Break the template boundary: Test of different protein scaffolds by genetic code expansion resulted in NADPH-dependent secondary amine catalysis

Thomas Williams
Cardiff University
Yu-Hsuan Tsai
Cardiff University https://orcid.org/0000-0003-0589-5088
Louis Luk (LukLY@cardiff.ac.uk)
Cardiff University

Article

Keywords: Artificial Enzyme, Genetic Code Expansion, Secondary Amine Organocatalysis, Iminium Catalysis, Protein Engineering

Posted Date: September 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-468406/v2

Abstract

Here, incorporation of secondary amine by genetic code expansion was used to expand the potential protein templates for artificial enzyme design. Pyrrolysine analogue containing a D-proline could be stably incorporated into proteins, including the multidrug-binding LmrR and nucleotide-binding dihydrofolate reductase (DHFR). Both modified scaffolds were catalytically active, mediating transfer hydrogenation with a relaxed substrate scope. The protein templates played a distinctive role in that, while the LmrR variants were confined to the biomimetic BNAH as the hydride source, the optimal DHFR variant favorably used the pro-*R* hydride from NADPH for reactions. Due to the cofactor compatibility, the DHFR secondary amine catalysis could also be coupled to an enzymatic recycling scheme. This work has illustrated the unique advantages of using proteins as hosts, and thus the presented concept is expected to find uses in enabling tailored secondary amine catalysis.

Introduction

The use of secondary amines as catalytic motifs has tremendous potentials in artificial enzyme design. In nature, primary amines including the amino groups of lysine residues and protein N-termini are widely used in enzyme catalysis, serving as an acid, a base or nucleophile.[1-2] In contrast, the use of secondary amines in enzyme catalysis is significantly rarer and their scope has not been fully explored.[3-7] However, while the basicity of cyclic secondary amines are similar to their primary counterparts, they are significantly more nucleophilic.[8-10] Additionally, when a secondary amine reacts with a carbonyl substrate, the resulting iminium ion intermediate does not contain a proton on the nitrogen atom; this can prompt reactions with a latent nucleophile or a base for enamine formation, driving the catalytic cycles forward.[11-13] Indeed, proline and its derivatives (e.g., prolyl-RNA) have been proposed to play roles in the prebiotic world catalyzing the formation of crucial building blocks such as carbohydrates and nucleotides.[14-16] Given the reactivity of secondary amines, their incorporation into protein scaffolds can generate catalytically active entities which have strong potentials to be transformed into highly active and selective artificial enzymes.

Currently, only two types of templates have been used to generate protein-based secondary amines for catalysis. 4-Oxalocrotonate tautomerase (4-OT) contains a N-terminal proline residue within its cavity, and it can be used to mediate iminium or enamine catalysis.[17-20] In the second approach, pyrrolidines covalently linked to biotin were introduced to streptavidin (Sav) affording a protein-hosted secondary amine catalytic system.[21-23] These protein catalysts could catalyze various types of reactions, including conjugate addition,[19-21] aldol condensation,[18, 22] transfer hydrogenation[23] and epoxidation.[24] However, since there are only a few proteins that contain a N-terminal proline within their cavity[3, 6] and also only a few that bind biotin with significant affinity,[25-26] the choice of protein templates for hosting secondary amine is limited. Indeed, the full benefit of using protein as host has not been revealed, and a generic approach that facilitates the incorporation of secondary amine is needed.

Genetic code expansion has the advantage of fewer restrictions on the choice of protein scaffolds and the site therein.[27-31] Previously, an aniline motif has been added to the multidrug binding protein LmrR by incorporating unnatural amino acid *p*-azidophenylalanine followed by chemical reduction.[27-30] Whilst being less nucleophilic than pyrrolidine,[8, 10] the aniline in LmrR was able to catalyze various carbon-heteroatom and carbon-carbon ligation reactions.[27-30] Another example is the introduction of a *N*-methyl histidine residue into the computationally designed protein BH32;[31] subsequent laboratory evolution afforded a catalyst capable of hydrolyzing fluorescein esters through formation of a covalent intermediate. Previously, secondary amines have been incorporated into protein scaffolds (e.g., β-galactosidase) as pyrrolysine analogues, in which proline and its derivatives were attached to lysine through isopeptide bonds.[32] However, their applications in artificial enzyme design have yet to be explored. Furthermore, stability of these proline derivatives is unknown, as they can be proteolytically hydrolyzed during recombinant preparation. Nevertheless, pyrrolysine analog incorporation is worth exploring because it can be an effective approach to screen a broad range of protein templates through introduction of stable, catalytically active secondary amines.

In this work, through genetic code expansion, unnatural amino acids **1-3** bearing a cyclic secondary amine were site-specifically introduced into different scaffolds, including the super folder green fluorescent protein (sfGFP), multidrug-binding protein LmrR and nucleotide-binding dihydrofolate reductase (DHFR) (**Fig. 1A**).[33] Hydrolysis of the isopeptide bond was observed for proteins incorporated with **2** (L-proline) but not detected for proteins incorporated with **1** (D-proline) or **3** (L-thioproline). Incorporation of **1** resulted in catalytically active entities, and the effect of the protein template has been illustrated in the reactions catalyzed by the corresponding LmrR and DHFR variants. While both systems could catalyze the model reduction of various α,β-unsaturated aldehydes (**Fig. 1B & C**), the LmrR-hosted catalytic system could only use the biomimetic 1-benzyl-1,4-dihydronicotinamide (BNAH) as a source of hydride. In contrast, a DHFR variant was found to recruit the pro-*R* hydride from the cofactor NADPH for transfer hydrogenation. Further investigations revealed that the hydride delivery in the DHFR system is not rate-limiting, and the catalysis can be coupled to an enzymatic NADPH regeneration scheme. This work has proved the concept that genetic code expansion is a viable tool to examine different protein templates for designed secondary amine catalysis.

Results And Discussion

Unnatural amino acids containing a secondary amine were tested for incorporation (**Fig. 1A**). Amino acids **1**, **2**, and **3** are derived from D-proline, L-proline and L-thioproline (L-thiazolidine-4-carboxylic acid), respectively (see SI for their synthesis). Based on previous reports, **1** and **3** are substrates of the wild-type *Methanosarcina bakeri* pyrrolysyl-tRNA synthetase (MbPylRS) and its engineered variant ThzKRS, respectively.[32, 34] Due to the structural similarity between **2** and **3**, we envisaged that **2** could be a substrate of ThzKRS. Incorporation of unnatural amino acids was tested using a version of sfGFP whose gene bears a TAG codon in the middle (i.e., 150th amino acid residue, Asn150TAG), and formation of the full-length protein implied successful incorporation of the unnatural amino acid as indicated by SDS-

PAGE analysis (**Fig. S1** and **Table S1**).[35] However, since the corresponding unnatural amino acid residue is located at a solvent exposed position, rather than the interior of the protein, their stability and catalytic activity were not further examined.

For catalyst design, the unnatural amino acids were incorporated into the multidrug binding protein LmrR,[36] a scaffold known to be catalytically competent after modification of its hydrophobic pocket including the insertion of an unnatural amino acid.[37-40] Hence, we selected four previously reported residues, Val15, Asn19, Met89 and Phe93 (**Fig. 1B**), located in the hydrophobic pocket for substitution with **1**, **2** or **3**. The 12 LmrR variants were successfully produced by *Escherichia coli* as revealed by SDS-PAGE analysis (**Fig. S2**). The dimeric nature of these variants was confirmed by size exclusion chromatography, with the exception of **LmrR-Met89-2** whose oligomeric nature appears to be hampered by the modification (see **SI**, Pg S10-12). However, liquid chromatography-mass spectrometry (LC-MS) investigation indicated that variants containing amino acid **2** were unstable and up to ~35% hydrolysis of the L-prolyl group was observed at position 89 (**Fig. 2**). Direct incorporation of lysine was ruled out as the full-length protein could not be obtained without supplementing **2** in the growth cultures. Hence, this observation was attributed to the hydrolysis of the L-prolyl group during protein production, as *E. coli* contains enzymes such as proline iminopeptidase which can cleave the N-terminal proline in polypeptides (e.g., UniProt: A0A6N7NN15_ECOLX). In contrast, truncation was not observed in proteins incorporated with either D-proline **1** or L-thioproline **3**, and hence they were concluded non-hydrolysable.

