Promiscuity guided evolution of a decarboxylative aldolase for synthesis of chiral tertiary alcohols

Meghan E. Campbell, Amanda R. Ohler, Matthew J. McGill, Andrew R. Buller*

Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706

* Andrew R. Buller

Email: arbuller@wisc.edu

Author Contributions: M.E.C., A.R.O, and M.J.M performed research; M.E.C. and A.R.B. designed research, analyzed data, and wrote the paper.

Competing Interest Statement: The authors have filed a patent related to this work.

Keywords: Protein engineering, directed evolution, non-canonical amino acid, biocatalysis, enamine

Abstract

Enzymes play an increasingly important role in synthetic biology and organic synthesis. Many potential applications benefit from promiscuous activity with a diverse array of substrates. Here, we show how to intentionally guide an enzyme towards generality through multi-generational directed evolution using substrate-multiplexed screening (SUMS). We demonstrate the advantages of promiscuity-guided evolution in a challenging context, engineering the decarboxylative aldolase UstD to perform a C-C bond forming reaction with ketone electrophiles. Mutations outside of the active site that impact catalytic function were immediately revealed by shifts in promiscuity, even when the overall activity was lower. By re-targeting these distal residues that couple to the active site with saturation mutagenesis, broadly activating mutations were readily identified. When analyzing active site mutants, SUMS identified both specialist enzymes that would have more limited utility as well as generalist enzymes with complementary activity on diverse substrates. These new UstD enzymes catalyze convergent synthesis of non-canonical amino acids bearing tertiary alcohol side chains. This methodology is easy to implement and enables the rapid and effective evolution of enzymes to catalyze desirable new functions.

Significance Statement

Many applications of enzymes benefit from activity on structurally diverse substrates. Traditional implementations of directed evolution generate catalysts that react with high selectivity for model substrates, particularly when multiple rounds of evolution are performed. Such specificity is often considered a hallmark of naturally evolved enzymes. We show how substrate-multiplexed screening can be used to intentionally engineer an enzyme for promiscuous activity with many substrates. This strategy simultaneously enables high-throughput identification of distal sites that influence catalytic activity, an historic challenge in enzymology. We demonstrate this promiscuity-based evolution to engineer a reaction that has no counterpart in Nature or traditional organic synthesis: decarboxylative aldol-type addition into unactivated ketone electrophiles, which enables the one-step synthesis of a rare class of amino acid.

Main Text

Introduction

Directed evolution has found widespread use enhancing enzymes for myriad industrial settings, including pharmaceuticals, fine chemicals, and bioremediation.[1–4] The power of directed evolution lies in iterative genetic diversification and selection to customize enzymatic properties.[5,6] This process is routinely implemented to increase enzyme activity, either by increasing catalyst concentration, stability, or reactivity under researcher-defined conditions.[7,8] The rapid identification of stabilizing mutations that do not compromise function still requires trial and error, but recent advances in computational design and machine learning are making rapid strides in this area.[9–12] In contrast, strategies to identify mutations that directly alter catalytic activity are more limited.[13,14] Mutagenesis of active site residues is reliable.[15] In cases where stereo- or regioselectivity is the focus, researchers may target distal sites that increase specificity, even if activity is compromised.[16,17] It remains to develop a generalizable strategy for experimental identification of distal sites that influence catalysis, which is a long-standing challenge in protein engineering.[18,19]

In standard implementations of directed evolution, a single model substrate is chosen for screening.[20–22] When activity on a single substrate is the desired goal, this process is uniquely powerful. When broad activity with diverse substrates is the goal, this method can also be successful, although it is typically unknown whether alternative variants that were screened might have activity with other substrates not under selection.[23–25] Indeed, there are many cases where directed evolution yields a catalyst that has high activity for the model substrate but struggles to

react with analogs.[26–30] This phenomenon is also observed in natural evolution, as modern enzymes are theorized to be less promiscuous than the ancient enzymes from which they evolved.[31] Ancestral sequence reconstruction has been successfully used to generate enzymes with altered promiscuity or thermostability for biocatalytic applications.[32–34] In some cases, intermediates along a directed evolution lineage are more promiscuous and evolution inadvertently limited the scope of the transformation.[26–30] More broadly, the inability to efficiently track substrate promiscuity during the initial screening phases hinders the ability to engineer enzymes towards generality.

Recently, we advanced substrate multiplexed screening (SUMS) as an alternative screening strategy for directed evolution. This method tracks enzyme promiscuity while minimizing researcher intervention and screening time.[35,36] By screening substrates in competition, changes to total activity and relative product abundance (promiscuity), are measured simultaneously (**Fig. 1A**). This method has been used in a few cases to discover large boosts in activity with defined subclasses of substrates using small active site libraries.[37–39] In other cases, a SUMS strategy was used to maintain previously existing promiscuity and to ensure mutations that turn on undesired activity are not accumulated during evolution.[40,41] We have reported the previous single example of SUMS combined with globally random mutagenesis.[35] It was observed that mutation at a residue flanking the active site of the tryptophan synthase β-subunit, TrpB, decreased activity but altered promiscuity. Subsequent site-saturation mutagenesis at this position then led to a variant with a boost in activity for many substrates, including substrates not in the original screens.[35] This intriguing result prompted us to explore whether SUMS could be leveraged reliably to identify mutations outside of the active site that directly impact enzyme promiscuity.[16]

Many uncertainties must be addressed to effectively traverse a protein fitness landscape using promiscuity information. When model systems that are poised to succeed are used for method development, one can overlook limitations to wider applicability.[42–45] For example, use of highly thermostable enzymes opens access to mutations that are activating, but destabilizing. Strategies that rely on such mutations may not translate to mesophilic homologs. For this reason, we opted not to continue using TrpB for the present study.[39–42] The assays for engineering methodology can also impose limitations on potential utility. While in vivo, continuous evolution is the most powerful protein engineering approach,[46] reliance on in vivo compatibility can limit the transformations being explored.[47] Similarly, engineering strategies that rely on fluorescence reporters to screen >10^4 sequences per round may not apply when the desired reaction must be measured in lower throughput.[48,49]

Here, we explore the effectiveness of multi-generational, promiscuity-guided engineering in the context of a historically challenging enzymatic reaction, aldol addition into ketones.[50–52] The PLP-dependent decarboxylative aldolase, UstD, was discovered to catalyze the final step of ustiloxin B biosyntheis. [53,54] This enzyme generates an enamine nucleophile that adds into an aldehyde in a convergent C–C bond forming reaction yielding a γ-hydroxy non-canonical amino acid (**Fig. 1B**). Subsequently, Zhang *et al.* utilized a homolog of UstD from *Aspergillus pseudonomius* to demonstrate the broad native activity of *Ap*UstD with benzaldehyde derivatives.[55] We reported the classical evolution of UstD from *Aspergillus flavus* for improved activity using benzaldehyde as a model substrate.[9] This evolution, serendipitously, generated a 'generalist' enzyme, dubbed UstD$^{2.0}$, capable of reacting with diverse aldehyde electrophiles to generate chiral secondary alcohols. Reactions with ketone electrophiles would yield chiral tertiary alcohols, a highly sought motif in medicinal chemistry.[56–58] A recent report by Zhang *et al.* demonstrated *Ap*UstD can react with di-ketones. However, the reaction was limited to the activated dione motif (**Fig. 1C**).[59] Aldol-type reactions with ketones are deceptively challenging compared to aldehydes, as they are thermodynamically more stable and kinetically slower to react.[60,61] Hence, aldol addition into unactivated ketones is a formidable testing ground for method development. Additionally, UstD and its homologs come from a niche area of secondary metabolism and the native sequence diversity is limited to mesophilic organisms. The two characterized homologs have low soluble expression in *E. coli*. We anticipated that the use of

such a modest enzyme to overcome a challenging transformation would provide a candid assessment of SUMS and promiscuity-guided evolutionary strategies.

Results and Discussion

Identification of distal 'hotspots' from global random mutagenesis

By screening on a substrate mixture, we hypothesized that changes in the relative rates of reactivity would aid in the identification of residues outside the active site that are influencing catalysis. We first considered the design of the substrate space by testing UstD$^{2.0}$ activity against a mixture of substrates with distinct steric and electronic properties. Two aldehydes, thiophene-3-carboxaldehyde (**1a**), *o*-tolualdehyde (**1b**) were chosen as well as an unactivated ketone (**1c**). In direct competition, robust signals for the secondary alcohol products (**2a**, **2b**) were observed via ultra-high pressure liquid chromotagraphy-mass spectrometry (UPLC-MS) and indicated each substrate reacts with similar efficiency. The ketone **1c** was added at a five-fold higher concentration, which enabled observation of a small signal corresponding to product formation (**2c**, **Fig. S1**). We found that assaying UstD as a whole cell catalyst provided reproducible data, further simplifying the screening process.

