal was synthesized by addition of 42.2 mL ethylene glycol, 0.02 g toluene *p*-sulfonic acid, and 25 g of 4-cyanobenzaldehyde to 200 mL of toluene. After refluxing overnight with a Dean–Stark trap to remove water generated during the reaction, the reaction mixture was cooled to room temperature and 40 mL of 5% NaHCO₃ aqueous solution was added. The organic layer was extracted, washed with deionized water three times, and evaporated to yield 4-(1,3-dioxacyclopent-2-yl) benzonitrile. 10 g of 4-(1,3-dioxacyclopent-2-yl)benzonitrile in 100 mL dry diethyl ether was added dropwise into 4.33 g of LiAlH₄ in 100 mL dry diethyl ether under nitrogen. The reaction mixture was stirred for 4 h at 0 °C and 12 h at room temperature, then quenched by 95% ethanol, and further quenched by 50% ethanol in water. The ether supernatant was dried under reduced pressure to yield Npal.

Npm: 4.9 g allyl chloroformate in 150 mL 1,4-dioxane was added into a solution of 4-(2-aminoethyl)aniline (5 g) in 150 mL 10% aq. acetic acid. The reaction mixture was stirred overnight at room temperature and then diluted with water prior to washing with diethyl ether. The aqueous layer was adjusted to pH 14 by 2 M NaOH (aq) and was extracted by diethyl ether. The organic fraction was evaporated to yield Npm.

Nsee: 50 mL of 6 M NaOH was added to 20 g of diethylene glycol monoethyl ether and 50 mL of tetrahydrofuran (THF) cooled to 0 °C, followed by dropwise addition of 54 g tosyl chloride in 80 mL THF under N₂. After stirring for 1 h at 0 °C, the reaction mixture was left for an additional hour at room temperature. The mixture was extracted with diethyl ether and the organic layer washed with 1 M NaOH and water before evaporating under vacuum to yield 2-(2-ethoxyethoxy) ethyl tosylate. 40 g of 2-(2-ethoxyethoxy)ethyl tosylate and 250 mL of DMF were charged to a 500 mL flask under N₂. 31.5 g of NaN₃ was added to the reaction mixture which was then stirred at 60 °C for 36 h before cooling to room temperature. The reaction mixture was extracted with diethyl ether and washed with water before evaporating under vacuum to produce 2-(2-ethoxyethoxy) ethyl azide. 2-(2-Ethoxyethoxy)ethyl azide (20 g) in 160 mL THF and 40 g