We chose the conversion of cinnamaldehyde **4** to its reduced counterpart **5** as the model reaction (**Fig. 3**) because the streptavidin (Sav)-hosted pyrrolidine was shown to be able to catalyze this reaction,[23] and hence direct comparison among the protein-based catalytic systems can be made. BNAH, unlike other biomimetics such as Hantzsch ester, is soluble in aqueous buffers and thus was tested as the hydride donor.[23] Conversion of the reaction was estimated using a gas chromatography-mass spectrometry (GC-MS) setup as previously described (see SI).[41] Though containing seven lysine residues and a free N-terminal amino group, the wild-type LmrR could only catalyze less than 3% of conversion implying minimal activity (**Fig. 3A** and **Table S2**). Substitution of Met89 with **1**, **2** or **3** also showed marginal improvement (~5% conversion). For all the variants incorporated with **2** of which truncation was detected, low conversion (<15%) was observed. Substitution of thioproline **3** at position 19 and D-proline **1** at position 15 resulted in a detectable amount of product conversion (> 20%). Nevertheless, the replacement of Phe93 with **1** (**LmrR-Phe93-1**) was most promising giving a yield 19-fold higher than that of the wild-type enzyme (58% vs. 3%). **LmrR-Phe93-2** differs by one stereocenter (D- vs L-proline) and, while a majority contain the secondary amine (~80%, see **Fig. 2**), it could hardly catalyze the model reaction (<2%). In contrast, **LmrR-Phe93-3**, the non-hydrolysable thiazolidine equivalent of **LmrR-Phe93-2**, was able to mediate a lower but detectable amount of conversion (12%). Accordingly, several factors, including the chirality, stability and electron density of the catalyst, appear to play roles in the extent of conversion. NADPH was tested as the hydride donor for catalysis by **LmrR-Phe93-1**, but no conversion was detected. Transfer hydrogenation with NADPH by unnatural amino acid **1** was also not detectable (**Fig. 3B**). These observations align with the literatures. The secondary amines hosted by Sav could not recruit NADH for

reaction in aqueous buffer.[23] Similarly, NADH was found to be a poor hydride donor for small organocatalytic secondary amines.[42-43] Indeed, while examples of metal catalysts and artificial metalloenzymes using NAD(P)H or $NAD(P)^+$ for reactions have been reported,[44-48] the equivalents in organocatalytic system are much rarer.

Seeing if NADPH can be used as a hydride source by a switch of template, the *E.coli* dihydrofolate reductase (DHFR) was tested. DHFR catalyzes the step of hydride transfer from the C4 pro-*R* hydride of NADPH to C6 of dihydrofolate.[49-51] Containing a Rossmann fold, DHFR is composed of a multi-stranded β sheet packed by two α helices on each side and binds nucleotide-containing cofactor NADPH with an association rate up to 10^5 $M^{-1}{\cdot}s^{-1}$ (**Fig. 1C**).[50-51] Additionally, it contains a mobile M20 loop (residue 9-23) which closes the active site upon binding NADPH,

thereby generating a shielded environment surrounding the nicotinamide motif.[52-53] These features present DHFR as a plausible scaffold for transfer hydrogenation by NADPH. Hence, Ala7, Phe31 and Ser49 that are in proximity to the nicotinamide motif were individually replaced with **1**, the non-hydrolysable unnatural amino acid that gave the highest conversion in the previously examined system (LmrR). The resulting DHFR secondary amines were verified by SDS-PAGE (**Fig. S3**) and mass spectrometry (**Fig. S4**). Similarly, hydrolysis of the D-prolyl group was not detected.

Activity tests revealed that DHFR variants incorporated with **1** could facilitate NADPH-dependent transfer hydrogenation (**Fig. 3B**). For the wild-type DHFR which contains six lysine residues and a N-terminal amino group, the reaction product **5** was undetectable by GC-MS analysis. Replacement of Ser49 with **1** led to considerable protein precipitation with no detectable products. Contrarily, replacement of Ala7 with **1** could effectively catalyze the reaction as evident by significant product conversion (**Fig. 3B** and **Table S3**). Though BNAH can be used as the hydride source, the extent of product conversion was noticeably higher when utilizing NADPH. With 0.1 equivalent (10 mol%) of **DHFR-Ala7-1**, 90% of starting material **4** was converted into product **5** in aqueous PBS buffer at pH 7.0 using 2.0 equivalents of NADPH after 18 hours, while ~34% conversion with BNAH was achieved under the same condition (**Table S4**).

Similar to many other secondary amine organocatalysts,[9] formation of iminium ion is critical towards carbonyl substrate activation, and thus its transient existence during the catalysis by **LmrR-Phe93-1** and **DHFR-Ala7-1** was probed. The two variants were incubated with the starting material **4**, treated with $NaCNBH_3$ and subjected to high resolution mass spectrometric analysis, similar to previously described (see SI).[28] The **LmrR-Phe93-1** sample predominantly yielded one protein species with an observed mass matching the calculated molecular weight of the reduced covalent protein-substrate intermediate (**Fig. S5A**). Similarly, a substantial portion of the **DHFR-Ala7-1** variant also yielded the reduced covalent intermediate under the same condition. However, there was a noticeable amount of unmodified **DHFR-Ala7-1**, suggesting this variant is less reactive and the secondary amine in **1** is less accessible (see below for kinetic measurement). Interestingly, hydrolysis of the D-prolyl group was also observed upon

treatment with $NaCNBH_3$ (**Fig. S5B**). Treatment of these protein species with chymotrypsin revealed digested peptides, whose elemental composition identified by liquid chromatography mass spectrometry (LC-MS) corresponds to the trapped intermediate (**Fig. S6** and **S7**). In contrast, the wild-type LmrR and DHFR that cannot catalyze the transfer hydrogenation did not show any evidence of the iminium intermediate formation. Accordingly, these results support a LUMO-lowering reaction mechanism,[9] in which the secondary amine activates the starting material through iminium ion formation for transfer hydrogenation (**Fig. 4**).

Contrary to the previous work which used a LmrR-hosted aniline for catalysis,[28] no lysine modification was detected in our protein scaffolds after $NaCNBH_3$ treatment. The chemoselectivity observed here implies the significantly higher reactivity of the secondary amine which forms iminium intermediate with relatively inert carbonyl substrate under aqueous conditions. Furthermore, in LmrR the aniline organocatalysis was found to be optimal at position 15,[28] whereas we obtained higher conversions with the secondary amine organocatalyst at position 93, indicating that activity test at different positions is essential during protein-based catalyst design, an experiment that can be readily achieved by genetic code expansion.

