Enzymatic synthesis of diverse oligoamide polymer precursors

Liangyu Qian[1,*], Isaiah T. Dishner[2,*], Dana L. Carper[1], Vilmos Kertesz[1], Nicholas T. Zolnierczuk[1], Nikki A. Thiele[2], John F. Cahill[1], Jeffrey C. Foster[2,†], Joshua K. Michener[1,†]

[1]Biosciences and [2]Chemical Sciences Divisions, Oak Ridge National Laboratory, Oak Ridge, TN

[*]These authors contributed equally

[†]To whom correspondence should be addressed: fosterjc@ornl.gov, michenerjk@ornl.gov

Notice: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

Abstract:
Polyamides are important natural and synthetic polymers, best exemplified by proteins and nylons respectively. Proteins demonstrate that novel polymers with emergent properties can be generated by combining diverse monomers in precisely defined sequences. However, commercial polyamides only represent a small fraction of the potential diversity in polyamide sequences. The relative simplicity of nylon sequences derives from the challenges of controlling the incorporation of complex monomer mixtures during polymerization. In this study, we demonstrate that amide synthetases can efficiently catalyze the synthesis of sequence-defined oligoamides, ligating a broad array of commercially-relevant dicarboxylic acids with diamines and omega-amino acids at molar yields up to 80% without the need for protecting groups. These oligoamides can be chemically polymerized to yield polyamides with higher molecular weights than can be obtained from traditional nylon salts. Using a cofactor regeneration system, oligoamide synthesis was scaled to produce sufficient quantities of material for laboratory-scale characterization and solid state polymerization. This study establishes a direct biocatalytic route for the facile synthesis of sequenced oligoamides from unprotected bifunctional substrates, opening new possibilities to synthesize sequenced copolyamides.

Introduction

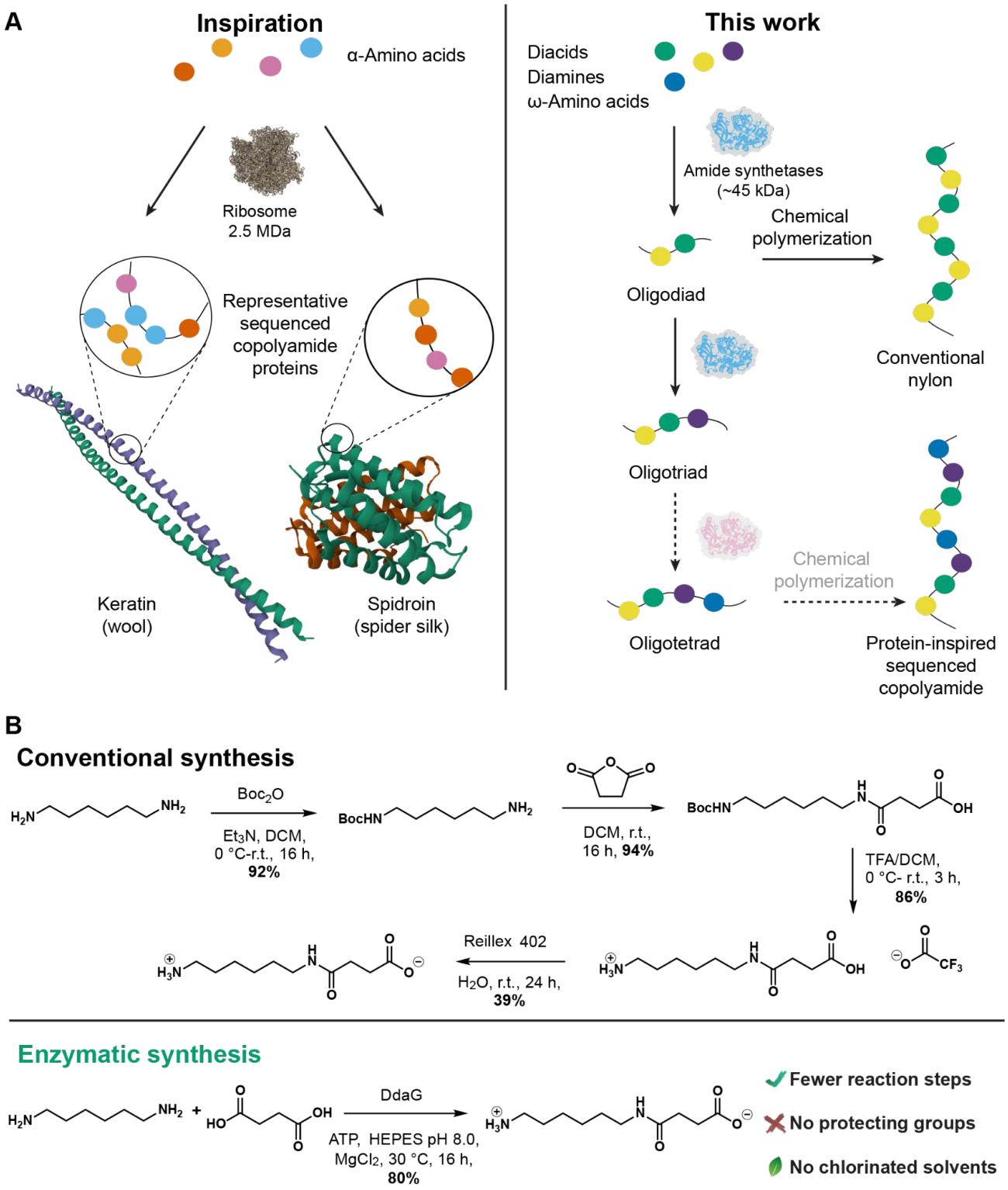
Polyamides are polymers linked by amide groups and synthesized by the reaction of diamines with dicarboxylic acids, by the condensation of ω-amino acids, or by ring-opening polymerization of cyclic lactams. Nylons, a prominent category of synthetic thermoplastic polyamides, were produced at an estimated scale of 9 million tons in 2020 (Global Industry Analysts, 2020). Major commercial applications of nylons include fibers for carpets and fabrics, filaments for fishing nets or additive manufacturing, and automotive parts. A variety of nylon backbones are commercially available, providing different properties based on the repeat units of the polymer. Practical synthetic considerations limit commercial nylons to either homopolymers of a single ω-amino acid (e.g. nylon 6), combinations of one diacid and one diamine (e.g. nylon 66), or block copolymers of the same. More complex repeat units could yield polyamide materials with novel properties, but preparing the prerequisite oligoamides necessary to produce sequenced copolyamides is synthetically laborious, requiring multiple protection and deprotection steps proportional to the number of coupled units.

The potential impact of polymer sequence control is best exemplified by the diverse properties of proteins (Figure 1A). Proteins are natural polyamides synthesized by ribosomes and primarily composed of twenty natural α-amino acids connected by peptide (amide) bonds and arranged in specific sequences. The resulting polyamides demonstrate emergent properties based on structured interactions between sequenced monomers. For example, natural materials such as wool and silk are assembled from proteins,[1] as are enzyme biocatalysts.[2] However, while ribosomes can synthesize polyamides with high molecular weight and exquisite sequence control, expanding the monomer scope beyond canonical α-amino acids is challenging.[3–9] Alternatively, using enzymes to selectively couple nylon-like substrates could lead to polymeric materials with novel sequences and properties.