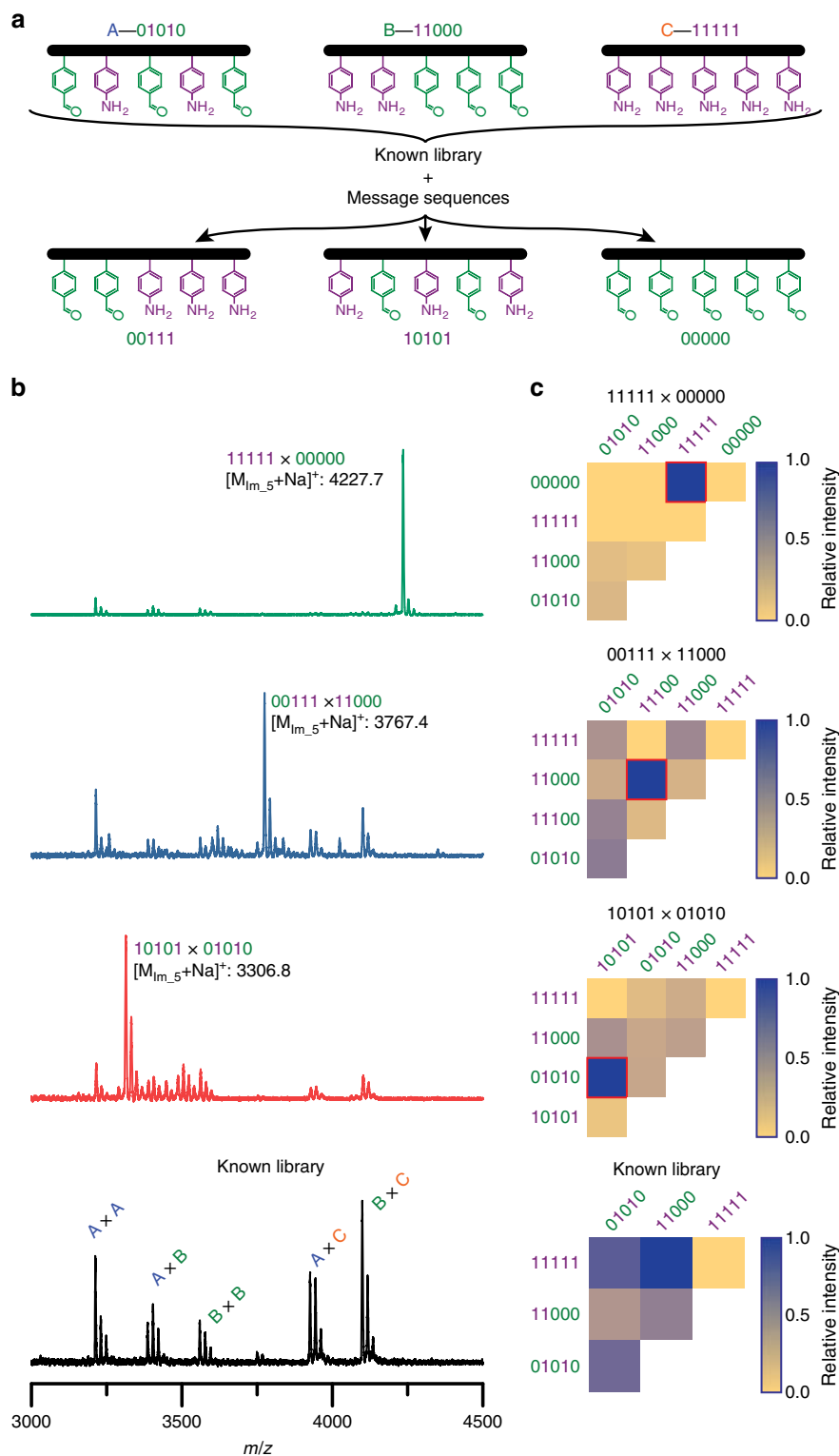


Fig. 4 Dynamic covalent information storage and retrieval. **a** Schematic of a known library composed of three unique, mass-labeled oligomeric sequences A (01010), B (11000), and C (11111), and individual messages sequences used to challenge the library. **b** MALDI mass spectrum of the known sequence library treated to the dissociation/extraction/annealing assembly process in the absence of a message strand (black, bottom), resulting in non-specific binding among the library strands, yielding all five possible sets of dimeric ladder species with varying numbers of rungs. Introduction of a message sequence to the known library and treatment of the mixture to the multi-step assembly process yielded MALDI mass spectra with dominant peaks attributable to the in-registry hybridization product of the message strand with its library complement (10101 message, second from bottom, red; 00111 message, second from top, blue; 00000 message, top, green). **c** Hybridization specificity of message strands with their library complements assessed by normalized MALDI mass spectrum intensities. Possible dimeric ladder species are shown; red boxes indicate expected hybrids between message and complementary library strands. Peaks at multiples of +18m/z values are attributable to ladders species with fewer rungs.

triphenylphosphine were charged to a 250 mL flask under nitrogen. Upon stirring overnight, the reaction mixture was quenched with 220 mL water and allowed to stir for another day. The resulting solution was washed with toluene and dichloromethane, and evaporated under vacuum to yield Nee.

Ndab: To a 20 mL vial equipped with magnetic stirrer, 500 mg of DABCYL succinimide ester and 10 mL (excess) of distilled ethylene diamine were added. The reaction was allowed to stir for 12 h and then excess ethylene diamine was removed by rotary evaporation. The product was extracted by dichloromethane and then washed with distilled water to remove excess amine and succinimide. The organic layer was dried over Na_2SO_4 , filtered and solvent removed by rotary evaporation to yield Ndab.

Oligo(peptoid) synthesis. Peptoids were synthesized via a sub-monomer approach to solid phase synthesis using Fmoc-Photolabile SS (0.1 mmol scale, 100–200 mesh, 1% DVB, Advanced ChemTech, Kentucky) as the resin. Syntheses were performed in an automated microwave synthesizer (Liberty Blue, CEM Corporation, North Carolina) adapted for peptoid synthesis. Resin was swelled at room temperature for 5 min with DMF before deprotection with 20% 4-methylpiperidine in DMF (v/v) for 30 s at 75 °C and 90 s at 90 °C. Subsequently, alternating additions of 1 mL of 1 M bromoacetic acid (coupled with 1 mL of 1 M *N,N'*-diisopropylcarbodiimide (DIC)) and 2.5 mL of 0.5 M primary amine monomer were repeated until desired sequence length was achieved. Primary amine monomers consisted of Nee as the principal spacer, an allyloxycarbonyl- (Alloc) protected amine (Npal protected, Nam unprotected), an ethylene acetal-protected aldehyde (Npal protected, Nal unprotected), and Nme as a mass label (Supplementary Table 1). The *N*-terminal of the oligomers were capped with 1 M acetic anhydride, activated by DIC. Fluorescein-functionalized oligomers were synthesized by omitting the acetic anhydride capping step and instead of introducing a final coupling step by adding 1 mL of a 0.4 M 5,6-carboxyfluorescein solution in DMF in addition to 1 mL 0.6 M HCTU in DMF and 1.5 mL of 0.8 M *N,N*-diisopropylethylamine in DMF.