The k_{cat} constants were estimated to be highest for the previously reported streptavidin-hosted secondary amine system (>20 fold, ~0.01275 s^{-1}),[23] followed by LmrR (~2 fold) and eventually DHFR which recruits NADPH as the hydride donor instead of BNAH (**Table S5** & **S6**, **Fig. S8** and **Pg S17-18** for assessment). These differences can be explained by the locations of the catalytic centers. The Sav-hosted secondary amine is most solvent exposed[21] and thus it can readily react with the carbonyl substrate contributing to a large k_{cat} constant. In contrast, in both **LmrR-Phe93-1** and **DHFR-Ala7-1** the secondary amines are less accessible locating within the protein cavities and lower k_{cat} constants were observed. However, the estimated Michaelis constant (K_M) value for the corresponding hydride donor was found to be lowest for DHFR (94 μM), resulting in an enhancement in the bimolecular rate constant that is approximately 6-fold higher than that of the LmrR system. Accordingly, though the secondary amine may be less accessible in the DHFR variant, its tight binding nature to NADPH facilitates the step of hydride transfer, hence the significant conversion observed (**Fig. 3** and **5**, also see below). On the other hand, the K_M for cinnamaldehyde could not be accurately determined due to precipitation (in aqueous buffer containing 10% methanol). Lastly, the estimated bimolecular rate constants for all three systems were considerably lower than those for natural enzymes that recruit BNAH or NADPH for reactions.[54-55] Nevertheless, this work proves the concept that incorporation of secondary amine by genetic code expansion facilitates reactions with the designated nucleophile (NADPH vs. BNAH), and performance of these first-generation catalysts can be readily enhanced through established techniques.[27, 56-57]

When [4*R*-^{2}H]-NADPH (NADPD) was used in the catalysis by **DHFR-Ala7-1**, only one product isotopologue with the deuteride located at the C_β position was identified by GC-MS (**Fig. S9**). This observation agrees with previous crystal and biochemical analysis which illustrated that the wild-type DHFR catalyzes the

transfer of the C4 pro-*R* hydride of NADPH.[58-59] Furthermore, this protein catalyst is regioselective favoring 1,4- over 1,2-addition, the latter was observed in the Sav-hosted secondary amine catalytic system as a side reaction.[22] Given its selectivity, primary kinetic isotope effect (KIE) of hydride transfer was assessed. The turnover rate constants under saturating conditions between the reactions that use NADPH and NADPD yielded a kinetic isotope effect (KIE) of 1.1 ± 0.2 (**Fig. S10**), implying that the step of hydride transfer is not rate-limiting. Other chemical step(s) such as iminium ion formation can be rate-limiting.[60] Alternatively, the physical step of releasing the oxidized cofactor from the enzyme can dictate the rate of catalytic turnover, as observed in the natural reaction catalyzed by the wild-type enzyme.[50, 53] Subsequently, the stereoselectivity of hydride transfer was investigated by use of the prochiral substrate β-methylated cinnamaldehyde. Though product formation was observed, its low conversion (24%) has prevented us from accurately accessing the enantioselectivity through chiral LC analysis (**Fig. 5**). Hence, for future work both the reactivity and stereoselectivity of reactions with β-alkylated cinnamaldehydes will be optimized through established protein engineering techniques such as directed evolution.[61-64]

Substrate scope of **LmrR-Phe93-1** and **DHFR-Ala7-1** were investigated. Aromatic and ketone analogues of cinnamaldehyde were evaluated as alternative substrates. Both systems were able to accept the aldehydes as substrates. Although different hydride donors were used, product conversions by **DHFR-Ala7-1** were generally higher than those by **LmrR-Phe93-1** with respect to per mol% of catalyst (**Fig. 5, Table S7 and S8**). However, the ketone substrate showed no turnover in either system.

Introducing catalytic motifs by genetic code expansion has allowed us to explore different protein scaffolds and hence discover a NADPH-dependent secondary amine catalyst. Although NADPH is structurally complex, catalysis by **DHFR-Ala7-1** could be coupled with an enzymatic cofactor regeneration scheme (**Fig. 6A**). Importantly, this experiment proves the concept that secondary amine organocatalysis, amongst other chemical catalytic systems,[65] can be driven by a well-established biocatalysis tool with the potential to be integrated into live cells. The conversion of starting material **4** to product **5** by **DHFR-Ala7-1** was performed in the presence of the reaction by glucose 6-phosphate dehydrogenase (G6PDH), a commercially available enzyme that oxidizes glucose 6-phosphate whilst reducing $NADP^+$ to regenerate NADPH.[66] The product conversion could reach 90% (10 μM NADPH) when the coupling reaction was included, whereas under the same reaction conditions the conversion was barely detectable (<1%) in the absence of G6PDH (**Fig. 6B**). A significant drop in conversion was observed when [NADPH] was less than 1 μM; a plausible explanation is that the DHFR variant becomes destabilized without cofactor binding. Nevertheless, the turnover number for NADPH (ratio of product formed to NADPH used) can reach as high as 10^4 (**Fig. 6C**).

Conclusion

Secondary amines are more nucleophilic than their primary counterparts, and their incorporations have converted poorly active, lysine-containing proteins into catalytically active entities. Previous approaches are limited to only a handful of protein templates.[17-23] Here, this limitation on protein templates has been

majorly alleviated by use of genetic code expansion, through which secondary amines were introduced to different scaffolds (sfGFP, LmrR and DHFR) at the position of choice. As a proof of concept, a secondary amine catalyst with unique cofactor dependence has been promptly established, indicating the benefit of using proteins as hosts. Provided the significant activity of secondary amines, it is expected that their incorporations by genetic code expansion will be further applied in the creation of artificial enzymes with designed activity.[61-64]

Declarations

Acknowledgements

We would like to thank EPSRC (EP/L027240/1) for the MS resources and Cardiff School of Chemistry for a PhD studentship to Thomas L. Williams. We would like to thank for the financial support provided by BBSRC (BB/T015799/1), Leverhulme Trust (RPG-2017-195) and Royal Society (RG170187) for Louis Y. P. Luk. We would like to thank Professor Dr Jason Chin at the MRC Laboratory of Molecular Biology for MbPylRS plasmid and Professor Dr Rudolf K Allemann at Cardiff University for the DHFR and TbADH plasmids. We also thank Professor Dr Anthony Green at the University of Manchester for the valuable discussion during manuscript preparation.

References

1. Rodrigues, M. J.; Windeisen, V.; Zhang, Y.; Guédez, G.; Weber, S.; Strohmeier, M.; Hanes, J. W.; Royant, A.; Evans, G.; Sinning, I.; Ealick, S. E.; Begley, T. P.; Tews, I., Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis. *Nat Chem Biol* **2017,** *13* (3), 290-294.

2. Liang, J.; Han, Q.; Tan, Y.; Ding, H.; Li, J., Current advances on structure-function relationships of pyridoxal 5'-phosphate-dependent enzymes. *Frontiers Molec Biosciences* **2019,** *6* (4).

3. Chen, L. H.; Kenyon, G. L.; Curtin, F.; Harayama, S.; Bembenek, M. E.; Hajipour, G.; Whitman, C. P., 4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per monomer. *J Biol Chem* **1992,** *267* (25), 17716-17721.

4. Subramanya, H. S.; Roper, D. I.; Dauter, Z.; Dodson, E. J.; Davies, G. J.; Wilson, K. S.; Wigley, D. B., Enzymatic ketonization of 2-hydroxymuconate: specificity and mechanism investigated by the crystal structures of two isomerases. *Biochemistry* **1996,** *35* (3), 792-802.

5. Whitman, C. P., The 4-oxalocrotonate tautomerase family of enzymes: how nature makes new enzymes using a β−α−β structural motif. *Arch Biochem Biophys* **2002,** *402* (1), 1-13.

6. Tchou, J.; Grollman, A. P., The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site. *J Biol Chem* **1995,** *270* (19), 11671-11677.

7. Sidorkina, O. M.; Laval, J., Role of the N-terminal proline residue in the catalytic activities of the Escherichia coli Fpg protein. *J Biol Chem* **2000,** *275* (14), 9924-9929.

8. Chamorro, E.; Duque-Noreña, M.; Pérez, P., A comparison between theoretical and experimental models of electrophilicity and nucleophilicity. *J Mol Struct THEOCHEM* **2009,** *896* (1-3), 73-79.

9. Erkkila, A.; Majander, I.; Pihko, P. M., Iminium catalysis. *Chem Rev* **2007,** *107* (12), 5416-70.

10. Brotzel, F.; Chu, Y. C.; Mayr, H., Nucleophilicities of primary and secondary amines in water. *J Org Chem* **2007,** *72* (10), 3679-3688.

11. MacMillan, D. W., The advent and development of organocatalysis. *Nature* **2008,** *455* (7211), 304-308.

12. Zou, Y. Q.; Hormann, F. M.; Bach, T., Iminium and enamine catalysis in enantioselective photochemical reactions. *Chem Soc Rev* **2018,** *47* (2), 278-290.