One potential biosynthetic route for sequenced polyamides uses non-ribosomal peptide synthetases (NRPSs). These multimodular enzymes offer greater flexibility in monomer scope and bond formation but are difficult to express and engineer.[10–12] Finally, soluble amide synthetases have previously been used to synthesize amide bonds, though primarily in small molecules such as pharmaceuticals rather than polymers.[13–15] These enzymes use ATP to adenylate a carboxylic acid, which is then condensed with an amine to form an amide bond. Typically, amide synthetases catalyze amide bond formation during the synthesis of complex secondary metabolites such as antibiotics and siderophores.[16–18] They are relatively small (400-500 amino acids), easy to express and purify, and active in heterologous hosts. However, the known substrate scope of natural or engineered amide synthetases is limited.[15,18–20]

In this study, we demonstrated that amide synthetases can catalyze the assembly of oligoamides using a broad array of commercially-relevant diacids, diamines, and ω-amino acids (Figure 1B). These oligoamides can then be chemically polymerized to generate high-molecular-weight polyamides. By coupling amide bond formation with an ATP-regeneration system, we scaled up this biocatalytic process and established it as a viable method for synthesizing new nylon monomers. Facile enzymatic synthesis of new monomers will enable characterization of polyamide sequence-function relationships and the design of new polymers.

82
83
84


A
Inspiration
α-Amino acids
Ribosome
2.5 MDa
Representative
sequenced
copolyamide
proteins
Keratin
(wool)
Spidroin
(spider silk)
This work
Diacids
Diamines
ω-Amino acids
Amide synthetases
(~45 kDa)
Chemical
polymerization
Oligodiad
Conventional
nylon
Oligotriad
Chemical
polymerization
Oligotetrad
Protein-inspired
sequenced
copolyamide
B
Conventional synthesis
Boc2O
Et3N, DCM,
0 °C-r.t., 16 h,
92%
DCM, r.t.,
16 h, 94%
TFA/DCM,
0 °C- r.t., 3 h,
86%
Reillex 402
H2O, r.t., 24 h,
39%
Enzymatic synthesis
DdaG
ATP, HEPES pH 8.0,
MgCl2, 30 °C, 16 h,
80%
Fewer reaction steps
No protecting groups
No chlorinated solvents

85

Figure 1: Enzymatic synthesis of sequenced polyamides. A) The emergent properties of ribosomally-synthesized sequenced copolyamide proteins provides inspiration for the use of amide synthetases to synthesize nylon-relevant sequenced copolyamides. B) Comparison of chemical and enzymatic routes for synthesis of the hexamethylene diamine – succinic acid diad.

Results and Discussion

Amide synthetases can combine polymer-relevant monads

To explore the substrate range of amide synthetases, five enzymes (DdaG, SfaB, SuCphA1, McbA and PylC) were heterologously expressed in *E.coli* BL21(DE3). Among these, four enzymes (DdaG, SfaB, SuCphA1 and McbA) were successfully purified as soluble hexahistidine-tagged proteins (Figures S1-S2), while His-tagged PylC was insoluble. To enhance PylC solubility, a SUMO tag was added to the C-terminal of the hexahistidine tag, resulting in a soluble His-SUMO-tagged PylC (Figure S3). We first confirmed the reported activities of DdaG to ligate 2,3-diaminopropionate with fumarate,[17] SfaB to ligate 5-chlorovaleric acid with cadaverine,[19] and SuCphA1 to ligate aspartate with arginine.[21] However, the His-SUMO-tagged PylC failed to exhibit its native activity of ligating lysine and ornithine.[22] Attempts to cleave the His-SUMO tag from the recombinant PylC were unsuccessful. Therefore, we next conducted *in vitro* biochemical assays using the soluble enzymes (DdaG, SfaB, SuCphA1, and McbA) with a diverse range of substrates, including traditional nylon constituents (adipic acid **A**, hexamethylenediamine **M**, succinic acid **S**, *p*-xylylenediamine **X**, and terephthalic acid **T**) and novel compounds varying in chain length, functional groups (e.g., thiol, ketone, and hydroxyl groups), aromaticity, and cyclicity, including dicarboxylic acids, diamines, and ω-amino acids (Figure S4).

We initially detected product formation using direct injection immediate droplet-on-demand /open port sampling interface-mass spectrometry (IDOT/OPSI-MS) for untargeted characterization of enzymatic reactions.[23] Using this approach, 41 out of 96 reaction combinations generated detectable products above the control using either DdaG or SfaB (Figures 2 and S5). However, we did not observe detectable product formation in reactions with SuCphA1 or McbA, consistent with findings of a previous report that McbA does not accept aliphatic substrates.[15] To more accurately identify these products, we confirmed 24 structures through analysis of fragmentation patterns by MS^2 (Figures S6-S29). Both enzymes displayed activities in ligating the carboxylic groups of the diacids with a wide spectrum of amine acceptors including linear diamines (cadaverine **C**, hexamethylenediamine **M**), an aromatic diamine (*p*-xylylenediamine **X**), a cyclic diamine (*cis*-1,4-cyclohexanediamine **N**), and the amine groups of ω-amino acids (6-aminohexanoic acid **H**).

