JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Acyl Transfer Catalytic Activity in De Novo Designed Protein with N-Terminus of α-Helix As Oxyanion-Binding Site

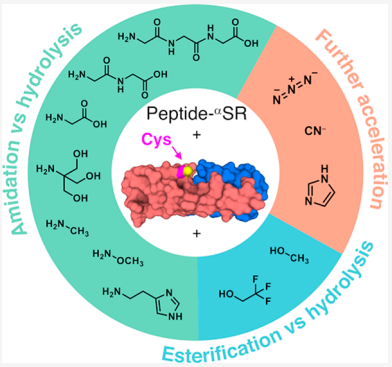
Elise A. Naudin, Alastair G. McEwen, Sophia K. Tan, Pierre Poussin-Courmontagne, Jean-Louis Schmitt, Catherine Birck, William F. DeGrado,* and Vladimir Torbeev*

Cite This: J. Am. Chem. Soc. 2021, 143, 3330−3339

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The design of catalytic proteins with functional sites capable of specific chemistry is gaining momentum and a number of artificial enzymes have recently been reported, including hydrolases, oxidoreductases, retro-aldolases, and others. Our goal is to develop a peptide ligase for robust catalysis of amide bond formation that possesses no stringent restrictions to the amino acid composition at the ligation junction. We report here the successful completion of the first step in this long-term project by building a completely de novo protein with predefined acyl transfer catalytic activity. We applied a minimalist approach to rationally design an oxyanion hole within a small cavity that contains an adjacent thiol nucleophile. The N-terminus of the α-helix with unpaired hydrogen-bond donors was exploited as a structural motif to stabilize negatively charged tetrahedral intermediates in nucleophilic addition−elimination reactions at the acyl group. Cysteine acting as a principal catalytic residue was introduced at the second residue position of the α-helix N-terminus in a designed three-α-helix protein based on structural informatics prediction. We showed that this minimal set of functional elements is sufficient for the emergence of catalytic activity in a de novo protein. Using peptide-$^{\alpha}$thioesters as acyl-donors, we demonstrated their catalyzed amidation concomitant with hydrolysis and proved that the environment at the catalytic site critically influences the reaction outcome. These results represent a promising starting point for the development of efficient catalysts for protein labeling, conjugation, and peptide ligation.



INTRODUCTION

Designing proteins with the desired catalytic properties provides a learning ground for understanding the fundamentals of enzyme catalysis and represents a major scientific challenge.[1−3] Moreover, the detailed knowledge of the enzyme mechanisms can be useful for deeper insights into chemical reactivity in general. In the recent years, there has been a significant progress in the design of enzymes using computational methods for a variety of reactions including retro-aldol reaction,[4] Kemp elimination,[5] Diels−Alder reaction,[6] and ester hydrolysis.[7,8] The design strategy, proven to be general, consists of a combinatorial search of protein sequence, as well as optimization of backbone conformation and side chain rotamers to find the best arrangement of functional groups that stabilize the transition state of the catalyzed reaction when it is docked into known protein scaffolds.[9] However, the catalytic efficiency of such computationally redesigned proteins turned out to be substantially lower in comparison to the observed catalytic rates for natural enzymes.[10] The currently available solution to achieve higher rates is to apply high-throughput methods such as directed evolution to optimize the sequences of initially designed scaffolds yielding man-made enzymes with higher catalytic efficiencies.[11,12] A recent highlight of the application of directed evolution strategy is the highly proficient Zn-containing hydrolase with catalytic efficiency surpassing the hydrolytic activities of the naturally occurring metalloenzymes.[13]

Installing a catalytic function increases the level of design complexity compared to a more tractable task of the design of three-dimensional protein structure.[14−16] The difficulty of recapitulating high catalytic efficiencies of natural enzymes by purely computational strategies is attributed to a lack of high enough precision of placement of critical catalytic residues and inability to control long-range inter-residue interactions including protein conformational dynamics. Several rounds of laboratory evolution are able to rectify these deficiencies by remodeling the structure of the active site,[17] as well as optimizing the surrounding second shell residues that contribute to an intricate balance between precise structural complementarity to transition state(s) of the catalyzed reaction and protein conformational flexibility mediating the efficient substrate binding and product release.[18,19]

Received: September 20, 2020
Published: February 26, 2021

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

At the present moment, methodologically more simple minimalist approach for de novo design of protein catalysts is still on par with computational methods.[20] Minimalist design is based on chemical logic and intuition and implies the incorporation of a minimal number of elements (such as amino acid residues or cofactors) into a de novo protein scaffold that is initially devoid of catalytic properties in order to convert it into the catalytically active variant.[14,16,20] The recent examples of de novo catalytic proteins include a thermostable hydrolase based on a 7-helix bundle and containing multiple copies of Cys-His-Glu triad as catalytic apparatus,[21] a promiscuous peroxidase that utilizes heme as a cofactor bound to a 4-helix bundle scaffold,[22] and a hydrolase based on calmodulin, where the activity is regulated allosterically upon binding to a Ca^{2+} ion.[23] In addition, it was shown that the catalytic proteins developed using minimalistic design can also be subjected to rounds of directed evolution to improve their properties.[24]

Our aim is to elaborate a protein catalyst that can be used to accelerate amide bond formation with applications for chemical or semisynthesis of proteins, or for protein modification. Chemical approaches for production of proteins hold several advantages in comparison to molecular biology tools such as the possibility to incorporate a diverse set of noncanonical building blocks[25,26] or cross-links[27] or to produce mirror-image protein forms,[28,29] proteins with nonlinear backbone topology,[30,31] or even constructs that are hybrids of polypeptides and other types of polymers.[32] These capabilities can enrich protein science with the concepts and tools from physical organic, supramolecular, polymer, and other chemistries.

We are inspired by the naturally occurring ligases, such as butelase-1,[33] sortase A,[34] or ubiquitin ligases[35] that illustrate the feasibility of high catalytic efficiencies for condensation of acyl donor and acceptor peptide substrates. In addition, re-engineered proteases have been demonstrated to perform condensation of peptides.[36] For instance, previously re-engineered variant of subtilisin was successfully used to assemble a 124-residue ribonuclease A.[37] Furthermore, newer and more efficient variants of proteolytic proteins with ligation catalytic activity were recently developed.[38,39] However, both natural ligases and re-engineered proteases possess restrictions to amino acids sequences that can bind to their substrate pockets; in other words, the catalytic efficiency is coupled to substrate specificity in these enzymes. Moreover, the overall complexity of these large and evolved structures complicates the improvement and further tuning of their enzymatic properties. To have a more general tool that is not limited by substrate scope and can be compatible with a variety of reaction conditions, in this study we introduced a catalytic ligation activity from first-principles into a de novo protein scaffold.

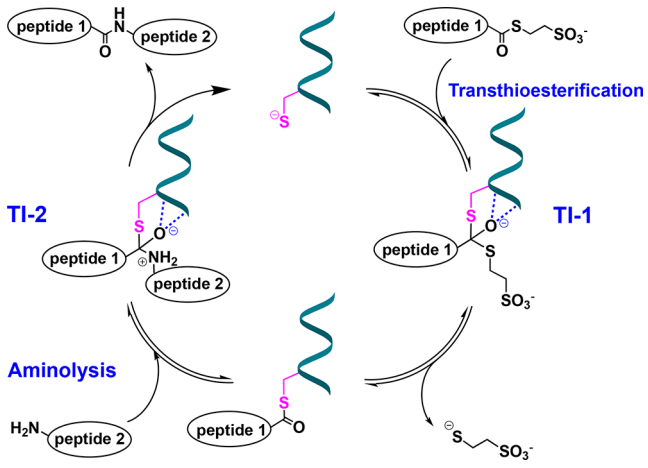
RESULTS AND DISCUSSION

Protein Design. The minimal functional elements required for the ligation reaction are a principal nucleophilic residue (such as Cys or Ser) and an "oxyanion hole" arrangement[40] of hydrogen-bond donors to stabilize the negative charge of the tetrahedral intermediates in this reaction. These functional motifs are present in both natural ligases and re-engineered proteases.[36] We selected Cys as a key catalytic residue in favor of Ser for the following reasons: (i) the thiol group of Cys has lower pK_a (8.5 in unstructured peptides)[41] and potentially can be used as a single residue instead of catalytic dyads or triads (e.g., classical Ser-His-Asp) that promote deprotonation of hydroxyl in Ser, because the correct placement of two or three residues can already complicate the design; (ii) the activation of peptide acyl donor in the form of peptide anchored to protein scaffold via Cys-thioester bond can be advantageous to the related Ser-oxoester, because the thioester group benefits from an enhanced reactivity toward most of the nucleophiles because of the poorer orbital interactions between the sulfur atom and carbonyl moiety.[42]