13. Xiang, S. H.; Tan, B., Advances in asymmetric organocatalysis over the last 10 years. *Nat Commun* **2020,** *11* (1), 3786.

14. Kubyshkin, V.; Budisa, N., The Alanine World Model for the Development of the Amino Acid Repertoire in Protein Biosynthesis. *Int. J. Molec. Sci.* **2019,** *20* (21), 5507.

15. Northrup, A. B.; MacMillan, D. W., Two-step synthesis of carbohydrates by selective aldol reactions. *Science* **2004,** *305* (5691), 1752-5.

16. Hein, J. E.; Blackmond, D. G., On the Origin of Single Chirality of Amino Acids and Sugars in Biogenesis. *Acc. Chem. Res.* **2012,** *45* (12), 2045-2054.

17. Zandvoort, E.; Baas, B.-J.; Quax, W. J.; Poelarends, G. J., Systematic Screening for Catalytic Promiscuity in 4-Oxalocrotonate Tautomerase: Enamine Formation and Aldolase Activity. *ChemBioChem* **2011,** *12* (4), 602-609.

18. Zandvoort, E.; Geertsema, E. M.; Quax, W. J.; Poelarends, G. J., Enhancement of the promiscuous aldolase and dehydration activities of 4-oxalocrotonate tautomerase by protein engineering. *Chembiochem* **2012,** *13* (9), 1274-1277.

19. Guo, C.; Saifuddin, M.; Saravanan, T.; Sharifi, M.; Poelarends, G. J., Biocatalytic asymmetric Michael additions of nitromethane to alpha,beta-unsaturated aldehydes via enzyme-bound iminium ion intermediates. *ACS Catal* **2019,** *9* (5), 4369-4373.

20. Biewenga, L.; Saravanan, T.; Kunzendorf, A.; van der Meer, J. Y.; Pijning, T.; Tepper, P. G.; van Merkerk, R.; Charnock, S. J.; Thunnissen, A. W. H.; Poelarends, G. J., Enantioselective synthesis of

pharmaceutically active gamma-aminobutyric acids using a tailor-made artificial Michaelase in one-pot cascade reactions. *ACS Catal* **2019,** *9* (2), 1503-1513.

21. Nodling, A. R.; Swiderek, K.; Castillo, R.; Hall, J. W.; Angelastro, A.; Morrill, L. C.; Jin, Y.; Tsai, Y. H.; Moliner, V.; Luk, L. Y. P., Reactivity and selectivity of iminium organocatalysis improved by a protein host. *Angew Chem Int Ed* **2018,** *57* (38), 12478-12482.

22. Santi, N.; Morrill, L. C.; Luk, L. Y. P., Streptavidin-Hosted Organocatalytic Aldol Addition. *Molecules* **2020,** *25* (10), 2457.

23. Santi, N.; Morrill, L. C.; Swiderek, K.; Moliner, V.; Luk, L. Y. P., Transfer hydrogenations catalyzed by streptavidin-hosted secondary amine organocatalysts. *Chem Commun* **2021,** *57* (15), 1919-1922.

24. Xu, G.; Crotti, M.; Saravanan, T.; Kataja, K. M.; Poelarends, G. J., Enantiocomplementary epoxidation reactions catalyzed by an engineered cofactor-independent non-natural peroxygenase. *Angew Chem Int Ed* **2020,** *59* (26), 10374-10378.

25. Dundas, C. M.; Demonte, D.; Park, S., Streptavidin-biotin technology: improvements and innovations in chemical and biological applications. *Appl Microbiol Biotechnol* **2013,** *97* (21), 9343-9353.

26. Le, Q.; Nguyen, V.; Park, S., Recent advances in the engineering and application of streptavidin-like molecules. *Appl Microbiol Biotechnol* **2019,** *103* (18), 7355-7365.

27. Mayer, C.; Dulson, C.; Reddem, E.; Thunnissen, A. W. H.; Roelfes, G., Directed evolution of a designer enzyme featuring an unnatural catalytic amino acid. *Angew Chem Int Ed* **2019,** *58* (7), 2083-2087.

28. Drienovska, I.; Mayer, C.; Dulson, C.; Roelfes, G., A designer enzyme for hydrazone and oxime formation featuring an unnatural catalytic aniline residue. *Nat Chem* **2018,** *10* (9), 946-952.

29. Zhou, Z.; Roelfes, G., Synergistic catalysis in an artificial enzyme by simultaneous action of two abiological catalytic sites. *Nat Catal* **2020,** *3* (3), 289-294.

30. Leveson-Gower, R. B.; Zhou, Z.; Drienovská, I.; Roelfes, G., Unlocking iminium catalysis in artificial enzymes to create a Friedel–Crafts alkylase. *ACS Catalysis* **2021,** *11* (12), 6763-6770.

31. Burke, A. J.; Lovelock, S. L.; Frese, A.; Crawshaw, R.; Ortmayer, M.; Dunstan, M.; Levy, C.; Green, A. P., Design and evolution of an enzyme with a non-canonical organocatalytic mechanism. *Nature* **2019,** *570* (7760), 219-223.

32. Polycarpo, C. R.; Herring, S.; Berube, A.; Wood, J. L.; Soll, D.; Ambrogelly, A., Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. *FEBS letters* **2006,** *580* (28-29), 6695-6700.

33. Nodling, A. R.; Spear, L. A.; Williams, T. L.; Luk, L. Y. P.; Tsai, Y. H., Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells. *Essays in biochemistry* **2019,** *63*

(2), 237-266.

34. Nguyen, D. P.; Elliott, T.; Holt, M.; Muir, T. W.; Chin, J. W., Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J Am Chem Soc* **2011,** *133* (30), 11418-11421.

35. Williams, T. L.; Iskandar, D. J.; Nodling, A. R.; Tan, Y.; Luk, L. Y. P.; Tsai, Y. H., Transferability of N-terminal mutations of pyrrolysyl-tRNA synthetase in one species to that in another species on unnatural amino acid incorporation efficiency. *Amino Acids* **2021,** *53* (1), 89-96.

36. Agustiandari, H.; Lubelski, J.; van den Berg van Saparoea, H. B.; Kuipers, O. P.; Driessen, A. J., LmrR is a transcriptional repressor of expression of the multidrug ABC transporter LmrCD in *Lactococcus lactis*. *J Bacteriol* **2008,** *190* (2), 759-763.

37. Roelfes, G., LmrR: A privileged scaffold for artificial metalloenzymes. *Acc Chem Res* **2019,** *52* (3), 545-556.

38. Bos, J.; Fusetti, F.; Driessen, A. J.; Roelfes, G., Enantioselective artificial metalloenzymes by creation of a novel active site at the protein dimer interface. *Angew Chem Int Ed* **2012,** *51* (30), 7472-7475.

39. Drienovska, I.; Alonso-Cotchico, L.; Vidossich, P.; Lledos, A.; Marechal, J. D.; Roelfes, G., Design of an enantioselective artificial metallo-hydratase enzyme containing an unnatural metal-binding amino acid. *Chem Sci* **2017,** *8* (10), 7228-7235.

40. Leveson-Gower, R. B.; Mayer, C.; Roelfes, G., The importance of catalytic promiscuity for enzyme design and evolution. *Nat Rev Chem* **2019,** *3* (12), 687-705.

41. Cattaneo, S.; Freakley, S. J.; Morgan, D. J.; Sankar, M.; Dimitratos, N.; Hutchings, G. J., Cinnamaldehyde hydrogenation using Au–Pd catalysts prepared by sol immobilisation. *Catalysis Science & Technology* **2018,** *8* (6), 1677-1685.

42. Ouellet, S. G.; Tuttle, J. B.; MacMillan, D. W., Enantioselective organocatalytic hydride reduction. *J Am Chem Soc* **2005,** *127* (1), 32-33.