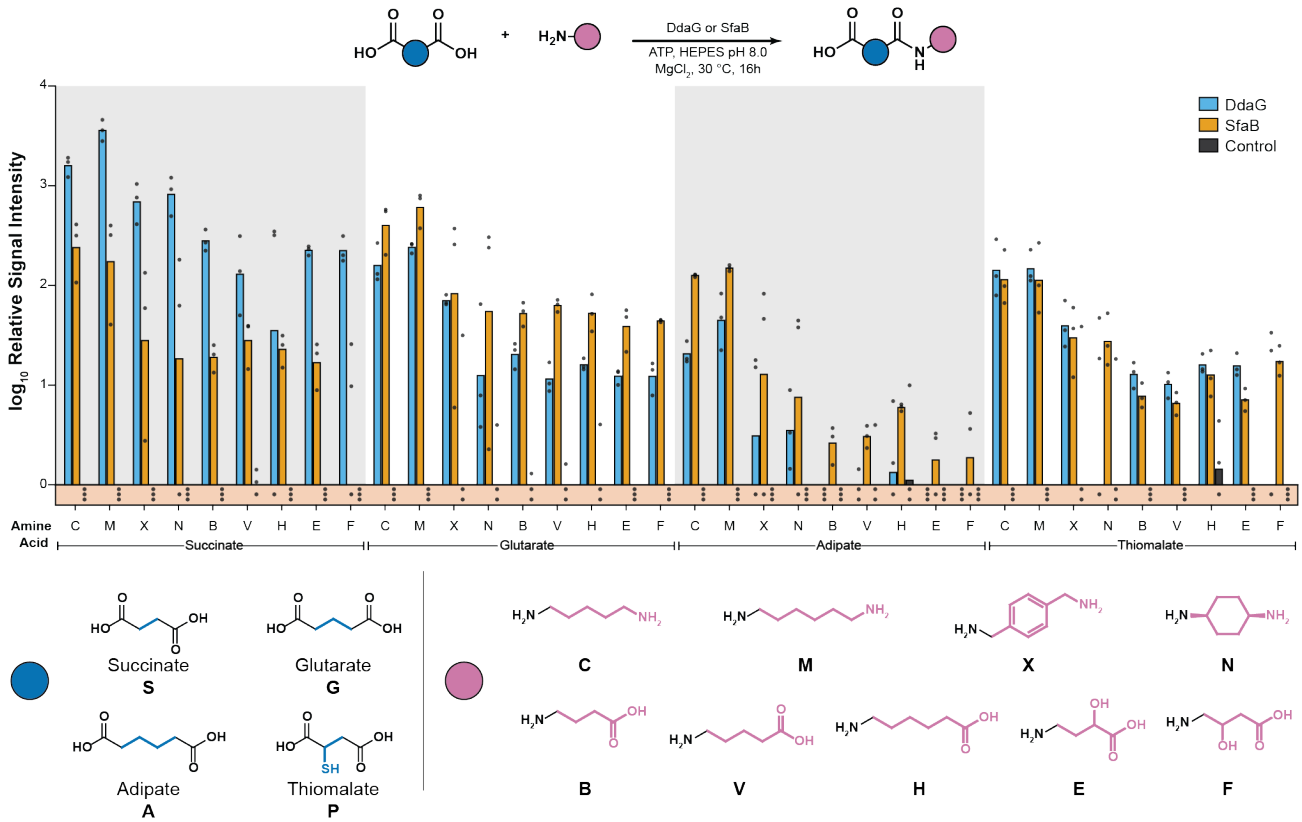


Figure 2: Amide synthetases generate polymer-relevant diads. *In vitro* biochemical assays were conducted by incubating different diacids (i.e., succinic acid, glutaric acid, adipic acid and thiomalic acid) with diamines or ω-amino acids in the presence of enzymes (i.e., DdaG or SfaB) or no-enzyme control. Products were assayed using I.DOT/OPSI-MS. The $\log_{10}$MS signal intensity was cut off at 5.0 to filter out noise. **C**: cadaverine; **M**: hexamethylenediamine; **X**: *p*-xylylenediamine; **N**: *cis*-1,4-cyclohexanediamine; **B**: 4-aminobutyrate; **V**: 5-aminovalerate; **H**: 6-aminohexanoic acid; **E**: 4-amino-2-hydroxybutanoic acid; **F**: 4-amino-3-hydroxybutanoic acid. Sample size is $N = 3$.

Notable differences between DdaG and SfaB were observed in their preference for carboxylic acid donors. DdaG exhibited excellent activity in ligating succinic acid (**S**) with a wide range of amine acceptors and showed lower activities with glutaric acid (**G**) and adipic acid (**A**). In contrast, SfaB displayed higher efficiency using **G** and **A** as carboxylic donors with diverse amine acceptors. In addition, DdaG can also catalyze reactions involving terephthalic acid (**T**) and α-ketoglutaric acid (**K**) with multiple diamines, while SfaB was inactive in these reactions. Alternatively, SfaB was capable of activating 1,4-cyclohexanedicarboxylic acid (**O**) with specific diamines (i.e., **M** and **X**), whereas DdaG showed negligible activity (Figure S5). It is also notable that both enzymes were unable to activate the carboxylic groups of ω-amino acids. Previous studies showed that SfaB cannot adenylate diacids for subsequent reactions with hydroxylamine, based on colorimetric detection of pyrophosphate released during adenylation. However, our observations reveal that SfaB can act on a range of diacids in combination with diamines and ω-amino acids. These differing results suggest that SfaB also exhibits specificity towards the amine acceptor. Interestingly, we noticed both enzymes could activate diacids bearing additional functional groups

such as thiomalic acid (**P**), suggesting the potential of using enzymes to produce novel value-added oligoamides.[24]

Given the distinct activities of these enzymes with diacids and their considerable tolerance for various diamines, we investigated the effect of diacid and diamine chain length on enzyme activity. We tested linear diacids and diamines with carbon lengths of 7 to 10, excluding C10 diamines due to solubility issues (Figure S30). The results for DdaG were consistent with our previous findings, showing a preference for **S** as the diacid and comparable activity across the full range of diamine lengths. In contrast, SfaB, although less active in ligating **S** with longer diamines, showed activity with C7-C10 diacids and **M**. We hypothesize that SfaB has a larger substrate-binding pocket to accommodate longer diacids.

To further characterize DdaG performance, we conducted kinetic studies using succinic acid and hexamethylenediamine as substrates. The k_{cat} and K_m values for DdaG were determined to be 3.2 ± 0.2 min^{-1} and 8.8 ± 1.8 mM (Figure S31), respectively. Compared to its native reaction with fumaric acid and 2,3-diaminopropionic acid,[17] the k_{cat} with non-native substrates is 4-fold lower and K_m is 15-fold higher. These findings indicate substantial potential for engineering DdaG to enhance its activity with non-native substrates.

Exploration of amide synthetases' potential for synthesizing complex oligoamides

Since DdaG and SfaB displayed a broad substrate scope of polymer-relevant monads, we next asked whether the resulting diads can serve as substrates for triad formation. Therefore, we selected four diads (i.e., **CS**, **MS**, **MG**, and **MA**; Figure S4) for chemical synthesis. **MS**, **MG**, and **CS** TFA salts were synthesized by reacting mono-Boc protected diamines with cyclic anhydrides to form α,ω-*N*-boc amino carboxylic acids that were subsequently deprotected by reaction with trifluoroacetic acid. Although we initially attempted to synthesize the diads directly from unprotected diamines using the procedure reported by Madhavachary *et al*.,[25] we ultimately found this method overly cumbersome as it required complex chromatographic separation to completely remove the difunctionalized byproduct. Additionally, the instability of adipic anhydride[26] required us to instead synthesize **MA** by coupling *N*-Boc-1,6-hexamethylenediamine with monomethyl adipate using EDC as the coupling agent.

We then tested whether DdaG and SfaB could ligate these synthesized diads with the diacids previously shown to act as carboxylic acid donors. Analysis by I.DOT/OPSI-MS demonstrated that all four diads could be ligated with diacid monads to form symmetric or asymmetric diacid triads (e.g. **GMG** and **SMG**) (Figure 3), and 6 out of 13 were confirmed by MS^2 spectra (Figures S32-37). This finding is consistent with our previous results, which indicated that the enzymes activate diacids but not the acid moiety of ω-amino acids and maintain their substrate preferences—DdaG preferentially uses **S**, while SfaB prefers **G** and **A**.

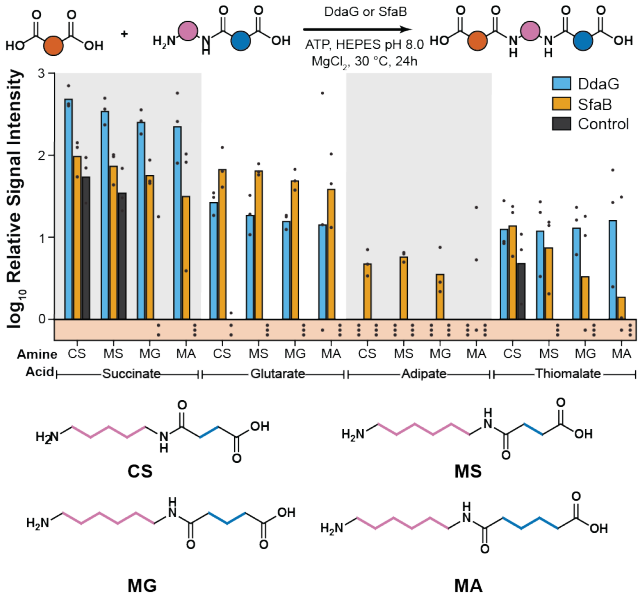


Figure 3: Enzymatic synthesis of diacid triads. *In vitro* biochemical assays were conducted by incubating different diacids (i.e., succinic acid, glutaric acid, adipic acid and thiomalic acid) with chemically synthesized diads (i.e., **CS**, **MS**, **MG** and **MA**) in the presence of enzymes (i.e., DdaG or SfaB) or no-enzyme control. Products were assayed using I.DOT/OPSI-MS. The $\log_{10}$MS signal intensity was cut off at 5.0 to filter out noise. **CS**: succinic acid-cadaverine diad; **MS**: succinic acid-hexamethylenediamine diad; **MG**: glutaric acid-hexamethylenediamine diad; **MA**: adipic acid-hexamethylenediamine diad. Sample size is $N = 3$.

