HHS Public Access

Author manuscript

Nat Chem. Author manuscript; available in PMC 2014 December 26.

Published in final edited form as:

Nat Chem. 2013 April ; 5(4): 282–292. doi:10.1038/nchem.1577.

Enzyme-Free Translation of DNA into Sequence-Defined Synthetic Polymers Structurally Unrelated to Nucleic Acids

Jia Niu, **Ryan Hili**, and **David R. Liu***

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Abstract

The translation of DNA sequences into corresponding biopolymers enables the production, function, and evolution of the macromolecules of life. In contrast, methods to generate sequence-defined synthetic polymers with similar levels of control have remained elusive. Here we report the development of a DNA-templated translation system that enables the enzyme-free translation of DNA templates into sequence-defined synthetic polymers that have no necessary structural relationship with nucleic acids. We demonstrate the efficiency, sequence-specificity, and generality of this translation system by oligomerizing building blocks including polyethylene glycol (PEG), α-(D)-peptides, and β-peptides in a DNA-programmed manner. Sequence-defined synthetic polymers with molecular weights of 26 kDa containing 16 consecutively coupled building blocks and 90 densely functionalized β-amino acid residues were translated from DNA templates using this strategy. We integrated the DNA-templated translation system developed here into a complete cycle of translation, coding sequence replication, template regeneration, and re-translation suitable for the iterated *in vitro* selection of functional sequence-defined synthetic polymers unrelated in structure to nucleic acids.

Nucleic acid-templated polymerization is the molecular essence of gene replication, transcription, and translation. The ability of nucleic acids to template protein synthesis in living systems also enables the evolution of proteins with new structures and functions. In contrast, synthetic polymers are generally not created in a manner that enables single monomer-level control over polymer length and sequence.[1,2] Despite significant progress in controlling the structure[3-5] and molecular weight distribution[6-8] of synthetic polymers, methods that enable precise control over synthetic polymer sequence and length have remained elusive.[9] In part because of this limitation, synthetic polymers have primarily served as bulk materials rather than as precisely folded molecules with the ability to bind a target molecule with high affinity and selectivity, or the ability to catalyze a chemical reaction.

*Correspondence: drliu@fas.harvard.edu.

Author contributions

J.N., R.H., and D.R.L. designed the research, analyzed the data and co-wrote the manuscript.
J.N. performed the experiments.

An alternative approach to generating synthetic polymers of defined sequence and length that parallels the biosynthesis of proteins is the translation of DNA or RNA into sequence-defined synthetic polymers. Crucially, such a translation capability would also enable the laboratory evolution of synthetic polymers with structures and functional properties not limited to those of natural biopolymers through iterated cycles of translation, selection, and template replication. Several laboratories have developed enzyme-mediated and non-enzymatic nucleic acid-templated polymerization strategies that effect the translation of DNA or RNA sequences into biopolymer analogs including modified DNA, peptide nucleic acid (PNA), threose nucleic acid (TNA), hexitol nucleic acid (HNA), non-natural peptides, and others (Figure 1b).[10-16] Our group and others have developed enzyme-free DNA-templated oligomerization strategies that use DNA oligonucleotides as templates to direct the oligomerization of PNA,[17-19] functionalized DNA oligonucleotides,[20] amine acylation substrates,[21] and Wittig olefination substrates.[22] We integrated DNA-templated PNA oligomerization with an *in vitro* selection system for synthetic PNAs, enabling the proof-of-principle iterated translation and selection of a streptavidin-binding PNA oligomer from a library of 10^8 sequence-defined PNAs.[23] Chaput and coworkers recently selected a thrombin-binding TNA aptamer from a TNA library generated by a DNA polymerase-mediated TNA translation (Figure 1b).[24] Most recently, using laboratory-evolved DNA polymerase enzymes that accept non-natural nucleotide analogs, Holliger and coworkers expanded the pool of nucleic acid polymers that can be enzymatically translated from DNA and reverse-transcribed back to DNA to include HNA, TNA, 2'-O,4'-methylene-β-D-ribonucleic acid (locked nucleic acids, LNA), cyclohexyl nucleic acid (CeNA), arabinonucleic acid (ANA), and 2'-fluoro-arabino-nucleic acid (FANA) (Figure 1b).[25]

While these advances establish a strong foundation for future efforts in synthetic nucleic acid analog evolution, all examples to date of non-ribosomal translation systems to generate macromolecules, beyond those that exploit unique features of the Wittig olefination reaction,[22] require that the polymeric product closely resemble natural nucleic acids and maintain the ability to hybridize directly with a nucleic acid template (Figure 1b). This requirement imposes major constraints on the structural and functional potential of synthetic polymers generated by existing artificial translation strategies. Here we report the development and implementation of a strategy that overcomes this limitation and enables the non-enzymatic translation of DNA templates into sequence-defined synthetic polymers unrelated to nucleic acids. This strategy can support a complete cycle of translation, template replication and regeneration, and re-translation, signifying the ability of the system developed in this work to support iterated cycles of *in vitro* selection of non-nucleic acid synthetic polymers.

Results

Translation Strategy Design

We sought to emulate the function of a transfer RNA (tRNA) as an adapter that recognizes a template codon and brings a cognate non-nucleic acid building block into reactive proximity of a growing peptide chain (Figure 1a). We designed each substrate molecule to contain: (*i*) a PNA pentamer, analogous to a tRNA anticodon, that sequence-specifically recognizes a

DNA template codon; (*ii*) a synthetic polymer building block, analogous to an aminoacyl ester, that bears no necessary structural relationship with nucleic acids; and (*iii*) cleavable linkers that connect each PNA anticodon with its cognate synthetic polymer building block in a macrocycle (Figure 1d). We designed the substrates to be macrocycles to decrease the entropic cost of the building block coupling reactions and to increase their regioselectivity by aligning otherwise freely rotating building blocks into conformations that juxtapose reactive functional groups.[26]

We designed the translation process to proceed in three stages (Figure 1c). First, substrates hybridize sequence-specifically to a DNA template that contains a 5' hairpin followed by consecutive DNA codons. Substrate-template hybridization increases the effective molarity of reactive groups on adjacent building blocks. Second, a catalyst or reagent initiates coupling between building blocks, resulting in their oligomerization in a sequence-programmed order. Since the 5' end of the DNA template contains a group capable of coupling with the first building block, the synthetic polymer emerges from the translation process covalently linked to its encoding template. As a result, translation products can undergo iterative rounds of *in vitro* selection, template replication, and retranslation. Finally, after the oligomerization reaction is complete, the linkers between the PNA anticodons and the synthetic polymer are cleaved, releasing the linear synthetic polymer-DNA template conjugate from the PNA adapters. Because the entire translation process does not require any structural or functional feature of the synthetic polymer building blocks beyond their ability to support coupling and linker cleavage, this strategy should be compatible with a wide variety of polymers, including those unrelated to nucleic acids.

Evaluation of Building Block Coupling Chemistries and Substrate Architectures

To identify an efficient coupling reaction for polymerization, we explored chemistries known to be compatible with DNA and with solid-phase peptide synthesis.[27] For each candidate chemistry, we considered two substrate architectures: ones with heterobifunctional "AB" building blocks that each contains both types of coupling reaction functional groups, and substrates with homobifunctional "AA" or "BB" building blocks that each contains only one type of reactive functional group (Figure 2a) AB building blocks, when properly aligned, are capable of reacting with any adjacent AB building blocks, while AA or BB building blocks can only react with adjacent BB or AA building blocks, respectively. While AB-type substrates are conceptually simpler and place fewer restrictions on the resulting polymer structures than AA/BB-type substrates, they may be prone to intramolecular cyclization, an undesired alternative to polymerization (Figure 2a and Supplementary Results Figure S11).[28,29] We tested the efficiency and selectivity of both types of substrates using five candidate coupling chemistries: amine acylation, reductive amination, oxime and hydrazone formation, and Cu(I)-catalyzed alkyne-azide 1,3-dipolar cycloaddition (CuAAC or "click" chemistry).[30,31]

Using a combination of solid- and solution-phase reactions, we synthesized macrocyclic substrates for all five chemistries (Figure 2c, see Supplementary Methods for detailed synthetic procedures). Two disulfide-containing amino acid residues were incorporated in each macrocyclic substrate as cleavable linkers between the synthetic polymer building

block and the PNA adapter (Figure 1d, Figure 2c). For our initial studies, polyethylene glycol (PEG) chains served as the synthetic polymer building blocks. Coupling chemistries were evaluated by attempting the oligomerization of six consecutive building blocks on a DNA template containing six codons. For testing AB substrates, all six template codons were identical (TCCTT); for testing AA/BB substrates, templates contained alternating AA codons (AATCC) and BB codons (ATACC) (Figure 2a). These codons were chosen because they met criteria of efficient and sequence-specific hybridization between DNA and PNA in our previous studies.[23] Although G was not used in template codons in this study, our previous work suggests that G can also be successfully incorporated into codon sets.[18,23] After oligomerization of the substrates, dithiolthreitol (DTT) was added to cleave the disulfide linkers and liberate the synthetic polymer products.