43. Brogan, A. P.; Dickerson, T. J.; Janda, K. D., Nornicotine-organocatalyzed aqueous reduction of alpha,beta-unsaturated aldehydes. *Chem Commun* **2007,** (46), 4952-4954.

44. Betanzos-Lara, S.; Liu, Z.; Habtemariam, A.; Pizarro, A. M.; Qamar, B.; Sadler, P. J., Organometallic Ruthenium and Iridium Transfer-Hydrogenation Catalysts Using Coenzyme NADH as a Cofactor. *Angew. Chem. Int. Ed.* **2012,** *51* (16), 3897-3900.

45. Maenaka, Y.; Suenobu, T.; Fukuzumi, S., Hydrogen Evolution from Aliphatic Alcohols and 1,4-Selective Hydrogenation of NAD+ Catalyzed by a [C,N] and a [C,C] Cyclometalated Organoiridium

Complex at Room Temperature in Water. *J. Am. Chem. Soc.* **2012,** *134* (22), 9417-9427.

46. Maenaka, Y.; Suenobu, T.; Fukuzumi, S., Efficient Catalytic Interconversion between NADH and NAD+ Accompanied by Generation and Consumption of Hydrogen with a Water-Soluble Iridium Complex at Ambient Pressure and Temperature. *J. Am. Chem. Soc.* **2012,** *134* (1), 367-374.

47. Okamoto, Y.; Köhler, V.; Ward, T. R., An NAD(P)H-Dependent Artificial Transfer Hydrogenase for Multienzymatic Cascades. *J. Am. Chem. Soc.* **2016,** *138* (18), 5781-5784.

48. Liang, A. D.; Serrano-Plana, J.; Peterson, R. L.; Ward, T. R., Artificial Metalloenzymes Based on the Biotin–Streptavidin Technology: Enzymatic Cascades and Directed Evolution. *Acc. Chem. Res.* **2019,** *52* (3), 585-595.

49. Luk, L. Y.; Javier Ruiz-Pernia, J.; Dawson, W. M.; Roca, M.; Loveridge, E. J.; Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J.; Tunon, I.; Moliner, V.; Allemann, R. K., Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. *Proc Natl Acad Sci USA* **2013,** *110* (41), 16344-16349.

50. Fierke, C. A.; Johnson, K. A.; Benkovic, S. J., Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from Escherichia coli. *Biochemistry* **1987,** *26* (13), 4085-92.

51. Zhang, Z.; Rajagopalan, P. T.; Selzer, T.; Benkovic, S. J.; Hammes, G. G., Single-molecule and transient kinetics investigation of the interaction of dihydrofolate reductase with NADPH and dihydrofolate. *Proc Natl Acad Sci USA* **2004,** *101* (9), 2764-2769.

52. Miller, G. P.; Benkovic, S. J., Strength of an interloop hydrogen bond determines the kinetic pathway in catalysis by Escherichia coli dihydrofolate reductase. *Biochemistry* **1998,** *37* (18), 6336-42.

53. Luk, L. Y. P.; Loveridge, E. J.; Allemann, R. K., Protein motions and dynamic effects in enzyme catalysis. *Physical chemistry chemical physics : PCCP* **2015,** *17* (46), 30817-27.

54. Knaus, T.; Paul, C. E.; Levy, C. W.; de Vries, S.; Mutti, F. G.; Hollmann, F.; Scrutton, N. S., Better than nature: nicotinamide biomimetics that outperform natural coenzymes. *J Am Chem Soc* **2016,** *138* (3), 1033-9.

55. Guarneri, A.; Westphal, A. H.; Leertouwer, J.; Lunsonga, J.; Franssen, M. C. R.; Opperman, D. J.; Hollmann, F.; Berkel, W. J. H.; Paul, C. E., Flavoenzyme-mediated regioselective aromatic hydroxylation with coenzyme biomimetics. *ChemCatChem* **2020,** *12* (5), 1368-1375.

56. Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R., Directed evolution of artificial metalloenzymes for in vivo metathesis. *Nature* **2016,** *537* (7622), 661-665.

57. Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Kohler, V.; Lewis, J. C.; Ward, T. R., Artificial metalloenzymes: reaction scope and optimization strategies. *Chem Rev* **2018,** *118* (1), 142-231.

58. Sawaya, M. R.; Kraut, J., Loop and subdomain movements in the mechanism of Escherichia coli dihydrofolate reductase: crystallographic evidence. *Biochemistry* **1997,** *36* (3), 586-603.

59. Stojkovic, V.; Perissinotti, L. L.; Willmer, D.; Benkovic, S. J.; Kohen, A., Effects of the donor-acceptor distance and dynamics on hydride tunneling in the dihydrofolate reductase catalyzed reaction. *J Am Chem Soc* **2012,** *134* (3), 1738-1745.

60. Swiderek, K.; Nodling, A. R.; Tsai, Y. H.; Luk, L. Y. P.; Moliner, V., Reaction mechanism of organocatalytic Michael addition of nitromethane to cinnamaldehyde: A case study on catalyst regeneration and solvent effects. *The journal of physical chemistry. A* **2018,** *122* (1), 451-459.

61. Althoff, E. A.; Wang, L.; Jiang, L.; Giger, L.; Lassila, J. K.; Wang, Z.; Smith, M.; Hari, S.; Kast, P.; Herschlag, D.; Hilvert, D.; Baker, D., Robust design and optimization of retroaldol enzymes. *Protein Sci* **2012,** *21* (5), 717-726.

62. Garrabou, X.; Beck, T.; Hilvert, D., A promiscuous de novo retro-aldolase catalyzes asymmetric Michael additions via Schiff base intermediates. *Angew Chem Int Ed* **2015,** *54* (19), 5609-5612.

63. Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R.; Baker, D.; Griffiths, A. D.; Hilvert, D., Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nat Chem* **2017,** *9* (1), 50-56.

64. Hilvert, D., Design of protein catalysts. *Annu Rev Biochem* **2013,** *82* (1), 447-470.

65. Kohler, V.; Wilson, Y. M.; Durrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knorr, L.; Haussinger, D.; Hollmann, F.; Turner, N. J.; Ward, T. R., Synthetic cascades are enabled by combining biocatalysts with artificial metalloenzymes. *Nat Chem* **2013,** *5* (2), 93-99.

66. Wong, C.-H.; Whitesides, G. M., Enzyme-catalyzed organic synthesis: NAD(P)H cofactor regeneration by using glucose-6-phosphate and the glucose-5-phosphate dehydrogenase from *Leuconostoc mesenteroides*. *J Am Chem Soc* **2002,** *103* (16), 4890-4899.

67. Sandner, D.; Krings, U.; Berger, R. G., Volatiles from *Cinnamomum cassia* buds. *Z Naturforsch C J Biosci* **2018,** *73* (1-2), 67-75.