It should be mentioned that in addition to ligases and proteases, the activation in the form of acyl-thioesters occurs in nonribosomal peptide or fatty acid synthases[43,44] and in protein splicing through self-catalyzed intein exclusion and ligation of two adjacent exteins[45] highlighting the utility of thioesters for acyl-transfer reactions. Another advantage of Cys as a catalytic residue is the possibility of using synthetic peptide-$^{\alpha}$thioesters as acyl donor substrates, which can undergo transthioesterication with Cys, thus forming covalently attached peptide−protein species. Peptide-$^{\alpha}$thioesters are accessible by both Boc/benzyl and Fmoc/tBu solid-phase peptide synthesis (SPPS)[46] and are commonly used in enzyme-free total chemical synthesis of proteins, either via silver-catalyzed condensation of partially protected peptide-$^{\alpha}$thioesters and amino-peptides[47] or with the help of native chemical ligation, which is a highly valuable chemoselective reaction of fully unprotected peptide-$^{\alpha}$thioesters and cysteinyl-peptides proceeding in aqueous solutions.[48]

Scheme 1 outlines the principal steps in the putative mechanism of catalysis of peptide ligation by a de novo protein. A peptide-$^{\alpha}$thioester reacts with the catalytic Cys and undergoes transthioesterification. After passing through the first tetrahedral intermediate (TI-1), the resulting branched peptide−protein thioester is posed to react with H_2N-peptide, the acyl acceptor substrate. The collapse of the second tetrahedral intermediate (TI-2) leads to the release of the ligated product and free catalytic protein. For this double-

Scheme 1. Mechanism for the Catalysis of the Acyl Transfer Reaction (e.g., Peptide Ligation) by De Novo Catalytic Protein (Depicted Schematically in Cyan)[a]



[a]TI-1 and TI-2 refer to the key tetrahedral intermediates of transthioesterification and amidation steps, respectively, where negatively charged oxyanions are stabilized by the H-bonding network (shown in blue).

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

displacement ping-pong mechanism to operate efficiently, both negatively charged intermediates of the two addition−elimination steps need to be stabilized, for example, through H-bonding with amides of protein scaffold serving as an "oxyanion hole". Therefore, the protein framework in a de novo protein must include nucleophilic Cys placed in such an environment.

Previously, it was demonstrated that phospho-serine (pSer) when introduced at the N-terminus of an α-helix can lead to its stabilization presumably due to productive H-bonding of unpaired amides of the α-helical N-terminus to the negatively charged phosphate group.[49] This stabilizing effect was most prominent when pSer was placed at the N2 (the second residue at the N-terminus of the α-helix) position.[49] By analyzing the structural data from PDB, we provide support to these experimental results: clustering and alignment of crystal structures containing pSer at Ncap (the residue preceding the first amino acid that adopts an α-helical conformation), N1, and N2 (the first and second positions at the N-terminus of α-helix) illustrate that backbone amides at the α-helical N-termini indeed can contribute 1-to-2 H-bonds to one of the distal oxygens of phospho-group (Figure 1). The tetrahedral intermediates of addition−elimination steps of the acyl-transfer catalytic mechanism shown in Scheme 1 are negatively charged and bear similarities in structure to pSer; therefore, the N-terminus of the α-helix represents a suitable structural motif for stabilization of the respective tetrahedral intermediates.

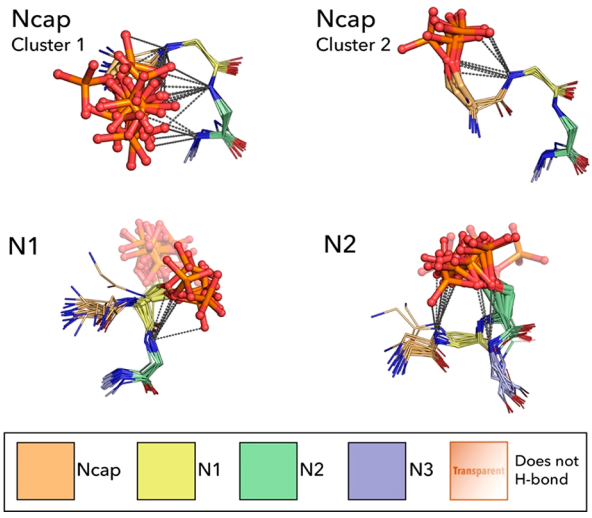
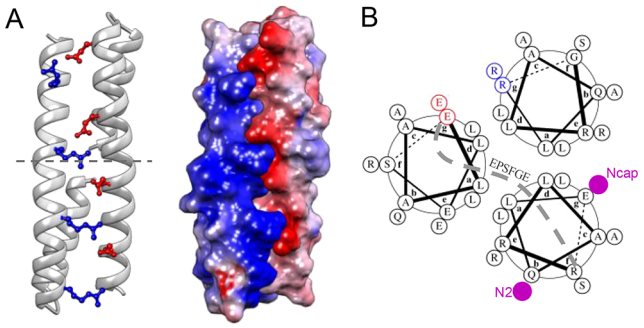


Figure 1. Structural alignment of pSer-containing fragments from structures in PDB (accessed in February 2020) with pSer located at Ncap, N1, and N2 positions of the corresponding α-helical fragments. The occurrences of pSer in a nonredundant version of the PDB were determined using the database and methods previously described.[50] pSers at Ncap positions form two clusters, due to a bimodal distribution of its ψ-dihedrals. In both clusters, all of the pSers form H-bonds with downstream amides. However, only 7 out of 23 pSers in the N1 position hydrogen bond to the helix backbone. The geometry of N2 pSers appears as stabilizing as Ncap pSers; 13 out of 16 fragments participate in H-bonds with their preceding and/or succeeding amides. pSers at the Ncap and N2 positions are observed contributing bivalent H-bonds, whereas only 1 H-bond is observed for pSer at the N1 position. Example H-bonds are represented by dashed lines, and phosphate groups that do not H-bond with the helix backbone are transparent. The PDB entries of the structures used in the alignment are listed in Table S1.

To test the feasibility of such catalytic strategy, we have chosen the de novo designed domain-swapped dimer (DSD) protein that is composed of two identical sequences that spontaneously dimerize to form a three-α-helical bundle.[51] In this scaffold, the two adjacent α-helical N-termini are facing each other, creating a cavity that represents a proper location to incorporate an active site. The environment at the juxtaposition of two N-termini is in the proximity of the hydrophobic interior of the helical bundle and, at the same time, solvent exposed; therefore, this part of the protein can permit access by various peptide substrates and allow for catalysis without introducing any steric restrictions in substrate structures.

Charge patterning of the two segments of DSD scaffold was used to redesign the homodimer into a heterodimer to distinguish the two N-termini for their further functionalization. Thus, one strand was patterned with negatively charged residues (Glu) ("negative" strand), whereas another one with positively charged (Arg) ("positive" strand) (Figure 2A, B). The Arg residues were used to substitute all Lys in the original DSD scaffold to prevent unwanted acylation of Lys side-chains.



C **Het-N2**

'Positive' strand:
Ac-NLAALRSELQALRREGFSPERLAALESRLQALERRLAALRSRLQALRG-NH_2
Cys-N2:
Ac-GLCALRSELQALRREGFSPEELAALESELQALERELAALRSELQALRG-NH_2

D **Het-Ncap**

'Positive' strand:
Ac-NLAALRSELQALRREGFSPERLAALESRLQALERRLAALRSRLQALRG-NH_2
Cys-Ncap:
Ac-CLAALRSELQALRREGFSPEELAALESELQALERELAALRSELQALRG-NH_2

Trunc-Cys-N2: Ac-GLCALRSELQAL-NH_2

Neg-N2

'Positive' strand:
Ac-NLAALRSELQALRREGFSPERLAALESRLQALERRLAALRSRLQALRG-NH_2
Ala-N2:
Ac-GLAALRSELQALRREGFSPEELAALESELQALERELAALRSELQALRG-NH_2

Figure 2. Design of heterodimers. (A) Left: charged residue patterning by grafting negative charges (Glu in red) on the strand containing the Cys catalytic residue and positive charges (Arg in blue) onto the other helix. Right: resulting coulombic surface is shown. (B) Helical wheel diagram of the proteins that corresponds to the half-structure depicted above the dashed line in A. Leucine residues constitute the hydrophobic core (*a* and *d* positions). Glu and Arg residues at *e* and *g* positions promote heterodimerization. (C) Sequences of the two monomers of **Het-N2**. (D) Sequences of the three control constructs with Cys at the Ncap position in **Het-Ncap**, a 12-residue N-terminal fragment **trunc-Cys-N2**, and negative control **Neg-N2** with Ala at the N2 position. The underlined sequence fragments correspond to that depicted on the wheel diagram in B.