Translation reactions were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE) and generated up to seven distinct PAGE bands. We hypothesized that the fastest-migrating (lowest) band corresponds to the starting template, the highest band corresponds to the full-length translation product, and the intermediate bands are truncated intermediates (Figure 2b, Supplementary Results Figure S5). When any of the three key components (the DNA template, the substrates, or the coupling reagent or catalyst) were omitted from the reaction mixture, no products were observed (Supplementary Results Figure S5). Mass spectrometry (MS) analysis confirmed that the material from the seventh (top-most) band has a molecular weight consistent with that of the full-length translation product (see below).

The DNA-templated oligomerization of AB- and AA/BB-type substrates using amine acylation as the coupling chemistry failed to couple more than two building blocks (Figure 2b, lanes 2 and 10). These results are consistent with previous findings that amine acylation is inefficient in DNA-templated oligomerizations.[17,32,33] Moreover, the intramolecular cyclization of AB-type substrates significantly decreased product yields (Supplementary Results Figure S11). Next, we tested reductive amination, which has been successfully used to polymerize analogs of DNA[34] and PNA.[17,18,23] Unfortunately, both AB- and AA/BB-type substrates coupled no more than twice under the conditions tested (Figure 2b, lanes 4 and 12). We hypothesized that the strict geometric requirements of reductive amination[17,35] are problematic in the context of these large macrocyclic substrates.

Next we explored chemistries that form products that are thermodynamically more stable than imines, such as hydrazone- and oxime-based coupling reactions. While hydrozone and oxime formation using AB building blocks did not deliver higher levels of desired products due to efficient competing intramolecular cyclization (Figure 2b, lanes 5 and 6), oligomerization of AA/BB building blocks using these chemistries successfully coupled up to five building blocks on the template (Figure 2b, lanes 13 and 14). Despite these promising observations, we were concerned with the potential instability of a polymer containing oxime linkages in basic pH conditions that may be needed during future *in vitro* selection procedures (Supplementary Results Figure S7).[36]

Finally, we tested the ability of the CuAAC reaction to effect translation.[19,37,38] Although inefficient coupling was again observed for AB-type substrates, presumably due to

competitive substrate cyclization (Figure 2b, lane 8), the CuAAC reaction resulted in full-length hexamer product when using AA/BB-type substrates (Figure 2b, lane 15). Further optimization of building block geometries significantly improved translation efficiency, such that up to 70 % of translation products were full-length (Figure 2b, lane 16, 18, and 19, Supplementary Results Table S4).

We determined the structural requirements for efficient translation using the CuAAC coupling chemistry. First, we varied the length of synthetic polymer building block. DNA-templated oligomerization of substrates containing synthetic polymer building blocks as short as four PEG units was efficient in generating full-length product (Supplementary Results Figure S6). Surprisingly, a moderate amount of full-length product is generated even when the synthetic polymer moiety is completely missing (Supplementary Results Figure S6, lanes 1 and 2). These results suggest that the DNA-templated translation strategy accommodates a wide range of lengths of the synthetic polymer building blocks. Next, we varied the lengths of the groups between the CuAAC coupling partners and the synthetic polymer building blocks. Alkynyl substrates with spacers of one PEG-unit and those with spacers of four PEG units both yielded full-length translation products efficiently (Supplementary Results Figure S6), suggesting that the translation is relatively insensitive to the length of these spacers. Interestingly, the oxygen atom at the β-position of the alkynyl group is required for efficient translation (compare Figure 2b, lanes 15 and 17 with lane 19). We speculate that the increased polarity of the oxygen atom may help reduce the tendency of the otherwise hydrophobic spacers to adopt compact conformations that are incompatible with oligomerization.

Finally, we tested if the macrocyclic nature of the substrates improved translation outcomes. Uncyclized substrates **6-uc** and **12-uc**, the linear analogues of **6** and **12**, result in prematurely terminated oligomerization when copolymerized with macrocyclic **12** and **6**, respectively (Figure 2b, lanes 20 and 21). When both **6-uc** and **12-uc** were used together in a DNA-templated translation reaction, the yield of full-length product was much lower than when macrocyclic substrates **6** and **12** were used (Figure 2b, lane 22 versus lane 16). In light of these findings, we pursued the CuAAC-mediated oligomerization of AA/BB-type macrocyclic substrates **6** and **12** (Figure 2c) as our model system for development.

Characterization of Translation Products

With an efficient coupling chemistry identified, we characterized the translation process in greater detail using modified substrates that contain only one functional group for coupling and therefore terminate oligomerization (**T6** and **T6-f**) (Figure 3a). **T6-f** is conjugated to the Alexa Fluor 647 fluorophore (Figure 3b) and therefore its incorporation can be readily quantitated. We prepared three templates in which the terminator codon (CATCA), analogous to a stop codon, was placed at either the second, fourth, or sixth codon position. Denaturing PAGE analysis revealed that oligomerization with substrates **6, 12**, and **T6-f** proceeded until the codon position complementary to the terminator substrate, and then stopped (Figure 3c). Fluorescent imaging revealed that **T6-f** was indeed incorporated in a sequence-programmed manner (Figure 3c, right gel, lane 3-5). In contrast, no Alexa Fluor 647 fluorescence was observed in a control reaction containing all substrates but using a

template lacking a terminator codon (Figure 3c, right gel, lane 2), indicating that the terminator is incorporated only when its corresponding codon is present in the template. Taken together, these results demonstrate that this DNA-templated translation strategy proceeds in a sequence-specific manner. An in-depth study of the sequence specificity of the system is presented below.

In order to unambiguously determine the presence and the identity of the full-length product, we analyzed translation reactions using high-resolution electrospray ionization (ESI) mass spectrometry. Products from the DNA-templated oligomerization of **6, 12**, and a terminator substrate (**T6**) (Figure 3b) along with a template containing the terminator codon were subjected to disulfide linker cleavage and analyzed by denaturing PAGE. The largest molecular weight band hypothesized to be the full-length translation product was excised and the material was extracted from the gel. The DNA portion of the product was digested using P1 nuclease and the remaining material was analyzed by ESI-LC-MS (Figure 3e). The mass spectrum revealed multiply charged species of a single mass consistent with the full-length synthetic polymer (observed mass = 9,979.2 Da; expected mass = 9,979.8 Da) (Figure 3d). Lower molecular weight products from the translation that were resolved by gel electrophoresis were also analyzed using ESI-LC-MS. The observed masses of these species are consistent with that of truncated translation products that failed to couple with downstream substrates (Supplementary Results Figure S8). These observations collectively support our interpretation of the PAGE data and suggest that the major product generated using this DNA-templated translation strategy is the sequence-programmed full-length synthetic polymer.

Detailed Study of Sequence Specificity

The sequence specificity of any synthetic polymer translation strategy is crucial to its suitability for synthetic polymer evolution. To test the sequence specificity of this system in greater depth, and to characterize its compatibility with a template containing multiple codons, we attempted the translation of a DNA template containing six different codons from two codon sets (Figure 4). Codons from set 1 encoded the azide building blocks, while codons from set 2 encoded the alkyne building blocks. Codons from set 1 and set 2 alternated along the coding region of the template. Each translation reaction used a mixture of PNA substrates comprising five AA/BB-type azide or alkyne substrates and one terminator substrate, each encoded by a different codon. Only if the terminator substrate is correctly installed opposite its complementary codon, and if the other azide and alkyne bifunctional substrates are incorporated sequence specifically, will oligomerization generate a product of the desired length.