Figures

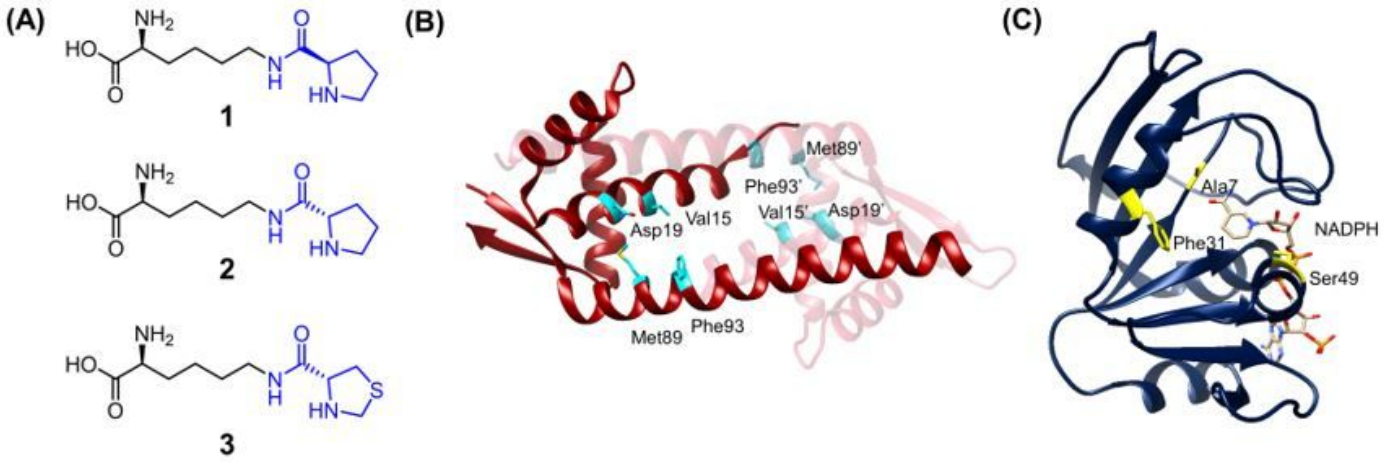
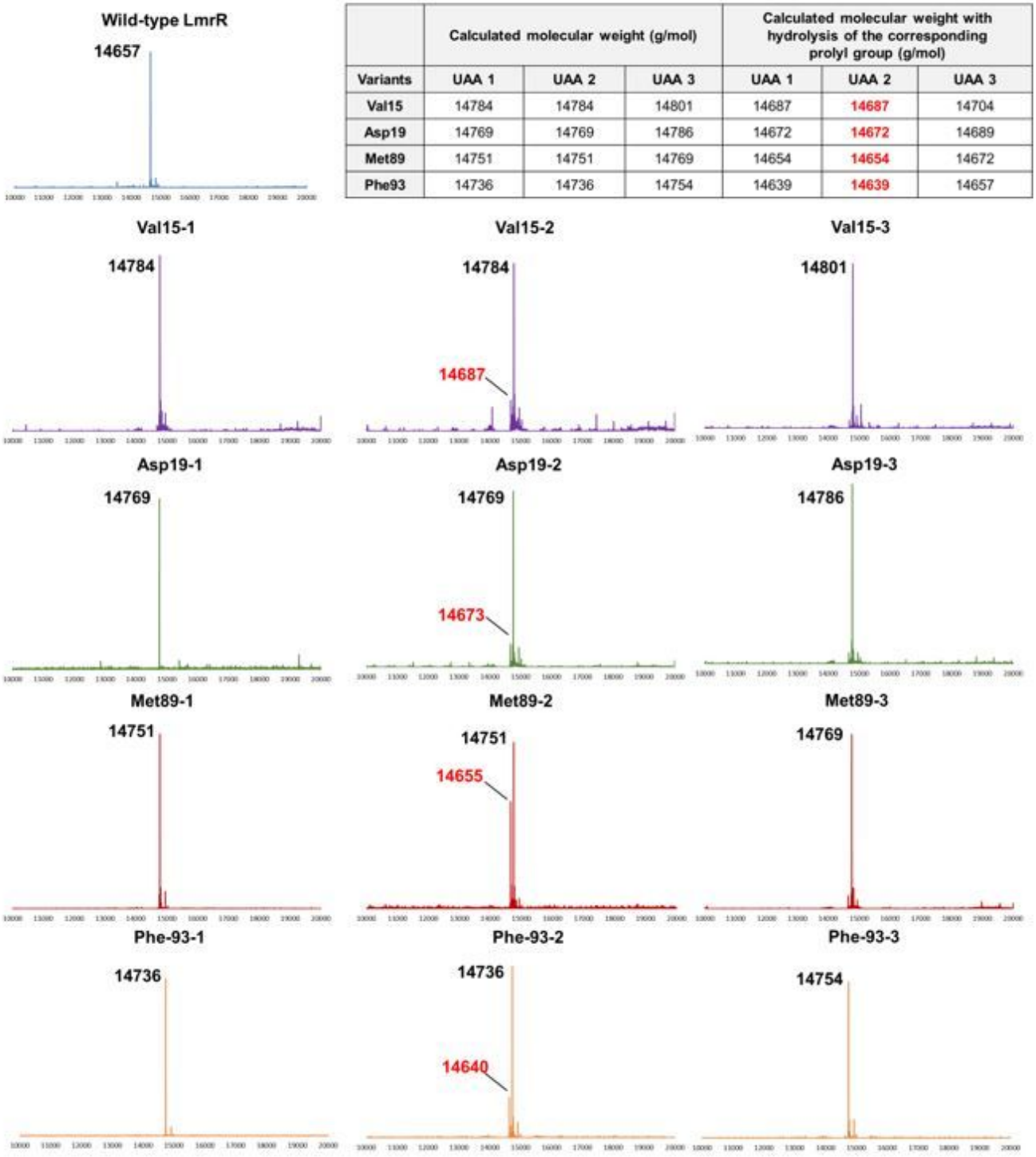


Figure 1

Components used in the design of the protein-based secondary amines. (A) The structures of the unnatural amino acids used include D-prolyl-L-lysine (1), L-prolyl-L-lysine (2) and L-thiazolidine-L-lysine (3). (B) The X-ray crystal structure of the Lactococcus multidrug resistant regulator LmrR (PDB: 3F8F). Two monomers are shown in red and light red. Residues targeted for unnatural amino acid incorporation are highlighted in cyan. (C) The X-ray crystal structure of E. coli dihydrofolate reductase (DHFR) with NADPH bound in the active site (PDB: 1RA1). Residues targeted for unnatural amino acid incorporation are highlighted in yellow.



	Calculated molecular weight (g/mol)			Calculated molecular weight with hydrolysis of the corresponding prolyl group (g/mol)		
Variants	UAA 1	UAA 2	UAA 3	UAA 1	UAA 2	UAA 3
Val15	14784	14784	14801	14687	14687	14704
Asp19	14769	14769	14786	14672	14672	14689
Met89	14751	14751	14769	14654	14654	14672
Phe93	14736	14736	14754	14639	14639	14657

Figure 2

Deconvoluted ESI mass spectra for the wild type LmrR (calculated to be 14657 g/mol) and its LmrR variants where Val15, Asp19, Met89 and Phe93 were individually replaced with unnatural amino acids 1, 2 or 3. All 13 proteins were found to have the N-terminal methionine removed. For variants incorporated with 2, profound peaks that match closely to hydrolysis of the L-prolyl group were observed (red fonts).