We screened 880 clones from a global random mutagenesis library (see SI for details). From this screen we identified no variants with general boosts in activity on all substrates. Some variants, however, appeared to have changes in promiscuity (**Fig. 2A**). Diverse strategies to quantitate promiscuity have been developed.[62,63] However, applications of these metrics to our data demonstrated they have limited utility because activity on **1c** is low. Most variants had reduced activity relative to parent and the noise associated with low-intensity measurements was indistinguishable from a change in promiscuity. Nevertheless, visual inspection suggested a range of promiscuity shifting effects in some variants. We selected 7 such variants for sequencing, from which we identified 12 mutations (**Fig. 2A**). Analysis of the UstD$^{2.0}$ crystal structure showed that no mutations were in the active site and the alpha carbons (Cα) of the mutations were an average 18 Å away from the cofactor (**Fig. 2B**). We hypothesized that these mutations may be 'hotspots' or altering activity. Although these specific mutations are deleterious under the screening conditions, some other mutation at the same site may be beneficial for catalysis, which we tested with site saturation mutagenesis (SSM).

Mutation of distal promiscuity-shifting sites reveals activating mutations

We selected P82 as an initial site to test this 'hotspot' hypothesis. This residue is 11 Å from the catalytic Lys and the crystal structure of UstD$^{2.0}$ showed that it forms a cis-peptide bond. Mutation at Pro therefore has the potential to introduce more pervasive structural changes. We also considered a site from random mutagenesis where there was a comparable decrease in activity, but no significant change in promiscuity, G373. To our surprise, SSM at both positions revealed generally activating mutations. When re-screened outside of a competition setting on just **1c**, the top variants were P82Q, and G373E, which had 9.6, and 1.2-fold higher activity compared to parent, respectively (**Fig. S2**). These mutations were additive and the new variant, QE, had 11.7-fold improvement in activity with ketone **1c** (**Fig. S2**). As the mutations are distal to the active site, it is difficult to identify the molecular mechanism through which these mutations operate. Nevertheless, this increase in activity is significant because UstD$^{2.0}$ activity with **1c** was nearly stoichiometric and crossed the threshold to catalytic, albeit a modest 17 turnovers with QE (**Fig. S3**).

Substrate space redesign can increase sensitivity to changes in reactivity

The initial substrate mixture was chosen for its simplicity. Spurred by the success of the promiscuity-guided evolution above, we considered whether more information might be accessible from SUMS when using a substrate mixture with more diverse electrophiles. To maintain some continuity between the global random mutagenesis and the SSM screens both **1b** and **1c** were retained in a new substrate mixture. We introduced two new ketones, 1,1,1-trifluoro-3-phenyl-2-propanone (**1d**) and 4'-nitroacetophenone (**1e**). In simple mixtures with equimolar concentrations,

the activated electrophiles (**1b**, **1d**) dominated the overall reactivity and the signal for the tertiary alcohol product**, 2e**, was on the order of experimental noise (**Fig. 3**). We therefore altered the substrate concentrations to ensure reproducible and robust signals for all products via UPLC-MS. The resulting mixture contained a 9:1 ratio of ketone to aldehyde substrates. In this straightforward way, substrate mixtures for multi-generational evolutionary campaigns can be re-tuned as reactivity is expanded.

We used QE as the parent enzyme for SSM at distal promiscuity-shifting sites identified from globally random mutagenesis (**Fig. 2B**). Although none of these mutations led to general increases in activity, SSM mutagenesis at eight of 12 sites led to variants with at least a 1.5-fold increase in total activity during screening (**Fig. S5**). We observed that mutations D86V and V330A both produced a 2-fold increase in activity. We also found, serendipitously, that synonymous codon changes at I141 and S371 led to boosts in whole-cell catalyst activity, presumably by increasing the soluble enzyme expression. We again screened a SSM library at a control site. The Y418H mutation was identified by globally random mutagenesis as one that decreased activity without a significant shift in promiscuity and SSM at Y418 did not lead to increases in activity (**Fig. S6**).

This two-step process of identifying 'hotspots' by promiscuity shifts followed by SSM led to activating mutations with no prior information about protein structure, dynamics, or evolution. This strategy is distinct from stereo- or regioselectivity-based screening, which only considers positions whose initial hits provide the desired change in selectivity.[16,17] The key addition made by our approach is that *any* shift in promiscuity, not just one in the desired direction, is now recognized as a 'hit' for subsequent mutagenesis. While future studies may untangle the basis of these activating effects, the present investigation is focused on the development of a practical engineering approach that incorporates promiscuity information. We continued by deploying another common step in enzyme evolution: recombination of activating mutations.

Identification of cooperative mutational effects in a recombination library

There are many successful strategies for designing and screening recombination libraries.[64–68] We considered five different positions (F75, D86, I141, V330, S407) that were distributed across the protein structure. Screening data was used to identify degenerate codons that limit the inclusion of mutations that are deleterious when introduced independently (**Fig. S5**, **S7**). The resulting library consisted of ~2,800 possible variants (See Table S1). Here we acknowledge a tradeoff between library size and screening intensity. It is not necessary to exhaustively screen all possible combinations, but rather to use a carefully crafted library to efficiently traverse the largest practically, accessible recombination space.

We maintained the substrate mixture from the previous round of evolution (**Fig. 3**) and increased reaction time from one hour to eight hours to reduce the risk of selecting destabilizing mutations or those that accelerate decomposition of product through retro-aldol cleavage. We screened 704 clones, representing a maximum of ~24% of the theoretical library space and observed increases in total activity up to 2.6-fold (**Fig. S8**). The top variants were validated using single substrate reactions leading to the quadruple variant AIIRQ (F75A, D86I, V330R, S407Q). This variant displayed increased activity with all ketone substrates (~1.5–2.5-fold change) and only a modest decrease in activity with the aldehyde (80% parent activity, **Fig. S8**). Additionally, AIIRQ had higher soluble expression (~80 mg protein/ L culture) compared to QE and the other recombination variants (~50 mg protein/ L culture). Therefore, AIIRQ was chosen as the new parent enzyme for subsequent evolution.

Active site engineering reveals two mutants with distinct promiscuity

Previously, we identified a loop region in the UstD active site encompassing residues 391-393 where several mutations were found that increased activity on aldehydes.[9] We hypothesized that combinatorial retargeting of these sites would reveal variants with higher activity. Two additional sites, M299 and T388 were also included because both have side chains that protrude into the active site (**Fig. 4**). We had screened a SSM library at M299 with QE as the parent, which showed conservative mutations were generally neutral to activating (**Fig. S9**). With this information, we designed a focused recombination library of 2,300 variants. Cooperative

effects are particularly common in enzyme active sites and there is a higher propensity for multiple active site mutations to be deleterious, raising the specter of laborious screening of predominately inactive sequence space.[49] We therefore biased the mutational load from 3.9 to 3.3 mutations per variant during library construction (see supporting information for details).

We screened 968 clones within this focused library, representing a maximum 42% of the theoretical library space. The promiscuity profiles of these variants showed more diversity than earlier screens, as expected (**Fig. S10**). From these libraries we identified seven variants with disparate reactivity profiles (**Fig. S11**). These activated enzymes had distinct amino acid sequences and contained an average of 3.1 mutations. None of these activated variants contained mutation at all 5 sites simultaneously, supporting the choice to decrease the mutational load. One notable variant, 7G04 (M299V, T388I) had much higher activity with **1c**, upwards of a 14-fold boost, but had minimal changes with other substrates. While this activity is impressive, this variant represents a specialist that one would want to *avoid* when the goal is to develop a catalyst with general synthetic utility. We therefore selected two variants with distinct and broad reactivity profiles, 7G11 (M299V, T391S, M393W) and 7B05 (T391S, M393F) for further exploration.

Lineage analysis of promiscuity guided evolution

Characterizing variants in an evolutionary lineage is often used to show how successive rounds of mutagenesis affect activity.[29,69] Such retrospective analyses are typically limited to the single transformation under selective pressure, with the notable exception of evolution that features a substrate walking strategy.[70] Prior to lineage analysis of UstD, we performed a brief survey of reaction conditions using the variant 7G11. Increasing the concentration of L-Asp from 50 to 250 mM increased yields of **2c** and **2d** (see SI for details, **Fig. S12**). While such high concentrations are not ideal, the amino acid is cheap and commercially available. Additional equivalents of PLP relative to enzyme (between 10–50-fold excess) were also beneficial for both substrates. To understand how promiscuity changed through evolution, we performed a lineage analysis with a small set of substrates without competition under these optimized conditions.