We also performed additional experiments to test the potential of DdaG and SfaB to form oligoamide tetrads. We synthesized two symmetrical diamine triads, **MSM** and **MAM** (Figure S4), and tested for ligation with diacid monads. We determined from I.DOT/OPSI-MS that both enzymes were capable of forming tetrads by coupling **MSM** with linear diacids including **S**, **G**, **A,** and **P** while preserving their diacid preferences (Figures S38-40). However, neither of the enzymes was active towards **MAM**. Additionally, similar to the findings for diad formation, DdaG also exhibited low activities with **K** and SfaB showed low activities using **O**.

Quantification and scale up of diad biosynthesis

To quantify amide dimer yield, we measured the production of representative diads (i.e., **CS**, **MS**, **MG**, and **MA**) in small-scale reactions by comparison to chemically-synthesized standards (Figures 4A and S41). The molar yield after overnight incubation ranged from 3% to 80%, with DdaG demonstrating the highest activity in synthesizing **MS** (~80%) and **CS** (~60%) but low production of **MG** and **MA**. In contrast, SfaB exhibited lower activity than DdaG for production of **MS** and **CS** but produced higher amounts of **MG** (~20%) and **MA** (~3%).

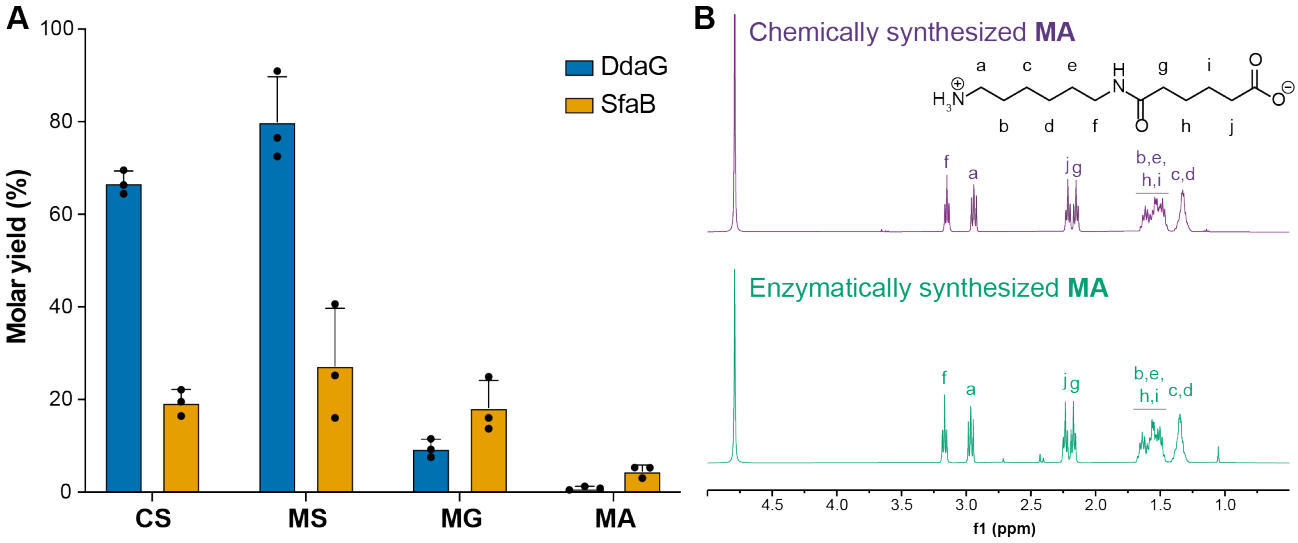


Figure 4: Diad quantification. (A) Small-scale quantification of representative diads. *In vitro* biochemical assays were conducted by incubating different diacids (i.e., succinic acid, glutaric acid, or adipic acid) with diamines (i.e., cadaverine or hexamethylenediamine) in the presence of enzymes (i.e., DdaG or SfaB). Products were quantified by comparing to the chemically-synthesized standards using OPSI-MS. **CS**: succinic acid-cadaverine diad; **MS**: succinic acid-hexamethylenediamine diad; **MG**: glutaric acid-hexamethylenediamine diad; **MA**: adipic acid-hexamethylenediamine diad. Sample size is $N = 3$. Error bars are shown as the mean ± standard deviation. (B) ^{1}H NMR spectra of enzymatic-synthesized and chemically-synthesized **MA** diad.

To demonstrate the utility of enzymatic synthesis for oligoamide production, we next focused on scale-up and isolation, using synthesis of **MS** by DdaG as a model reaction. We increased the concentration of each substrate to 50 mM and increased the reaction volume to 50 mL, while keeping the concentration of DdaG at 10 µM. Based on OPSI-MS analysis, the reaction yield was 44% (Figure 5A). We then developed a purification protocol wherein the crude reaction mixture was reacted with 2-acetyldimedone (dde-OH) to provide a dde-functionalized adduct that was easily isolated chromatographically, and subsequently deprotected via hydrazinolysis (Scheme S1 and Figures S42-45) to provide the target diad. Using this method, we isolated 60 mg (~11%) of spectroscopically pure **MS.** We also performed a similar scale-up of **MA** production by SfaB. 10 mg **MA** (~5%, Figure 4B) was isolated from a 5 mL reaction with 50 mM of each substrate and 10 µM SfaB.

