ChemPubSoc Europe

DOI: 10.1002/cbic.201800733

CHEMBIOCHEM Concepts

Selection, Addiction and Catalysis: Emerging Trends for the Incorporation of Noncanonical Amino Acids into Peptides and Proteins in Vivo

Clemens Mayer*[a]

Expanding the genetic code of organisms by incorporating noncanonical amino acids (ncAAs) into target proteins through the suppression of stop codons in vivo has profoundly impacted how we perform protein modification or detect proteins and their interaction partners in their native environment. Yet, with genetic code expansion strategies maturing over the past 15 years, new applications that make use—or indeed repurpose—these techniques are beginning to emerge. This Concept article highlights three of these developments: 1) The incorporation of ncAAs for the biosynthesis and *selection* of bioactive macrocyclic peptides with novel ring architectures, 2) synthetic biocontainment strategies based on the *addiction* of microorganisms to ncAAs, and 3) enzyme design strategies, in which ncAAs with unique functionalities enable the *catalysis* of new-to-nature reactions. Key advances in all three areas are presented and potential future applications discussed.

Background: Stop Codon Suppression

The site-selective incorporation of ncAAs into peptides and proteins in vivo can be achieved through the suppression of a stop codon by the action of an orthogonal translation system (OTS, Figure 1 A).[1–3] An OTS is comprised of an engineered aminoacyl tRNA synthetase (aaRS), which loads a ncAA onto its corresponding tRNA, while none of the OTS components interacts with endogenous amino acids, aaRSs or tRNAs (=orthogonal, Figure 1 B).[4] The incorporation of ncAAs by an OTS is achieved if: 1) protein engineering efforts have changed the cognate substrate preference of the aaRS to a ncAA of choice, and 2) the tRNA features the anticodon sequence (e.g., CUA) to a stop codon, for example, UAG (Figure 1 A). Charged tRNAs are then recruited to the ribosome, where in-frame UAG stop codons in mRNAs are suppressed, resulting in the site-selective incorporation of a ncAA into the nascent peptide chain. Despite typically modest suppression efficiencies and the fact that not every ncAA is genetically encodable (metabolic stability, limited uptake, etc.), this strategy has proven exceptionally versatile.[3,4] Indeed, more than 150 ncAAs have been successfully incorporated into peptides and proteins of interest in a variety of model organisms. Since its conception more than 15 years ago, stop codon suppression as a strategy has predominantly been applied for introducing ncAAs with functional groups that enable site-selective protein modification and/or elucidating, altering or regulating protein function.[3,5] However, more recently, OTSs have also been repurposed for other tasks. This Concept article will highlight recent developments for which ncAA incorporation has proven particularly impactful and is divided into three sections: 1) Selection, 2) Addiction, and 3) Catalysis.

Selection: ncAAs in Peptide Macrocyclization

Macrocyclic peptides (MPs) are privileged scaffolds for the development of chemical probes and therapeutics.[6,7] Combining a high degree of functional complexity with a restricted conformational flexibility make MPs well-suited to achieve tight binding to notoriously difficult targets, such as biomolecular interfaces.[8] Moreover, peptide macrocyclization is a straightforward means to reduce protease degradation and can facilitate cellular uptake.[9] Lastly, peptides are genetically encodable and lend themselves to massive parallel screening and selection efforts that allow for identifying tight binders from randomized populations. As a result, methods to genetically encode MP libraries are sought after, yet the number of cyclization strategies—and therefore, the accessible ring geometries—remain limited.[10]

Bioactive MPs found in nature often feature ncAAs and utilize unique functionalities placed in their side chains to promote macrocyclization.[11] In an effort to mimic such natural products, the Suga group has led efforts to encode MPs containing a wide variety of ncAAs in vitro.[12] Taking advantage of well-defined, reconstituted translation systems, in which natural aaRS/tRNA pairs are replaced with synthetically pre-charged tRNAs featuring non-standard building blocks,[13] the group

[a] Dr. C. Mayer
Stratingh Institute for Chemistry, University of Groningen
Nijenborgh 4, 9747 AG Groningen (The Netherlands)
E-mail: c.mayer@rug.nl

The ORCID identification numbers for the author of this article can be found under https://doi.org/10.1002/cbic.201800733.

This article is part of the young researchers' issue ChemBioTalents. To view the complete issue, visit http://chembiochem.org/chembiotalents

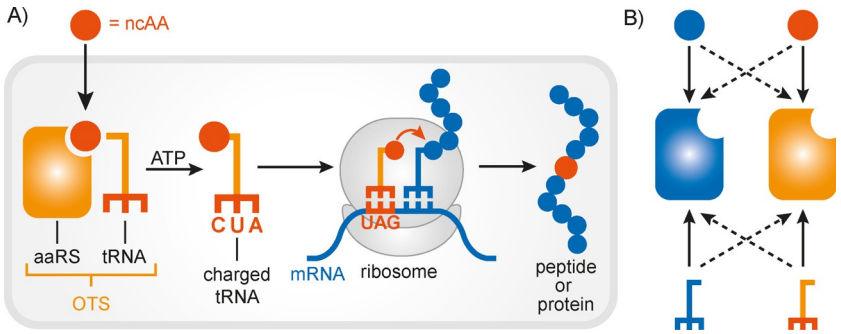


Figure 1. A) A ncAA is taken up into the cell, where it is charged onto an orthogonal tRNA through the action of an engineered aaRS. Once charged, this tRNA, which features the complementary sequence of a stop codon (e.g., CUA), is recruited to the ribosome, where it suppresses an in-frame stop codon (UAG) located on an mRNA. B) Schematic representation of orthogonality with respect to the OTS. Note that engineered aaRS and tRNA (orange) do not interact (dashed arrows) with the endogenous aaRS/tRNA pairs (blue).