We observed steady increases in activity with substrates that were under direct selective pressure (**1c** and **1d**), leading to the two enzymes 7B05 and 7G11 (**Fig. 5,** green and grey bars). The variant 7B05 performed 2,120 turnovers with **1d**, representing a ~4-fold increase in activity. The activity of UstD$^{2.0}$ with **1c** was negligible. The final variant, 7G11, performed a modest 67 turnovers with **1c**. While this reaction moved from stoichiometric to catalytic, further improvements could come from additional rounds of evolution.

We probed the effects of evolution on substrates not under selective pressure and found unexpected results. The parent enzyme, UstD$^{2.0}$, was previously evolved for activity with benzaldehyde (**1g**) on which it performed ~9,700 turnovers.[9] However, UstD activity with **1g** *decreased* steadily throughout the evolution (**Fig 5**, blue bars). Because aldehydes are intrinsically more reactive than ketones, we had originally anticipated that any molecular changes that accelerate the C–C bond forming step could be expected to work with both classes of substrate. In contrast, activity with furylacetone (**1f**) increased ~6-fold (**Fig. 5**, red bars), indicating that evolution enhanced activity on a substrate not under direct selective pressure.

We considered how the activity might increase with ketone substrates, but not an aldehyde. One possibility is that distal mutations increased activity by tuning the lifetime of the reactive enamine nucleophile (**Fig 1B**). Once the enamine forms, it experiences a kinetic competition between protonation, which quenches the nucleophile to form L-alanine (Ala), and C–C bond formation with an electrophile.[53] With highly reactive electrophiles, changes to this competition would have a negligible effect. We measured the ratio of protonation vs C–C bond formation with **1d** and found that protonation was favored by a factor of 2 in the parent enzyme, consistent with the challenging nature of the C–C bond formation. This rate of protonation was suppressed with the AIIRQ variant, such that C–C bond formation was preferred (**Fig. 5**). This ratio further shifted with the active site recombinant 7B05, which had lower Ala formation and favored C–C bond formation by 4-fold.

Sequence-activity relationships indicate that complex molecular effects are influencing activity. The variants 7B05 and 7G11 share the T391S mutation and differ in residues 299 and 393 (**Fig. 4**). Evidence for cooperativity within the active site comes from the recombination library, where we serendipitously acquired sequence-function data that corresponds to a stepwise mutational walk between the two activated enzymes. The two possible intermediate mutants have <15% of parent activity (**Fig. S13**). Hence, active site recombination captured epistatic interactions between these sites. Further assessment of enzyme mechanism and the impacts of particular mutations or collections thereof may be fertile ground for future inquiry.

Convergent biocatalytic synthesis of chiral tertiary alcohols

The synthetic utility of 7G11 and 7B05 was assessed on both analytical scale and preparative scale. Beginning with analytical scale analysis, we observed 7B05 had higher activity for some highly activated trifluoromethyl ketone substrates (**Fig. S14**). In most others, 7G11 was the more proficient catalyst (**Fig. S14**). A subset of these amino acids was chosen for preparative-scale reactions in order to demonstrate isolation strategies for chemically diverse products. Efficient isolation of aromatic amino acids was possible using reverse-phase flash chromatography (**Fig. 6,** see SI for further details) and stereoselectivity was assessed using Marfey's analysis.[71,72]

Trifluoroacetophenone (**1h**) reacted efficiently with 7G11 and **2h** was isolated in 55% yield and excellent diastereoselectivity (>20:1). 7B05 converted **1d**, one of the substrates under selection during evolution, to **2d** in 46% yield with excellent d.r. (>20:1). Excitingly, the heterocyclic amino acid, **2f** was isolated in 39% yield and excellent d.r. (>20:1). Product **2i** was isolated in 31% yield and high d.r. (18:1). A crystal structure of **2i** revealed the absolute configuration as 2*S*,4*S*,5*R*. Hence, 7G11 maintains the same configurational preference shown with aldehyde substrates wherein the α-amine and γ-hydroxy are *anti* to one another in a standard linear depiction (**Fig. 5, S15, S16**). When the ketone is positioned within a ring, we observed exceptional activity and **2j** was isolated in 96% yield as a 1:1 mixture of diastereomers.

The aliphatic ncAA, **2k** was sufficiently hydrophobic to be isolated using reverse phase chromatography in good yield (59%) as a 1:1 mixture of diastereomers. Both diastereomers were structurally characterized by small molecule crystallography (**Fig. S17**). The more hydrophilic amino acids **2l** and **2m** were isolated by addition of Fmoc in a telescoped fashion. The protected amino acids were subsequently isolated using normal-phase chromatography. This process led to **2l** in 62% yield and excellent e.r. (98:2). Yield of the trihydroxyleucine **2m**, derived from reaction with 1,3-dihydroxyacetone (**1m**), was initially low (36%). Higher yield was straightforward to obtain by increasing the catalyst loading to 0.4 mol% 7G11, which gave **2m** in a 71% isolated yield (99:1 e.r.) as the Fmoc-protected amino acid. Trihydroxyleucine is found in a fungal natural product[73] and was previously synthesized in four steps and isolated in 32% yield as the lactone.[74] For all other amino acids, UstD enabled the first reported synthesis.

Conclusions

We report a multigenerational protein engineering campaign that uses changes in promiscuity and activity information together to guide evolution. Distal residues that were coupled to the active site were discovered in high-throughput via SUMS, without the need for detailed structural or mechanistic studies. This promiscuity information effectively increases the utility of global-random mutagenesis by an order of magnitude, as the hit rate for promiscuity-shifting mutations was approximately 1:100 in a mesostable protein backbone for a challenging chemical reaction. Previously, distal 'hotspots' could only reliably be found from detailed sequence[75] or structural analysis[76] or experimentally with kinetic screening using fluorescence reporters, or specialized microfluidic devices.[48,77,78] The SUMS approach is a simple experimental extension of plate-based directed evolution.

Promiscuity-guided evolution helped to both find generally activating mutations and to avoid active site mutations that led to specialization. We identified two variants with complementary increases in activity that we used to synthesize a small set of chemically diverse amino acids. These new decarboxylative aldolases represent a significant advance as they

provide a new strategy for the convergent synthesis of chiral tertiary alcohols. Future work will encompass mechanistic investigations into the specific molecular determinants for UstD activity and the application of promiscuity-guided evolution to new classes of enzymes.

While SUMS engineering provides benefits traditional engineering cannot, it also has drawbacks. There is modestly more reaction optimization prior to screening and the data analysis is more complex. Promiscuity profiles allow researchers to directly observe reaction scope, a boon when pursuing generalists. However, it becomes more difficult to define the "best" variant, as activated variants can encompass changes in promiscuity, activity, or both. We emphasize, however, that such tradeoffs have always been occurring in directed evolution, researchers were simply blind to these effects during screening.

Materials and Methods

Protein expression and purification

An overnight culture of *E. coli* BL21(DE3) harboring a pET-22b(+) plasmid encoding a given UstD$^{2.0}$ variant was inoculated with a single colony and grown overnight. The overnight culture was then used to inoculate 1 L of TB_{amp}, which was grown to an optical density (OD) of 0.4–0.6 before induction with 100 μM IPTG. After overnight incubation, cells were then harvested by centrifugation and the cell pellets were stored at -20 ^{o}C until purification. Following lysis and sonication the cell lysate supernatant was purified using nickel affinity chromatography. Additional details can be found in SI Appendix.

Library generation and screening for directed evolution

Mutagenesis was carried out via PCR and the DNA product was purified using a preparative agarose gel. The purified DNA fragment was inserted into a pET-22b(+) vector by the Gibson Assembly method. BL21 (DE3) *E. coli* cells were subsequently transformed with the resulting cyclized DNA product via electroporation. Following recovery, cells were plated onto LB plates with 100 μg/mL Ampicillin (amp) and incubated overnight. Single colonies were used to inoculate a 96-well plate containing TB_{amp} in each well was inoculated with single colonies. The plates were grown overnight. Expression plates were prepared with TB_{amp} in each well and inoculated from the overnight culture. The expression plates were grown for 3 h then placed on ice before induction with 0.1 mM IPTG. The expression plates were grown overnight then centrifuged to harvest the cell pellets. Expression plates were stored at -20 °C until screening. Variants were evaluated for changes in activity and promiscuity by SUMS. Subsequently, reactions were quenched, clarified, and analyzed by UPLC-MS. Additional details can be found in SI Appendix.