For all six terminator substrates, the predominant translation product was the polymer of expected length (Figure 4, Supplementary Results Table S5). These results establish that all 12 substrates tested (six bifunctional substrates and six terminators) containing six different PNA anticodons are incorporated in a template sequence-programmed manner, even in the presence of a stoichiometric excess of non-cognate substrates.

Translation of DNA into Longer And Structurally More Diverse Synthetic Polymers

Since DNA hybridization in this approach is spatially separated from synthetic polymer building blocks, this system in theory should support the translation of DNA templates into a variety of synthetic polymers beyond the PEG-based polymers used in our initial studies. To test this possibility, we designed macrocyclic substrates that contain synthetic polymer building blocks of greater structural diversity. We synthesized macrocyclic substrates containing β-peptide and α-(D)-peptide backbones with a variety of proteinogenic and non-proteinogenic aryl, alkyl, amino, and carboxyl side chains (Figure 5a). We performed translation reactions with these structurally diverse building blocks as described above with the terminator substrates programmed to be incorporated at the sixth position of each oligomer.

Denaturing PAGE analysis revealed that **13**, an alkynyl substrate with a β-peptide building block co-oligomerized successfully with azido-PEG substrate **6** in the presence of a DNA template to provide full-length β-peptide-containing hexamer products (Figure 5a, lane 2, Supplementary Results Table S6). Similarly, substrate **14**, an alkyne-linked α-(D)-peptide substrate, co-oligomerized successfully with azido-PEG substrate **6** to provide full-length products in comparable yield (Figure 5a, lane 3, Supplementary Results Table S6). An azide-linked α-(D)-peptide substrate (**15**) also co-oligomerized with α-(D)-peptide substrate **14** to yield full-length α-(D)-peptide synthetic polymer as the major product (Figure 5a, lane 4, Supplementary Results Table S6).

Finally, an azide-linked, fully side-chain functionalized β-peptide substrate (**22**) co-oligomerized with an alkynyl, fully side-chain functionalized β–peptide substrate (**23**) to yield full-length β–peptide polymers with excellent efficiency (Figure 5a, lane 5, Supplementary Results Table S6). A notable feature of **22** and **23** is that they both contain minimized spacers between the synthetic polymer building blocks and the coupling functional groups and therefore result in translation products with a high ratio of encoded, densely functionalized synthetic polymer building blocks to non-encoded groups. ESI-LC-MS analysis after linker cleavage, gel purification, and P1 nuclease digestion confirmed the mass of the full-length products (for **6** + **13**, observed mass = 10,487.0 Da; expected mass = 10,487.2 Da; for **6** + **14**, observed mass = 11,072.0 Da; expected mass = 11,071.6 Da; for **14** + **15**, observed mass = 12,195.0 Da; expected mass = 12,194.6 Da; for **22** + **23**, observed mass = 9,645.3 Da; expected mass = 9643.5 Da) (Supplementary Results Figure S9). Taken together, these examples demonstrate that the strategy developed in this work can sequence-specifically translate DNA templates into synthetic polymers of uniform length containing a variety of backbone structures. To our knowledge, these results also represent the first enzyme-free translation of nucleic acids into synthetic polymers that have no ability to hybridize to DNA or RNA.

We characterized the ability of this translation system to generate longer synthetic polymer products. We performed the DNA-templated translation of templates containing six, eight, ten, 12, 14, and 16 pentameric codons using substrates **22, 23**, and terminator **T7-f**. The fluorescent terminator **T7-f** was structurally similar to **T6-f** but contains a shorter synthetic polymer building block and a minimized spacer. **T7-f** was included in each translation

reaction to enable full-length products to be easily visualized since the last codon in each template uniquely encodes **T7-f** and no other substrates contain fluorophores. Fluorescent imaging of the denaturing PAGE gel revealed that all translation reactions tested generated full-length polymer product. Although the translation of progressively longer templates yielded less full-length product as expected, the fraction of full-length products among total translation products remained above 60 % in all cases, and 66% of the translation products using the longest template tested were full-length polymers containing 16 consecutive building blocks (Figure 5b, Supplementary Results Table S7). This 16-mer product, not including the DNA template or PNA adapters, has a molecular weight of 26 kDa and contains 90 β-amino acid residues. These observations indicate the feasibility of translating DNA sequences into sequence-defined synthetic polymer products of molecular weights comparable to those of functional biological polymers.

A Complete Cycle of Translation, Coding Sequence Amplification, Template Regeneration, and Re-Translation

Iterative cycles of *in vitro* selection and amplification (SELEX) can allow extremely rare but highly functional biopolymers to emerge from a sequence-diverse library.[39-41] In order for sequence-defined synthetic polymers to undergo iterative *in vitro* selection cycles, the DNA template encoding these polymers must be amplified and re-translated. We designed a DNA template that can be amplified using two modified primers (Figure 6). One of the primer contains a 5' hairpin loop that installs a 5' hexylnyl group for coupling with the synthetic polymer, while the other primer is equipped with a 5' biotin group to be used in a downstream strand-separation step.

After translation of the DNA template to a synthetic polymer that contains five densely functionalized β-peptide building blocks and a terminator using substrates **22, 23**, and **T7** (Figure 5c), the full-length DNA-synthetic polymer conjugate was purified by PAGE (Figure 6, lanes 2 and 3). A minute fraction ($1/50{,}000^{th}$) of the resulting material, simulating the amount that might survive an *in vitro* selection, was subjected to PCR amplification (Figure 6, lane 4). The sense (template) strand of the resulting double-stranded PCR product was isolated by capturing the anti-sense strand using immobilized streptavidin (Figure 6, lane 5). The resulting single-stranded DNA was subjected to re-translation using the same substrates to generate translation products that are indistinguishable by PAGE electrophoresis from the translation products generated by the starting template (Figure 6, lane 6). DNA sequencing of the antisense strand of the PCR product further confirmed that sequence of the regenerated DNA template was identical to that of the starting template (Supplementary Results Figure S13). These results collectively demonstrate that the DNA-templated translation strategy developed here can support a complete cycle of translation, template amplification, and template regeneration and therefore provides a foundation for the *in vitro* selection of sequence-defined synthetic polymers.

Discussion

We have designed and implemented a DNA-templated translation system capable of generating sequence-defined synthetic polymers that have no necessary structural similarity

to nucleic acids and that do not need to directly hybridize to DNA or RNA. We identified the CuAAC reaction and the AA/BB substrate architecture as key factors to achieve efficient translation. The use of a unique terminator substrate facilitated the analysis of the sequence specificity of the process and also enabled the identification of full-length translation products when coupled with mass spectrometry. This system can be used to generate synthetic polymers containing diverse backbone structures including PEG, β-peptides, and α-(D)-peptides. Building blocks containing different backbones (such as **6**+**14**) can be co-polymerized sequence-specifically, and full-length products as large as 26 kDa (arising from 16 consecutive substrate couplings and containing 90 side chain-functionalized β-amino acid residues) were generated using this strategy. Together, these features enabled a complete cycle of translation, template amplification, template regeneration, and re-translation of a densely functionalized sequence-defined β-peptide, laying the foundation for the discovery of functional non-nucleic acid synthetic polymers through iterated *in vitro* selection.