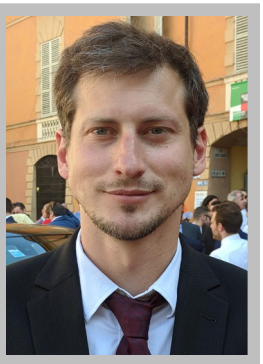
generated MPs with up to eleven ncAAs in their sequence.[14] Placing uniquely reactive functional groups in their side chains facilitated the installation of novel ring geometries that could be interfaced with traditional mRNA display[15] and allowed for selection of MP inhibitors against a variety of target proteins.[12]

Performing analogous experiments in vivo is significantly more challenging, as multiple, chemically distinct ncAAs cannot reliably be incorporated simultaneously. Moreover, cyclization strategies need to be compatible with the plethora of reactive groups inside the producing organism. Nevertheless, in vivo strategies are desirable, as they can be linked to phenotypic screens or selection against potential targets in their native environment.[12,16,17] Efforts toward developing such strategies have been led by the Fasan group, who began utilizing ncAAs with unique functionalities to access MPs from genetically encoded, yet purified peptides.[18–21] Guided by these efforts, the group established a set of criteria a ncAA needs to fulfill to enable cyclization in vivo.[22] First, suitable ncAAs should be amenable to incorporation through stop codon suppression, a fact that limits their structures tyrosine and pyrrolysine analogues, as OTSs based on aaRSs for these amino acids are currently most advanced. Moreover, functionalities introduced by ncAAs need to be sufficiently reactive to undergo cyclization with a nearby canonical amino acid in the same peptide, but not too reactive to undergo side reactions with competing molecules in the cellular milieu. With these criteria in mind the Fasan group identified aromatic ncAAs featuring either an appropriate nucleophile or leaving group as candidates to promote peptide macrocyclization in vivo (Figure 2A and B).[22,23]

For example, introducing 3-(2-mercaptoethyl)aminophenylalanine (MeaF, Figure 2C) allowed cyclization in the presence of a C-terminal GyrA intein.[23] Inteins transiently form thioesters, which are prone to undergo substitution reactions in the presence of appropriate nucleophiles (Figure 2A). Specifically, MP formation in the presence of MeaF is a two-step process: first, a transthioesterification, promoted by the nucleophilic sulfhydryl group of MeaF (pK_a 8.5), yields a thiolactone, which then undergoes an irreversible S→N acyl transfer to yield the desired *N*-alkyl macrolactam (see Flag-*cyclo*(Strep3) in Figure 2D for an example of the structure formed). Rather unexpected, the latter step was found to be rate-determining in vivo and MPs ($N>4$ in Figure 2A) could only be isolated after extended production times in *Escherichia coli* (24 hours). Critically, Fasan and co-workers showed that MPs obtained through such a biosynthesis can be used in a model selection strategy. More specifically, based on a previously identified streptavidin binding peptide motif (HPQ) 480 MeaF-containing MPs variants were produced in 96-well plates and their binding properties assessed in parallel. One peptide (Flag-*cyclo*(Strep3) in Figure 2D) demonstrated binding to streptavidin ($K_D = 1.1$ μM), providing a proof-of-concept that this methodology is applicable to identify bioactive MPs.

Another ncAA that allowed for peptide macrocyclization in vivo was *O*-2-bromomethyltyrosine (O2beY, Figure 2C).[22] O2beY features a bromine as leaving group, which is readily displaced by a nearby cysteine residue to yield small to medium-sized MPs ($n=1$ to 9 in Figure 2B). Again, the authors were not only able to demonstrate efficient macrocyclization in vivo, but also showcased the potential of this approach by the design and affinity maturation of a MP inhibitor for the sonic hedgehog/patched interaction.[24] The matured peptide HL2-m5 (Figure 2D) underwent quantitative macrocyclization

Clemens Mayer received his Bachelor's (2007) and Master's (2009) degrees from Graz University of Technology. He then joined Donald Hilvert at ETH Zürich to pursue a PhD, after which he moved to the University of Cambridge (2014) to work as a postdoctoral fellow with Sir Shankar Balasubramanian. In 2016, he joined Gerard Roelfes at the University of Groningen. He was appointed Assistant Professor for Biomolecular Chemistry and Catalysis at the University of Groningen in 2018. His research interests center around repurposing established genetic code expansion techniques for bioprocess strategies that enable the production of value-added compounds from readily available precursors.