Lineage analysis

In each well of a 96-well plate one electrophile, L-aspartate, pyridoxal-5'-phosphate and potassium phosphate buffer were added. Reactions were initiated by addition of UstD. After 4 h, reactions were quenched and denatured enzyme was removed by filtration. Marfey's derivatization of the amino acid products allowed for quantitation by UPLC-MS at 340 nm using a standard curve. Additional details can be found in SI Appendix.

Preparative scale biocatalysis

A 100-mL round bottom flask was charged with the desired ketone, methanol, potassium phosphate buffer, L-aspartate sodium salt monohydrate, and PLP. The reaction was initiated upon addition of enzyme (7G11 or 7B05). After 4 h, the reaction mixture was quenched with acetonitrile and centrifuged to remove aggregated protein. The decanted supernatant was then concentrated and purified via flash chromatography.

Acknowledgments

We thank the members of Buller group for many helpful conversations. This work was supported by the Office of the Vice Chancellor for Research and Graduate Education at the University of

Wisconsin-Madison with funding from the NIH (DP2- GM137417) to A.R.B. We thank NIH Grant 1S10OD020022-1 for providing funding for the Q Extractive Plus Orbitrap used for high-resolution mass spectrometry analysis of prepared compounds. The purchase of the Bruker D8 VENTURE Photon III X-ray diffractometer was partially funded by NSF Award #CHE-1919350 to the UW–Madison Department of Chemistry.

References

(1) Huffman, M. A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K. R.; Canada, K. A.; Devine, P. N.; Duan, D.; Forstater, J. H.; Grosser, S. T.; Halsey, H. M.; Hughes, G. J.; Jo, J.; Joyce, L. A.; Kolev, J. N.; Liang, J.; Maloney, K. M.; Mann, B. F.; Marshall, N. M.; McLaughlin, M.; Moore, J. C.; Murphy, G. S.; Nawrat, C. C.; Nazor, J.; Novick, S.; Patel, N. R.; Rodriguez-Granillo, A.; Robaire, S. A.; Sherer, E. C.; Truppo, M. D.; Whittaker, A. M.; Verma, D.; Xiao, L.; Xu, Y.; Yang, H. Design of an in Vitro Biocatalytic Cascade for the Manufacture of Islatravir. *Science (1979)* **2019**, *366* (6470), 1255–1259. https://doi.org/10.1126/science.aay8484.

(2) France, S. P.; Lewis, R. D.; Martinez, C. A. The Evolving Nature of Biocatalysis in Pharmaceutical Research and Development. *JACS Au* **2023**, *3* (3), 715–735. https://doi.org/10.1021/JACSAU.2C00712/ASSET/IMAGES/LARGE/AU2C00712_0037.JPEG.

(3) Radley, E.; Davidson, J.; Foster, J.; Obexer, R.; Bell, E. L.; Green, A. P. Engineering Enzymes for Environmental Sustainability. *Angewandte Chemie International Edition* **2023**, *62* (52), e202309305. https://doi.org/10.1002/ANIE.202309305.

(4) Wu, S.; Snajdrova, R.; Moore, J. C.; Baldenius, K.; Bornscheuer, U. T. Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angewandte Chemie International Edition* **2021**, *60* (1), 88–119. https://doi.org/10.1002/ANIE.202006648.

(5) Chen, K.; Arnold, F. H. Tuning the Activity of an Enzyme for Unusual Environments: Sequential Random Mutagenesis of Subtilisin E for Catalysis in Dimethylformamide. *Proceedings of the National Academy of Sciences* **1993**, *90* (12), 5618–5622. https://doi.org/10.1073/PNAS.90.12.5618.

(6) Arnold, F. H. The Nature of Chemical Innovation: New Enzymes by Evolution *. *Q Rev Biophys* **2015**, *48* (4), 404–410. https://doi.org/10.1017/S003358351500013X.

(7) Reetz, M. T. Laboratory Evolution of Stereoselective Enzymes: A Prolific Source of Catalysts for Asymmetric Reactions. *Angewandte Chemie International Edition* **2011**, *50* (1), 138–174. https://doi.org/10.1002/ANIE.201000826.

(8) Buller, R.; Lutz, S.; Kazlauskas, R. J.; Snajdrova, R.; Moore, J. C.; Bornscheuer, U. T. From Nature to Industry: Harnessing Enzymes for Biocatalysis. *Science (1979)* **2023**, *382* (6673). https://doi.org/10.1126/SCIENCE.ADH8615/ASSET/F4FDD437-C48A-4541-A9C1-09C12EC9B06D/ASSETS/IMAGES/LARGE/SCIENCE.ADH8615-FA.JPG.

(9) Ellis, J. M.; Campbell, M. E.; Kumar, P.; Geunes, E. P.; Bingman, C. A.; Buller, A. R. Biocatalytic Synthesis of Non-Standard Amino Acids by a Decarboxylative Aldol Reaction.

Nature Catalysis 2022 5:2 **2022**, *5* (2), 136–143. https://doi.org/10.1038/s41929-022-00743-0.

(10) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. Robust Deep Learning–Based Protein Sequence Design Using ProteinMPNN. *Science (1979)* **2022**, *378* (6615), 49–56. https://doi.org/10.1126/SCIENCE.ADD2187/SUPPL_FILE/SCIENCE.ADD2187_SM.PDF.

(11) Dantas, G.; Correnti, C.; Reichow, S. L.; Havranek, J. J.; Eletr, Z. M.; Isern, N. G.; Kuhlman, B.; Varani, G.; Merritt, E. A.; Baker, D. High-Resolution Structural and Thermodynamic Analysis of Extreme Stabilization of Human Procarboxypeptidase by Computational Protein Design. *J Mol Biol* **2007**, *366* (4), 1209–1221. https://doi.org/10.1016/J.JMB.2006.11.080.

(12) Rapp, J. T.; Bremer, B. J.; Romero, P. A. Self-Driving Laboratories to Autonomously Navigate the Protein Fitness Landscape. *Nature Chemical Engineering* **2024**, *1* (1), 97–107. https://doi.org/10.1038/s44286-023-00002-4.

(13) Honda Malca, S.; Duss, N.; Meierhofer, J.; Patsch, D.; Niklaus, M.; Reiter, S.; Hanlon, S. P.; Wetzl, D.; Kuhn, B.; Iding, H.; Buller, R. Effective Engineering of a Ketoreductase for the Biocatalytic Synthesis of an Ipatasertib Precursor. *Communications Chemistry 2024 7:1* **2024**, *7* (1), 1–11. https://doi.org/10.1038/s42004-024-01130-5.

(14) Thorpe, T. W.; Marshall, J. R.; Turner, N. J. Multifunctional Biocatalysts for Organic Synthesis. *J Am Chem Soc* **2024**. https://doi.org/10.1021/JACS.3C09542.

(15) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angewandte Chemie International Edition* **2005**, *44* (27), 4192–4196. https://doi.org/10.1002/ANIE.200500767.

(16) Andorfer, M. C.; Park, H. J.; Vergara-Coll, J.; Lewis, J. C. Directed Evolution of RebH for Catalyst-Controlled Halogenation of Indole C–H Bonds. *Chem Sci* **2016**, *7* (6), 3720–3729. https://doi.org/10.1039/C5SC04680G.

(17) Reetz, M. T. Controlling the Enantioselectivity of Enzymes by Directed Evolution: Practical and Theoretical Ramifications. *Proc Natl Acad Sci U S A* **2004**, *101* (16), 5716–5722. https://doi.org/10.1073/PNAS.0306866101/ASSET/78F6A5F4-076A-4E3B-A072-9441F5AA24B8/ASSETS/GRAPHIC/ZPQ0140444650009.JPEG.

(18) Gu, J.; Xu, Y.; Nie, Y. Role of Distal Sites in Enzyme Engineering. *Biotechnol Adv* **2023**, *63*, 108094. https://doi.org/10.1016/J.BIOTECHADV.2023.108094.

(19) Karamitros, C. S.; Murray, K.; Winemiller, B.; Lamb, C.; Stone, E. M.; D'arcy, S.; Johnson, K. A.; Georgiou, G. Leveraging Intrinsic Flexibility to Engineer Enhanced Enzyme Catalytic Activity. *330 MEMORIAL LIBRARY* **2022**. https://doi.org/10.1073/pnas.

(20) Boville, C. E.; Romney, D. K.; Almhjell, P. J.; Sieben, M.; Arnold, F. H. Improved Synthesis of 4-Cyanotryptophan and Other Tryptophan Analogues in Aqueous Solvent Using Variants of TrpB from Thermotoga Maritima. *Journal of Organic Chemistry* **2018**. https://doi.org/10.1021/acs.joc.8b00517.