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

Subsequently, catalytic Cys residue was engineered into the "negative" strand (containing excess of Glu) at N2 position, based on the most efficient stabilizing effects with pSer-containing peptides,[49] resulting in the **Cys-N2** peptide fragment. The sequences of the corresponding peptide fragments are depicted in Figure 2C. The heterodimer resulting from complexation of **Cys-N2** with "positive" strand is abbreviated as **Het-N2**. Concurrently, three control scaffolds were designed to test the importance of the N-terminal position of the catalytic cysteine and to evaluate the relevance of the entire structure for the catalytic activity. First, a second heterodimer (named as **Het-Ncap**) was designed where the "negative" strand (abbreviated as **Cys-Ncap**) contained the Cys residue at the Ncap position and the "positive" strand remained the same as for **Het-N2** (sequences are shown in Figure 2D). Second, a truncated variant of the **Cys-N2** strand was considered that corresponds to its 12 N-terminal residues and is abbreviated as **trunc-Cys-N2** (see Figure 2D). Finally, a negative control **Neg-N2** heterodimer consisting of an **Ala-N2** strand with Ala replacing Cys in **Cys-N2** and the "positive" strand was engineered.

Synthesis and Folding. The chemical synthesis of the 48-residue "positive" and "negative" strands was realized by "in situ neutralization" Boc/benzyl-SPPS,[52] whereas the shorter 12-residue **trunc-Cys-N2** was synthesized by Fmoc/tBu-SPPS[53] (analytical data are presented in Table S2). The assembly and folding of the **Het-N2** construct and the controls **Het-Ncap** and **Neg-N2** proceeded spontaneously by dissolving the corresponding "positive" and "negative" strands in phosphate buffer at near-neutral pH and mixing the solutions in a 1:1 ratio. The anticipated α-helical structure was confirmed by circular dichroism (CD) spectroscopy for **Het-N2** protein and stability was accessed by thermal denaturation by monitoring ellipticity at 222 nm, a strong indicator of α-helicity (Figures S1 and S2). On the basis of these measurements, in the absence of denaturant, **Het-N2** remained folded from 5 to 90 °C. Furthermore, the high helical content was observed with up to 2 M Gn·HCl (guanidine hydrochloride) in pH 6.9 buffer solution at temperature below 40 °C (Figure S2). Increasing the concentration of denaturant led to lowering of an apparent melting temperature and at 6 M Gn·HCl the **Het-N2** was found to be unfolded. The dimeric composition of **Het-N2** was confirmed by analytical ultracentrifugation (AUC) (Figure S3).

To demonstrate that the dimer detected for the **Het-N2** sample by sedimentation velocity experiment corresponded to a heterodimer rather than a mixture of the corresponding homodimers we turned to urea denaturation (Figure S4). Urea was preferred as a chemical denaturant, because the electrostatic interactions between solvent-exposed side chains were used to stabilize the heterodimer, and the favorability of these interactions would decrease significantly at the very high ionic strength required to denature the protein with Gn·HCl. CD-monitored denaturation curves for **Het-N2** as well as for **Cys-N2** and the "positive" strand homodimers were therefore used to determine their thermodynamic stabilities (Figure S4D). **Het-N2** proved to be approximately 10-times more stable than **Cys-N2** homodimer and 100-times more stable than the "positive" strand homodimer. Furthermore, predominantly heterodimeric species were observed by native electrospray mass-spectrometry with minor quantities of homodimers (<5%) (Figure S5).

Crystallographic Structures Show the Successful Design of the Oxyanion Hole. Finally, the designed **Het-N2** as well as the control **Het-Ncap** were crystallized. The X-ray data were collected at 2.33 Å (from anisotropic data, 4.43 Å in the worst direction) and 1.45 Å resolution for **Het-N2** and **Het-Ncap**, respectively, and the structures were solved by molecular replacement using the known structure of the DSD protein (PDB ID: 1G6U).[51] In the crystal structures both **Het-N2** and **Het-Ncap** were found to form the intended heterodimers with a three-α-helical bundle arrangement (Figure 3 and Figures S6 and S7). In comparison with the

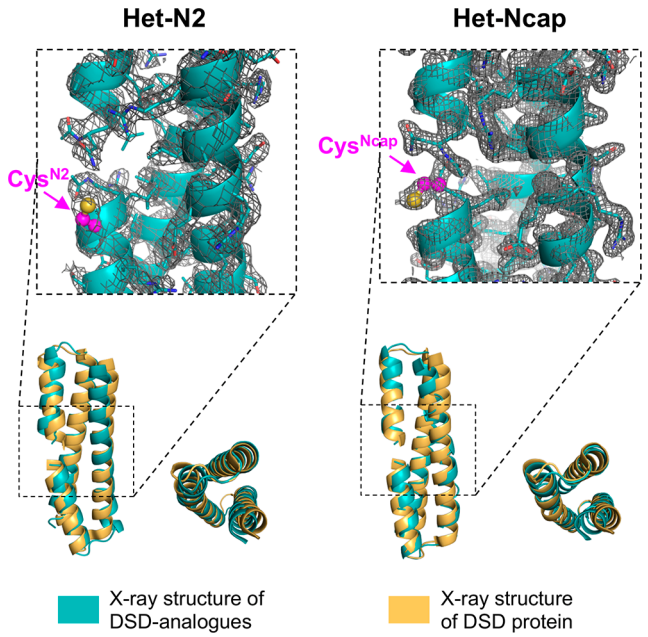


Figure 3. X-ray structures of the synthesized DSD analogues **Het-N2** (left) and **Het-Ncap** (right). Catalytic sites are shown revealing solvent-exposed cysteine residues depicted in magenta with sulfur atom in yellow in ball-and-sticks representation. The simulated annealing $2F_o-F_c$ electron density map is shown at the 1.0σ level. The X-ray structures of **Het-N2** and **Het-Ncap** (depicted in cyan) are superimposed onto the original DSD structure (in yellow). PDB IDs: 6Z0L (**Het-N2**) and 6Z0M (**Het-Ncap**). Table S3 provides the refinement statistics. Figures S6 and S7 show crystal packing and conformational variability of N-termini.

original crystal structure of the DSD protein, small shifts of the backbone were observed (on average, 1.84 Å of RMSD on Cα for **Het-N2** and 1.79 Å for **Het-Ncap**), resulting in changes of the coiled-coil pitch[54] from 218 Å for DSD protein to 199 Å for **Het-N2** and 177 Å for **Het-Ncap**. Substitution of Arg residues for Lys resulted in rotamer adjustments to maintain salt bridges between the two monomers.

Importantly, in **Het-N2** the Cys residue lies adjacent to a thin crevasse formed by the abutting N-terminal ends of the two helices (Figure 4). To validate the anticipated interactions of negatively charged tetrahedral intermediates with oxyanion hole, we solved an X-ray structure at 1.5 Å resolution for the **Het-N2** construct where the Cys^{N2} residue was converted to Cys-S-sulfonate to mimic the tetrahedral intermediates structures (Figure 5). Two **Het-N2-SO$_3^-$** dimers were observed in an asymmetric unit differing in conformation of N-termini and orientation of sulfonate moiety. In the first conformation, one of the oxygens of the sulfonate group is H-

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

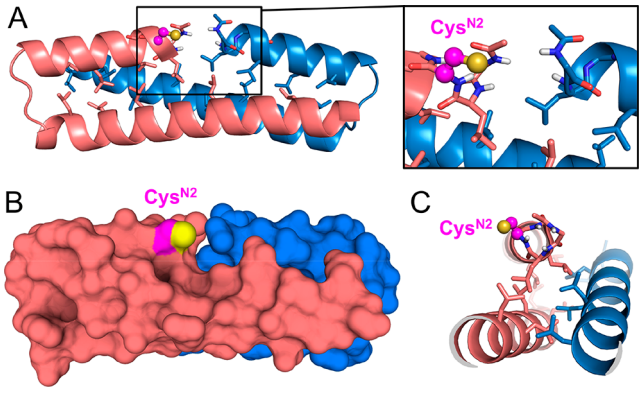


Figure 4. Active site in **Het-N2** is located at the interface of two N-termini of α-helices with nucleophilic Cys^{N2} positioned next to a small cavity. (A) Horizontal view illustrating the helix–helix interface with the unpaired N–H groups pointing toward the cavity; (B) surface representation; (C) clipped section along the helical axis. Sulfur atom of Cys^{N2} is in yellow.