Oligo(peptoids) were subjected on-resin to 0.1 equivalents of tetrakis (triphenylphosphine) palladium(0) and 25 equivalents of phenylsilane per Alloc group in dry DCM for one hour to deprotect the Alloc-amines. The deprotection solution was filtered off, and peptoids were introduced to a fresh deprotection solution for another hour. Upon filtration of the second solution, the peptoids on-resin were submerged in 5 mL DMF and cleaved for 36 h under irradiation at $\sim 25 \text{ mW cm}^{-2}$ with 405 nm. Resin was separated from liberated peptoids with the use of a syringe filter. Solvent was removed under vacuum, and the oligomers were reconstituted in acetonitrile and further purified in preparative HPLC using a linear gradient of acetonitrile (MeCN) and water: (1) 30% MeCN, 0.1–2.1 min; (2) 30–95% MeCN, 2.1–16.1 min; (3) 95% MeCN, 16.1–23.1 min; (4) 95% MeCN, 23.1–26.1 min. Purified peptoids were lyophilized to yield off-white powder.

Peptoid hybridization. Conventional single-step method: A vial was charged with 20 μL of 10 mM solutions of each of the desired oligo(peptoid) sequences in acetonitrile, 0.2 eq. of $\text{Sc}(\text{OTf})_3$ per potential imine bond from a 10 mM stock solution, and sufficient solvent to afford 200 μL total of a 2% (v/v) of water/acetonitrile solution. The mixture stirred gently for 2 h at 70 °C.

Dissociation/extraction/annealing method: Dissociation of peptoid sequences was achieved by adding a rare-earth metal triflate (either 1.5 eq. of $\text{Sc}(\text{OTf})_3$, $\text{Yb}(\text{OTf})_3$, or $\text{Lu}(\text{OTf})_3$, or 3.0 eq. of $\text{Y}(\text{OTf})_3$ or $\text{Sm}(\text{OTf})_3$ per potential imine bond) from a 10 mM stock solution to a 200 μL 2% (v/v) of water/acetonitrile solution containing 0.2 μmol of each desired oligo(peptoid) sequence and heating at 70 °C. After heating for 2 h, 200 μL of chloroform and 2 mL of water were added to the vial, followed by gentle shaking. The mixture was allowed to stand and, upon complete phase separation³⁰, the organic layer was extracted and stirred in a new vial at 70 °C for 2 h for oligomer annealing.

Sequence identification method: Sequence identification hybridizations were similarly performed by adding 15 μL of a 10 mM solution of ‘message’ peptoid to a vial containing 200 μL of a 2% (v/v) of water/acetonitrile solution, 0.2 μmol of each peptoid strand constituting the known library (01010, 11000, and 11111) and 0.225 μmol of $\text{Sc}(\text{OTf})_3$. As before, this solution was heated at 70 °C for 2 h, then diluted with 200 μL of chloroform. After heating, 2 mL of water was added and the solution was gently shaken. Upon isolation of the chloroform layer, annealing was carried out at 70 °C for 2 h.

Förster resonance energy transfer. Quantification of hybridization selectivity was achieved via the dissociation/extraction/annealing of 1 mM hybridization mixtures initially consisting of distinct combinations of 01010-FAM, 10101, 10001, 10101-DABCYL, and 10001-DABCYL strands in stoichiometric ratios and 1.5 eq. $\text{Sc}(\text{OTf})_3$ in 200 μL of a 2% (v/v) of water/acetonitrile solution. After annealing at 70 °C for 6 h and room temperature for one day, 200 μL of a pH 9 bicarbonate-carbonate buffer solution was added to each hybridization mixture. The hybridized solutions were further diluted to 1.2 mL with acetonitrile, then the fluorescence of the resultant solutions was recorded with a Horiba Quanta Master fluorimeter by exciting at 495 nm and measuring the fluorescence intensity at 525 nm. Triplicates of each solution were tested and normalized to a control mixture of 01010-FAM that was subject to the same dissociation/extraction/annealing process. Results are reported with standard error.