(21) Murciano-Calles, J.; Romney, D. K.; Brinkmann-Chen, S.; Buller, A. R.; Arnold, F. H. A Panel of TrpB Biocatalysts Derived from Tryptophan Synthase through the Transfer of Mutations That Mimic Allosteric Activation. *Angewandte Chemie - International Edition* **2016**. https://doi.org/10.1002/anie.201606242.

(22) Sarai, N. S.; Fulton, T. J.; O'Meara, R. L.; Johnston, K. E.; Brinkmann-Chen, S.; Maar, R. R.; Tecklenburg, R. E.; Roberts, J. M.; Reddel, J. C. T.; Katsoulis, D. E.; Arnold, F. H. Directed Evolution of Enzymatic Silicon-Carbon Bond Cleavage in Siloxanes. *Science (1979)* **2024**, *383* (6681), 438–443. https://doi.org/10.1126/SCIENCE.ADI5554/SUPPL_FILE/SCIENCE.ADI5554_MDAR_REPRODUCIBILITY_CHECKLIST.PDF.

(23) Romney, D. K.; Murciano-Calles, J.; Wehrmüller, J. E.; Arnold, F. H. Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *J Am Chem Soc* **2017**. https://doi.org/10.1021/jacs.7b05007.

(24) Wetzl, D.; Gand, M.; Ross, A.; Müller, H.; Matzel, P.; Hanlon, S. P.; Müller, M.; Wirz, B.; Höhne, M.; Iding, H. Asymmetric Reductive Amination of Ketones Catalyzed by Imine Reductases. *ChemCatChem* **2016**, *8* (12), 2023–2026. https://doi.org/10.1002/cctc.201600384.

(25) Prier, C. K.; Camacho Soto, K.; Forstater, J. H.; Kuhl, N.; Kuethe, J. T.; Cheung-Lee, W. L.; Di Maso, M. J.; Eberle, C. M.; Grosser, S. T.; Ho, H. I.; Hoyt, E.; Maguire, A.; Maloney, K. M.; Makarewicz, A.; McMullen, J. P.; Moore, J. C.; Murphy, G. S.; Narsimhan, K.; Pan, W.; Rivera, N. R.; Saha-Shah, A.; Thaisrivongs, D. A.; Verma, D.; Wyatt, A.; Zewge, D. Amination of a Green Solvent via Immobilized Biocatalysis for the Synthesis of Nemtabrutinib. *ACS Catal* **2023**, *13* (12), 7707–7714. https://doi.org/10.1021/acscatal.3c00941.

(26) Romney, D. K.; Sarai, N. S.; Arnold, F. H. Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids. *ACS Catal* **2019**, *9* (9), 8726–8730. https://doi.org/10.1021/acscatal.9b02089.

(27) Blikstad, C.; Dahlström, K. M.; Salminen, T. A.; Widersten, M. Substrate Scope and Selectivity in Offspring to an Enzyme Subjected to Directed Evolution. *FEBS J* **2014**, *281* (10), 2387–2398. https://doi.org/10.1111/FEBS.12791.

(28) Matsumura, I.; Ellington, A. D. In Vitro Evolution of Beta-Glucuronidase into a Beta-Galactosidase Proceeds Through Non-Specific Intermediates. *J Mol Biol* **2001**, *305* (2), 331–339. https://doi.org/10.1006/JMBI.2000.4259.

(29) Crawshaw, R.; Crossley, A. E.; Johannissen, L.; Burke, A. J.; Hay, S.; Levy, C.; Baker, D.; Lovelock, S. L.; Green, A. P. Engineering an Efficient and Enantioselective Enzyme for the

Morita–Baylis–Hillman Reaction. *Nat Chem* **2022**, *14* (3), 313–320. https://doi.org/10.1038/s41557-021-00833-9.

(30) Almhjell, P. J.; Johnston, K. E.; Porter, N. J.; Kennemur, J. L.; Bhethanabotla, V. C.; Ducharme, J.; Arnold, F. H. The β-Subunit of Tryptophan Synthase Is a Latent Tyrosine Synthase. *Nat Chem Biol* **2024**. https://doi.org/10.1038/s41589-024-01619-z.

(31) Jensen, R. A. Enzyme Recruitment in Evolution of New Function. *Annu Rev Microbiol* **1976**, *30* (Volume 30,), 409–425. https://doi.org/10.1146/ANNUREV.MI.30.100176.002205/CITE/REFWORKS.

(32) Jones, B. S.; Ross, C. M.; Foley, G.; Pozhydaieva, N.; Sharratt, J. W.; Kress, N.; Seibt, L. S.; Thomson, R. E. S.; Gumulya, Y.; Hayes, M. A.; Gillam, E. M. J.; Flitsch, S. L. Engineering Biocatalysts for the C−H Activation of Fatty Acids by Ancestral Sequence Reconstruction**. *Angewandte Chemie International Edition* **2024**, *63* (18), e202314869. https://doi.org/10.1002/ANIE.202314869.

(33) Livada, J.; Vargas, A. M.; Martinez, C. A.; Lewis, R. D. Ancestral Sequence Reconstruction Enhances Gene Mining Efforts for Industrial Ene Reductases by Expanding Enzyme Panels with Thermostable Catalysts. *ACS Catal* **2023**, *13* (4), 2576–2585. https://doi.org/10.1021/ACSCATAL.2C03859/ASSET/IMAGES/LARGE/CS2C03859_0007.JPEG.

(34) Wilding, M.; Peat, T. S.; Kalyaanamoorthy, S.; Newman, J.; Scott, C.; Jermiin, L. S. Reverse Engineering: Transaminase Biocatalyst Development Using Ancestral Sequence Reconstruction. *Green Chemistry* **2017**, *19* (22), 5375–5380. https://doi.org/10.1039/C7GC02343J.

(35) Mcdonald, A. D.; Higgins, P. M.; Buller, A. R. Substrate Multiplexed Protein Engineering Facilitates Promiscuous Biocatalytic Synthesis. *Nat. Commun.* **2022**, *13*, 5242. https://doi.org/10.1038/s41467-022-32789-w.

(36) Zmich, A.; Perkins, L. J.; Bingman, C.; Acheson, J. F.; Buller, A. R. Multiplexed Assessment of Promiscuous Non-Canonical Amino Acid Synthase Activity in a Pyridoxal Phosphate-Dependent Protein Family. *ACS Catal* **2023**, 11644–11655. https://doi.org/10.1021/ACSCATAL.3C02498/SUPPL_FILE/CS3C02498_SI_002.CIF.

(37) Jakoblinnert, A.; Wachtmeister, J.; Schukur, L.; Shivange, A. V.; Bocola, M.; Ansorge-Schumacher, M. B.; Schwaneberg, U. Reengineered Carbonyl Reductase for Reducing Methyl-Substituted Cyclohexanones. *Protein Eng Des Sel* **2013**, *26* (4), 291–298. https://doi.org/10.1093/PROTEIN/GZT001.

(38) Junker, S.; Roldan, R.; Joosten, H. J.; Clapés, P.; Fessner, W. D. Complete Switch of Reaction Specificity of an Aldolase by Directed Evolution In Vitro: Synthesis of Generic Aliphatic Aldol Products. *Angewandte Chemie International Edition* **2018**, *57* (32), 10153–10157. https://doi.org/10.1002/ANIE.201804831.

(39) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J. Simultaneous Screening of Multiple Substrates with an Unspecific Peroxygenase

Enabled Modified Alkane and Alkene Oxyfunctionalisations. *Catal Sci Technol* **2021**, *11* (18), 6058–6064. https://doi.org/10.1039/D0CY02457K.

(40) Stanišić, A.; Hüsken, A.; Kries, H. HAMA: A Multiplexed LC-MS/MS Assay for Specificity Profiling of Adenylate-Forming Enzymes. *Chem Sci* **2019**, *10* (44), 10395–10399. https://doi.org/10.1039/C9SC04222A.

(41) McDonald, A. D.; Bruffy, S. K.; Kasat, A. T.; Buller, A. R. Engineering Enzyme Substrate Scope Complementarity for Promiscuous Cascade Synthesis of 1,2-Amino Alcohols. *Angewandte Chemie - International Edition* **2022**, *61* (46). https://doi.org/10.1002/anie.202212637.

(42) Steinberg, B.; Ostermeier, M. Environmental Changes Bridge Evolutionary Valleys. *Sci Adv* **2016**, *2* (1). https://doi.org/10.1126/SCIADV.1500921.