Efficiency and fidelity are essential features of any translation strategy, and in this system arise from at least three design considerations. First, the substrates are organized and subsequently polymerized in a DNA-templated manner. This approach transfers the sequence specificity of Watson-Crick base pairing into the ordering of synthetic polymer building blocks, and also enables reactive groups of adjacent substrates to obtain effective molarities sufficient to drive efficient product formation. In contrast, mixing the substrates in the presence of catalyst but in the absence of a matched DNA template inefficiently generates higher molecular weight products in a non-specific manner (Supplementary Results Figure S10). Second, the macrocyclic nature of the substrate significantly enhance their reactivity by aligning synthetic polymer building blocks in close proximity on the same side of the DNA double helix. The AA/BB architecture of the substrates further minimize undesired reactions including cyclization.[42] Finally, the CuAAC reaction is highly efficient and tolerant of wide variety of functional groups in neutral aqueous solution, and the triazole linkage formed by this reaction is known to be stable under these conditions.[30,43]

While this translation system does not rely on cells or enzymes and therefore is liberated from the structural requirements associated with ribosomal or polymerase-mediated polymerization, achieving these performance characteristics in the absence of enzymes required that substrates meet a different and substantial set of criteria. The substrates used in this work are macrocyclic, a feature that is necessary to reduce the entropic cost of desired polymerization pathways by limiting conformations unfavorable to reaction (Figure 2b, lanes 16 versus lanes 20-22).[26] Moreover, unlike ribosomal translation in which three-base codon:anticodon hybridization templates the polymerization of single amino acids, this system requires PNA adapters that each contain five nucleobases to enable efficient and sequence-specific hybridization. To ensure hybridization at 4 °C, the T_m of substrates **6** and **12** hybridized to a DNA template is 22.4 °C in the reaction buffer used for polymerization (Supplementary Results Figure S12), compared to a T_m < 10 °C of a macrocyclic substrate with a PNA adapter sequence containing only four nucleobases.

Sequence-defined polymers in the molecular weight range already achieved by this system include many naturally occurring proteins with remarkable binding and catalytic activities. That said, the extent to which the above requirements limit the functional potential of the

resulting synthetic polymers remains to be seen, and it may be necessary to explore additional substrate structures, coupling chemistries, or coding schemes in order to enable the successful evolution of synthetic polymers with desirable binding or catalytic properties. Towards this end, the highly modular nature of the substrates used in this work facilitates the modification of each substrate component including the PNA adapter, synthetic polymer building block, and linkers.

The translation system developed here has the potential to enable the laboratory evolution of a wide range of synthetic polymers and to reveal the evolutionary potential of macromolecules beyond the reach of previous translation systems (Figure 1b). For example, the sequence-specific polymerization of β-peptides that are predisposed to form secondary structures may enable the evolution of **β**-proteins[44-46] with novel structures and functions including protease resistance, [47] improved cell penetration,[48] and antibiotic activity.[49]

Methods

For complete experimental methods see the Supplementary Information.

Synthesis of substrate 6

The synthesis of the linear precursor of **6** was performed on 10 μmol scale on a Protein Technologies Tribute-UV automated peptide synthesizer using Fmoc chemistry, with DMF as the solvent and 20% piperidine in DMF as the deprotection reagent. Peptide couplings were performed on Rink amide low-loading resin using 5 eq. Fmoc-protected amino acid activated with 4.5 eq. HATU and 10 eq. *N*-methylmorpholine. The sequence of the linear precursor is NH_2-Lk-*G*-*G*-*A*-*T*-*T*-Lk-Lys(N_3)-PEG_{16}-Glu(ODmab)-Lys(N_3)-$CONH_2$, where Lk is the disulfide linker; Lys(N3) is ε-azido lysine; PEG_{16} is coupled using Fmoc-NH-PEG_{16}-COOH as one residue; Glu(ODmab) is coupled using Fmoc-NH-Glu-ODmab; and italicized letters represent PNA nucleotides. The resulting resin was placed in a glass peptide reaction vessel, and 2 % hydrazine in DMF (2 mL × 5) was added to deprotect the Dmab group. The resin was washed with DMF (2 mL × 5), 5 % DIPEA in DMF (2 mL × 2), and DMF again (2 mL). The linear precursor was cyclized on resin using diisopropylcarbodiimide (DIC, 20 μL, 125 μmol) and hydroxybenzotriazole (HOBt, 16.2 mg, 120 μmol) in 3 mL DMF for 48 h at room temperature. The product was cleaved from the resin by treatment of a mixture (2 mL × 2) of 95 % trifluoroacetic acid (TFA), 2.5 % water, and 2.5 % triisopropylsilane (TIS). The crude macrocycle was precipitated in diethyl ether and purified by C18 reverse-phase HPLC using 0.1 % TFA in water and 0.1 % TFA in acetonitrile in a linear gradient from 10 % to 40 % acetonitrile as the mobile phase. HPLC fractions were characterized by MALDI mass spectrometry and fractions containing pure macrocycle were combined and lyophilized.

DNA-templated oligomerization

Translation reactions were prepared by combining 40 μL of degassed 0.1 M aqueous HEPES pH 8.0 buffer, 1 μL of 4 M NaCl, 10 pmol of DNA template, 120 pmole of azide substrate (4 eq. per template codon), 80 pmol of alkyne substrate (4 eq. per template codon), 40 pmol of the terminator substrate, 0.5 μmol tris(hydroxypropyltriazolyl)amine, THPTA[50] and water to

a total volume of 50 μL. The reaction was heated to 95 °C and slowly cooled to 5 °C. The CuAAC reaction was initiated by the addition of 1 μL 50 mM $CuSO_4$, followed by the addition of 1 μL 0.5 M sodium ascorbate. The reaction mixture was maintained at 4 °C overnight. After incubation, the reaction was desalted using a Sephadex minicolumn (Princeton Separations) and analyzed by 10 % denaturing PAGE. The gel was stained with ethidium bromide and analyzed by UV illumination and densitometry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Howard Hughes Medical Institute and the NIH/NIGMS (R01GM065865). J.N. was partially supported by an Eli Lilly Organic Chemistry Graduate Fellowship. R.H. was supported by a postdoctoral fellowship from Canada's National Science and Engineering Council (NSERC). We are grateful to Jennifer Heemstra, Yevgeny Brudno, Christoph Dumelin, Yanyan Lu, and Sunia Trauger for helpful discussions.

References

1. Badi N, Lutz J-F. Sequence control in polymer synthesis. Chem Soc Rev. 2009; 38:3383–90. [PubMed: 20449057]
2. Lutz J-F. Polymer chemistry: a controlled sequence of events. Nature Chem. 2010; 2:84–5. [PubMed: 21124395]
3. Pfeifer S, Lutz J-F. A facile procedure for controlling monomer sequence distribution in radical chain polymerizations. J Am Chem Soc. 2007; 129:9542–3. [PubMed: 17636902]
4. Nakatani K, Ogura Y, Koda Y, Terashima T, Sawamoto M. Sequence-regulated copolymers via tandem catalysis of living radical polymerization and in situ transesterification. J Am Chem Soc. 2012; 134:4373–83. [PubMed: 22296320]
5. Soeriyadi AH, Boyer C, Nyström F, Zetterlund PB, Whittaker MR. High-order multiblock copolymers via iterative Cu(0)-mediated radical polymerizations (SET-LRP): toward biological precision. J Am Chem Soc. 2011; 133:11128–31. [PubMed: 21707082]
6. Lo PK, Sleiman HF. Nucleobase-templated polymerization: copying the chain length and polydispersity of living polymers into conjugated polymers. J Am Chem Soc. 2009; 131:4182–3. [PubMed: 19275231]
7. McHale R, Patterson JP, Zetterlund PB, Reilly RKO. Biomimetic radical polymerization via cooperative assembly of segregating templates. Nature Chem. 2012; 4:491–7. [PubMed: 22614385]
8. Belowich ME, et al. Positive Cooperativity in the Template-Directed Synthesis of Monodisperse Macromolecules. J Am Chem Soc. 2012; 134:5243–61. [PubMed: 22303894]
9. Lutz J-F. Sequence-controlled polymerizations: the next Holy Grail in polymer science? Poly Chem. 2010; 1:55.
10. Leitzel JC, Lynn DG. Template-directed ligation: from DNA towards different versatile templates. Chem Rec. 2001; 1:53–62. [PubMed: 11893058]
11. Ura Y, Beierle JM, Leman LJ, Orgel LE, Ghadiri MR. Self-assembling sequence-adaptive peptide nucleic acids. Science. 2009; 325:73–7. [PubMed: 19520909]
12. Li X, Zhan Z-YJ, Knipe R, Lynn DG. DNA-catalyzed polymerization. J Am Chem Soc. 2002; 124:746–7. [PubMed: 11817938]
13. Schrum JP, Ricardo A, Krishnamurthy M, Blain JC, Szostak JW. Efficient and rapid template-directed nucleic acid copying using 2'-Amino-2′3′-dideoxyribonucleoside--5′-phosphorimidazolide monomers. J Am Chem Soc. 2009:552–562.
14. Henry A, Romesberg FE. The evolution of DNA polymerases with novel activities. Curr Opin Biotechnol. 2005; 16:370–7. [PubMed: 16006114]