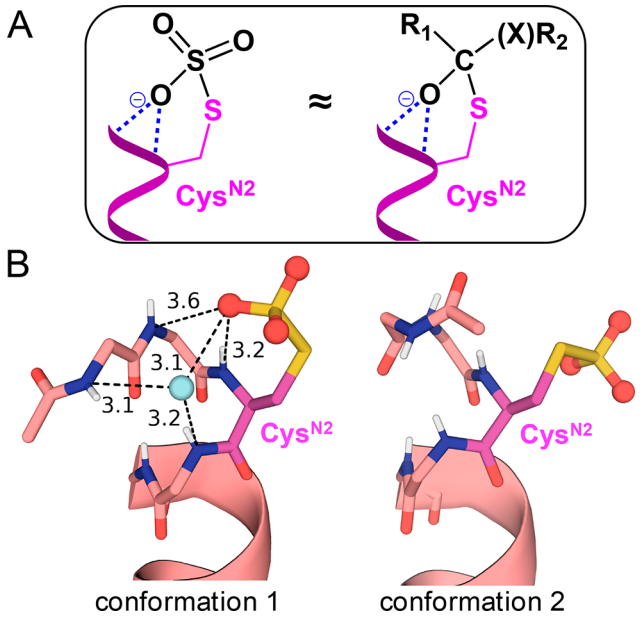


Figure 5. Cys-S-sulfonate mimicking tetrahedral intermediates. (A) Similarity of geometry and H-bonding interactions with catalytic protein for Cys-SO_3^- and tetrahedral intermediates for trans-thioesterification and amidation steps (X = S or NH_2^+). (B) Two distinct conformations of Cys-SO_3^- and helical N-termini observed in the crystal structure of **Het-N2-SO_3^-** (PDB ID: 7BEY). Conformation 1 contains Cys-SO_3^- H-bonded to adjacent amides. An H_2O molecule mediating additional H-bonding interactions is in cyan. The H-bond donor–acceptor distances are shown in Å. More information is shown in Table S3 and Figures S8 and S9.

bonded to backbone amides, confirming the predictions shown in Figure 1. In the second conformation the sulfonate group points toward the solvent. Such conformational isomerism can be important for catalytic activity, providing key stabilizing interactions present in the first conformation and being available for reaction or facilitating product release illustrated by the second conformation. Thus, the **Het-N2-SO_3^-** structure validates the design and sets the groundwork for future optimization of the sequence surrounding the binding site.

Acyl-Transferase Activity. The evaluation of the acyl-transferase catalytic activity of **Het-N2** and its comparison to the controls (i.e., **Het-Ncap**, **trunc-Cys-N2**, **Neg-N2**, glutathione, and MPAA (4-mercaptophenylacetic acid)) were performed at fixed conditions (acyl donor, 200 μM; acyl acceptor, 200 mM; catalyst, 100 μM in 50 mM sodium phosphate buffer at pH 7.5) using liquid chromatography electrospray ionization mass-spectrometry. Sequences and analytical data for acyl donor peptide-$^{\alpha}$thioester substrates are provided in Table S4. Tris (tris(hydroxymethyl)-aminomethane) was used as the standard acyl acceptor because of the pK_a 8.1 similar to N-terminal amine in peptides.[55] Background reaction (without protein catalyst) was also recorded. As an example, Figure 6A shows the kinetics of consumption of Ac-VALENF-$^{\alpha}$thioester substrate in the presence of different catalysts, and Figure 6B illustrates the evolution of branched adduct resulting from the trans-thioesterification between the substrate-thioester and the catalytic cysteine. In the presence of **Het-N2**, peptide-thioester transthioesterified most rapidly forming peptide-catalytic protein branched thioester adduct accompanied by the release of Mes (2-mercaptoethanesulfonate).

The accumulation of this covalently bound Cys-thioester intermediate supports a ping-pong mechanism as expected in the hypothesized catalytic mechanism outlined in Scheme 1. For comparison, the same reaction performed with **Het-Ncap**, **trunc-Cys-N2**, and Cys-containing glutathione proceeded much slower (Figure 6A, B), did not occur with **Neg-N2**, and when MPAA was added at c 100 μM the product of thioester exchange was not observed. It should be mentioned that MPAA is a common aryl-thiol catalyst for native chemical ligation, however, used at much higher concentrations (10–200 mM).[56] The subsequent rate-limiting step involved amidation by Tris that occurred in competition with hydrolysis. Thus, the ratio of amidation over hydrolysis (A/H) and the observed pseudo first order rate constant (k_{obs}) calculated from the depletion of initial peptide-$^{\alpha}$thioester substrate concentration were used as quantifiers to characterize and compare reactivity of the catalysts at different conditions (Table 1).

For the **Het-N2** construct showing appreciable catalytic effects in Figure 6A, k_{obs} is $(1.1 \pm 0.1) \times 10^{-5}$ s^{-1} (at concentrations 200 μM for Ac-VALENF-$^{\alpha}$thioester and 100 μM for **Het-N2**), which is 5.5-times higher than k_{obs} $(2.0 \pm 0.3) \times 10^{-6}$ s^{-1} for the background reaction without catalyst. The other catalysts influenced the rate of the reaction less significantly with only 2-fold acceleration over background for **Het-Ncap** (Table S5), highlighting the importance of the N2 Cys position. Furthermore, Cys^{N2}-to-Ala substitution in negative control **Neg-N2** resulted in kinetics being similar to background level.

Next, for **Het-N2** the reaction was performed with different concentrations of Ac-VALENF-$^{\alpha}$thioester (from 0.2 to 2 mM) (examples of kinetic traces are displayed in Figure S10). The corrected initial rates of catalyzed peptide-$^{\alpha}$thioester consumption obtained by subtracting the observed initial rates of background reaction from catalyzed reaction were fitted to Michaelis–Menten eq (Figure 6C). The estimated k_{cat} and K_m values are $(5.4 \pm 0.9) \times 10^{-5}$ s^{-1} and 0.3 ± 0.2 mM, respectively (standard error of the fitting is indicated), resulting in catalytic efficiency (k_{cat}/K_m) of ~0.2 M^{-1} s^{-1}. In addition, the catalytic turnover was confirmed by performing the experiment were fixed amounts of substrate (corresponding to starting concentration of 200 μM) were added two times

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330–3339

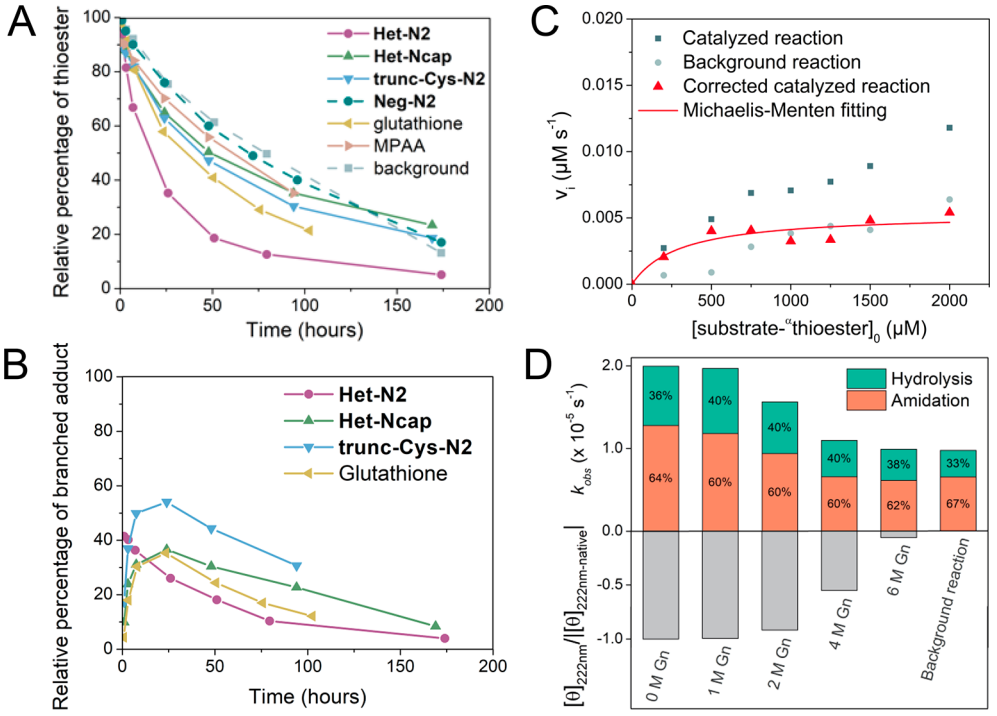


Figure 6. Characterization of catalytic activity. (A) Ac-VALENF-$^{\alpha}$thioester consumption in the presence of **Het-N2**, **Het-Ncap**, **trunc-N2-Cys**, glutathione, and MPAA (4-mercaptophenylacetic acid) in comparison to background reaction and negative control **Neg-N2**. (B) Changes in concentration of peptide−protein branched thioester resulting from the transthioesterification between Ac-VALENF-$^{\alpha}$thioester and **Het-N2**, **Het-Ncap**, **trunc-Cys-N2**, or glutathione. (C) Michaelis−Menten kinetics for the catalyzed reaction of Ac-VALENF-$^{\alpha}$thioester with **Het-N2**. (D) Observed rate constants and outcome of the reaction between Ac-GRLEEIDR-$^{\alpha}$thioester and Tris in the presence of **Het-N2** as a function of the concentration of Gn·HCl and the degree of folding of **Het-N2** based on CD measurements.

sequentially (Figure S11). The **Het-N2** processed the two batches of substrate with similar efficiency.