(43) Amitai, G.; Gupta, R. D.; Tawfik, D. S. Latent Evolutionary Potentials under the Neutral Mutational Drift of an Enzyme. *HFSP J* **2007**, *1* (1), 67–78. https://doi.org/10.2976/1.2739115/10.2976/1.

(44) Aharoni, A.; Gaidukov, L.; Khersonsky, O.; Gould, S. M. Q.; Roodveldt, C.; Tawfik, D. S. The "evolvability" of Promiscuous Protein Functions. *Nature Genetics 2004 37:1* **2004**, *37* (1), 73–76. https://doi.org/10.1038/ng1482.

(45) Markin, C. J.; Mokhtari, D. A.; Sunden, F.; Appel, M. J.; Akiva, E.; Longwell, S. A.; Sabatti, C.; Herschlag, D.; Fordyce, P. M. Revealing Enzyme Functional Architecture via High-Throughput Microfluidic Enzyme Kinetics. *Science (1979)* **2021**, *373* (6553). https://doi.org/10.1126/SCIENCE.ABF8761.

(46) Rix, G.; Watkins-Dulaney, E. J.; Almhjell, P. J.; Boville, C. E.; Arnold, F. H.; Liu, C. C. Scalable Continuous Evolution for the Generation of Diverse Enzyme Variants Encompassing Promiscuous Activities. *Nature Communications 2020 11:1* **2020**, *11* (1), 1–11. https://doi.org/10.1038/s41467-020-19539-6.

(47) Jones, K. A.; Snodgrass, H. M.; Belsare, K.; Dickinson, B. C.; Lewis, J. C. Phage-Assisted Continuous Evolution and Selection of Enzymes for Chemical Synthesis. *ACS Cent Sci* **2021**, *7* (9), 1581–1590. https://doi.org/10.1021/acscentsci.1c00811.

(48) Obexer, R.; Pott, M.; Zeymer, C.; Griffiths, A. D.; Hilvert, D. Efficient Laboratory Evolution of Computationally Designed Enzymes with Low Starting Activities Using Fluorescence-Activated Droplet Sorting. *Protein Engineering, Design and Selection* **2016**, *29* (9), 355–366. https://doi.org/10.1093/PROTEIN/GZW032.

(49) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angewandte Chemie - International Edition* **2005**, *44* (27), 4192–4196. https://doi.org/10.1002/anie.200500767.

(50) Laurent, V.; Gourbeyre, L.; Uzel, A.; Hélaine, V.; Nauton, L.; Traïkia, M.; De Berardinis, V.; Salanoubat, M.; Gefflaut, T.; Lemaire, M.; Guérard-Hélaine, C. Pyruvate Aldolases

Catalyze Cross-Aldol Reactions between Ketones: Highly Selective Access to Multi-Functionalized Tertiary Alcohols. *ACS Catal* **2020**, *10* (4), 2538–2543. https://doi.org/10.1021/ACSCATAL.9B05512/SUPPL_FILE/CS9B05512_SI_001.PDF.

(51) Laurent, V.; Darii, E.; Aujon, A.; Debacker, M.; ean-Louis Petit, J.; HØlaine, V.; ibor Liptaj, T.; Breza, M.; Mariage, A.; Nauton, L.; Traïkia, M.; Salanoubat, M.; arielle Lemaire, M.; hristine GuØrard-HØlaine, C.; de Berardinis, Vø. Synthesis of Branched-Chain Sugars with a DHAP-Dependent Aldolase: Ketones Are Electrophile Substrates of Rhamnulose-1-Phosphate Aldolases. *Angewandte Chemie* **2018**, *130* (19), 5565–5569. https://doi.org/10.1002/ANGE.201712851.

(52) Liu, Z. Q.; Xiang, Z. W.; Shen, Z.; Wu, Q.; Lin, X. F. Enzymatic Enantioselective Aldol Reactions of Isatin Derivatives with Cyclic Ketones under Solvent-Free Conditions. *Biochimie* **2014**, *101* (1), 156–160. https://doi.org/10.1016/J.BIOCHI.2014.01.006.

(53) Ye, Y.; Minami, A.; Igarashi, Y.; Izumikawa, M.; Umemura, M.; Nagano, N.; Machida, M.; Kawahara, T.; Shin-ya, K.; Gomi, K.; Oikawa, H. Unveiling the Biosynthetic Pathway of the Ribosomally Synthesized and Post-Translationally Modified Peptide Ustiloxin B in Filamentous Fungi. *Angewandte Chemie International Edition* **2016**, *55* (28), 8072–8075. https://doi.org/10.1002/anie.201602611.

(54) Umemura, M.; Nagano, N.; Koike, H.; Kawano, J.; Ishii, T.; Miyamura, Y.; Kikuchi, M.; Tamano, K.; Yu, J.; Shin-ya, K.; Machida, M. Characterization of the Biosynthetic Gene Cluster for the Ribosomally Synthesized Cyclic Peptide Ustiloxin B in Aspergillus Flavus. *Fungal Genetics and Biology* **2014**, *68*, 23–30. https://doi.org/10.1016/j.fgb.2014.04.011.

(55) Zhang, R.; Tan, J.; Luo, Z.; Dong, H.; Ma, N.; Liao, C. Stereo-Selective Synthesis of Non-Canonical γ-Hydroxy-α-Amino Acids by Enzymatic Carbon–Carbon Bond Formation. *Catal Sci Technol* **2021**, *11* (22), 7380–7385. https://doi.org/10.1039/D1CY00955A.

(56) Fujimori, D. G.; Hrvatin, S.; Neumann, C. S.; Strieker, M.; Marahiel, M. A.; Walsh, C. T. Cloning and Characterization of the Biosynthetic Gene Cluster for Kutznerides. *Proc Natl Acad Sci U S A* **2007**, *104* (42), 16498–16503. https://doi.org/10.1073/PNAS.0708242104/SUPPL_FILE/08242FIG4.JPG.

(57) Morokuma, K.; Watari, H.; Mori, M.; Sakai, R.; Irie, R.; Oikawa, M. Total Synthesis of Lycoperdic Acid and Stereoisomers (Part 2): Completion of the Synthesis and Biological Evaluation. *Tetrahedron* **2023**, *145*, 133623. https://doi.org/10.1016/J.TET.2023.133623.

(58) Maharaj, V. J.; Moodley, N.; Vahrmeijer, H. Characterization of Natural Monatin Isomers, a High Intensity Sweetener from the Plant Sclerochiton Ilicifolius from South Africa. *South African Journal of Botany* **2018**, *115*, 37–43. https://doi.org/10.1016/J.SAJB.2017.12.007.

(59) Zhang, R.; Zhang, C.; Tan, J.; He, Y.; Zhuo, D.; Zhang, J.; Luo, Z.; Li, Q.; Yao, J.; Ke, C.; Tang, C.; Ye, Y.; He, S.; Sheng, X.; Liao, C. Enzymatic Synthesis of Noncanonical α-Amino Acids Containing γ-Tertiary Alcohols. *Angewandte Chemie International Edition* **2024**, *63* (7), e202318550. https://doi.org/10.1002/ANIE.202318550.

(60) Li, Z.; Jangra, H.; Chen, Q.; Mayer, P.; Ofial, A. R.; Zipse, H.; Mayr, H. Kinetics and Mechanism of Oxirane Formation by Darzens Condensation of Ketones: Quantification of the Electrophilicities of Ketones. *J Am Chem Soc* **2018**, *140* (16), 5500–5515. https://doi.org/10.1021/JACS.8B01657/SUPPL_FILE/JA8B01657_SI_003.CIF.

(61) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; Murdzek, J., Ed.; AIP Publishing, 2006.

(62) Weeks, A. M.; Wells, J. A. Engineering Peptide Ligase Specificity by Proteomic Identification of Ligation Sites. *Nature Chemical Biology 2017 14:1* **2017**, *14* (1), 50–57. https://doi.org/10.1038/nchembio.2521.

(63) Nath, A.; Atkins, W. M. A Quantitative Index of Substrate Promiscuity. *Biochemistry* **2008**, *47* (1), 157–166. https://doi.org/10.1021/BI701448P/SUPPL_FILE/BI701448P-FILE007.PDF.

(64) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angewandte Chemie International Edition* **2005**, *44* (27), 4192–4196. https://doi.org/10.1002/ANIE.200500767.

(65) Reetz, M. T.; Carballeira, J. D. Iterative Saturation Mutagenesis (ISM) for Rapid Directed Evolution of Functional Enzymes. *Nature Protocols 2007 2:4* **2007**, *2* (4), 891–903. https://doi.org/10.1038/nprot.2007.72.

(66) Stemmer, W. P. C. Rapid Evolution of a Protein in Vitro by DNA Shuffling. *Nature 1994 370:6488* **1994**, *370* (6488), 389–391. https://doi.org/10.1038/370389a0.