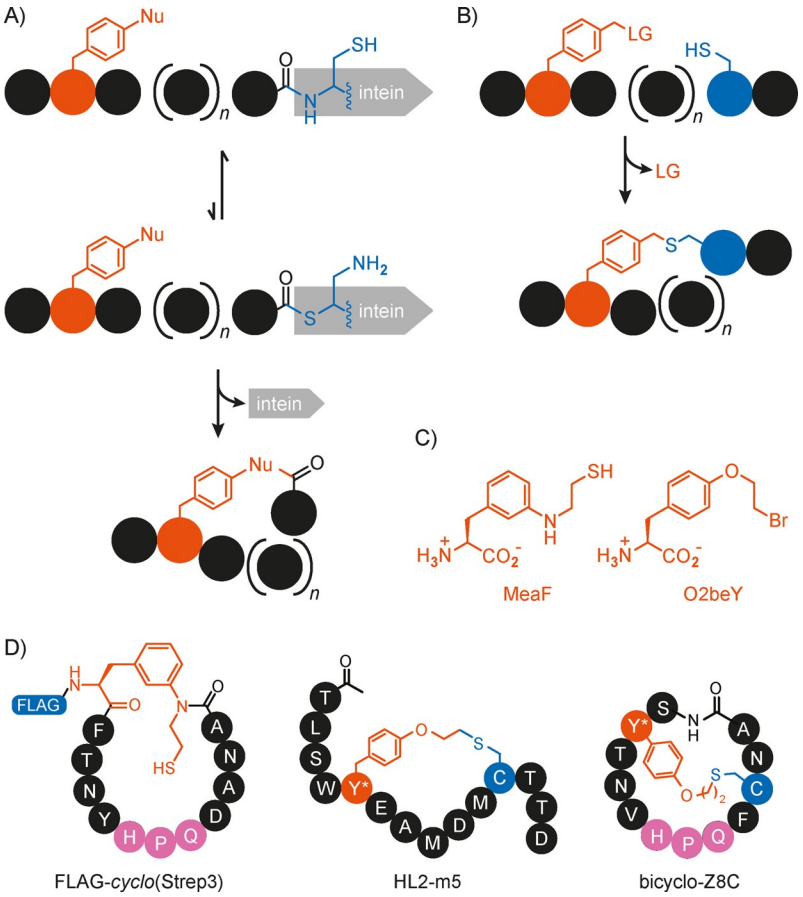


Figure 2. Schematic representations of in vivo peptide macrocyclization strategies based on ncAAs. A) C-terminal inteins transiently form thioesters, which can undergo irreversible substitution reactions with proximal nucleophiles. B) Conversely, ncAAs equipped with appropriate leaving groups can undergo a substitution reaction with a nearby cysteine residue. For both strategies, a successful macrocyclization is critically dependent on fine-tuning the reactivity of the functional group placed in the side chain of the ncAA. C) Structures of the ncAAs that are used to produce MPs in vivo. D) Schematic representation of MPs discussed in this section of the article. ncAAs are highlighted in red, potential reaction partners in blue and amino acids identified to bind to streptavidin in pink. FLAG = MDYKDDDDKGSGS.

in vivo and showed a 120-fold increased affinity, when compared to the parent, linear peptide.

Lastly, the Fasan group also demonstrated that O2beY-mediated macrocyclization is compatible with split intein circular ligation of peptides and proteins (SICLOPPS), a widely employed technique to obtain MPs in vivo.[17,25,26] When combined with SICLOPPS, O2beY incorporation allowed for the biosynthesis of bicyclic peptides. A proof-of-principle that the resulting bicyclic peptides are attractive scaffolds for identifying bioactive MPs is provided by the biosynthesis of bicyclo-Z8C (Figure 2D), which was based on the streptavidin-binding peptide described above. Critically, bicyclo-Z8C showed a 2.5-fold lower IC_{50} value, when compared to a monocyclic variant, which was attributed to the lower conformational flexibility of bicyclo-Z8C.[26]

Together, these examples demonstrate that ncAAs are a promising means to access and identify bioactive MPs with novel ring geometries. As the presented strategies enable peptide macrocyclization in vivo, they are now readily available for interfacing with phenotypic screens and selection strategies. Thus, future efforts will focus on expanding this methodology to select mono- and bicyclic peptides for biologically relevant targets. Moreover, the introduction of ncAAs other than MeaF and O2beY is likely to make other cyclization strategies available in the near future. Specifically, ncAAs that enable macrocyclization strategies with side chains other than cysteines and transiently formed thioesters are desirable to access "natural product-like" MPs with new ring geometries.

Addiction: ncAAs in Biocontainment

Synthetic biology aims to deploy genetically modified organisms (GMOs) as common and valued solutions in clinical, industrial and environmental settings.[27] However, such real-world applications necessitate the development of biocontainment strategies, reminiscent to those outlined in the 1975 Asilomar conference for recombinant DNA.[28] Generally, effective biocontainment strategies must protect against GMO escape mechanisms, including mutagenic drift, environmental supplementation and horizontal gene transfer.[29]

In principle, constructing GMOs with an alternative genetic code by introducing a ncAA into an essential gene (Figure 3A) would advance the barrier between an engineered and a natural organism, as survival of the former would depend on an

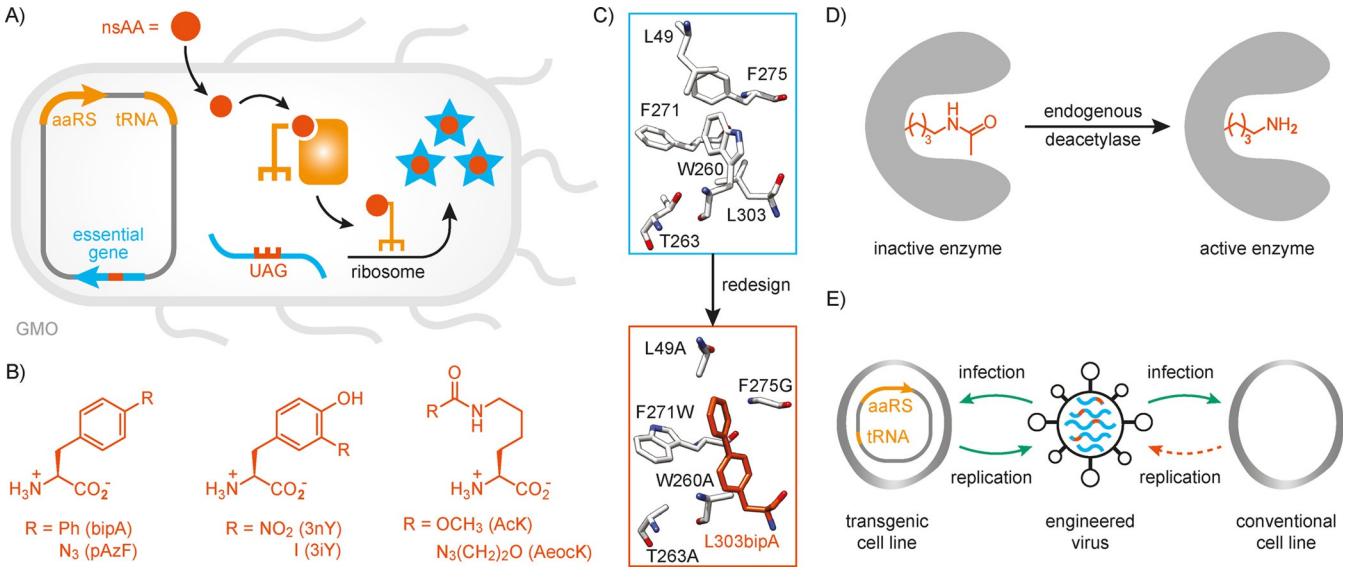


Figure 3. A) Schematic representation of a biocontainment strategy, in which a GMO is addicted to a ncAA. An OTS, provided on a plasmid or integrated in the genome, enables the introduction of a ncAA into an essential protein. One way to ensure low escape frequencies is making the function of an essential protein dependent on the incorporation of the ncAA (top right). B) Structures of ncAAs that have been employed to create addicted organisms. C) Dependence on a ncAA can be achieved by redesigning (either by computation or directed evolution) the hydrophobic packing of an essential protein. The computational redesign of the hydrophobic core of TyrS is shown as an example [PBD IDs: 2YXN (top) and 4OUD (bottom)]. D) Replacing a catalytic lysine residue with AcK in BCAT from *E. coli* results in an inactive enzyme that undergoes activation by endogenous deacetylases. E) Schematic representation of a live but replication-incompetent virus. In the viral genome, multiple conserved residues are replaced by UAG stop codons. As a result, the engineered virus can infect and replicate in a transgenic cell line harboring an OTS. Conversely, it maintains full infectivity for conventional cell lines but cannot replicate, eliciting a strong immune response in these cells.