15. Josephson K, Hartman MCT, Szostak JW. Ribosomal synthesis of unnatural peptides. J Am Chem Soc. 2005; 127:11727–35. [PubMed: 16104750]
16. Datta B, Schuster GB. DNA-directed synthesis of aniline and 4-aminobiphenyl oligomers: programmed transfer of sequence information to a conjoined polymer nanowire. J Am Chem Soc. 2008; 130:2965–73. [PubMed: 18266359]
17. Rosenbaum DM, Liu DR. Efficient and sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes. J Am Chem Soc. 2003; 125:13924–5. [PubMed: 14611205]
18. Kleiner RE, Brudno Y, Birnbaum ME, Liu DR. DNA-templated polymerization of side-chain-functionalized peptide nucleic acid aldehydes. J Am Chem Soc. 2008; 130:4646–59. [PubMed: 18341334]
19. Peng X, Li H, Seidman M. A Template-Mediated Click-Click Reaction: PNA-DNA, PNA-PNA (or Peptide) Ligation, and Single Nucleotide Discrimination. Eur J Org Chem. 2010; 2010:4194–4197.
20. Chen W, Schuster GB. DNA-Programmed modular assembly of cyclic and linear nanoarrays for the synthesis of two-dimensional conducting polymers. J Am Chem Soc. 2012; 134:840–3. [PubMed: 22242713]
21. He Y, Liu DR. Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. Nature Nanotech. 2010; 5:778–82.
22. Mckee ML, et al. Programmable one-Pot multistep organic synthesis using DNA junctions. J Am Chem Soc. 2011:1146–9.
23. Brudno Y, Birnbaum ME, Kleiner RE, Liu DR. An in vitro translation, selection and amplification system for peptide nucleic acids. Nature Chem Biol. 2010; 6:148–55. [PubMed: 20081830]
24. Yu H, Zhang S, Chaput JC. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. Nature Chem. 2012; 4:183–7. [PubMed: 22354431]
25. Pinheiro VB, et al. Synthetic Genetic Polymers Capable of Heredity and Evolution. Science. 2012; 336:341–344. [PubMed: 22517858]
26. Driggers EM, Hale SP, Lee J, Terrett NK. The exploration of macrocycles for drug discovery--an underexploited structural class. Nat Rev Drug Discov. 2008; 7:608–24. [PubMed: 18591981]
27. Li X, Liu DR. DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. Angew Chem Int Ed. 2004; 43:4848–70.
28. Hud NV, Jain SS, Li X, Lynn DG. Addressing the problems of base pairing and strand cyclization in template-directed synthesis: a case for the utility and necessity of "molecular midwives" and reversible backbone linkages for the origin of proto-RNA. Chem Biodiv. 2007; 4:768–83.
29. Winnik A. End-to-End Cyclization of Polymer Chains. Acc Chem Res. 1985; 18:73–79.
30. Kolb HC, Finn MG, Sharpless KB. Click chemistry: diverse chemical function from a few Good reactions. Angew Chem Int Ed. 2001; 40:2004–2021.
31. Lallana E, Riguera R, Fernandez-Megia E. Reliable and Efficient Procedures for the Conjugation of Biomolecules through Huisgen Azide-Alkyne Cycloadditions. Angew Chem Int Ed. 2011; 50:8794–8804.
32. Schmidt JG, Christensen L, Nielsen PE, Orgel LE. Information transfer from DNA to peptide nucleic acids by template-directed syntheses. Nucleic acids research. 1997; 25:4792–6. [PubMed: 9365258]
33. Boehler C, Nielsen PE, Orgel LE, Böhler C. Template switching between PNA and RNA oligonucleotides. Nature. 1995; 376:578–581. [PubMed: 7543656]
34. Goodwin JT, Lynn DG. Template-directed synthesis: use of a reversible reaction. Chem Commun. 1992:9197–8.
35. Gartner ZJ, Kanan MW, Liu DR. Expanding the reaction scope of DNA-templated synthesis. Angew Chem Int Ed. 2002; 783:1796–1800.
36. Cho BR, Cho NS, Song KS, Son KN, Kim YK. Elimination reactions of (E) -2, 4-dinitrobenzaldehyde O-aryloximes promoted by RO - / ROH buffers in EtOH. J Org Chem. 1998; 2:3006–9.

37. Kumar R, et al. Template-directed oligonucleotide strand ligation, covalent intramolecular DNA circularization and catenation using click chemistry. J Am Chem Soc. 2007; 129:6859–64. [PubMed: 17488075]

38. Chouikhi D, Barluenga S, Winssinger N. Clickable peptide nucleic acids (cPNA) with tunable affinity. Chem Commun. 2010; 46:5476–8.

39. Wright MC, Joyce GF. Continuous in Vitro Evolution of Catalytic Function. Science. 1997; 276:614–617. [PubMed: 9110984]

40. Joyce GF. Nonenzymatic Template-directed Synthesis of Informational Macromolecules. Cold Spring Harbor Symp Quant Biol. 1987; 52:41–51. [PubMed: 2456883]

41. Daguer JP, et al. DNA-templated combinatorial assembly of small molecule fragments amenable to selection/amplification cycles. Chem Sci. 2011; 2:625.

42. Andersen CS, Yan H, Gothelf KV. Bridging one helical turn in double-stranded DNA by templated dimerization of molecular rods. Angew Chem Int Ed. 2008; 47:5569–72.

43. El-Sagheer AH, Brown T. Click chemistry with DNA. Chem Soc Rev. 2010; 39:1388–405. [PubMed: 20309492]

44. Petersson EJ, Schepartz A. Toward beta-amino acid proteins: design, synthesis, and characterization of a fifteen kilodalton beta-peptide tetramer. J Am Chem Soc. 2008; 130:821–3. [PubMed: 18166055]

45. Goodman CM, Choi S, Shandler S, DeGrado WF. Foldamers as versatile frameworks for the design and evolution of function. Nature Chem Biol. 2007; 3:252–62. [PubMed: 17438550]

46. Martinek, Ta; Fulop, F. Side-chain control of beta-peptide secondary structures. Design principles. Eur J Biochem. 2003; 270:3657–3666. [PubMed: 12950249]

47. Frackenpohl J, Arvidsson PI, Schreiber JV, Seebach D. The outstanding biological stability of beta- and gamma-peptides toward proteolytic enzymes: an in vitro investigation with fifteen peptidases. Chembiochem. 2001; 2:445–55. [PubMed: 11828476]

48. Bautista AD, Appelbaum JS, Craig CJ, Michel J, Schepartz A. Bridged beta(3)-peptide inhibitors of p53-hDM2 complexation: correlation between affinity and cell permeability. J Am Chem Soc. 2010; 132:2904–6. [PubMed: 20158215]

49. Olsen C, et al. Antimicrobial, hemolytic, and cytotoxic activities of beta-peptoid-peptide hybrid oligomers: improved properties compared to natural AMPs. Chembiochem. 2010; 11:1356–60. [PubMed: 20503219]

50. Chan TR, Hilgraf R, Sharpless KB, Fokin VV. Polytriazoles as copper(I)-stabilizing ligands in catalysis. Org Lett. 2004; 6:2853–5. [PubMed: 15330631]