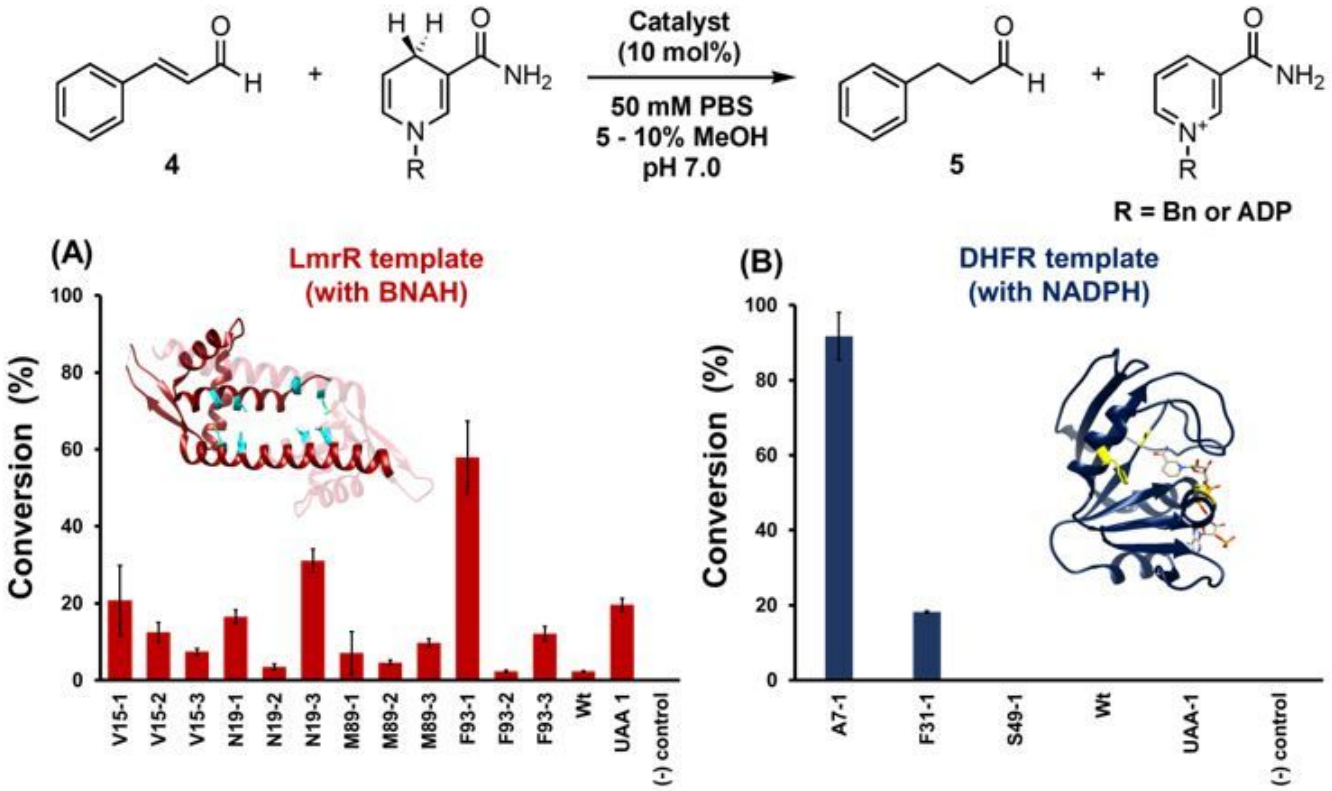


Figure 3

Assessment of the catalytic efficiency in the transfer hydrogenation reaction. (A) Conversion of the model reaction catalyzed by the LmrR variants using BNAH as determined by GC-MS (see SI). The model organocatalytic transfer hydrogenation reaction contained cinnamaldehyde (68 μM, 1 eq) and LmrR variants (6.8 μM), and then BNAH (136 μM, 2 eq) was added and stirred for 18 h in PBS buffer (pH 7.0, 10% methanol) at 25 °C. (B) Conversion of the model reaction catalyzed by the DHFR variants using NADPH as determined by GC-MS. The model organocatalytic transfer hydrogenation reaction contained cinnamaldehyde (52 μM, 1 eq) and DHFR variants (5.1 μM), and then NADPH (104 μM, 2 eq) was added and stirred for 18 h in PBS buffer (pH 7.0, 5% methanol) at 25 °C. The template experiments (A) and (B) were tested alongside the wild-type proteins, 1 and a negative (-) control where the reaction was performed without any catalyst. Each reaction was performed in triplicate and the mean value (± standard deviation) is shown.

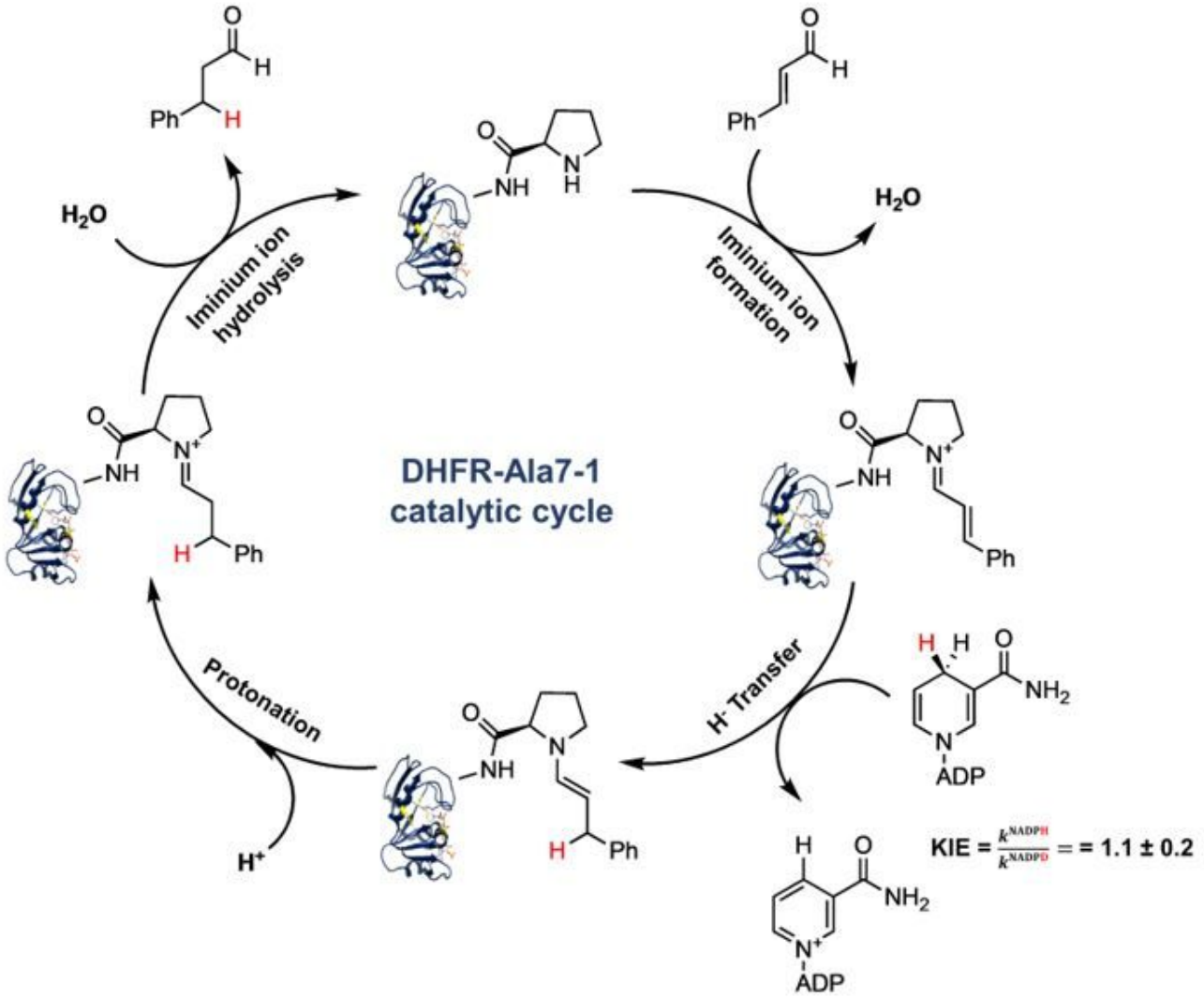


Figure 4

Catalytic cycle of the organocatalytic transfer hydrogenation reaction by DHFR-Ala7-1 proposed based on the observed trapped iminium intermediate (see Fig. S5B and S7). The unnatural amino acid forms an iminium ion with the α,β-unsaturated carbonyl substrate, and hydride transfer occurs from the pro-R position of NADPH to Cβ of the iminium intermediate. A kinetic isotope effect of 1.1 ± 0.2 was measured using NADPD and NADPH during catalysis.