exogenously supplied synthetic molecule. Therefore, the resulting *synthetic auxotrophs* could not be rescued by cross-feeding with complex media and horizontal gene transfer would be hampered due to the dependence of the GMO on an alternative genetic code and a matching OTS. However, a synthetic biocontainment constructed by introducing a stop codon in an essential gene can easily be breached by reversion of the introduced nonsense codon or the introduction of a canonical amino acid instead of the ncAA. In fact, escape frequencies (EFs) for GMOs with a single UAG codon at a permissive site in an essential protein are $\approx 10^{-6}$–10^{-7},[29] which does not meet the National Institutes of Health suggested maximum EF of 10^{-8}.[30] Thus, to achieve tighter containment the need for the synthetic building block has to be reinforced, for example, by making protein function strictly dependent on the incorporation of the ncAA (Figure 3A).

Such *addiction* of protein function to a synthetic amino acid can be achieved by various means. For example, Church and co-workers designed synthetic auxotrophs by computationally redesigning hydrophobic interactions in protein cores to exclusively accommodate 4,4′-biphenylalanine (bipA, Figure 3B, C).[31] For the genes of adenylate kinase and tyrosyl-tRNA synthetase this redesign resulted in two separate organisms with low EFs ($\approx 10^{-8}$). Moreover, engineering a single GMO that harbored the identified mutations for both genes amplified the effect and resulted in a synthetic auxotroph for which no escape variants could be detected (EF $< 10^{-12}$).

Directed evolution provides another means to redesign the hydrophobic packing of protein cores to make it depend on a ncAA. Specifically, the Ellington group selected TEM-1 β-lactamase variants, the ability of which to confer carbenicillin resistance was dependent on the introduction of 3-nitrotyrosine (3nY) or 3-iodotyrosine (3iY, Figure 3B).[32] For a promising engineered variant, TEM-1-B9, phenylalanine was the only canonical amino acid that could rescue the activity in absence of these ncAAs. However, codons for phenylalanine (UUU and UUC) cannot be accessed by a single mutation from UAG, thus making this reversion unlikely. The group confirmed that this is an unlikely escape mechanism by culturing *E. coli* strains, which harbored TEM-1.B9 and the OTS on a single plasmid, continuously in liquid or solid media without detecting any escape variants (EF $< 10^{-11}$). The single plasmid setup is particularly notable, as it allowed for transformation of other enterobacteriae, which all became dependent on 3nY in the presence of ampicillin and did not escape the containment (EF $< 10^{-9}$).

Another means for making protein function dependent on the presence of a ncAA involves replacement of a natural active site residue with a non-standard one. For example, the Schultz group reported the incorporation of N_ε-acetyllysine (AcK, Figure 3B) into the essential branched chain aminotransferase (BCAT) of *E. coli* (Figure 3D).[33] Specifically, replacing a catalytic lysine with AcK will first produce an inactive BCAT variant, which is then activated upon deacetylation by endogenous acetyltransferases in *E. coli*. As such, synthetic auxotrophs can only breach this containment through mutations that allow for the incorporation of lysine in response to the UAG nonsense codon. Indeed, the authors identified this mechanism as the common feature in escape mutants (EF $> 10^{-8}$). Se-

quence analysis of escape variants revealed a point mutation in the anticodon of an *E. coli* lysine tRNA (3′-UUU-5′ to 3′-AUU-5′), which resulted in an anticodon that could suppress UAG stop codons through G/U wobble pair formation. To counter this escape mechanism, a barnase gene featuring two in-frame UAA stop codons in its mRNA was added to the plasmid. Barnase production is lethal and full-length protein will only be produced in escape variants featuring the mutated lysine tRNA anticodon (3′-AUU-5′), which is complementary to the in-frame UAA stop codons. With this conditional kill switch in place escape mutants effectively committed suicide, which resulted in a tight biocontainment (EF $< 10^{-11}$).

While making protein function dependent on the presence of a ncAA is desirable, it requires a certain degree of design or evolution. An alternative strategy involves the replacement of conserved residues with a ncAA across multiple essential genes. This strategy was showcased by the Isaacs group, who employed multiplex automated genome engineering (MAGE)[34] to identify a synthetic auxotroph that featured three TAG sites in the MurG, DnaS and SerS genes, as well as an OTS decoding *p*-azidophenylalanine (pAzF, Figure 3B).[35] In the identified GMO, addiction to pAzF is achieved by the ability of the ncAA to replace conserved aromatic residues in these three genes. Similar to the escape mechanism for AcK-dependent organisms mentioned above though, all escape mutants in this work featured point mutations in the anticodon of one of the three endogenous tyrosine tRNAs. The identified mutations converted them into suppressor tRNAs, resulting in the incorporation of tyrosine instead of pAzF in response to UAG codons. To overcome this mechanism, the Isaacs group deleted two of the three *E. coli* tyrosine tRNAs; the remaining one was therefore essential for tyrosine incorporation in the GMO, preventing it from accumulating mutations and becoming a suppressor tRNA. In addition to a stringent biocontainment (EF $< 10^{-12}$), the authors demonstrated that the resulting synthetic auxotroph could not be rescued by cross-feeding and that the alternative genetic code indeed impeded horizontal gene transfer.[35]

Related to these efforts, introducing multiple stop codons in the genome of viruses is an emerging strategy to generate live but replication-incompetent virus (Figure 3E).[36] These can serve as live-attenuated vaccines, as they retain their full infectivity and elicit a strong immune response. Zhou and co-workers, for example, have recently created a replication-incompetent influenza A virus, in which multiple UAG codons were introduced at conserved residues throughout the viral genome.[37] Viruses created by this strategy were highly reproductive in transgenic cell lines that featured an OTS to decode UAG codons with N_ε-2-azidoethyloxycarbonyllysine (AeocK), yet did not show any replication in absence of AeocK (EF $< 10^{-11}$). As anticipated, these engineered influenza A strains generated a strong immune response against both parental and antigenically distinct strains. It is conceivable that this strategy can be applied to a wide variety of viruses and replication-incompetent viruses could not only be used prophylactically as vaccines, but also in diagnostic and therapeutic applications.