The catalysis of peptide-$^{\alpha}$thioester amidation/hydrolysis by **Het-N2** can be compared with hydrolytic activity using a continuous assay with *para*-nitrophenyl acetate (*p*NPA) as a substrate,[57] which moreover was previously tested for other de novo designed protein catalysts.[20,21] The data obtained at pH 7 and 22 °C and presented in Figures S12−S14 and Table S6 allowed to derive k_{cat} of 3.5 × 10^{-4} s^{-1} (corresponding to rate enhancement k_{cat}/k_{uncat} of 138) and K_m of 1.4 mM, resulting in a catalytic efficiency of 0.25 M^{-1} s^{-1}, which is close to the value obtained with the thioester substrate. Compared to previously reported most efficient de novo hydrolases, the catalytic efficiency of **Het-N2** is nearly 10 times lower;[20,21] however, no additional catalytic residues (e.g., histidine) are present in **Het-N2**, and thus the catalytic effect emerges mainly from the interactions with N-terminus of α-helix.

In addition to the Ac-VALENF-$^{\alpha}$thioester, other thioester substrates with different sequences (Table S4) and C-terminal amino acids (Arg, Gly) were assayed. Both Arg and Gly thioesters were found to undergo faster reactions in the presence and absence of **Het-N2** than Ac-VALENF-$^{\alpha}$thioester (Table 1 and Figure S15), which is most likely caused by intramolecular catalytic effect of Arg side chain and favorable steric effect due to the absence of side chain in Gly. However, in comparison to Ac-VALENF-$^{\alpha}$thioester, the k_{obs} was only two times faster compared to the respective background reaction at standard reactant concentrations (Table 1). Another difference is the A/H ratio, which for Phe thioester was found to be 0.8 indicating hydrolysis to be a major outcome, whereas for Arg and Gly thioester A/H was in the range of 1.7−2.4, which corresponds to a more pronounced amidation. Two distinct sequences that were studied for Arg thioester showed similar results (Table 1) suggesting that it is the C-terminal amino acid on the acyl donor that plays a decisive role in influencing the rate and its outcome with **Het-N2** catalyst.

The fact that the 12-residue **trunc-Cys-N2** is less efficient than **Het-N2** (Figure 6A) confirms that the overall stable structure of DSD is required for stabilization of oxyanion at N-terminus of α-helix. Using Ac-GRLEEIDR-$^{\alpha}$thioester and Tris as the acyl acceptor, we further proved the significance of the folded structure for the catalytic activity by increasing the concentration of a denaturant in the reaction mixture (Figure 6D). The k_{obs} was found to correlate to the degree of folding based on circular dichroism (CD) data at different Gn·HCl concentrations: when the protein **Het-N2** was highly α-helical, the catalytic activity was at a maximum, whereas at a high concentration (6 M) of Gn·HCl that caused unfolding of the protein, the rate decreased to a background level. The same effect on the catalytic activity of **Het-N2** was also observed upon increasing the concentration of urea (Figure S16).

The pH of the buffer also modulated the rate of reaction and A/H ratio. Upon increasing the pH, the rate of Ac-GRLEEIDR-$^{\alpha}$thioester consumption in the presence of **Het-N2** increased with an inflection point at 8.6, which was similar for the reaction without protein catalyst (Figure S17). The maximum of the amidation product was observed at pH 8 with a relative final product percentage of 66%. To rule out any effect of the buffer on catalysis, we performed assays with different concentrations of phosphate (i.e., 50, 200, and 500 mM) and in two other buffer systems (50 mM MES or 50 mM MOPS) at pH 7.5. The catalytic activity of **Het-N2** remained

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

Table 1. Kinetic Parameters for Various Acyl Transfer Reactions Catalyzed by Het-N2 (100 μM)[a]

Amidation vs Hydrolysis (A vs H)			
acyl donor with Tris	k_{obs} (s^{-1})[b]	A/H ratio[b]	acceleration[c]
Ac-GRLEEIDR-$^{\alpha}$SR	1.8×10^{-5}	2.4	2
Ac-VALENR-$^{\alpha}$SR	2.0×10^{-5}	2.0	1.8
Ac-VALENF-$^{\alpha}$SR	1.1×10^{-5}	0.8	5.5
Ac-LYRAG-$^{\alpha}$SR	2.4×10^{-5}	1.7	2.4
acyl acceptor with Ac-GRLEEIDR-$^{\alpha}$SR	k_{obs} (s^{-1})[b]	A/H ratio[b]	acceleration[c]
Tris	1.8×10^{-5}	2.4	2
$MeONH_2$	5.7×10^{-5}	2.3	3.2
$MeNH_2$	2.6×10^{-5}	0.4	3.1
Gly	1.5×10^{-5}	0.1	2.6
Gly-Gly	2.3×10^{-5}	0.15	2.9
Gly-Gly-Gly	2.5×10^{-5}	0.3	2.7
histamine	9.6×10^{-5}	0.5	1.2
Esterification vs Hydrolysis (E vs H)			
acyl acceptor with Ac-GRLEEIDR-$^{\alpha}$SR	k_{obs} (s^{-1})	E/H ratio	acceleration[c]
trifluoroethanol	5.4×10^{-5}	max.: 1.8	4.4
methanol	2.2×10^{-5}	0.2	3
ethanol	no reaction		
Further Acceleration			
additives	k_{obs}(s^{-1})	A/H ratio	acceleration[d]
NaN_3 (154 mM)	15.9×10^{-5}	0.1	8.8 (17.7)
KCN (200 mM)	15.2×10^{-5}	0.4	8.4 (16.9)
imidazole (200 mM)	10.3×10^{-5}	1.0	5.7 (11.4)

[a]Additives were used for reaction between Tris (200 mM) and Ac-GRLEEIDR-$^{\alpha}$thioester (200 μM). [b]Estimated standard errors for k_{obs} and A/H are within 16 and 25% of their values, respectively. [c]Acceleration compared to background reaction without catalyst. [d]Acceleration compared to the reaction without additives and with catalyst, and in brackets to the reaction without additives and catalyst.

similar in different buffer conditions, with a slight increase of hydrolysis at higher concentration of phosphate (500 mM) as shown in Figure S18.

Catalytic Promiscuity. To test the ability of **Het-N2** to catalyze acyl transfer reactions other than amidation with Tris, we carried out the catalytic assays with several other acyl acceptors. Reactions of miscellaneous amines and few alcohols with Ac-GRLEEIDR-$^{\alpha}$thioester were studied (Table 1). Methoxyamine with lower pK_a of 4.7 provided with the fastest kinetics ($k_{obs} = 5.7 \times 10^{-5}$ s^{-1}) that represents more than three times acceleration over background and amidation was more favored (A/H = 2.33). In this particular case, the rate enhancement is likely resulting from the α-effect well-known for methoxyamine, where nucleophilicity of amine is increased because of the adjacent oxygen atom containing lone electron pairs.[58] In contrast, substrates with higher pK_a like methylamine (10.6) resulted in the excess of hydrolyzed product (A/H = 0.4) even if the rate of thioester consumption ($k_{obs} = 2.6 \times 10^{-5}$ s^{-1}) was similar to the reaction with Tris. We also investigated catalysis of peptide bond formation by **Het-N2** using glycine (pK_a = 9.6), a dipeptide Gly-Gly (pK_a = 8.2) and a tripeptide Gly-Gly-Gly (pK_a = 7.8)[59] as acyl acceptors. The Ac-GRLEEIDR-$^{\alpha}$thioester was less reactive with glycine and more prone to hydrolysis than with dipeptide (A/H = 0.1 and 0.15, respectively) and tripeptide (A/H = 0.3), illustrating an influence of the pK_a of acyl acceptor on the reactivity (Table 1). A peculiar change of reactivity was noticed for histamine (pK_a (amine) = 9.8 and pK_a (imidazole) = 6.04). A rather high observed rate constant was detected ($k_{obs} = 9.6 \times 10^{-5}$ s^{-1}); however, in the background reaction without **Het-N2**, the consumption of thioester was almost as fast and amidation was more pronounced compared to the reaction with **Het-N2** (A/H for uncatalyzed reaction is 0.9 versus 0.5 for catalyzed). This may indicate an additional catalytic effect of the imidazole ring on the initial thioester that is more accessible in solution than the Cys-thioester intermediate covalently bound to **Het-N2**. The influence of imidazole used as separate additive (i.e., in synergy with the catalytic protein) will be presented hereafter.