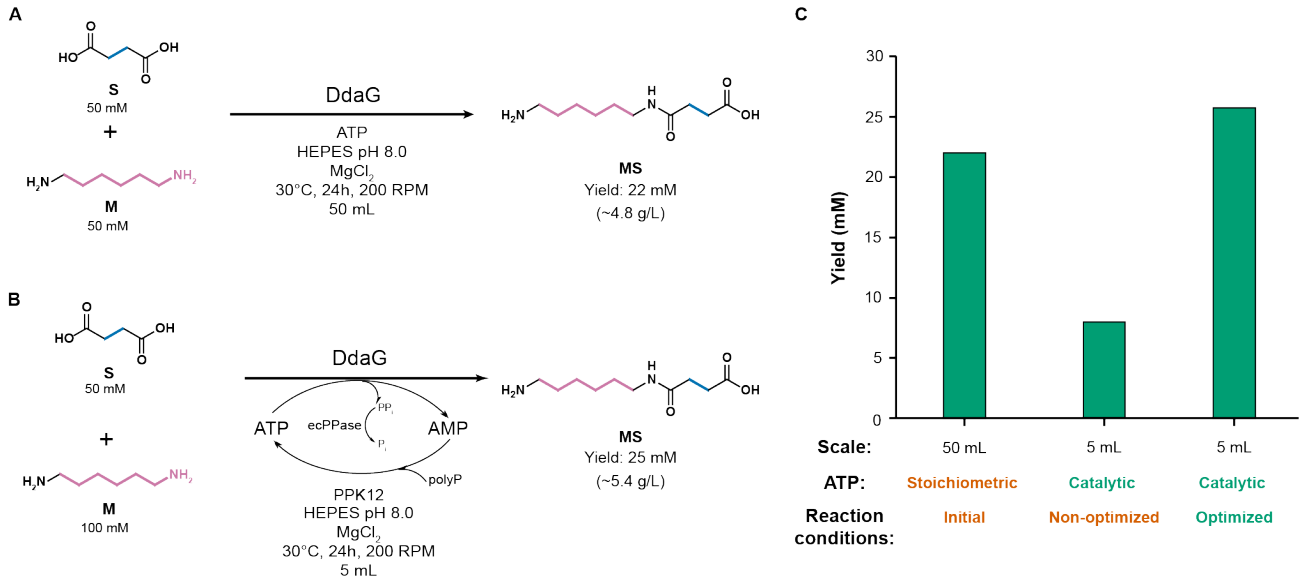


Figure 5: Diad scale up and optimization. (A) Scale up of **MS** production. A 50 mL reaction with 10-fold increased substrates loading was performed, and the yield was quantified by mass spectrometry. (B) ATP-recycled scale-up of **MS** production. A 5 mL reaction with increased substrate loading and an ATP regeneration system was performed, and the yield was quantified by mass spectrometry. (C) Comparison of **MS** production under different approaches. The 50 mL reaction was performed with initial setup: 50 mM succinic acid, 50 mM hexamethylenediamine, 10 mM $MgCl_2$, 10 µM DdaG, 50 mM ATP and 100 mM HEPES (pH 8). For the non-optimized 5 mL ATP-recycled reaction, 50 mM ATP was replaced with 1 mg/mL PPK12, 1 mM ATP, 1 mM AMP, 25 mM SHMP and 0.2 U ecPPase with other components unchanged. For the optimized condition, 100 mM hexamethylenediamine, 100 mM $MgCl_2$ and 5 mM AMP were used with other components unchanged from the non-optimization. All the reactions were incubated for 24 hours at 30 ºC with shaking at 200 rpm. These experiments were conducted in *N*=1.

DdaG and SfaB are ATP-dependent amide synthetases, and providing super-stoichiometric ATP to ensure effective conversion is cost-prohibitive. As an alternative, we implemented an ATP recycling system using a purified type 2-III polyphosphate kinase (PPK12).[27,28] In addition, we included inorganic pyrophosphatase from *E. coli* (ecPPase) to prevent potential inhibition by pyrophosphate. We performed reaction optimization using **MS** production by DdaG. We iteratively optimized the concentrations of the substrates, polyphosphate, magnesium, AMP, and pH (Figures S46-50). Based on these results, the finalized optimal conditions for **MS** production were 50 mM succinic acid, 100 mM hexamethylenediamine, 100 mM $MgCl_2$, 10 µM DdaG, 1 mg/mL PPK12, 5 mM AMP, 0.2 U ecPPase and 100 mM HEPES (pH 8). With these optimal conditions, we subsequently scaled up our reaction system with ATP regeneration from 50 µL to a 5 mL reaction volume. After 24 h incubation, a small volume (50 µL) of sample was analyzed by I.DOT/OPSI-MS, indicating a promising yield of ~50% (~5 g/L) and demonstrating a more than 3-fold improvement from the initial conditions (Figure 5B).

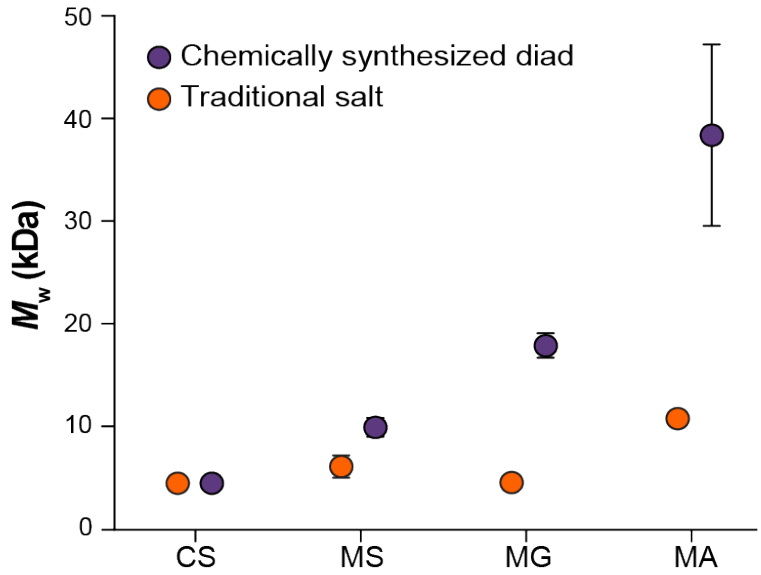


Figure 6: Diad Polymerization. Molecular weight comparison between polymers synthesized with freebased diad or traditional nylon salt. Polymers were synthesized through solid-state polymerization, using either the chemically-synthesized diad or traditional nylon salt as the starting material. **CS**: succinic acid-cadaverine diad; **MS**: succinic acid-hexamethylenediamine diad; **MG**: glutaric acid-hexamethylenediamine diad; **MA**: adipic acid-hexamethylenediamine diad. Sample size is $N = 3$. Error bars are shown as the mean ± standard deviation.

Next, we aimed to compare the polymers synthesized from these diads with the corresponding polymer obtained from the traditional salt. To this end, **MA**, **CS**, **MG**, and **MS** were synthesized through conventional methods to provide diads in sufficient quantity for polymerization (see Supplemental Methods and Figures S51-72 for details). Our initial efforts focused on the polymerization of **MA** to PA6,6. We polymerized the **MA** TFA salt using a two-step solid-state polymerization (SSP) by first pre-polymerizing the diad salt under an Ar atmosphere at 220 °C for 1 h and then further reacting under vacuum for an additional 7 h. Our initial trials generated PA6,6 with low MW, reaching a number average degree of polymerization ($\bar{X}_n$) of ~10 as measured by size exclusion chromatography (SEC) (Figure S73 and Table S1). This polymerization was also notably sluggish, containing a significant amount of unreacted diad after the initial pre-polymerization step (Figure S73). We initially suspected that the MW may have been limited by a stoichiometric imbalance of functional groups caused by partial trifluoroacetylation of the amine end groups. However, exchanging the TFA counterion for chloride provided no significant increase in polymer MW (Figure S74 and Table S1). We thus concluded that the rate of polymerization of the diad salts was limited by slow dissociation of the ammonium ion. We overcame this limitation by neutralizing the **MA** TFA salt with a commercially-available crosslinked poly(4-vinylpyridine) resin (Reillex® 402).[29] Removal of the TFA counterion was confirmed by ^{13}C nuclear magnetic resonance (NMR) spectroscopy, noting the disappearance of the TFA counterion peaks at 163 and 120 ppm (Figures S56 and S58). The neutralized **MA** was subsequently polymerized in triplicate via SSP, and the polymer MW was compared to polyamide samples prepared from a traditional PA6,6 salt. Polyamides synthesized from **MA** reached

significantly higher MW than those prepared from a traditional nylon salt (Figure 6). We hypothesize that the improved polymerization of **MA** could be due to several factors, most notably by improving the stoichiometric balance of the diamine and diacid components, minimizing evaporation of hexamethylenediamine, and starting the polymerization with 50% of the potential amide bonds preformed.