(67) Voigt, C. A.; Martinez, C.; Wang, Z. G.; Mayo, S. L.; Arnold, F. H. Protein Building Blocks Preserved by Recombination. *Nature Structural Biology 2002 9:7* **2002**, *9* (7), 553–558. https://doi.org/10.1038/nsb805.

(68) Zhao, H.; Giver, L.; Shao, Z.; Affholter', J. A.; Arnold, F. H. *Molecular Evolution by Staggered Extension Process (StEP) in Vitro Recotnbination*; 1998. http://www.nature.com/naturebiotechnology.

(69) Villalona, J.; Higgins, P. M.; Buller, A. R. Engineered Biocatalytic Synthesis of β-N-Substituted-α-Amino Acids. *Angewandte Chemie - International Edition* **2023**, *62* (43). https://doi.org/10.1002/anie.202311189.

(70) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. **2010**. https://doi.org/10.1126/science.1188934.

(71) Bhushan, R.; Brückner, H. Marfey's Reagent for Chiral Amino Acid Analysis: A Review. *Amino Acids* **2004**, *27* (3–4), 231–247. https://doi.org/10.1007/s00726-004-0118-0.

(72) Marfey, P. Determination of D-Amino Acids. II. Use of a Bifunctional Reagent, 1,5-Difluoro-2,4-Dinitrobenzene. *Carlsberg Res Commun* **1984**, *49* (6), 591–596. https://doi.org/10.1007/BF02908688.

(73) Faulstich, H.; Buku, A.; Bodenmüller, H.; Wieland, T. Virotoxins: Actin-Binding Cyclic Peptides of Amanita Virosa Mushrooms. *Biochemistry* **1980**, *19* (14), 334–343. https://doi.org/10.1021/BI00555A036.

(74) Weygand, F.; Mayer, F. Synthese von γ.δ.Δ'-Trihydroxy-L-Leucin. *Chem Ber* **1968**, *101* (6), 2065–2068. https://doi.org/10.1002/CBER.19681010623.

(75) Kaján, L.; Hopf, T. A.; Kalaš, M.; Marks, D. S.; Rost, B. FreeContact: Fast and Free Software for Protein Contact Prediction from Residue Co-Evolution. *BMC Bioinformatics* **2014**, *15* (1), 1–6. https://doi.org/10.1186/1471-2105-15-85/TABLES/2.

(76) Karamitros, C. S.; Murray, K.; Winemiller, B.; Lamb, C.; Stone, E. M.; D'Arcy, S.; Johnson, K. A.; Georgiou, G. Leveraging Intrinsic Flexibility to Engineer Enhanced Enzyme Catalytic Activity. *Proc Natl Acad Sci U S A* **2022**, *119* (23), e2118979119. https://doi.org/10.1073/PNAS.2118979119/SUPPL_FILE/PNAS.2118979119.SD01.CSV.

(77) Markin, C. J.; Mokhtari, D. A.; Sunden, F.; Appel, M. J.; Akiva, E.; Longwell, S. A.; Sabatti, C.; Herschlag, D.; Fordyce, P. M. Revealing Enzyme Functional Architecture via High-Throughput Microfluidic Enzyme Kinetics. *Science (1979)* **2021**, *373* (6553). https://doi.org/10.1126/SCIENCE.ABF8761/SUPPL_FILE/SCIENCE.ABF8761_SM.PDF.

(78) Romero, P. A.; Arnold, F. H. Exploring Protein Fitness Landscapes by Directed Evolution. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group December 2009, pp 866–876. https://doi.org/10.1038/nrm2805.

Figures

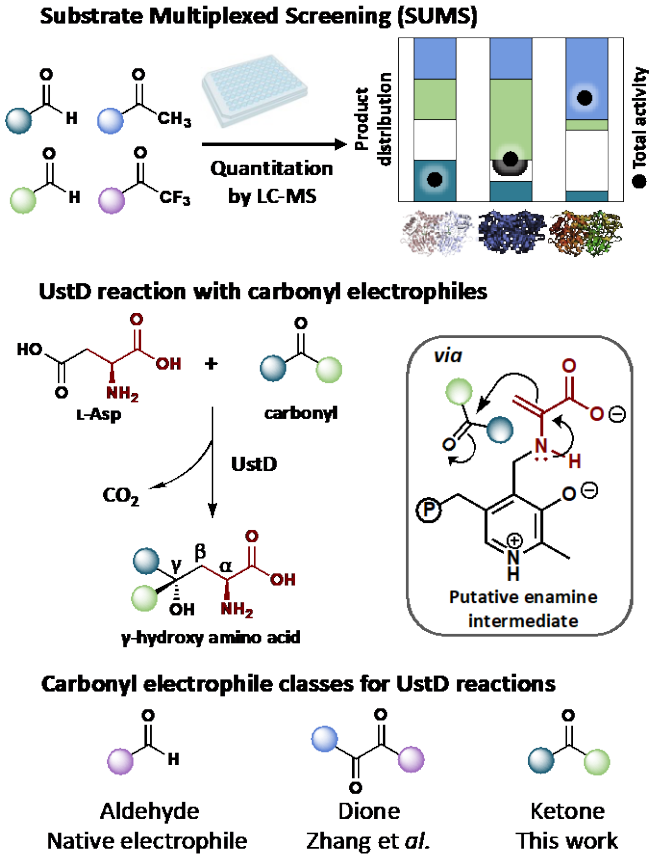

Figure 1. Application of SUMS. A. General SUMS procedure. **B.** UstD reaction scheme depicting the native reactivity with aldehyde electrophiles. **C.** The electrophile classes compatible with UstD reactions.

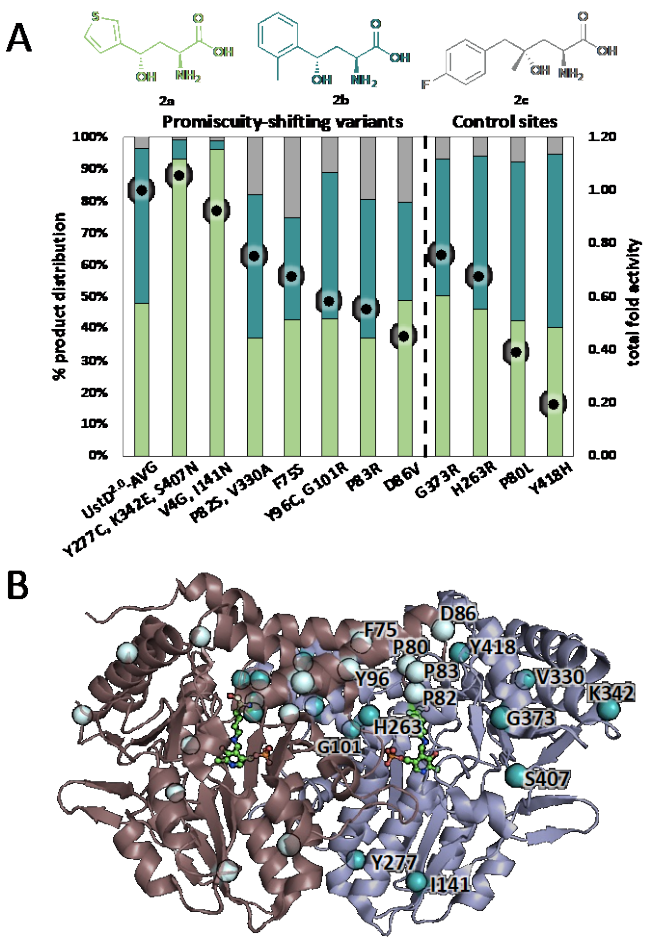

Figure 2. Distal putative allosteric sites identified by globally random mutagenesis. A. Combined results showing all promiscuity-shifting variants of interest from global random mutagenesis. Total activity for all products is represented by the dots. Product distribution is represented by the bars. The mutations found in each variant are displayed on the x-axis. **Conditions:** 50 mM L-asp, 4.2 mM **1a**, 4.2 mM **1b**, 41.6 mM **1c**, 50 μM PLP, 5% DMSO, 100 mM NaCl, 100 mM potassium phosphate pH 7.0, *E. coli* whole cells over-expressing UstD$^{2.0}$ variants, 37 °C, 200 rpm, 1 h reaction time. **B.** UstD$^{2.0}$ structure showing the location of the distal residues in each monomer. The teal spheres are located on chain B (purple) while the light cyan spheres are located on chain C (pink). The PLP cofactor (green) is shown in the internal aldimine form.