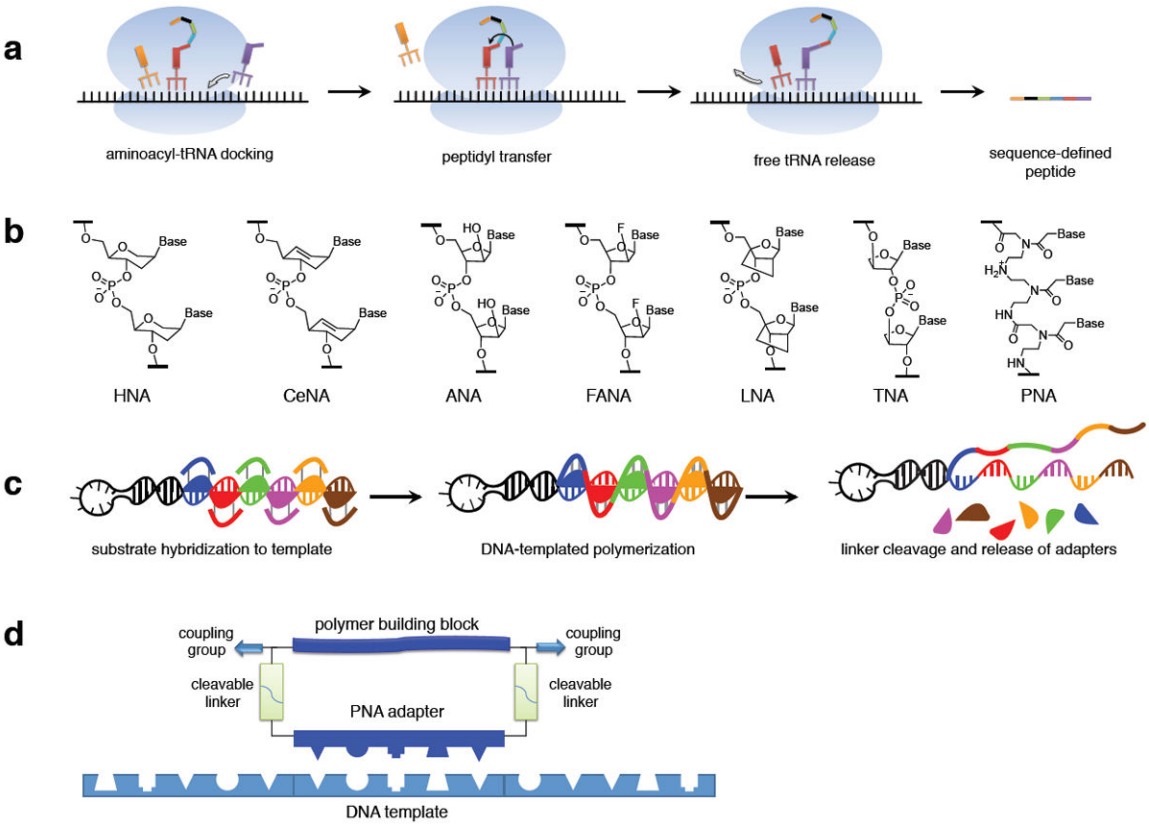

Figure 1. Natural and laboratory translation of nucleic acids into non-nucleic acid polymers

(a) In living systems, mRNA-templated, tRNA-mediated amine acylation catalyzed by the ribosome translates transcripts into sequence-defined proteins. (b) Synthetic polymers with non-natural backbones that can be translated from nucleic acid templates by current strategies are all necessarily analogs of DNA and RNA that retain their ability to base pair with templates. (c) The enzyme-free, DNA-templated polymerization strategy developed in this work translates DNA templates into sequence-defined non-nucleic acid polymers. Macrocyclic substrates hybridize with codons on a DNA template, organizing synthetic polymer building blocks along the template. Coupling reactions then oligomerize these pre-organized substrates. Finally, linker cleavage releases the PNA adapters and liberates the synthetic polymer product. (d) Representation of a macrocyclic substrate for the translation system in (c).

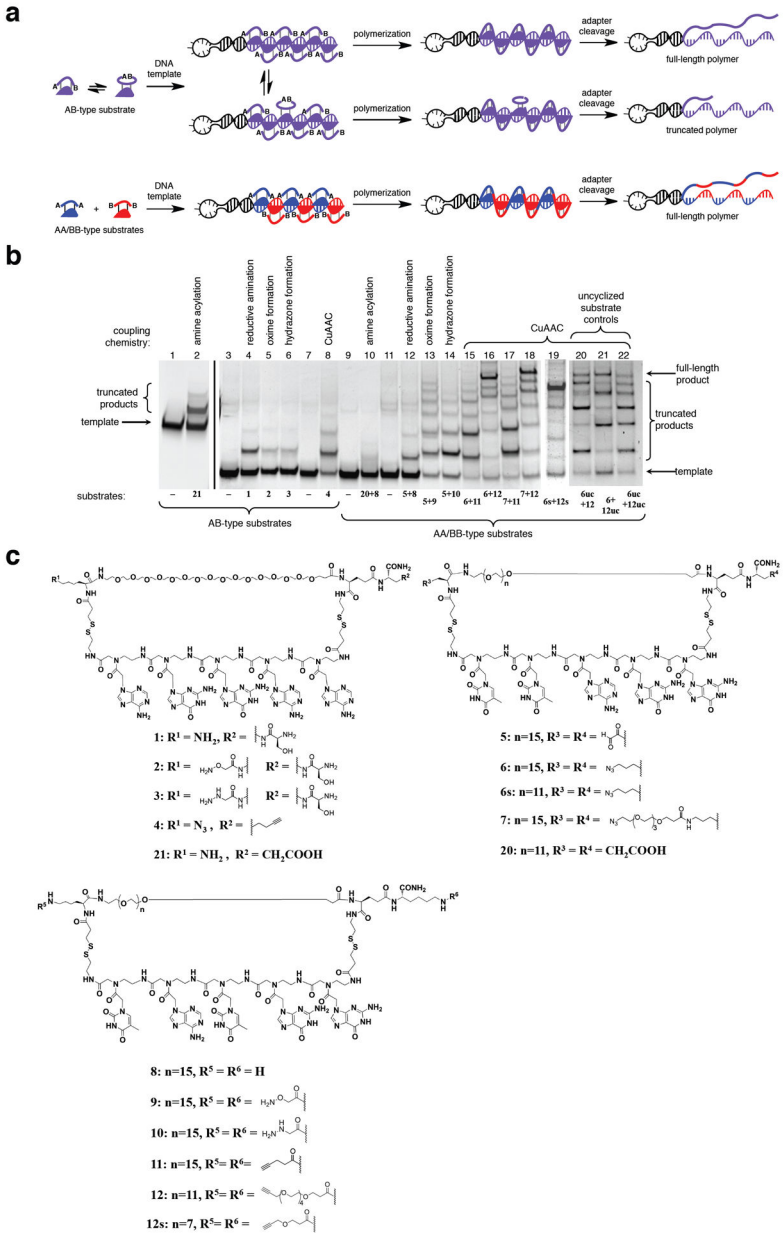

Figure 2. Evaluation of building block coupling chemistries

(a) Translation using AB-type and AA/BB-type substrates. While both types of substrates can participate in the translation process, AB-type substrates are prone to intramolecular cyclization. (b) Denaturing PAGE analysis of translation reactions to evaluate five candidate coupling chemistries. The templates used in lanes 1 and 2 contain eight codons, while all other templates contain six codons. The denaturing 10 % PAGE-urea gel was electrophoresed in TBE buffer at 200 V for 1 hour at room temperature, then stained with ethidium bromide. AA/BB-type substrates coupled by CuAAC offered the highest overall yields of full-length translation products (lanes 16, 18, and 19). In contrast, AB-type substrates failed to yield full-length translation products (lanes 1-8), potentially due to competitive intramolecular cyclization. Notably, the macrocyclic nature of the substrate is

required for efficient translation, as non-cyclized substrates showed significantly impaired oligomerization (lanes 20-22). (c) Structures of substrates **1-12, 6s, 12s, 20**, and **21**. See the Supplementary Methods for the structures of **6-uc** and **12-uc** and reaction conditions.