These experiments were repeated for alternative polyamides by varying the length of the diamine and/or diacid segments. Similar trends were observed for PA6,5 and PA6,4, with **MG** and **MS** diads producing higher MW polyamides than their respective nylon salts. However, no statistically significant difference in molecular weight was observed for PA5,4 synthesized from **CS** or the corresponding salt. We attribute this discrepancy to competing imidization and amidation pathways limiting the polymer MW. Consistent with this hypothesis, Fourier-transform infrared (FT-IR) spectroscopy of the PA6,4 polymer showed an imide C=O stretch at ~1770 cm^{-1} (Figure S75). Thermal imidization was also observed in PA5,4 samples, explaining the lower MW of polyamides containing succinic acid compared to PA6,6 and PA6,5. We also polymerized the enzymatically-synthesized **MS** to PA6,4 (Figure S76), though additional methods optimization is needed.

We used thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) to compare the degradation temperatures (T_d), glass transition temperatures (T_g), melting temperatures (T_m), and enthalpies of melting (ΔH_m) of polyamides prepared from diads to those prepared from the corresponding nylon salts (Figures S77-84 and Table S2). No considerable difference was observed in T_d for PA6,6, PA6,5, or PA5,4 prepared from either nylon salts or diads. However, a somewhat lower T_d was observed for PA6,4 synthesized from **MS** compared to its corresponding salt (Figure S81). Additionally, PA6,4 prepared from **MS** showed an apparent two step decomposition profile, whereas the salt-derived polyamide decomposed in a single step. Despite this difference in thermal stability, both PA6,4 samples display nearly identical thermal properties, consistent with the hypothesis that the two samples are chemically identical. Likewise, similar agreement in T_g, T_m, and ΔH_m for salt-derived and diad-derived polyamides were observed for PA6,6, PA6,5, and PA5,4.

Conclusion

In this work, we demonstrated that soluble amide synthetases can selectively couple unprotected, commercially-relevant diacids, diamines, and ω-amino acids into diacid or ω-amino acid amide diads, diacid triads, and ω-amino acid tetrads suitable for polymerization into conventional nylons or, ultimately, into sequenced copolyamides. Reaction molar yields reached up to 80%, while reaction kinetics demonstrated substantial room for further improvement through enzyme engineering. Reaction scale-up and optimization highlighted the potential for simple and efficient enzymatic synthesis of sufficient quantities of oligoamides for further characterization. Polymerization of amide diads yielded polymers with higher molecular weight than from conventional nylon salts, suggesting that oligoamides are viable building blocks for polymer synthesis. Cascades of engineered amide synthetases could be used *in vitro* to generate defined oligoamides for polymerization into sequenced co(polyamides) and analysis of sequence/property relationships, while the same enzymes could be applied *in vivo* for industrial-scale biosynthesis of desired oligoamides from unprotected substrates.

326 **Methods**
327 *Materials:* LB agar, LB medium and Terrific Broth were purchased from BD Difco. All chemicals
328 were purchased from ThermoFisher, Alfa Aesar, or Tokyo Chemical Industry (TCI).
329 Pyrophosphatase was purchased from New England Biolabs.
330
331 *Synthesis and Expression of Enzymes:* Genes encoding DdaG[17](GenBank: ADN39487.1),
332 SfaB[16](GenBank: ATY72527.1), McbA[18](GenBank: AGL76720.1), and PylC[22](GenBank:
333 AAZ69813.1) were synthesized by GenScript (Piscataway, NJ) with a N-terminal His_6 tag for
334 purification and cloned into pET28a(+). The gene encoding SuCphA1[21](GenBank: BAA17890.1)
335 was synthesized by GenScript (Piscataway, NJ) with a N-terminal His6 tag for purification and
336 cloned into pETDuet-1. For the His-SUMO-PylC construct, PCR amplified PylC was inserted into
337 Plasmid #112793 purchased from Addgene and replaced the location of EfaCas1 using Gibson
338 Assembly. Plasmids containing His-DdaG, His-SfaB, His-McbA, His-PylC, His-SUMO-PylC,
339 His-SuCphA1, His-PPK12 were transformed into chemically competent BL21 (DE3) *Escherichia*
340 *coli* cells (Sigma-Aldrich, St Louis, MO) according to the manufacturer's specifications.
341
342 His-PPK12 was produced as described previously.[27] Expression and purification of His-DdaG,
343 His-SfaB, His-McbA, His-PylC, His-SUMO-PylC and His-SuCphA1 were conducted following
344 previously published methods. A colony or glycerol stock of BL21(DE3) *E. coli* containing
345 plasmid DNA was used to inoculate 10 mL of LB medium supplemented with appropriate
346 antibiotics. The culture was incubated at 37 °C and 250 rpm overnight. The overnight culture was
347 then diluted 100-fold into 200 mL of Terrific Broth. The subculture was incubated at 37 °C at 250
348 rpm until an OD of 0.4-0.6 was reached. Protein expression was induced by adding IPTG to a final
349 concentration of 0.1 mM. The temperature was decreased to 16 °C and the culture was grown with
350 shaking overnight. The cells were harvested by centrifugation (4000 rpm, 40 min, 4 °C),
351 resuspended in 5 mL of lysis buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, 10mM
352 $MgCl_2$, 10% glycerol, pH 8.0), and sonicated on ice. Cell debris was removed by centrifugation
353 (10,000 rpm, 30 min, 4 °C). The supernatant was then filtered with an MCE filter (0.22 µm) and
354 loaded at 3 mL/min onto an ÄKTA Start FPLC (Cytiva Marlborough, MA) equipped with a 5-mL
355 His-Trap™ column (Cytiva). The column was washed with wash buffer (50 mM HEPES, 300 mM
356 NaCl, 30 mM imidazole, 10 mM $MgCl_2$, 10% glycerol, pH 8.0) before eluting with elution buffer
357 (50 mM HEPES, 300 mM NaCl, 250 mM imidazole, 10 mM $MgCl_2$, 10% glycerol, pH 8.0) at 3
358 mL/min. Purified proteins were concentrated with an Amicon® Ultra centrifugal filter at 10 kDa
359 cutoff, and buffer exchanged into exchange buffer (50 mM HEPES, pH 8.0, 100 mM NaCl, 10mM
360 $MgCl_2$, 10% glycerol). The final purified product was analyzed by SDS-PAGE (Figures S1-S3)
361 and enzyme concentration was measured using a NanoDrop™ 1000 Spectrophotometer (Thermo
362 Scientific) with exchange buffer as a blank, and the extinction coefficient of each protein was
363 calculated using ProtParam from the ExPASy Proteomics Server. The molecular weight of each
364 protein was calculated using ProtParam from the ExPASy Proteomics Server, and further
365 confirmed by the SDS-PAGE method. Pure enzymes were stored at -80 °C for subsequent activity
366 assays.
367
368 *In vitro Enzyme Activity Assays:* A typical enzymatic reaction was carried in triplicate at 50 µL
369 containing 100 mM HEPES (pH 8.0), 10 mM ATP, 10 mM $MgCl_2$, 10 µM enzyme, 5 mM carboxyl
370 group-containing compounds and 5 mM amine group-containing compounds, and was incubated
371 at 30 °C for 16 h with shaking at 200 rpm. The reaction was quenched by adding 50 µL methanol.