Overall, the highlighted examples demonstrate that addiction to ncAAs is a promising means to create tight biocontainment for synthetic biology applications.[29] Future efforts are likely to expand the number of ncAAs that can be used to create such GMOs and will see their integration into real-world applications. To avoid the use of an expensive ncAA, it is conceivable that endogenous metabolic pathways could be taken advantage of in order to shift the addiction to a more cost effective, yet still exogenously-supplied synthetic precursor of a ncAA.[38,39]

Catalysis: ncAAs in Designer Enzymes

The enzyme orotidine-5′-phosphate decarboxylase accelerates its target transformation by a factor of 10^{17} with respect to the uncatalyzed reaction and does so by using only canonical amino acids.[40] However, the scope of enzyme catalysis accessible by standard amino acids is fundamentally limited by the diversity of functional groups present in their side chains. Therefore, enzymes in nature routinely recruit electrophiles, redox equivalents, metal ions, transfer agents, etc., in the form of cofactors or co-substrates. Similarly, natural side chains in active sites can undergo posttranslational modifications to install uniquely reactive functional groups that promote a desired transformation (for example, converting cysteine to formylglycine in type I sulfatases).[41]

Enzyme design, in part, aims to create proficient protein catalysts with new-to-nature activities.[42,43] In analogy to the natural strategies described above, the incorporation of ncAAs into protein scaffolds has emerged as an attractive means to expand the reaction scope of designer enzymes beyond what is possible with canonical side chains.[44,45] For example, ncAAs that feature bioorthogonal handles in their side chains can be employed for covalent modification with abiological, transition metal catalysts (an *artificial metalloenzyme*, Figure 4A). Amongst others, the Lewis group has reported the creation of a designer cyclopropanation enzyme by first introducing pAzF (Figure 4B) into the binding pocket of prolyl oligopeptidase (POP) and then recruiting a dirhodium catalyst through a strain-promoted azide–alkyne cycloaddition.[46] Introducing a bioorthogonal handle (i.e., the azide in pAzF) for protein modification is more desirable than other bioconjugation strategies that rely on natural side chains (for example, cysteines), as the modification reaction can be carried out in complex mixtures.[5] This aspect is critical when attempting to evolve artificial metalloenzymes, as anchoring strategies are typically not compatible with the complex cellular milieu.[47] Indeed, the Lewis group has demonstrated the utility of pAzF in the directed evolution of the BOP-derived cyclopropanases.[48] By performing the cofactor anchoring reaction in bacterial lysates, the group was able to screen libraries of BOP-variants that were generated by random mutagenesis and could successfully identify more proficient cyclopropanases. Additionally, these engineered designer enzymes also showed superior reactivity and selectivity in related N–H, S–H and Si–H insertion reactions.

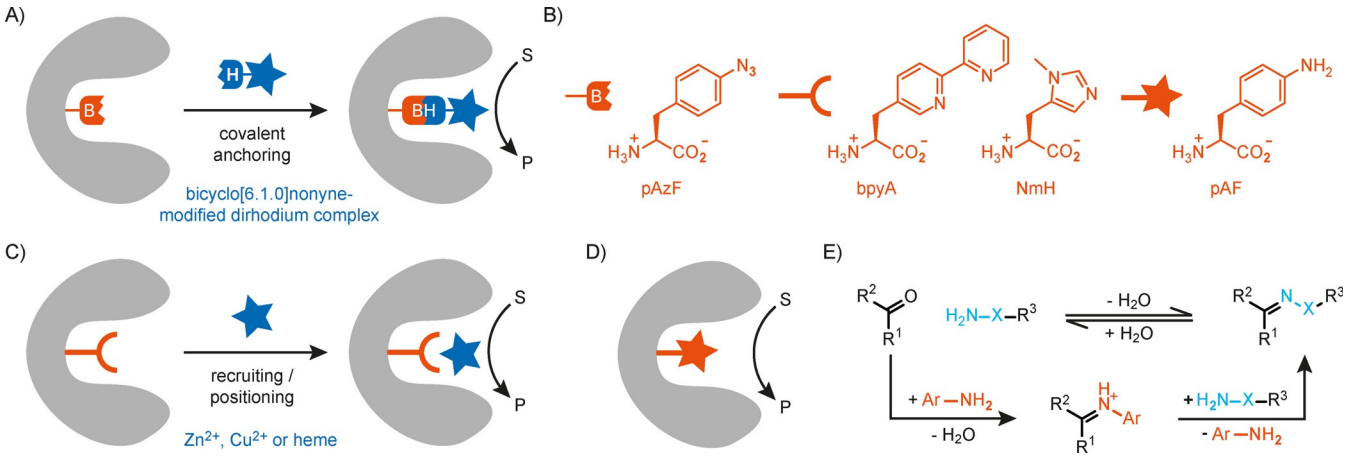


Figure 4. A) Introducing a ncAA harboring a biorthogonal handle (B) allows for covalent anchoring of artificial metal cofactors featuring the counterpart of the biological handle (H). A covalent bond is formed (BH) and the cofactor recruited to a protein binding site. B) Structures of ncAAs which have been used in enzyme design and schematic representation for which strategy they were used to are provided. C) Introducing a metal-chelating ncAA into a protein scaffold results in metalloenzymes by recruiting or positioning natural or synthetic metal cofactors. D) Introducing a ncAA with unique reactivity results in designer enzymes that do not need additional modification to become active. E) pAF-containing enzyme variants accelerate hydrazone (X = NH) and oxime (X = O) formation reaction through population of an iminium ion (= covalent catalysis) with a carbonyl moiety in the substrate.