Alcohols as acyl acceptors demonstrated moderate reactivity (Table 1). Esterification with trifluoroethanol, an alcohol with a reduced pK_a of 12.4, was observed in the presence of catalytic protein **Het-N2** with $k_{obs} = 5.4 \times 10^{-5}$ s^{-1} and up to 61% of detected ester formation compared to $k_{obs} = 1.2 \times 10^{-5}$ s^{-1} and 39% of ester for background reaction, but the resulting ester product was not stable and decomposed through hydrolysis after reaching a maximum of ester production. Methanol (pK_a = 15.2) reached a conversion of 17% of ester product with a 3-fold faster rate compared to the uncatalyzed reaction. Interestingly, ethanol (pK_a = 15.9) did not react in the acyl transfer with Ac-GRLEEIDR-$^{\alpha}$thioester.

Furthermore, we demonstrated that addition of azide and cyanide was able to modify kinetics of the acyl transfer reaction between Ac-GRLEEIDR-$^{\alpha}$thioester and Tris. The increasing concentrations of these additives with and without catalytic proteins led to higher observed rates of reactions and higher percentage of hydrolysis (see Table 1 and Figure S19). The most contrasting effect was discovered in the presence of azide, where the high concentration of NaN_3 (154 mM) in synergy with **Het-N2** contributed to 18-fold acceleration compared to background reaction (without additives and catalyst) and a 9-fold acceleration compared to conditions without additives and with catalyst. However, the A/H ratio dropped from 2.4 to 0.1. A similar tendency was observed with KCN in the reaction mixture (Figure S19). Presumably, transient acyl azides and acyl cyanides are the reactive intermediates that facilitate hydrolysis,[60,61] although these intermediates were not directly detected by LC-MS.

After observing the unusual reactivity of histamine as acyl acceptor, we evaluated the effect of imidazole as an additive. When imidazole was at low concentrations (2 and 20 mM), no considerable reactivity difference was observed, whereas at higher concentration (200 mM), observed rate constant increased nearly 6-fold to 10.3×10^{-5} s^{-1}, with amidation being more favored compared to with the other additives with a final A/H ratio of 1.0 (Figure S19). Previously, imidazole added at high concentrations (2.5 M) was reported to catalyze native chemical ligation.[62]

CONCLUSIONS

We describe the successful completion of the first step toward the design of a stable catalyst for thioester ligation. Our initial goals were to build a protein with a small active site cavity capable of activating a Cys thiol for attack of thioesters, and stabilizing the anionic high-energy tetrahedral intermediates in the reaction mechanism. The designed **Het-N2** was found to modestly accelerate a variety of acyl transfer reactions with a wide range of substrates. Thus, **Het-N2** can be compared to primitive (or primordial) enzymes possessing moderate catalytic efficiency and substrate promiscuity.[63] Such modest catalytic efficiency coupled to catalytic promiscuity is believed

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

to be essential for evolution of enzymes with distinct specificities.[63] The next step is to apply computational design and combinatorial library approaches taking advantage of the modular architecture of **Het-N2** scaffold to vary the composition at the active site and improve its catalytic properties.

The successful design of a catalyst also requires that the multiple chemical steps along the reaction path be understood, so that they can be optimized separately as needed. Enabling selective catalysis of amidation versus hydrolysis by design is a particularly challenging aspect. Recently, several X-ray structures were reported for proteolytic enzymes that also possess peptide macrocyclization or ligation activities and provide insights into how such selectivity can be achieved.[64−69] The principal catalytic residue is either Ser[64,65] or Cys[66−69] and catalysis is assisted by few other residues such as His, Asp/Glu, Tyr, and backbone amides. The enhancement of ligation activity over proteolysis is believed to occur by exclusion of water from the acyl-enzyme covalent adduct via additional structural elements such as "capping" lid domain,[64] binding of follower peptide[65] or hydrophobicity of the active site.[66−69] Remarkably, structural comparison of the enzymes with ligation (amidation) activity with the homologous variants possessing enhanced proteolytic (hydrolysis) activity showed that the structures of the related ligases and proteases are very similar and only subtle changes were observed in their catalytic pockets.[65−69] Furthermore, minimal changes in active site composition are sufficient to repurpose a ligase into a protease and vice versa.[66−69]

Thus, it is not surprising that the branched thioester adduct observed in our work where electrophilic thioester moiety is exposed to aqueous buffer undergoes fast hydrolysis in competition with amidation. The choice to use Cys and not Ser as the principle nucleophilic residue is not sufficient to bias the reactivity toward the reaction with amines, despite the reported higher relative reaction rates of thioesters toward nitrogen nucleophiles than oxygen nucleophiles in comparison to oxoesters.[42,70] A possible way to further enhance selectivity toward amidation would be to incorporate selenocysteine as a principle catalytic residue because selenoesters are known to be ca. 100-times more reactive with amines than thioesters, whereas hydrolysis rates are on the same order of magnitude;[71,72] however, it would require inert oxygen-free conditions due to the associated selenol's redox chemistry.[73]

Our data indicate that the environment nearby catalytic Cys can slightly modify the amidation versus hydrolysis ratio. Thus, with the main goal being to achieve higher catalytic activities, our future work will be directed toward understanding the structural requirements for amidation versus hydrolysis by optimizing the sequence of the N-terminal residues of the two helices for a particular reaction path and developing new practical tools for peptide ligation and protein modification.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c10053.

Materials and methods, details on chemical synthesis, crystallization, data collection and structure refinement, catalytic assays, and supplementary tables and figures (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Vladimir Torbeev − *Institut de Science et d'Ingénierie Supramoléculaires (ISIS), International Center for Frontier Research in Chemistry (icFRC), University of Strasbourg, CNRS (UMR 7006), Strasbourg 67000, France;* orcid.org/0000-0002-7865-1044; Email: torbeev@unistra.fr

William F. DeGrado − *Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, University of California San Francisco, San Francisco, California 94158-9001, United States;* orcid.org/0000-0003-4745-263X; Email: william.degrado@ucsf.edu

Authors

Elise A. Naudin − *Institut de Science et d'Ingénierie Supramoléculaires (ISIS), International Center for Frontier Research in Chemistry (icFRC), University of Strasbourg, CNRS (UMR 7006), Strasbourg 67000, France*

Alastair G. McEwen − *Integrated Structural Biology Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS (UMR 7104), INSERM (U1258), University of Strasbourg, Illkirch 67404, France*

Sophia K. Tan − *Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, University of California San Francisco, San Francisco, California 94158-9001, United States;* orcid.org/0000-0002-6181-7810

Pierre Poussin-Courmontagne − *Integrated Structural Biology Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS (UMR 7104), INSERM (U1258), University of Strasbourg, Illkirch 67404, France*

Jean-Louis Schmitt − *Institut de Science et d'Ingénierie Supramoléculaires (ISIS), International Center for Frontier Research in Chemistry (icFRC), University of Strasbourg, CNRS (UMR 7006), Strasbourg 67000, France;* orcid.org/0000-0001-8667-7941

Catherine Birck − *Integrated Structural Biology Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS (UMR 7104), INSERM (U1258), University of Strasbourg, Illkirch 67404, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c10053

Author Contributions

E.A.N., A.G.M., P.P.-C., J.-L.S., C.B., and V.T. performed the experiments. E.A.N., A.G.M., S.K.T., C.B., W.F.D., and V.T. analyzed the data. E.A.N., W.F.D., and V.T. wrote the manuscript with the consultation from other authors.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been funded by the European Research Council (ERC-2016-StG, grant 715062-HiChemSynPro) and the LabEx CSC (ANR-10-LABX-0026_CSC). E.A.N. is grateful for an MRT doctoral fellowship from the French government. The European Synchrotron Radiation Facility and Synchrotron Soleil are acknowledged for provision of synchrotron beamtime (proposal numbers MX-1837, MX-1945, and 20181001) and we thank David Flot, Bart Van Laer, and Pierre Montaville for assistance in using beamlines ID23-1,

ID30A-3, and Proxima 1, respectively. We also acknowledge the support and the use of resources of the French Infrastructure for Integrated Structural Biology FRISBI ANR-10-INBS-05 and of Instruct-ERIC as well as the grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the program Investissements d'Avenir ANR-10-IDEX-0002-02.