The samples were centrifuged at 4,000 rpm for 10 min, and the supernatant was subjected to MS analysis to determine the formation of corresponding products. For the 50 ml reaction, 100 mM HEPES (pH 8.0), 50 mM ATP, 50 mM $MgCl_2$, 10 µM enzyme, 50 mM disodium succinic acid and 50 mM hexamethylenediamine were incubated at 30 °C for 16 h with shaking at 200 rpm. The reaction was quenched by adding 50 mL methanol. The samples were centrifuged at 4,000 rpm for 20 min, and the supernatant was subjected for product isolation .

Enzymatic reaction optimization coupling with ATP regeneration system: Reactions with ATP regeneration were performed in 100 mM HEPES pH 8, 1 mM ATP, 1 or 5 mM AMP, 10-100 mM $MgCl_2$, 5-25 mM SHMP, 50 mM disodium succinic acid, 50-250 mM hexamethylene diamine,10 µM DdaG, 0.2 U ecPPase, and 1 mg/mL PPK12. These reactions were then incubated at 30 ºC with shaking at 200 rpm. After 24h, reaction mix was quenched with 1 volume of methanol and samples were collected after centrifugation to clear insoluble precipitates.

Kinetic analysis of DdaG on reaction of succinate and hexamethylenediamine: The ligation of succinate and hexamethylenediamine by purified DdaG was monitored by observation of the product MS formation according to OPSI-MS. Kinetic assays were performed in triplicate with a reaction mixture containing 0.5 µM DdaG, 10 mM ATP, 100 mM HEPES (pH 8), 10 mM $MgCl_2$, 0.5 mM to 50 mM disodium succinate, 50 mM hexamethylenediamine in total volume of 50 µl reaction. Samples were taken after incubation at 30 °C for 5 min, 15 min, 25 min, and 35 min. The reaction was quenched by adding 50 µL methanol. The samples were centrifuged at 4,000 rpm for 10 min, and the supernatant was subjected to OPSI-MS analysis to determine the formation of corresponding products. Initial velocities were determined by linear regression analysis of the data. The kinetic parameters k_{cat}, and K_m and standard error were calculated by Graphpad Prism 9.

I.DOT/OPSI-MS Analysis of Enzyme Activity: The immediate drop-on-demand technology (Dispendix GmbH, Stuttgart, Germany) coupled with open port sampling interface-mass spectrometry (I.DOT/OPSI-MS) was used to analyze enzymatic reactions as previously described in detail.[23] Briefly, single droplets were mixed with solvent in an on-line fashion, enabling sampling throughputs of up to 2 s/sample. Here, enzymatic reactions were diluted 1:100 (*v/v*) in high performance liquid chromatography (HPLC) grade water with 500 nM propranolol acting as an internal standard for droplet capture. 40 µL of diluted reactions were transferred to a I.DOT S.100 96-well plate and analyzed without further separations. The I.DOT system was used to eject 20 nL of sample into the OPSI, into a flow of 75/25/0.1 (*v/v/v*) acetonitrile/water/formic acid and transported the sample to the electrospray ion source of the mass spectrometer. Dispensing throughput was 4 s/sample and each sample was measured in triplicate. Peak widths were ~1.2 s wide.

For identification of products by high mass resolution, a Thermo Q-Exactive HF mass spectrometer (ThermoFisher Scientific) in positive ion mode was used for characterization. Scan settings were sheath gas = 80, auxiliary gas = 40, electrospray voltage = 4 kV, ion injection time = 50 ms, automatic gain control = $3e^6$, capillary temperature = 200 °C, mass/charge (*m/z*) range = 100-750 *m/z*, and OPSI solvent flow of 250 µL/min. A 60,000 mass resolution was used for quantitative analyses while 240,000 resolution was used for compound identification. Tandem MS data were acquired using the Q-Exactive HF, but water adduct formation was commonly observed in the ion trap. To avoid water adducts in tandem mass spectra, tandem mass spectrometry data was collected using a Sciex 5600 TripleTOF time of flight mass spectrometer (Sciex) with the

following settings: GS1 = 75, GS2 = 35, electrospray voltage = 5.5 kV, capillary temperature = 200 °C, and OPSI solvent flow of 150 µL/min. Collision energies across 10-50 eV were evaluated for each ion.

Control of I.DOT settings (positioning, well selection, timing and droplet dispensing cycles), OPSI peak finding, data extraction from vendor file formats, and signal integration were enabled through custom software packages written in Delphi, Python, and C#, available upon request. For each droplet sampling event, the average mass spectrum from the resulting mass spectral peak was normalized to propranolol internal standard signal and background subtracted. Log-log matrix matched calibration curves were used for quantitation. Mass spectrometric signals typically spanned several orders of magnitude, thus, log-log calibrations resulted in reduced error across the full concentration range.

Tandem mass spectra: I.DOT/OPSI-MS^2 data were acquired for the possible observed compounds shown in Figures 2 and 3 in the main text. MS^2 data were acquired on a Sciex 5600 TripleTOF+ instrument using the same OPSI-MS probe as described previously. A time-of-flight system was preferred as in trap water adducts were observed for some compounds when fragmented with the orbitrap system. Putative fragment identifications are given for each compound using hypothesized fragmented pathways. All tandem mass spectra shown used a collision energy of 20eV and a declustering potential of 100. Additional collision energies were used to help with fragment identifications as necessary (data not shown). For brevity, a single tandem mass spectrum is shown for each possible compound. For compounds without a tandem mass spectrum, signals for those compounds were either too weak to generate clear tandem MS or were obfuscated by overlapping background ions in the 1 *m/z* isolation window.

Nuclear magnetic resonance (NMR) spectroscopy: ^{1}H and ^{13}C nuclear magnetic resonance NMR spectroscopy was performed on a Bruker Avance III 400 NMR spectrometer operating at 400 MHz (^{1}H) or 100 MHz (^{13}C). Chemical shifts are reported in parts per million (ppm) relative to residual protonated solvent.

Size exclusion chromatography (SEC): SEC analysis was performed on an Agilent 1260 Infinity II LC system equipped with an Agilent PL HFIPgel guard column (9 µm, 50 x 4.6 mm) and two Agilent PL HFIPgel columns (9 µm, 250 x 4.6 mm). A 0.02 M CF_3COONa HFIP solution was used as the mobile phase at 40 °C and a flow rate of 0.300 mL min^{-1}. Elution time was monitored using a differential refractive index (dRI) detector, a light scattering detector operating with two angles at 90° and 15°, and a differential viscometer. Number-average molecular weights (M_n), weight-average molecular weights (M_w) and dispersities ($Đ = M_w/M_n$) were calculated based on calibrations with PMMA standards using the Agilent GPC/SEC software.

Thermogravimetric analysis (TGA): TGA was performed on a TA Instruments Discovery TGA 55 under a nitrogen atmosphere. Sample mass was recorded as a function of temperature while the sample was heated from room temperature to 600 °C at a heating rate of 20 °C min^{-1}. Degradation temperature (T_d) was taken as the temperature at which the sample had lost 5% of its initial mass.