Instead of relying on a biorthogonal handle to recruit a synthetic cofactor, the incorporation of metal-chelating ncAAs that can directly bind metal ions or complexes is another means to create artificial metalloenzymes (Figure 4C).[49] For example, the Schultz group installed (2,2′-bipyridin-5-yl)alanine (bpyA, Figure 4B) into the *E. coli* catabolite activator protein, which upon binding of copper or iron ions endowed the protein with nuclease activity.[50] More recently, the Roelfes group has expanded this approach by introducing bpyA into the promiscuous, hydrophobic binding pockets of multidrug resistance regulators (MDRs). Upon binding copper ions, a number of MDR-based artificial enzymes were created that could catalyze abiological Friedel–Crafts alkylation[51,52] and hydration reactions.[53] The high activities and selectivities observed in these designer enzymes are the result of embedding the copper ion (through coordination to bpyA) into the MDR binding pockets, which aides in recruiting hydrophobic substrates. Incorporating genetically encodable, metal-chelating ncAAs, such as bpyA, into protein binding pockets is also an attractive strategy for future directed evolution campaigns. A metal-chelating ncAA alleviates the need for a posttranslational synthetic step to recruit the catalysts species and therefore could facilitate artificial metalloenzyme formation in complex media or even living cells. As a result, this strategy could significantly increase the throughput, when screening for improved designer enzymes.

In a variation of this theme, the Hilvert and Schultz groups have replaced histidine with N_δ-methyl histidine (NmH) to position a heme prosthetic group in myoglobin[54] and an essential zinc ion an mannose-6-phosphate isomerase.[55] While the latter resulted in a GMO addicted to NmH (EF $< 10^{-11}$), for the former the introduction of NmH was shown to subtly alter the electronic properties of the bound heme. Notably, installing NmH as axial heme ligand in a previously engineered myoglobin[56] did not only boost the peroxidase activity[57,58] but also its promiscuous cyclopropanation activity.[59] Moreover, in the presence of NmH, cyclopropanation reactions could be carried out in absence of a reducing agent, conditions under which the parent histidine variant was largely inactive.[56]

Besides recruiting, positioning and fine-tuning of (synthetic) metal cofactors, the Roelfes group has recently described a new strategy for enzyme design, in which a ncAA with a unique functionality is incorporated to act as a catalytic residue (Figure 4E). Specifically, the incorporation of *p*-aminophenylalanine (pAF) at position 15 in the MDR from *Lactococcus lactis* (LmrR), resulted in LmrR_V15pAF, which promoted hydrazone and oxime formation reactions.[60] This activity was ascribed to the unique ability of the aniline side chain of pAF to form an iminium ion (covalent catalysis) with an aldehyde substrate, which then undergoes a transimidation reaction in the presence of appropriate hydrazine or hydroxylamine substrates (Figure 4E).[61,62] Critically, the inherent catalytic activity of anilines to promote this reaction was boosted by placing pAF in the hydrophobic pore of LmrR. Taking advantage of both the unique reactivity of pAF and the ability of LmrR to recruit substrates, LmrR_V15pAF outperformed aniline in solution by a factor of ≈560. Lastly, in a follow-up study the authors demonstrated that the catalytic contribution of the ncAA can be boosted through consecutive rounds of directed evolution.[63] A total of four synergistic mutations were identified that, when combined, increased the turnover frequency (k_{cat}) of the parent designer enzyme by almost 100-fold and gave rise to variants that outperformed aniline in solution by more than four orders of magnitude.

Overall, these examples demonstrate that the incorporation of ncAAs into protein scaffolds has already begun to significantly expand the reaction scope of designer enzymes. In the future, placing metal-chelating amino acids into protein scaffolds will continue to create new and/or improved metal-binding environments that will give access to new reactivities. For

the incorporation of ncAAs that directly promote a target transformation, the creation and directed evolution of LmrR_V15pAF opens up questions that need to be answered. For example, is it a general phenomenon that the performance of organocatalysts, which are versatile yet notoriously slow, is boosted by placing them in protein scaffolds by ncAA incorporation?[64,65] And, are the resulting designer enzymes generally privileged starting points for directed evolution efforts, as it was demonstrated for LmrR_pAF?[63] Independent of the answers to these questions, designer enzymes will continue to make use of ncAAs to catalyze new-to-nature transformations.

Future Directions: Orthogonal Ribosomes

With each section closing with a brief outline of future directions, this section aims to highlight developments that could impact all areas discussed in this Concept article. As mentioned briefly in the *Selection* part, the simultaneous incorporation of multiple, chemically distinct ncAAs is typically not efficient.[66] This stems from the fact that adding codons to the existing genetic code is not straightforward and the number of codons readily available for recoding is limited. For example, universal reassignment of more than one nonsense codons is difficult, due to their native function to signal for termination of translation. To meet this challenge, the Chin group has begun to install a *parallel genetic code* into model organisms, by making use of orthogonal ribosomes (ORs).[67] In brief, ORs feature mutations in the 16S ribosomal RNA that enables them to recognize mRNAs, which are not translated by endogenous ribosomes.[68] Conversely, ORs do not recognize native mRNAs and as a result, do not participate in the synthesis of endogenous proteins. Being now not essential for the survival of the organism, an OR is free to accumulate mutations that alter its interaction with tRNAs and/or release factors.[69] Such engineering efforts have allowed for the addition of codons,[70] the selection of new OTS that are specific for ORs,[71] and the efficient encoding of two ncAAs simultaneously.[70,72] Thus, it is conceivable that applying ORs to the strategies discussed in this article could allow for novel peptide macrocyclization strategies, reinforce synthetic biocontainment, and further expand the scope of reactions catalyzed by designer enzymes.