REFERENCES

(1) Zanghellini, A. De Novo Computational Enzyme Design. *Curr. Opin. Biotechnol.* **2014**, *29*, 132−138.
(2) Kries, H.; Blomberg, R.; Hilvert, D. De Novo Enzymes by Computational Design. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221−228.
(3) Hilvert, D. Design of Protein Catalysts. *Annu. Rev. Biochem.* **2013**, *82*, 447−470.
(4) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. De Novo Computational Design of Retro-Aldol Enzymes. *Science* **2008**, *319*, 1387−1391.
(5) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008**, *453*, 190−195.
(6) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St. Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010**, *329*, 309−313.
(7) Bolon, D. N.; Mayo, S. L. Enzyme-like Proteins by Computational Design. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 14274−14279.
(8) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J. T.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A.; Xiao, R.; Seetharaman, J.; Su, M.; Vorobiev, S.; Lew, S.; Forouhar, F.; Kornhaber, G. J.; Hunt, J. F.; Montelione, G. T.; Tong, L.; Houk, K. N.; Hilvert, D.; Baker, D. Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis. *J. Am. Chem. Soc.* **2012**, *134*, 16197−16206.
(9) Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N. Computational Enzyme Design. *Angew. Chem., Int. Ed.* **2013**, *52*, 5700−5725.
(10) Korendovych, I. V.; DeGrado, W. F. Catalytic Efficiency of Designed Catalytic Proteins. *Curr. Opin. Struct. Biol.* **2014**, *27*, 113−121.
(11) Zeymer, C.; Hilvert, D. Directed Evolution of Protein Catalysts. *Annu. Rev. Biochem.* **2018**, *87*, 131−157.
(12) Bunzel, H. A.; Garrabou, X.; Pott, M.; Hilvert, D. Speeding up Enzyme Discovery and Engineering with Ultrahigh-Throughput Methods. *Curr. Opin. Struct. Biol.* **2018**, *48*, 149−156.
(13) Studer, S.; Hansen, D. A.; Pianowski, Z. L.; Mittl, P. R. E.; Debon, A.; Guffy, S. L.; Der, B. S.; Kuhlman, B.; Hilvert, D. Evolution of a Highly Active and Enantiospecific Metalloenzyme from Short Peptides. *Science* **2018**, *362*, 1285−1288.
(14) Nanda, V.; Koder, R. L. Designing Artificial Enzymes by Intuition and Computation. *Nat. Chem.* **2010**, *2*, 15−24.
(15) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537*, 320−327.
(16) Korendovych, I. V.; DeGrado, W. F. De Novo Protein Design, a Retrospective. *Q. Rev. Biophys.* **2020**, *53*, No. e3.
(17) Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Grütter, M. G.; Privett, H. K.; Mayo, S. L.; Hilvert, D. Precision Is Essential for Efficient Catalysis in an Evolved Kemp Eliminase. *Nature* **2013**, *503*, 418−421.
(18) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. Evolution of a Designed Retro-Aldolase Leads to Complete Active Site Remodeling. *Nat. Chem. Biol.* **2013**, *9*, 494−498.
(19) Preiswerk, N.; Beck, T.; Schulz, J. D.; Milovník, P.; Mayer, C.; Siegel, J. B.; Baker, D.; Hilvert, D. Impact of Scaffold Rigidity on the Design and Evolution of an Artificial Diels-Alderase. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 8013−8018.
(20) Marshall, L. R.; Zozulia, O.; Lengyel-Zhand, Z.; Korendovych, I. V. Minimalist de Novo Design of Protein Catalysts. *ACS Catal.* **2019**, *9*, 9265−9275.
(21) Burton, A. J.; Thomson, A. R.; Dawson, W. M.; Brady, R. L.; Woolfson, D. N. Installing Hydrolytic Activity into a Completely de Novo Protein Framework. *Nat. Chem.* **2016**, *8*, 837−844.
(22) Watkins, D. W.; Jenkins, J. M. X.; Grayson, K. J.; Wood, N.; Steventon, J. W.; Le Vay, K. K.; Goodwin, M. I.; Mullen, A. S.; Bailey, H. J.; Crump, M. P.; MacMillan, F.; Mulholland, A. J.; Cameron, G.; Sessions, R. B.; Mann, S.; Anderson, J. L. R. Construction and in Vivo Assembly of a Catalytically Proficient and Hyperthermostable de Novo Enzyme. *Nat. Commun.* **2017**, *8*, 358.
(23) Moroz, Y. S.; Dunston, T. T.; Makhlynets, O. V.; Moroz, O. V.; Wu, Y.; Yoon, J. H.; Olsen, A. B.; McLaughlin, J. M.; Mack, K. L.; Gosavi, P. M.; van Nuland, N. A. J.; Korendovych, I. V. New Tricks for Old Proteins: Single Mutations in a Nonenzymatic Protein Give Rise to Various Enzymatic Activities. *J. Am. Chem. Soc.* **2015**, *137*, 14905−14911.
(24) Moroz, O. V.; Moroz, Y. S.; Wu, Y.; Olsen, A. B.; Cheng, H.; Mack, K. L.; McLaughlin, J. M.; Raymond, E. A.; Zhezherya, K.; Roder, H.; Korendovych, I. V. A Single Mutation in a Regulatory Protein Produces Evolvable Allosterically Regulated Catalyst of Nonnatural Reaction. *Angew. Chem., Int. Ed.* **2013**, *52*, 6246−6249.
(25) Kent, S. B. H. Novel Protein Science Enabled by Total Chemical Synthesis. *Protein Sci.* **2019**, *28*, 313−328.
(26) Bondalapati, S.; Jbara, M.; Brik, A. Expanding the Chemical Toolbox for the Synthesis of Large and Uniquely Modified Proteins. *Nat. Chem.* **2016**, *8*, 407−418.
(27) Dang, B.; Wu, H.; Mulligan, V. K.; Mravic, M.; Wu, Y.; Lemmin, T.; Ford, A.; Silva, D.-A.; Baker, D.; DeGrado, W. F. De Novo Design of Covalently Constrained Mesosize Protein Scaffolds with Unique Tertiary Structures. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 10852−10857.
(28) Mandal, K.; Uppalapati, M.; Ault-Riché, D.; Kenney, J.; Lowitz, J.; Sidhu, S. S.; Kent, S. B. H. Chemical Synthesis and X-Ray Structure of a Heterochiral {D-Protein Antagonist plus Vascular Endothelial Growth Factor} Protein Complex by Racemic Crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14779−14784.
(29) Weinstock, M. T.; Jacobsen, M. T.; Kay, M. S. Synthesis and Folding of a Mirror-Image Enzyme Reveals Ambidextrous Chaperone Activity. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 11679−11684.
(30) Mandal, K.; Pentelute, B. L.; Bang, D.; Gates, Z. P.; Torbeev, V. Y.; Kent, S. B. H. Design, Total Chemical Synthesis, and X-Ray Structure of a Protein Having a Novel Linear-Loop Polypeptide Chain Topology. *Angew. Chem., Int. Ed.* **2012**, *51*, 1481−1486.
(31) Boehringer, R.; Kieffer, B.; Torbeev, V. Total Chemical Synthesis and Biophysical Properties of a Designed Soluble 24 KDa Amyloid Analogue. *Chem. Sci.* **2018**, *9*, 5594−5599.
(32) Kochendoerfer, G. G.; Chen, S. Y.; Mao, F.; Cressman, S.; Traviglia, S.; Shao, H.; Hunter, C. L.; Low, D. W.; Cagle, E. N.; Carnevali, M.; Gueriguian, V.; Keogh, P. J.; Porter, H.; Stratton, S. M.; Wiedeke, M. C.; Wilken, J.; Tang, J.; Levy, J. J.; Miranda, L. P.; Crnogorac, M. M.; Kalbag, S.; Botti, P.; Schindler-Horvat, J.; Savatski, L.; Adamson, J. W.; Kung, A.; Kent, S. B.; Bradburne, J. A. Design and Chemical Synthesis of a Homogeneous Polymer-modified Erythropoiesis Protein. *Science* **2003**, *299*, 884−887.
(33) Nguyen, G. K. T.; Wang, S.; Qiu, Y.; Hemu, X.; Lian, Y.; Tam, J. P. Butelase 1 Is an Asx-Specific Ligase Enabling Peptide Macrocyclization and Synthesis. *Nat. Chem. Biol.* **2014**, *10*, 732−738.
(34) Piotukh, K.; Geltinger, B.; Heinrich, N.; Gerth, F.; Beyermann, M.; Freund, C.; Schwarzer, D. Directed Evolution of Sortase A Mutants with Altered Substrate Selectivity Profiles. *J. Am. Chem. Soc.* **2011**, *133*, 17536−17539.
(35) Zheng, N.; Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu. Rev. Biochem.* **2017**, *86*, 129−157.