Inductively coupled plasma mass spectrometry. Post-dissociation/extraction solutions of hybridized 11111 \times 00000, 00111 \times 11000, or 10101 \times 01010 (initially 1 mM in 200 μL 2% (v/v) of water/acetonitrile and either 1.5 μmol of $\text{Sc}(\text{OTf})_3$, $\text{Yb}(\text{OTf})_3$, or $\text{Lu}(\text{OTf})_3$, or 3.0 μmol of $\text{Y}(\text{OTf})_3$ or $\text{Sm}(\text{OTf})_3$ in CHCl_3 were dried under a steady stream of nitrogen and reconstituted in 1 mL of 2% nitric acid (aq). Samples were diluted 4 \times 10⁶-fold with Milli-Q water and the residual scandium concentrations were determined by ICP-MS using a Perkin-Elmer NexION 2000. Experiments on each solution were performed in triplicate and results are reported with standard errors.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 29 August 2019; Accepted: 21 January 2020;

Published online: 07 February 2020

References

- Lehn, J.-M. Perspectives in supramolecular chemistry—from molecular recognition towards molecular information processing and self-organization. *Angew. Chem. Int. Ed.* **29**, 1304–1319 (1990).
- Lutz, J.-F., Ouchi, M., Liu, D. R. & Sawamoto, M. Sequence-controlled polymers. *Science* **341**, 1238149 (2013).
- Chen, Z. et al. Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111 (2018).
- Coppock, M. B., Miller, J. R. & Williams, M. E. Assembly of a trifunctional artificial peptide into an anti-parallel duplex with three Cu(II) cross-links. *Inorg. Chem.* **50**, 949–955 (2011).
- Gong, B. Molecular duplexes with encoded sequences and stabilities. *Acc. Chem. Res.* **45**, 2077–2087 (2012).
- Yang, Y. & Chen, C. -F. in *Hydrogen Bonded Supramolecular Structures* (eds. Li, Z. -T. & Wu, L. -Z.) 115–136 (Springer, Berlin, Heidelberg, 2015).
- Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99**, 237–247 (1982).
- Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
- Tørring, T., Voigt, N. V., Nangreave, J., Yan, H. & Gothelf, K. V. DNA origami: a quantum leap for self-assembly of complex structures. *Chem. Soc. Rev.* **40**, 5636–5646 (2011).
- Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485**, 623–626 (2012).
- Clausen-Schaumann, H., Rief, M., Tolksdorf, C. & Gaub, H. E. Mechanical stability of single DNA molecules. *Biophys. J.* **78**, 1997–2007 (2000).
- Nielsen, P. E. Peptide nucleic acid. A molecule with two identities. *Acc. Chem. Res.* **32**, 624–630 (1999).
- Eriksson, M. & Nielsen, P. E. PNA-nucleic acid complexes. Structure, stability and dynamics. *Q. Rev. Biophys.* **29**, 369–394 (1996).
- Zhang, L., Peritz, A. & Meggers, E. A simple glycol nucleic acid. *J. Am. Chem. Soc.* **127**, 4174–4175 (2005).
- Schweitzer, B. A. & Kool, E. T. Aromatic nonpolar nucleosides as hydrophobic isosteres of pyrimidine and purine nucleosides. *J. Org. Chem.* **59**, 7238–7242 (1994).
- Piccirilli, J. A., Krauch, T., Moroney, S. E. & Benner, S. A. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **343**, 33–37 (1990).
- Abdallah, H. O. et al. Stabilisation of self-assembled DNA crystals by triplex-directed photo-cross-linking. *Chem. Commun.* **52**, 8014–8017 (2016).
- Takasugi, M. et al. Sequence-specific photo-induced cross-linking of the two strands of double-helical DNA by a psoralen covalently linked to a triple helix-forming oligonucleotide. *Proc. Natl Acad. Sci. USA* **88**, 5602–5606 (1991).
- Ren, F., Day, K. J. & Hartley, C. S. Two- and three-tiered stacked architectures by covalent assembly. *Angew. Chem. Int. Ed.* **55**, 8620–8623 (2016).
- Tozawa, T. et al. Porous organic cages. *Nat. Mater.* **8**, 973–978 (2009).
- Huang, N., Wang, P. & Jiang, D. Covalent organic frameworks: A materials platform for structural and functional designs. *Nat. Rev. Mater.* **1**, 1–19 (2016).
- Reuther, J. F. et al. Dynamic covalent chemistry enables formation of antimicrobial peptide quaternary assemblies in a completely abiotic manner. *Nat. Chem.* **10**, 45–50 (2018).
- Reuther, J. F., Dahlhauser, S. D. & Anslyn, E. V. Tunable orthogonal reversible covalent (TORC) bonds: dynamic chemical control over molecular assembly. *Angew. Chem. Int. Ed.* **58**, 74–85 (2019).
- Hebel, M. et al. Sequence programming with dynamic boronic acid/catechol binary codes. *J. Am. Chem. Soc.* **141**, 14026–14031 (2019).