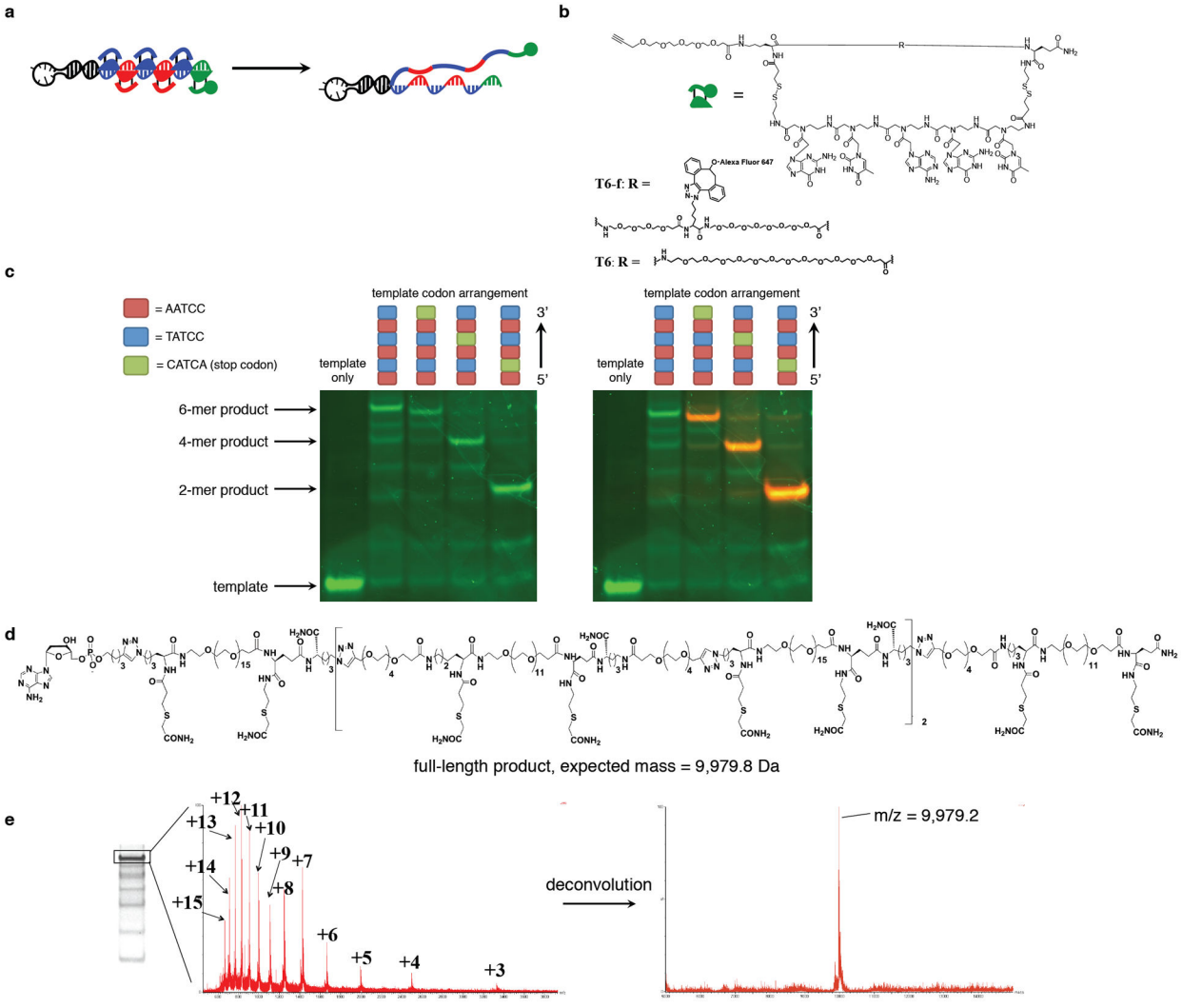

Figure 3. Identification and analysis of full-length translation products

(a) A fluorescent "terminator" substrate was used to detect full-length translation products. (b) The structure of terminator substrates **T6-f** and **T6**. (c) Dual-channel fluorescent image of a denaturing PAGE analyzing translation reactions containing **T6-f**. Single-stranded DNA was stained with SYBR gold. The colored rectangles above each lane indicate the arrangement of the codons on the template used in that experiment. The left image shows SYBR gold fluorescence. The right image shows both SYBR gold (green) and Alexa Fluor 647 fluorescence (red). (d) Structure of the full-length product with **T6** incorporated as the terminator after P1nuclease digestion. (e) Mass spectrometric characterization of the gel-purified full-length product. The left spectrum shows the original multiple charged states of a single molecular species in the ESI analysis, with the numbers indicating the charge of each ion; the right spectrum is deconvoluted from the left spectrum.

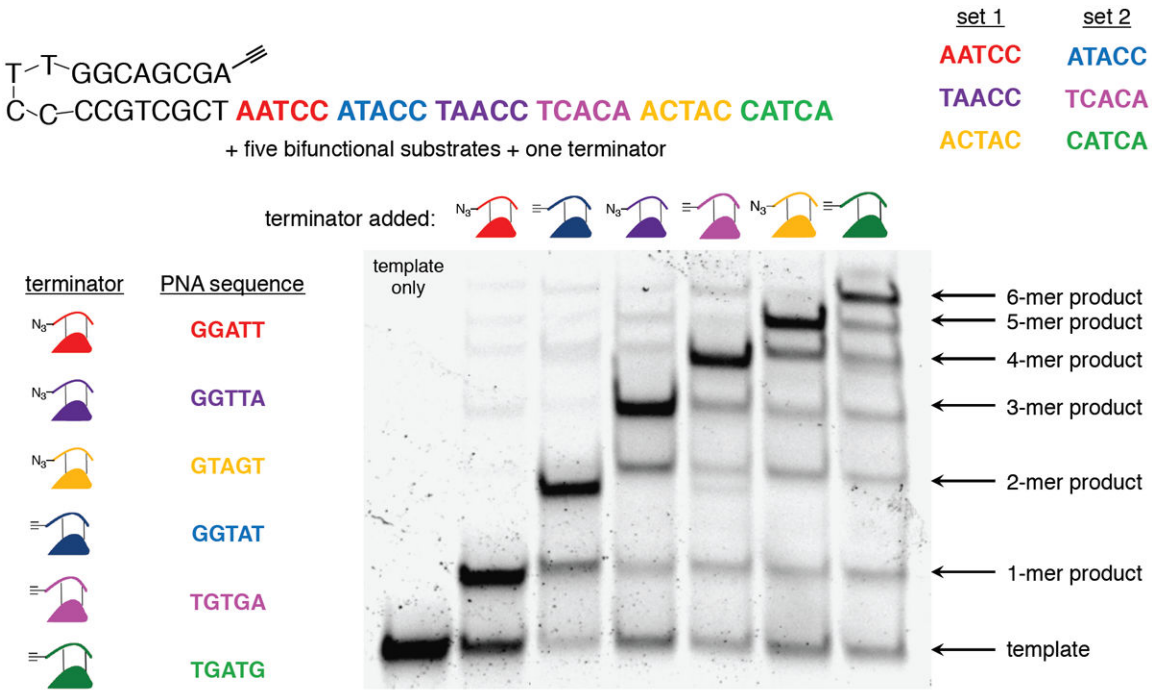

Figure 4. Sequence specificity of translation using templates containing mixtures of codons

Each template contains a coding region with six different codons from two codon sets. Every codon from one set is followed by a codon from the other set. Each translation reaction contained five bifunctional AA/BB-type substrates and one terminator substrate. Because incorporation of the terminator ends oligomerization at the codon position in which it is incorporated, polymer lengths reflect the sequence specificity of the translation process.