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339

(36) Weeks, A. M.; Wells, J. A. Subtiligase-Catalyzed Peptide Ligation. *Chem. Rev.* **2020**, *120*, 3127−3160.
(37) Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A. A Designed Peptide Ligase for Total Synthesis of Ribonuclease A with Unnatural Catalytic Residues. *Science* **1994**, *266*, 243−247.
(38) Toplak, A.; Nuijens, T.; Quaedflieg, P. J. L. M.; Wu, B.; Janssen, D. B. Peptiligase, an Enzyme for Efficient Chemoenzymatic Peptide Synthesis and Cyclization in Water. *Adv. Synth. Catal.* **2016**, *358*, 2140−2147.
(39) Schmidt, M.; Toplak, A.; Quaedflieg, P. J. L. M.; Ippel, H.; Richelle, G. J. J.; Hackeng, T. M.; van Maarseveen, J. H.; Nuijens, T. Omniligase-1: A Powerful Tool for Peptide Head-to-Tail Cyclization. *Adv. Synth. Catal.* **2017**, *359*, 2050−2055.
(40) Simón, L.; Goodman, J. M. Enzyme Catalysis by Hydrogen Bonds: The Balance between Transition State Binding and Substrate Binding in Oxyanion Holes. *J. Org. Chem.* **2010**, *75*, 1831−1840.
(41) Poole, L. B. The Basics of Thiols and Cysteines in Redox Biology and Chemistry. *Free Radical Biol. Med.* **2015**, *80*, 148−157.
(42) Yang, W.; Drueckhammer, D. G. Understanding the Relative Acyl-Transfer Reactivity of Oxoesters and Thioesters: Computational Analysis of Transition State Delocalization Effects. *J. Am. Chem. Soc.* **2001**, *123*, 11004−11009.
(43) Finking, R.; Marahiel, M. A. Biosynthesis of Nonribosomal Peptides. *Annu. Rev. Microbiol.* **2004**, *58*, 453−488.
(44) Smith, S.; Witkowski, A.; Joshi, A. K. Structural and Functional Organization of the Animal Fatty Acid Synthase. *Prog. Lipid Res.* **2003**, *42*, 289−317.
(45) Perler, F. B. Protein Splicing Mechanisms and Applications. *IUBMB Life* **2005**, *57*, 469−476.
(46) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoet, M.; Monbaliu, J.-C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **2019**, *119*, 7328−7442.
(47) Hojo, H.; Aimoto, S. Polypeptide Synthesis Using the S-Alkyl Thioester of a Partially Protected Peptide Segment. Synthesis of the DNA-Binding Domain of c-Myb Protein (142−193)-NH_2. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111−117.
(48) Dawson, P. E.; Muir, T.; Clark-Lewis, I.; Kent, S. B. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266*, 776−779.
(49) Andrew, C. D.; Warwicker, J.; Jones, G. R.; Doig, A. J. Effect of Phosphorylation on α-Helix Stability as a Function of Position. *Biochemistry* **2002**, *41*, 1897−1905.
(50) Tan, S. K.; Fong, K. P.; Polizzi, N. F.; Sternisha, A.; Slusky, J. S. G.; Yoon, K.; DeGrado, W. F.; Bennett, J. S. Modulating Integrin αIIbβ3 Activity through Mutagenesis of Allosterically Regulated Intersubunit Contacts. *Biochemistry* **2019**, *58*, 3251−3259.
(51) Ogihara, N. L.; Ghirlanda, G.; Bryson, J. W.; Gingery, M.; DeGrado, W. F.; Eisenberg, D. Design of Three-Dimensional Domain-Swapped Dimers and Fibrous Oligomers. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1404−1409.
(52) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. In Situ Neutralization in Boc-Chemistry Solid Phase Peptide Synthesis. *Int. J. Pept. Res. Ther.* **2007**, *13*, 31−44.
(53) Behrendt, R.; White, P.; Offer, J. Advances in Fmoc Solid-Phase Peptide Synthesis. *J. Pept. Sci.* **2016**, *22*, 4−27.
(54) Grigoryan, G.; DeGrado, W. F. Probing Designability via a Generalized Model of Helical Bundle Geometry. *J. Mol. Biol.* **2011**, *405*, 1079−1100.
(55) Sievers, A.; Beringer, M.; Rodnina, M. V.; Wolfenden, R. The Ribosome as an Entropy Trap. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 7897−7901.
(56) Johnson, E. C. B.; Kent, S. B. H. Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction. *J. Am. Chem. Soc.* **2006**, *128*, 6640−6646.
(57) Bender, M. L.; Kezdy, F. J.; Wedler, F. C. Alpha-Chymotrypsin: Enzyme Concentration and Kinetics. *J. Chem. Educ.* **1967**, *44*, 84−88.
(58) Jencks, W. P.; Carriuolo, J. Reactivity of Nucleophilic Reagents toward Esters. *J. Am. Chem. Soc.* **1960**, *82*, 1778−1786.
(59) Daragan, V. A.; Mayo, K. H. Peptide Dynamics in Triglycine: Coupling of Internal Bond Rotations and Overall Molecular Tumbling. *J. Phys. Chem.* **1994**, *98*, 10949−10956.
(60) Hünig, S.; Schaller, R. The Chemistry of Acyl Cyanides. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 36−49.
(61) Eggerer, H.; Stadtman, E. R.; Poston, J. M. On the Role of Acetyl Cyanide in the Cyanide-Induced Acetylation of Amino Acids by Enzymes of Clostridium Kluyveri. *Arch. Biochem. Biophys.* **1962**, *98*, 432−443.
(62) Sakamoto, K.; Tsuda, S.; Mochizuki, M.; Nohara, Y.; Nishio, H.; Yoshiya, T. Imidazole-Aided Native Chemical Ligation: Imidazole as a One-Pot Desulfurization-Amenable Non-Thiol-Type Alternative to 4-Mercaptophenylacetic Acid. *Chem. - Eur. J.* **2016**, *22*, 17940−17944.
(63) Khersonsky, O.; Roodveldt, C.; Tawfik, D. S. Enzyme Promiscuity: Evolutionary and Mechanistic Aspects. *Curr. Opin. Chem. Biol.* **2006**, *10*, 498−508.
(64) Koehnke, J.; Bent, A.; Houssen, W. E.; Zollman, D.; Morawitz, F.; Shirran, S.; Vendome, J.; Nneoyiegbe, A. F.; Trembleau, L.; Botting, C. H.; Smith, M. C. M.; Jaspars, M.; Naismith, J. H. The Mechanism of Patellamide Macrocyclization Revealed by the Characterization of the PatG Macrocyclase Domain. *Nat. Struct. Mol. Biol.* **2012**, *19*, 767−772.
(65) Chekan, J. R.; Estrada, P.; Covello, P. S.; Nair, S. K. Characterization of the Macrocyclase Involved in the Biosynthesis of RiPP Cyclic Peptides in Plants. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 6551−6556.
(66) Zauner, F. B.; Elsässer, B.; Dall, E.; Cabrele, C.; Brandstetter, H. Structural Analyses of Arabidopsis Thaliana Legumain γ Reveal Differential Recognition and Processing of Proteolysis and Ligation Substrates. *J. Biol. Chem.* **2018**, *293*, 8934−8946.
(67) Haywood, J.; Schmidberger, J. W.; James, A. M.; Nonis, S. G.; Sukhoverkov, K. V.; Elias, M.; Bond, C. S.; Mylne, J. S. Structural Basis of Ribosomal Peptide Macrocyclization in Plants. *eLife* **2018**, *7*, No. e32955.
(68) Hemu, X.; El Sahili, A.; Hu, S.; Wong, K.; Chen, Y.; Wong, Y. H.; Zhang, X.; Serra, A.; Goh, B. C.; Darwis, D. A.; Chen, M. W.; Sze, S. K.; Liu, C. F.; Lescar, J.; Tam, J. P. Structural Determinants for Peptide-bond Formation by Asparaginyl Ligases. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 11737−11746.
(69) James, A. M.; Haywood, J.; Leroux, J.; Ignasiak, K.; Elliott, A. G.; Schmidberger, J. W.; Fisher, M. F.; Nonis, S. G.; Fenske, R.; Bond, C. S.; Mylne, J. S. The Macrocyclizing Protease Butelase 1 Remains Autocatalytic and Reveals the Structural Basis for Ligase Activity. *Plant J.* **2019**, *98*, 988−999.
(70) Connors, K. A.; Bender, M. L. The Kinetics of Alkaline Hydrolysis and n-Butylaminolysis of Ethyl p-Nitrobenzoate and Ethyl p-Nitrothiolbenzoate. *J. Org. Chem.* **1961**, *26*, 2498−2504.
(71) Chu, S.-H.; Mautner, H. G. Analogs of Neuroeffectors. V. Neighboring-Group Effects in the Reactions of Esters, Thiolesters, and Selenolesters. The Hydrolysis and Aminolysis of Benzoylcholine, Benzoylthiolcholine, Benzoylselenolcholine, and of their Dimethylamino Analogs. *J. Org. Chem.* **1966**, *31*, 308−312.
(72) Wu, Z.-P.; Hilvert, D. Conversion of a Protease into an Acyl Transferase: Selenolsubtilisin. *J. Am. Chem. Soc.* **1989**, *111*, 4513−4514.
(73) Wu, Z.-P.; Hilvert, D. Selenosubtilisin as a Glutathione Peroxidase Mimic. *J. Am. Chem. Soc.* **1990**, *112*, 5647−5648.

https://dx.doi.org/10.1021/jacs.0c10053
J. Am. Chem. Soc. 2021, 143, 3330−3339