25. Strom, K. R., Szostak, J. W. & Prywes, N. Transfer of sequence information and replication of diimine duplexes. *J. Org. Chem.* **84**, 3754–3761 (2019).
26. Dunn, M. F., Wei, T., Zuckermann, R. N. & Scott, T. F. Aqueous dynamic covalent assembly of molecular ladders and grids bearing boronate ester rungs. *Polym. Chem.* **10**, 2337–2343 (2019).
27. Elliott, E. L., Hartley, C. S. & Moore, J. S. Covalent ladder formation becomes kinetically trapped beyond four rungs. *Chem. Commun.* **47**, 5028–5030 (2011).
28. Hartley, C. S., Elliott, E. L. & Moore, J. S. Covalent assembly of molecular ladders. *J. Am. Chem. Soc.* **129**, 4512–4513 (2007).
29. Mattia, E. & Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **10**, 111–119 (2015).
30. Zuckermann, R. N. Peptoid origins. *Biopolymers* **96**, 545–555 (2011).
31. Zuckermann, R. N., Kerr, J. M., Moos, W. H. & Kent, S. B. H. Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* **114**, 10646–10647 (1992).
32. Sun, J. & Zuckermann, R. N. Peptoid polymers: a highly designable bioinspired material. *ACS Nano* **7**, 4715–4732 (2013).
33. Edison, J. R. et al. Conformations of peptoids in nanosheets result from the interplay of backbone energetics and intermolecular interactions. *Proc. Natl Acad. Sci. USA* **115**, 5647–5651 (2018).
34. Wei, T., Furgal, J. C. & Scott, T. F. In situ deprotection and dynamic covalent assembly using a dual role catalyst. *Chem. Commun.* **53**, 3874–3877 (2017).
35. Wei, T., Furgal, J. C., Jung, J. H. & Scott, T. F. Long, self-assembled molecular ladders by cooperative dynamic covalent reactions. *Polym. Chem.* **8**, 520–527 (2017).
36. Giuseppone, N., Schmitt, J. L., Schwartz, E. & Lehn, J. M. Scandium(III) catalysis of transimination reactions. Independent and constitutionally coupled reversible processes. *J. Am. Chem. Soc.* **127**, 5528–5539 (2005).
37. Meng, W. et al. An autonomous molecular assembler for programmable chemical synthesis. *Nat. Chem.* **8**, 542–548 (2016).
38. Furgal, J. C., Goodson, T. & Laine, R. M. D_{5h} [PhSiO_{1.5}]₁₀ synthesis via F⁻ catalyzed rearrangement of [PhSiO_{1.5}]_n. An experimental/computational analysis of likely reaction pathways. *Dalt. Trans.* **45**, 1025–1039 (2016).
39. Wei, T., Jung, J. H. & Scott, T. F. Dynamic covalent assembly of peptoid-based ladder oligomers by vernier templating. *J. Am. Chem. Soc.* **137**, 16196–16202 (2015).
40. Marras, S. A. E., Kramer, F. R. & Tyagi, S. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res.* **30**, e122 (2002).
41. Botstein, D., Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95**, 14863–14868 (1998).
42. Pollack, J. R. et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23**, 41–46 (1999).

Acknowledgements

This work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award #DE-SC0012479 (hybridization selectivity), the National Science Foundation under Award #1462267 (oligomer synthesis and purification), the National Science Foundation Graduate Research Fellowship Program, and the National Institutes of Health under Grant #R01 DE026116 (amine/aldehyde condensation reaction equilibrium).

Author contributions

S.C.L. and T.F.S. conceived the project. S.C.L. designed and performed the experiments. T.F.S. directed the project. S.C.L. and T.F.S. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-14607-3>.

Correspondence and requests for materials should be addressed to T.F.S.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020