Acknowledgements

C.M. is thankful to Prof. Gerard Roelfes for his careful reading of this manuscript and for many helpful suggestions. C.M. also acknowledges the Netherlands Organisation for Scientific Research (NWO, Veni grant 722.017.007) and a Marie Skłodowska Curie Individual Fellowship (project no. 751509) for financial support.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: cyclic peptides · genetic code expansion · protein engineering · synthetic biology · unnatural amino acids

[1] L. Wang, A. Brock, P. G. Schultz, *Science* **2001**, *292*, 498 – 500.
[2] A. Dumas, L. Lercher, C. D. Spicer, B. G. Davis, *Chem. Sci.* **2015**, *6*, 50 – 69.
[3] D. D. Young, P. G. Schultz, *ACS Chem. Biol.* **2018**, *13*, 854 – 870.
[4] J. W. Chin, *Nature* **2017**, *550*, 53 – 60.
[5] K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764 – 4806.
[6] T. Morioka, N. D. Loik, C. J. Hipolito, Y. Goto, H. Suga, *Curr. Opin. Chem. Biol.* **2015**, *26*, 34 – 41.
[7] A. Zorzi, K. Deyle, C. Heinis, *Curr. Opin. Chem. Biol.* **2017**, *38*, 24 – 29.
[8] T. A. F. Cardote, A. Ciulli, *ChemMedChem* **2016**, *11*, 787 – 794.
[9] L. J. Walport, R. Obexer, H. Suga, *Curr. Opin. Biotechnol.* **2017**, *48*, 242 – 250.
[10] R. Obexer, L. J. Walport, H. Suga, *Curr. Opin. Chem. Biol.* **2017**, *38*, 52 – 61.
[11] E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discovery* **2008**, *7*, 608 – 624.
[12] T. Passioura, H. Suga, *Chem. Commun.* **2017**, *53*, 1931 – 1940.
[13] J. Morimoto, Y. Hayashi, K. Iwasaki, H. Suga, *Acc. Chem. Res.* **2011**, *44*, 1359 – 1368.
[14] T. Passioura, W. Liu, D. Dunkelmann, T. Higuchi, H. Suga, *J. Am. Chem. Soc.* **2018**, *140*, 11551 – 11555.
[15] D. S. Wilson, A. D. Keefe, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3750 – 3755.
[16] W. Zheng, N. Thorne, J. C. McKew, *Drug Discovery Today* **2013**, *18*, 1067 – 1073.
[17] A. Tavassoli, *Curr. Opin. Chem. Biol.* **2017**, *38*, 30 – 35.
[18] J. M. Smith, F. Vitali, S. A. Archer, R. Fasan, *Angew. Chem. Int. Ed.* **2011**, *50*, 5075 – 5080; *Angew. Chem.* **2011**, *123*, 5181 – 5186.
[19] J. R. Frost, J. M. Smith, R. Fasan, *Curr. Opin. Struct. Biol.* **2013**, *23*, 571 – 580.
[20] J. M. Smith, J. R. Frost, R. Fasan, *Chem. Commun.* **2014**, *50*, 5027 – 5030.
[21] J. M. Smith, N. C. Hill, P. J. Krasniak, R. Fasan, *Org. Biomol. Chem.* **2014**, *12*, 1135 – 1142.
[22] N. Bionda, A. L. Cryan, R. Fasan, *ACS Chem. Biol.* **2014**, *9*, 2008 – 2013.
[23] J. R. Frost, N. T. Jacob, L. J. Papa, A. E. Owens, R. Fasan, *ACS Chem. Biol.* **2015**, *10*, 1805 – 1816.
[24] A. E. Owens, I. De Paola, W. A. Hansen, Y. W. Liu, S. D. Khare, R. Fasan, *J. Am. Chem. Soc.* **2017**, *139*, 12559 – 12568.
[25] C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638 – 13643.
[26] N. Bionda, R. Fasan, *ChemBioChem* **2015**, *16*, 2011 – 2016.
[27] O. Wright, G. B. Stan, T. Ellis, *Microbiology* **2013**, *159*, 1221 – 1235.
[28] P. Berg, D. Baltimore, S. Brenner, R. O. Roblin, M. F. Singer, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1981 – 1984.
[29] A. J. Simon, A. D. Ellington, *F1000Research* **2016**, *5*, 2118.
[30] National Institutes of Health: *NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules*, Office of Science Policy, April **2016**.
[31] D. J. Mandell, M. J. Lajoie, M. T. Mee, R. Takeuchi, G. Kuznetsov, J. E. Norville, C. J. Gregg, B. L. Stoddard, G. M. Church, *Nature* **2015**, *518*, 55 – 60.
[32] D. S. Tack, J. W. Ellefson, R. Thyer, B. Wang, J. Gollihar, M. T. Forster, A. D. Ellington, *Nat. Chem. Biol.* **2016**, *12*, 138 – 140.
[33] W. Xuan, P. G. Schultz, *Angew. Chem. Int. Ed.* **2017**, *56*, 9170 – 9173; *Angew. Chem.* **2017**, *129*, 9298 – 9301.
[34] H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest, G. M. Church, *Nature* **2009**, *460*, 894 – 898.
[35] A. J. Rovner, A. D. Haimovich, S. R. Katz, Z. Li, M. W. Grome, B. M. Gassaway, M. Amiram, J. R. Patel, R. R. Gallagher, J. Rinehart, F. J. Isaacs, *Nature* **2015**, *518*, 89 – 93.
[36] N. Wang, Y. Li, W. Niu, M. Sun, R. Cerny, Q. Li, J. Guo, *Angew. Chem. Int. Ed.* **2014**, *53*, 4867 – 4871; *Angew. Chem.* **2014**, *126*, 4967 – 4971.
[37] L. Si, H. Xu, X. Zhou, Z. Zhang, Z. Tian, Y. Wang, Y. Wu, B. Zhang, Z. Niu, C. Zhang, G. Fu, S. Xiao, Q. Xia, L. Zhang, D. Zhou, *Science* **2016**, *354*, 1170 – 1173.
[38] R. A. Mehl, J. C. Anderson, S. W. Santoro, L. Wang, A. B. Martin, D. S. King, D. M. Horn, P. G. Schultz, *J. Am. Chem. Soc.* **2003**, *125*, 935 – 939.