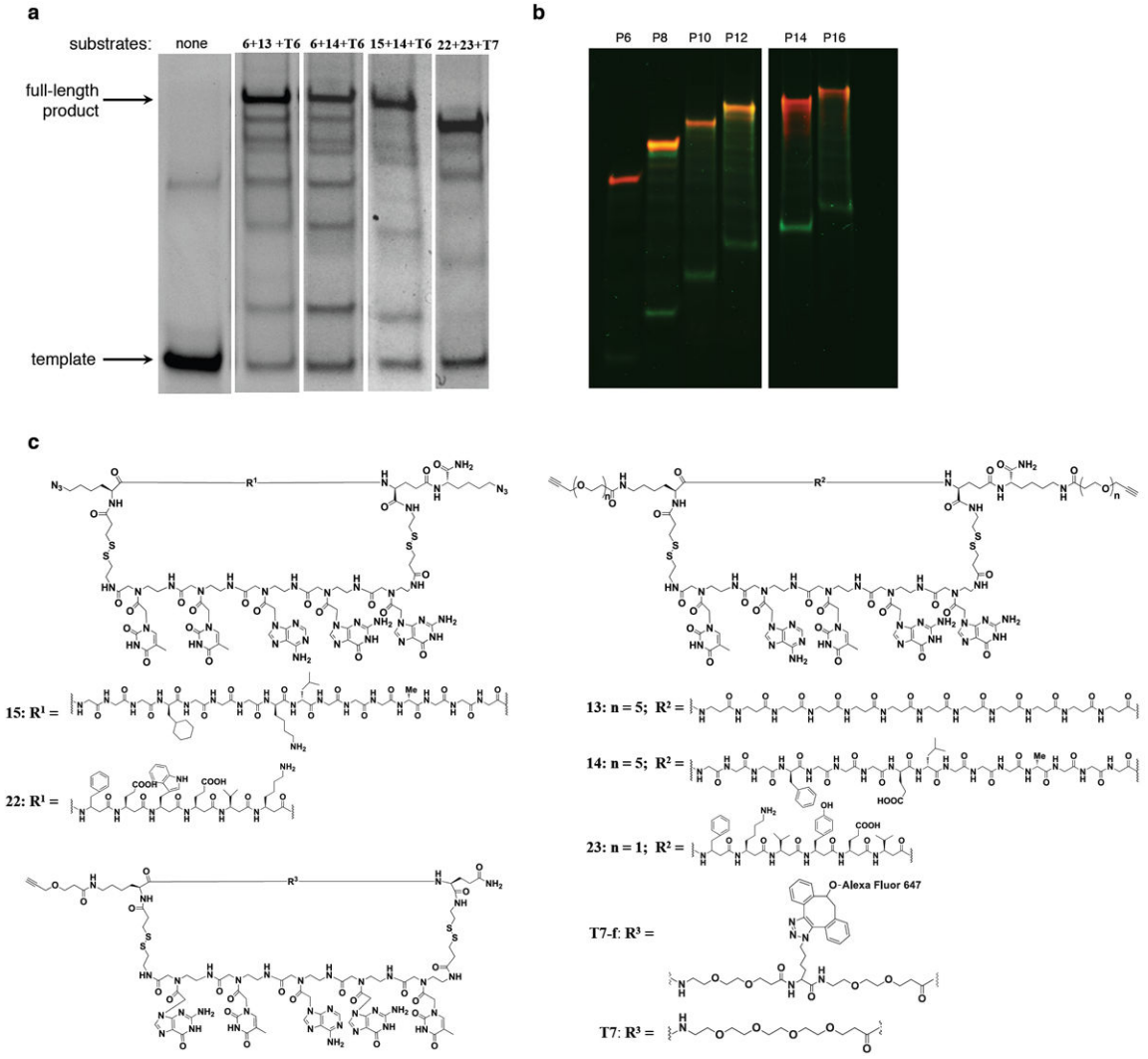

Figure 5. Translation of DNA sequences into longer polymers and polymers of different backbone structures

(a) Denaturing PAGE analysis of PEG substrate **6** co-polymerizing with β-peptide substrate **13**, PEG substrate **6** co-oligomerizing with α-(D)-peptide substrate **14**, substrate **14** co-oligomerizing with α-(D)-peptide substrate **15**, and densely functionalized β-peptide substrate **22** co-oligomerizing with densely functionalized β-peptide substrate **23**. The DNA template used in these experiments contained six codons, ending with the "stop" codon that recruits the terminator substrate **T6** or **T7**. (b) Translation of DNA into longer non-nucleic acid polymers. Translation reactions of DNA templates containing six, eight, ten, 12, 14, and 16 pentamer codons were performed using substrates **22** and **23**. The last codon of each template encoded the incorporation of the Alexa Fluor 647-linked terminator substrate **T7-f**. The denaturing PAGE was stained with SYBR gold. The green channel shows SYBR gold fluorescence and the red channel shows Alexa Fluor 647 fluorescence. Lanes are marked P6, P8, etc. reflecting the number of the codons in the template. All lanes contained 10 pmol template for translation reaction. (c) Structures of substrates **13, 14, 15, 22, 23**, and terminators **T7-f** and **T7**.

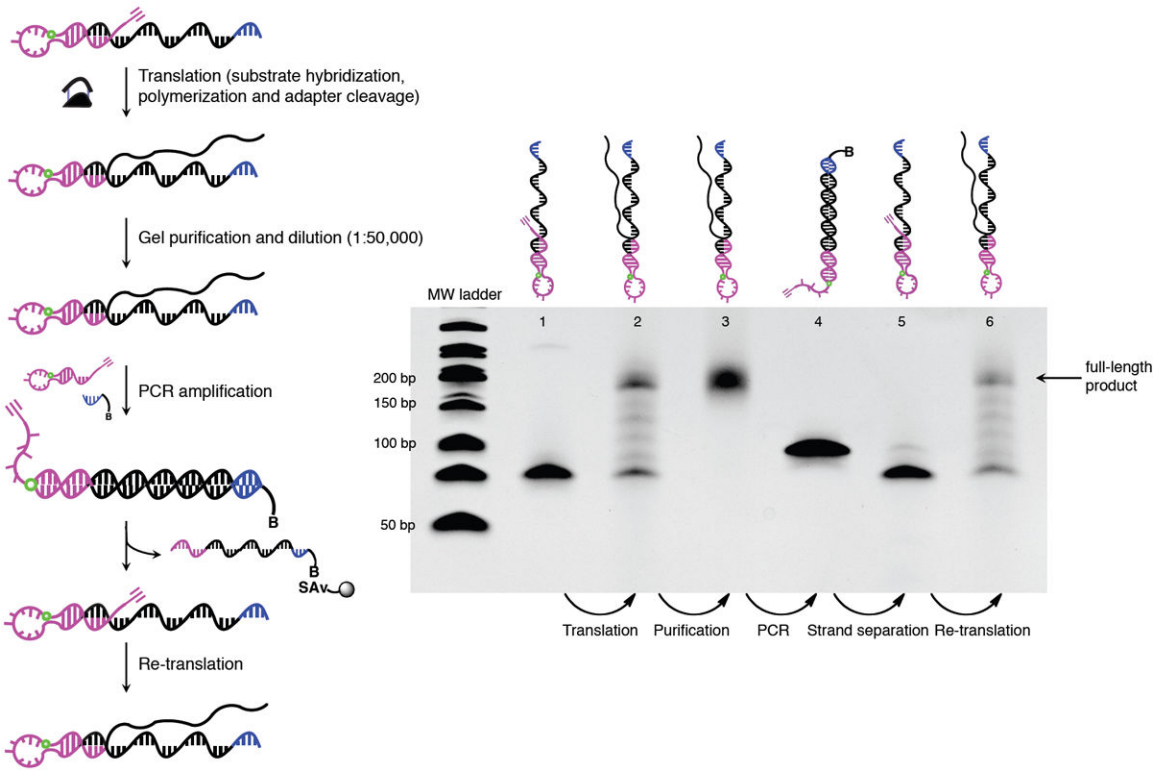

Figure 6. A complete cycle of translation, PCR amplification, strand separation, and re-translation

100 pmol starting template was translated into synthetic β-peptide polymer using substrates **22, 23**, and **T7** (lanes 1 and 2). The translation reaction was purified by denaturing PAGE (lane 3). A small amount (1/50,000th) of the translation product, simulating the amount surviving a typical *in vitro* selection, was subjected to PCR with two primers, generating a double-stranded template (lane 4). The PCR product was immobilized on streptavidin-linked beads and the "sense" strand was eluted with 150 mM NaOH (lane 5). Finally, this single-stranded template was re-translated into synthetic polymer using the same substrates (lane 6). The non-denaturing PAGE gel shown was visualized by ethidium bromide staining. B = biotin; SAv = streptavidin. We note that because the translation product includes a non-nucleic acid synthetic polymer, PAGE mobilities of the DNA template-synthetic polymer conjugates in lanes 2, 3, and 6 are not directly relatable to the molecular weight ladder.