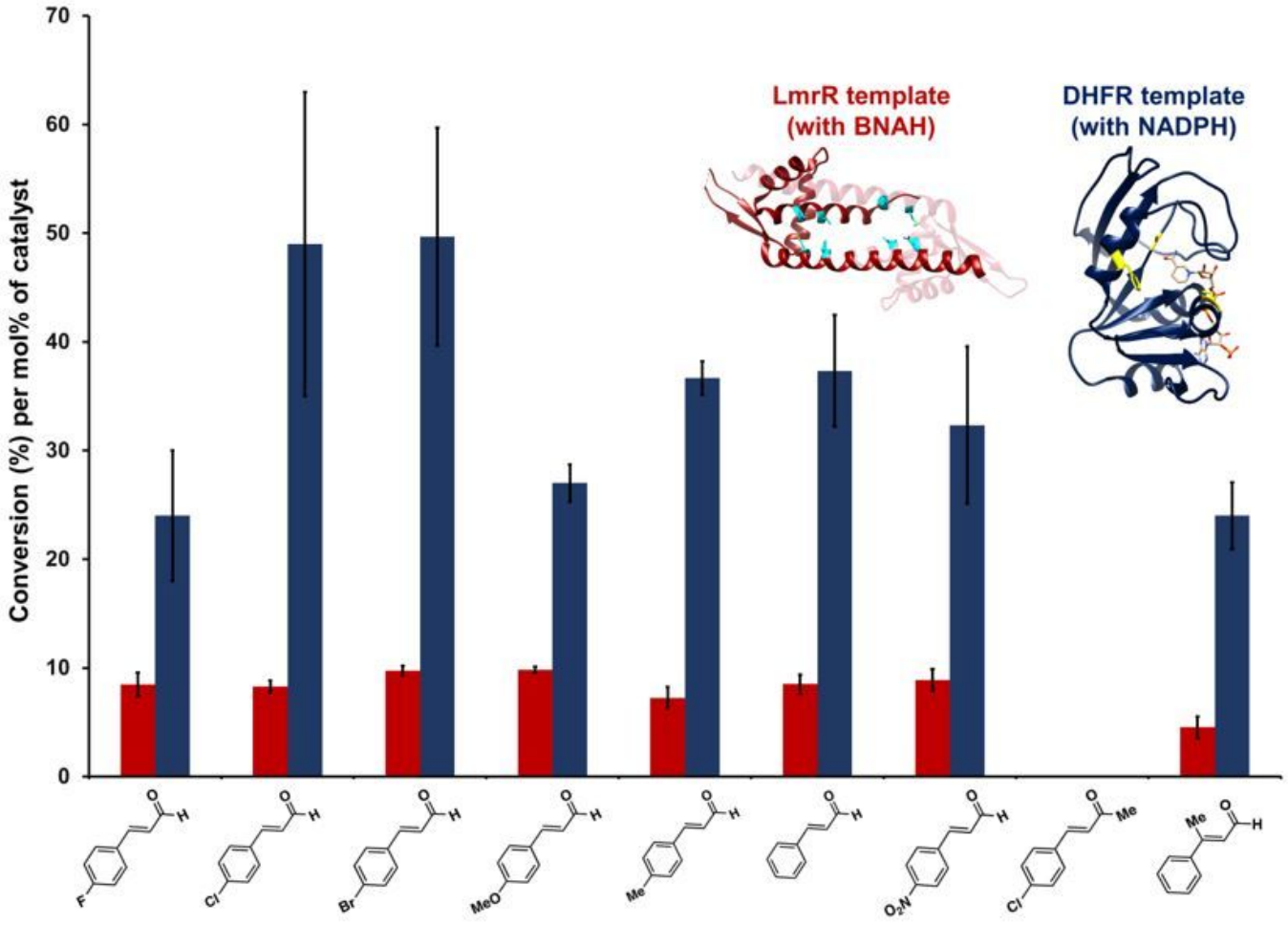
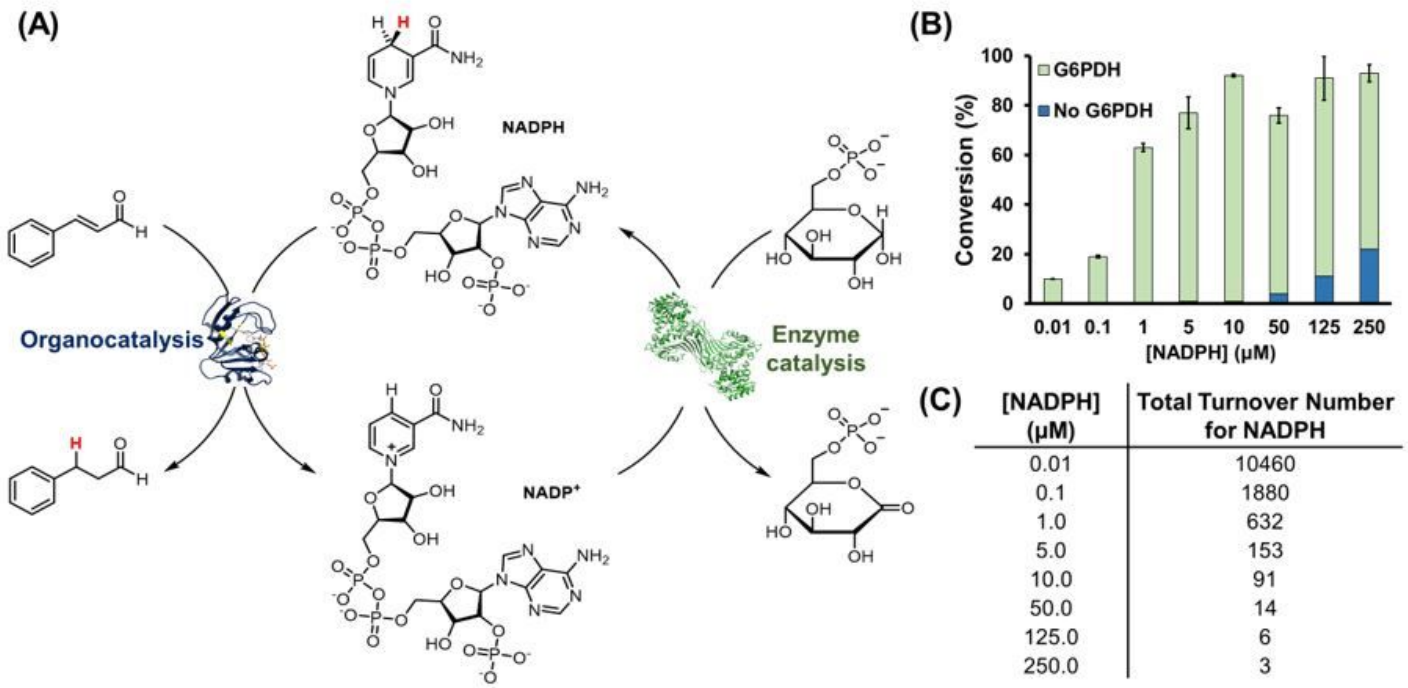


Figure 5

Substrate scope analysis of LmrR-Phe93-1 and DHFR-Ala7-1. The LmrR or DHFR variant (0.1 eq) was incubated with the indicated α,β-unsaturated carbonyl compound (1 eq) and their respective hydride donors (5 eq of BNAH and NADPH, respectively) for 48 hours. Conversion (%) per mol% was estimated by use of a 1H NMR spectroscopy as previously described (see SI).21-23 For the ketone, no change in the substrate could be observed. Each reaction was performed in triplicate and the mean conversion (± standard deviation) reported.



(C)

[NADPH] (μM)	Total Turnover Number for NADPH
0.01	10460
0.1	1880
1.0	632
5.0	153
10.0	91
50.0	14
125.0	6
250.0	3

Figure 6

Reaction scheme for the coupling between the organocatalytic DHFR-Ala7-1 and the enzymatic glucose-6-phosphate dehydrogenase (G6PDH) reactions. (A) The organocatalytic transfer hydrogenation reaction of cinnamaldehyde 4 (1 mM) by DHFR-Ala7-1 (10 mol%) was driven by the enzymatic G6PDH reaction (50 nM) which oxidizes glucose-6-phoshate (2 mM) to the corresponding lactone with the associated NADPH regeneration. (B) Product conversions were estimated by a GC-FID assay similar to previously described (see SI).67 (C) The total turnover number for NADPH refers the ratio of mole of product formed to mole of NADPH used. Each reaction was performed in triplicate and the mean reported.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.docx