[39] J. E. Jung, S. Y. Lee, H. Park, H. Cha, W. Ko, K. Sachin, D. W. Kim, D. Y. Chi, H. S. Lee, *Chem. Sci.* **2014**, *5*, 1881 – 1885.
[40] A. Radzicka, R. Wolfenden, *Science* **1995**, *267*, 90 – 93.
[41] N. M. Okeley, W. A. Van Der Donk, *Chem. Biol.* **2000**, *7*, 159 – 171.
[42] D. Hilvert, *Annu. Rev. Biochem.* **2013**, *82*, 447 – 470.
[43] S. C. Hammer, A. M. Knight, F. H. Arnold, *Curr. Opin. Green Sustain. Chem.* **2017**, *7*, 23 – 30.
[44] J. C. Lewis, *Curr. Opin. Chem. Biol.* **2015**, *25*, 27 – 35.
[45] F. Agostini, J. S. Völler, B. Koksch, C. G. Acevedo-Rocha, V. Kubyshkin, N. Budisa, *Angew. Chem. Int. Ed.* **2017**, *56*, 9680 – 9703; *Angew. Chem.* **2017**, *129*, 9810 – 9835.
[46] P. Srivastava, H. Yang, K. Ellis-Guardiola, J. C. Lewis, *Nat. Commun.* **2015**, *6*, 7789.
[47] F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis, T. R. Ward, *Chem. Rev.* **2018**, *118*, 142 – 231.
[48] H. Yang, A. M. Swartz, H. J. Park, P. Srivastava, K. Ellis-Guardiola, D. M. Upp, G. Lee, K. Belsare, Y. Gu, C. Zhang, R. E. Moellering, J. C. Lewis, *Nat. Chem.* **2018**, *10*, 318 – 324.
[49] T. Hayashi, D. Hilvert, A. P. Green, *Chem. Eur. J.* **2018**, *24*, 11821 – 11830.
[50] H. S. Lee, P. G. Schultz, *J. Am. Chem. Soc.* **2008**, *130*, 13194 – 13195.
[51] I. Drienovská, A. Rioz-Martínez, A. Draksharapu, G. Roelfes, *Chem. Sci.* **2015**, *6*, 770 – 776.
[52] M. Bersellini, G. Roelfes, *Org. Biomol. Chem.* **2017**, *15*, 3069 – 3073.
[53] I. Drienovská, L. Alonso-Cotchico, P. Vidossich, A. Lledós, J. D. Maréchal, G. Roelfes, *Chem. Sci.* **2017**, *8*, 7228 – 7235.
[54] T. Schwander, L. Schada von Borzyskowski, S. Burgener, N. S. Cortina, T. J. Erb, *Science* **2016**, *354*, 900 – 904.
[55] F. Gan, R. Liu, F. Wang, P. G. Schultz, *J. Am. Chem. Soc.* **2018**, *140*, 3829 – 3832.
[56] M. Bordeaux, V. Tyagi, R. Fasan, *Angew. Chem. Int. Ed.* **2015**, *54*, 1744 – 1748; *Angew. Chem.* **2015**, *127*, 1764 – 1768.
[57] A. P. Green, T. Hayashi, P. R. E. Mittl, D. Hilvert, *J. Am. Chem. Soc.* **2016**, *138*, 11344 – 11352.
[58] M. Pott, T. Hayashi, T. Mori, P. R. E. Mittl, A. P. Green, D. Hilvert, *J. Am. Chem. Soc.* **2018**, *140*, 1535 – 1543.
[59] T. Hayashi, M. Tinzl, T. Mori, U. Krengel, J. Proppe, J. Soetbeer, D. Klose, G. Jeschke, M. Reiher, D. Hilvert, *Nat. Catal.* **2018**, *1*, 578 – 584.
[60] I. Drienovská, C. Mayer, C. Dulson, G. Roelfes, *Nat. Chem.* **2018**, *10*, 946 – 952.
[61] A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* **2006**, *128*, 15602 – 15603.
[62] A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem. Int. Ed.* **2006**, *45*, 7581 – 7584; *Angew. Chem.* **2006**, *118*, 7743 – 7746.
[63] C. Mayer, C. Dulson, E. Reddem, A.-M. Thunnissen, G. Roelfes, *Angew. Chem. Int. Ed.* **2019**, *58*, 2083 – 2087; *Angew. Chem.* **2019**, *131*, 2105 – 2109.
[64] Y. Miao, M. Rahimi, E. M. Geertsema, G. J. Poelarends, *Curr. Opin. Chem. Biol.* **2015**, *25*, 115 – 123.
[65] A. R. Nödling, K. Świderek, R. Castillo, J. W. Hall, A. Angelastro, L. C. Morrill, Y. Jin, Y. H. Tsai, V. Moliner, L. Y. P. Luk, *Angew. Chem. Int. Ed.* **2018**, *57*, 12478 – 12482; *Angew. Chem.* **2018**, *130*, 12658 – 12662.
[66] M. G. Hoesl, N. Budisa, *Angew. Chem. Int. Ed.* **2011**, *50*, 2896 – 2902; *Angew. Chem.* **2011**, *123*, 2948 – 2955.
[67] C. C. Liu, M. C. Jewett, J. W. Chin, C. A. Voigt, *Nat. Chem. Biol.* **2018**, *14*, 103 – 106.
[68] O. Rackham, J. W. Chin, *Nat. Chem. Biol.* **2005**, *1*, 159 – 166.
[69] K. Wang, H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Biotechnol.* **2007**, *25*, 770 – 777.
[70] H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **2010**, *464*, 441 – 444.
[71] H. Neumann, A. L. Slusarczyk, J. W. Chin, *J. Am. Chem. Soc.* **2010**, *132*, 2142 – 2144.
[72] K. Wang, A. Sachdeva, D. J. Cox, N. W. Wilf, K. Lang, S. Wallace, R. A. Mehl, J. W. Chin, *Nat. Chem.* **2014**, *6*, 393 – 403.

Manuscript received: November 29, 2018
Accepted manuscript online: January 8, 2019
Version of record online: March 13, 2019