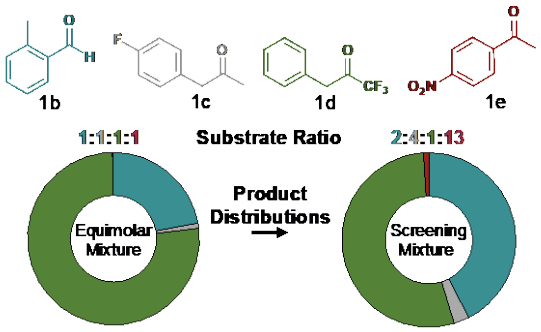

Figure 3. Substrate mixture design. Substrates included in the mixtures are shown above the plots. The plots show the amino acid product distribution from two different substrate mixtures with the color-coded ratio of substrates displayed above each. On the left, substrates are added in equal amounts (12.5 mM each). On the right, substrates are added in differing amounts according to their electrophilicity (5 mM **1b**, 10 mM **1c**, 2.5 mM **1d**, 32.5 mM **1e**). **Conditions:** 50 mM L-asp, 50 mM total electrophiles, 5 μM PLP, 5% DMSO, 100 mM NaCl, 100 mM potassium phosphate pH 7.0, QE (0.01 mol% catalyst, 10,000 Max TON), 37 °C, 1 h reaction time.

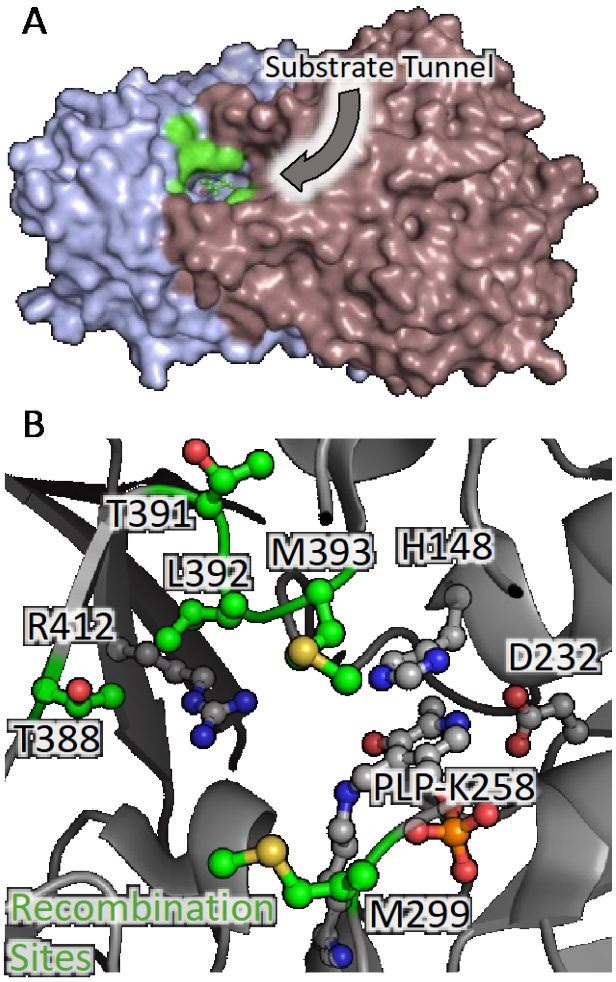

Figure 4. UstD active site targeted for mutagenesis. A. Space filling view of overall UstD structure. Individual monomers are colored purple (chain B) and pink (chain C) while the active site and PLP are shown in green. **B.** Close up view of UstD active site. The PLP complex and anchoring residues are shown in grey. The residues targeted for mutagenesis are shown in green.

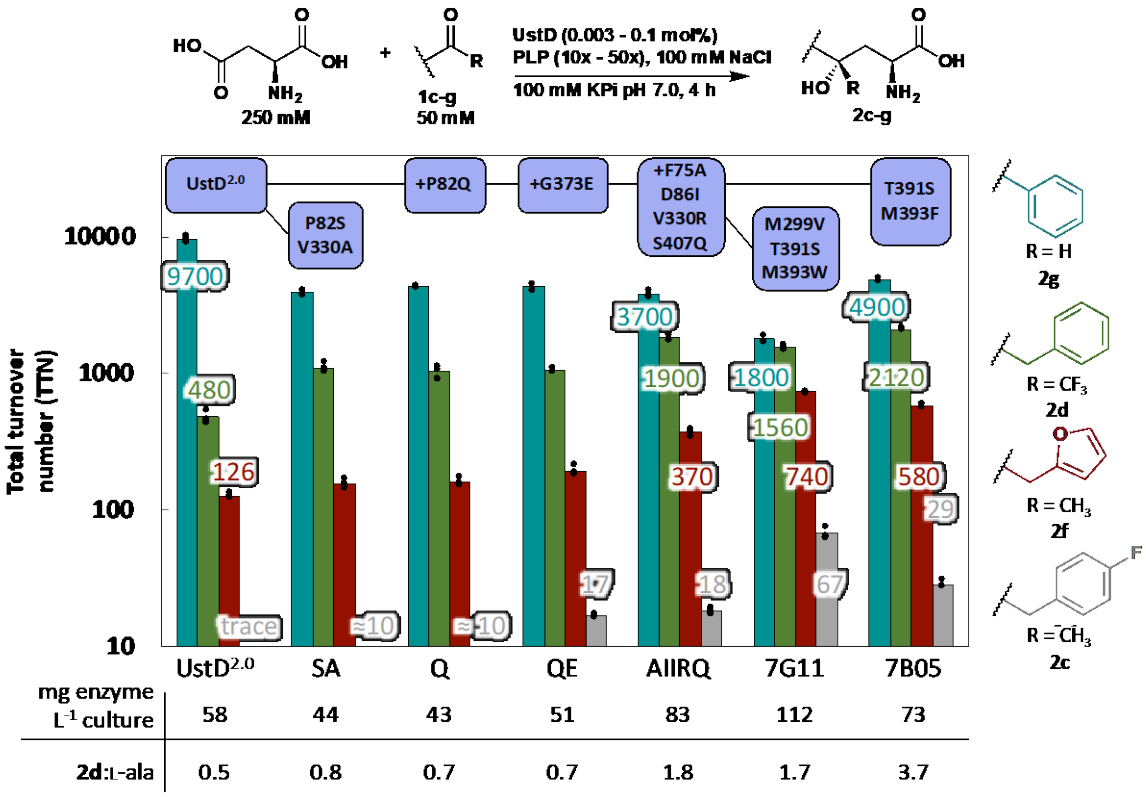

Figure 5. Lineage analysis of promiscuity guided evolution. Reactions were performed with single substrates in triplicate. The substrates are displayed to the right of the chart. The total turnover (TTN) number is displayed on the y-axis on a logarithmic scale. The bar represents the average of the technical replicates, and the dots represent the TTN measurement of each individual replicate. The color of each bar corresponds to its product. The mutations associated with each variant are displayed at the top of the graph with the connecting lines representing the relationship between them. Below the chart, the yield of protein per liter of culture is displayed for each variant. The ratio of **2d** to the shunt pathway product, L-ala, is displayed below the chart to demonstrate that the ratio changes over the lineage.

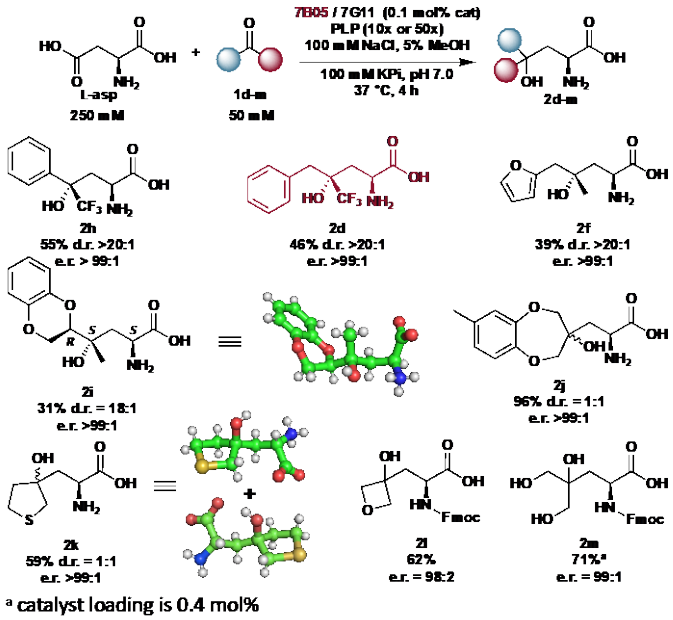

Figure 6. Preparative scale biocatalytic reactions.