Differential scanning calorimetry (DSC): DSC was performed on a TA Instruments Discovery DSC 250. Measurements were carried out in standard aluminum pans using a heat-cool-heat cycle with 10 and 5 °C heating and cooling rates, respectively. Melting (T_m) temperature and enthalpy

of melting were measured from the first heat cycle while glass transition (T_g) temperatures were measured from the second heating cycle.

Fourier-transform Infrared (FTIR) Spectroscopy: Attenuated total reflectance (ATR) Fourier-transform infrared spectroscopy was performed using a PerkinElmer Spectrum Two FTIR spectrometer. Transmission spectra were collected from 4000 to 500 cm^{-1}.

Acknowledgements
Research was sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725, for the U. S. Department of Energy. Work by N.T.Z was supported by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships (SULI) program. We thank Madan Gopal and Wilfred Chen (University of Delaware) for the PPK12 expression plasmid. LQ, ITD, JCF, and JKM are inventors on a patent application related to this work.

Author contributions
L.Q. performed the majority of the enzymatic characterization and optimization. I.T.D.: oligomer synthesis, polymerization, and characterization. D.L.C: Analytical characterization. V.K.: Analytical method development. N.T.Z.: Characterization of long-chain substrates. N.A.T.: Oligomer synthesis. J.F.C.: Analytical methods development and analysis, supervision of analytical research. J.C.F.: Conceptualization, supervision of chemistry research. J.K.M: Conceptualization; funding acquisition; supervision of biochemistry research.

References

[1] T. Schiller, T. Scheibel, *Commun. Mater.* **2024**, *5*, 1–18.
[2] E. L. Bell, W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niikura, S. Osuna, E. Romero, K. S. Ryan, N. J. Turner, S. L. Flitsch, *Nat. Rev. Methods Primers* **2021**, *1*, 1–21.
[3] T. Katoh, T. Sengoku, K. Hirata, K. Ogata, H. Suga, *Nat. Chem.* **2020**, *12*, 1081–1088.
[4] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498–500.
[5] R. Fricke, C. V. Swenson, L. T. Roe, N. X. Hamlish, B. Shah, Z. Zhang, E. Ficaretta, O. Ad, S. Smaga, C. L. Gee, A. Chatterjee, A. Schepartz, *Nat. Chem.* **2023**, *15*, 960–971.
[6] Z. L. Watson, I. J. Knudson, F. R. Ward, S. J. Miller, J. H. D. Cate, A. Schepartz, A. M. Abramyan, *Nat. Chem.* **2023**, *15*, 913–921.
[7] J. Lee, K. J. Schwarz, D. S. Kim, J. S. Moore, M. C. Jewett, *Nat. Commun.* **2020**, *11*, 4304.
[8] M. Spinck, C. Piedrafita, W. E. Robertson, T. S. Elliott, D. Cervettini, D. de la Torre, J. W. Chin, *Nat. Chem.* **2023**, *15*, 61–69.
[9] D. L. Dunkelmann, C. Piedrafita, A. Dickson, K. C. Liu, T. S. Elliott, M. Fiedler, D. Bellini, A. Zhou, D. Cervettini, J. W. Chin, *Nature* **2024**, *625*, 603–610.
[10] I. B. Folger, N. F. Frota, A. Pistofidis, D. L. Niquille, D. A. Hansen, T. M. Schmeing, D. Hilvert, *Nat. Chem. Biol.* **2024**, *20*, 761–769.
[11] K. A. J. Bozhüyük, F. Fleischhacker, A. Linck, F. Wesche, A. Tietze, C.-P. Niesert, H. B. Bode, *Nat. Chem.* **2018**, *10*, 275–281.
[12] K. A. J. Bozhüyük, L. Präve, C. Kegler, L. Schenk, S. Kaiser, C. Schelhas, Y.-N. Shi, W.

Kuttenlochner, M. Schreiber, J. Kandler, M. Alanjary, T. M. Mohiuddin, M. Groll, G. K. A. Hochberg, H. B. Bode, *Science* **2024**, *383*, eadg4320.
[13] M. Lubberink, W. Finnigan, S. L. Flitsch, *Green Chem.* **2023**, *25*, 2958–2970.
[14] M. Winn, S. M. Richardson, D. J. Campopiano, J. Micklefield, *Curr. Opin. Chem. Biol.* **2020**, *55*, 77–85.
[15] G. M. Landwehr, J. W. Bogart, C. Magalhaes, E. G. Hammarlund, A. S. Karim, M. C. Jewett, *Nat. Commun.* **2025**, *16*, 865.
[16] L. Wang, M. Zhu, Q. Zhang, X. Zhang, P. Yang, Z. Liu, Y. Deng, Y. Zhu, X. Huang, L. Han, S. Li, J. He, *ACS Chem. Biol.* **2017**, *12*, 3067–3075.
[17] M. A. Hollenhorst, J. Clardy, C. T. Walsh, *Biochemistry* **2009**, *48*, 10467–10472.
[18] Q. Tang, M. Petchey, B. Rowlinson, T. J. Burden, I. J. S. Fairlamb, G. Grogan, *ACS Catal.* **2024**, *14*, 1021–1029.
[19] M. Zhu, L. Wang, J. He, *Angew. Chem. Int. Ed Engl.* **2021**, *60*, 2030–2035.
[20] M. Winn, M. Rowlinson, F. Wang, L. Bering, D. Francis, C. Levy, J. Micklefield, *Nature* **2021**, *593*, 391–398.
[21] I. Sharon, S. Pinus, M. Grogg, N. Moitessier, D. Hilvert, T. M. Schmeing, *Nat. Commun.* **2022**, *13*, 3923.
[22] M. A. Gaston, L. Zhang, K. B. Green-Church, J. A. Krzycki, *Nature* **2011**, *471*, 647–650.
[23] J. F. Cahill, V. Kertesz, P. Saint-Vincent, H. Valentino, E. Drufva, N. Thiele, J. K. Michener, *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 1383–1391.
[24] N. A. Rorrer, S. F. Notonier, B. C. Knott, B. A. Black, A. Singh, S. R. Nicholson, C. P. Kinchin, G. P. Schmidt, A. C. Carpenter, K. J. Ramirez, C. W. Johnson, D. Salvachúa, M. F. Crowley, G. T. Beckham, *Cell Rep. Phys. Sci.* **2022**, *3*, 100840.
[25] R. Madhavachary, E. M. M. Abdelraheem, A. Rossetti, A. Twarda-Clapa, B. Musielak, K. Kurpiewska, J. Kalinowska-Tłuścik, T. A. Holak, A. Dömling, *Angew. Chem. Int. Ed Engl.* **2017**, *56*, 10725–10729.
[26] J. W. Hill, *J. Am. Chem. Soc.* **1930**, *52*, 4110–4114.
[27] M. R. Gopal, R. M. Dickey, N. D. Butler, M. R. Talley, D. T. Nakamura, A. Mohapatra, M. P. Watson, W. Chen, A. M. Kunjapur, *ACS Catal.* **2023**, *13*, 4778–4789.
[28] M. Tavanti, J. Hosford, R. C. Lloyd, M. J. B. Brown, *Green Chem.* **2021**, *23*, 828–837.
[29] D. M. Jewett, R. L. Ehrenkaufer, *Anal. Biochem.* **1982**, *122*, 319–321.